Journal of proteome research

pubs.acs.org/jpr

Article

Molecular Simulations and Network Modeling Reveal an Allosteric Signaling in the SARS-CoV-2 Spike Proteins

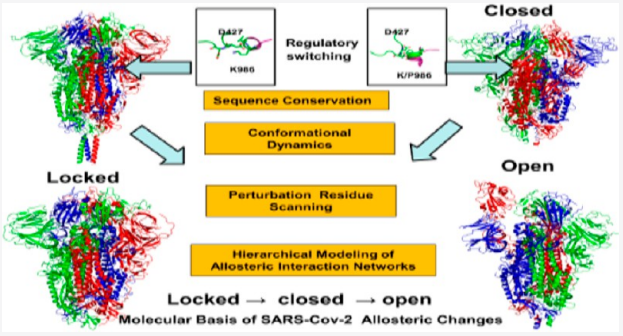
Gennady M. Verkhivker*

Cite This: J. Proteome Res. 2020, 19, 4587−4608

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The development of computational strategies for the quantitative characterization of the functional mechanisms of SARS-CoV-2 spike proteins is of paramount importance in efforts to accelerate the discovery of novel therapeutic agents and vaccines combating the COVID-19 pandemic. Structural and biophysical studies have recently characterized the conformational landscapes of the SARS-CoV-2 spike glycoproteins in the prefusion form, revealing a spectrum of stable and more dynamic states. By employing molecular simulations and network modeling approaches, this study systematically examined functional dynamics and identified the regulatory centers of allosteric interactions for distinct functional states of the wild-type and mutant variants of the SARS-CoV-2 prefusion spike trimer. This study presents evidence that the SARS-CoV-2 spike protein can function as an allosteric regulatory engine that fluctuates between dynamically distinct functional states. Perturbation-based modeling of the interaction networks revealed a key role of the cross-talk between the effector hotspots in the receptor binding domain and the fusion peptide proximal region of the SARS-CoV-2 spike protein. The results have shown that the allosteric hotspots of the interaction networks in the SARS-CoV-2 spike protein can control the dynamic switching between functional conformational states that are associated with virus entry to the host receptor. This study offers a useful and novel perspective on the underlying mechanisms of the SARS-CoV-2 spike protein through the lens of allosteric signaling as a regulatory apparatus of virus transmission that could open up opportunities for targeted allosteric drug discovery against SARS-CoV-2 proteins and contribute to the rapid response to the current and potential future pandemic scenarios.



KEYWORDS: *SARS-CoV-2 spike proteins, prefusion trimer, biophysical modeling, network analysis, allosteric signaling, regulatory hotspots, dynamic switching, allosteric drug discovery*

INTRODUCTION

The coronavirus SARS-CoV-2 is associated with acute respiratory distress syndrome[1,2] and is similar to the severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) viruses.[3] The genomic sequences of SARS-CoV-2 showed a high level of sequence similarity between the SARS-CoV-2, SARS, and MERS proteins involved in the replication cycle.[4−6] SARS-CoV-2 has four main structural proteins: spike (S) glycoprotein, small envelope (E) glycoprotein, membrane (M) glycoprotein, and nucleocapsid (N) protein, along with several accessory proteins.[7−9] Recent studies have identified that SARS-CoV-2 uses angiotensin-converting enzyme 2 (ACE2)[10−12] and the cellular protease transmembrane protease serine 2 (TMPRSS2) as cell entry receptors.[10] SARS spike proteins are trimeric fusion proteins with two main domains, an amino (N)-terminal S1 subunit and a carboxyl (C)-terminal S2 subunit. SARS-CoV-2 binds to the ACE2 receptor on the surface of the host cell using binding of the S1 region of the virus spike (S) protein followed by the fusion of the viral and cellular membranes mediated by the S2 subunit of the S protein.[13] The interaction between S and ACE2 results in fusion of the viral and cell membranes, releasing the viral genome and allowing for viral replication and host cell takeover. The crystal structure of the receptor-binding domain (RBD) of SARS-CoV-2 S protein bound to human ACE2 showed a common binding mode shared with SARS-CoV-S protein, and yet the SARS-CoV-2-RBD binding affinity is four times stronger.[14,15] Binding to the ACE2 receptor is a critical initial step for SARS-CoV to enter into target cells, and the increased infectiousness of SARS-CoV-2 is believed to arise from subtle structural and dynamic alterations in the RBD regions of the SARS-CoV2-S protein, which can increase the binding affinity for ACE2 as compared with SARS-CoV-S.

Special Issue: Proteomics in Pandemic Disease

Received: August 22, 2020
Published: October 2, 2020

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

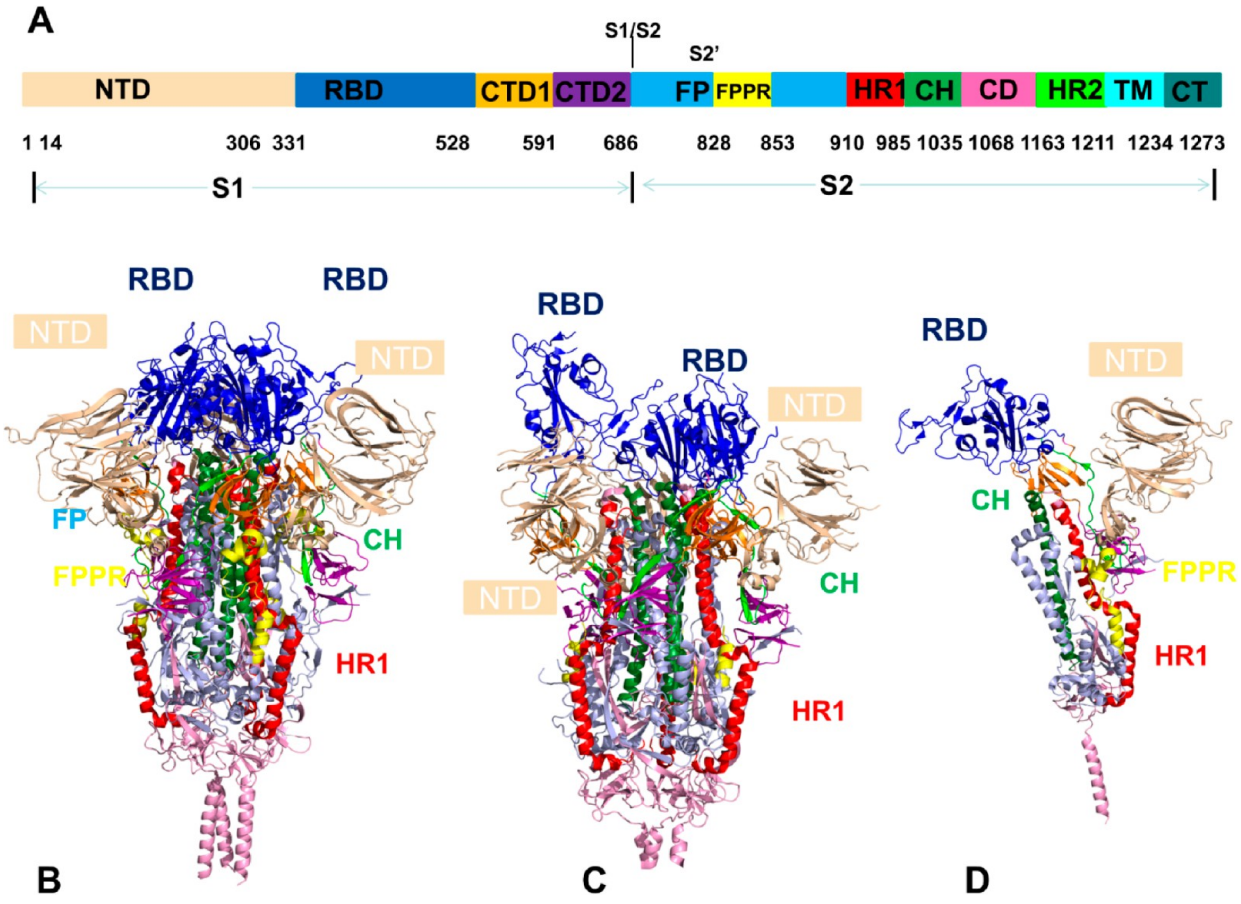


Figure 1. Domain and structural organization of the full-length SARS-CoV-2 spike (S) protein. (A) Schematic representation of domain organization and residue range for the full-length SARS-CoV-2 spike (S) protein. The subunits S1 and S2 include the N-terminal domain (NTD), receptor-binding domain (RBD), C-terminal domain 1(CTD1), C-terminal domain 2 (CTD2), S1/S2 cleavage site (S1/S2), S2′ cleavage site (S2′), fusion peptide (FP), fusion peptide proximal region (FPPR), heptad repeat 1 (HR1), central helix region (CH), connector domain (CD), heptad repeat 2 (HR2), transmembrane anchor (TM), and cytoplasmic tail (CT). (B) Structure of the wild-type full-length SARS-CoV-2 spike (S) protein trimer in the locked closed ("RBD-down") prefusion conformation (PDB ID 6XR8). The structure of the S protein is shown as ribbons, and structural components of the trimer are annotated and colored according to the color scheme shown in panel A. The annotated structural components include the NTD, RBD, FP, FPPR, HR1, and CH. (C) Structure of the SARS-CoV-2 spike (S) protein trimer in the open ("RBD-up") prefusion conformation state (PDB ID 6VYB) and the cryo-EM structure of the SARS-CoV-2 spike trimer (K986P/V987P) in the open state (PDB ID 6VSB). (D) Monomer form of the SARS-CoV-2 spike (S) protein trimer in the prefusion conformation state. Structural components of the trimer are annotated and colored according to the color scheme shown in panel A.

Strikingly, however, the full-length SARS-CoV and SARS-CoV-2 S proteins bind ACE2 with comparable affinities, which was attributed to the less-exposed and semioptimal orientation of the SARS-CoV-2-RBD regions to interactions with ACE2 within the full-length protein.[16] The high-resolution crystal structures of the RBD of the spike protein for SARS-CoV[17−19] and SARS-CoV-2 bound to the cell receptor ACE2[20,21] have identified residues and two major hotspots in the RBD that may be important for the stronger binding of SARS-CoV-2 with ACE2 binding. Despite the sequence and structure similarities between these complexes, a number of residues in the highly variable receptor binding motif (RBM) region are different between SARS-CoV-2 RBD and SARS-CoV RBD.[17−21] It was argued that although the binding affinity changes are important, these differences may not be sufficient to adequately explain the unusual transmissibility of the SARS-CoV-2 virus to the host receptor protein.[20,21]

The full-length SARS-CoV-2 spike (S) protein includes the N-terminal domain (NTD), RBD, C-terminal domain 1 (CTD1), C-terminal domain 2 (CTD2), S1/S2 cleavage site, S2′ cleavage site, fusion peptide (FP), fusion peptide proximal region (FPPR), heptad repeat 1 (HR1), central helix region (CH), connector domain (CD), heptad repeat 2 (HR2), transmembrane anchor, and cytoplasmic tail (CT) (Figure 1). The molecular basis of the virus transmission implies that the process starts with the binding of the spike protein to ACE2, which leads to a conformational change in the S1 and S2 domains of the spike protein and the exposure of the "fusion peptide" of the S2 domain, mediating the fusion of the viral and host cell membranes. The cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins revealed that the spike protein trimer can exist in a dynamic equilibrium between a metastable closed ("RBD-down") prefusion conformation that undergoes large structural rearrangements and hinge-based functional motions to adapt a receptor-accessible open ("RBD-up") state in which the S protein can fuse the viral membrane with the host cell membrane.[22−24]

The cryo-EM structures of the SARS-CoV-2 spike ectodomain trimer in two distinct conformations detailed the conformational changes of the SARS-CoV-2 S trimer and characterized the receptor-accessible epitopes, confirming that SARS-CoV-2 and SARS-CoV bind with similar affinities to

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

human ACE2.[25] The recent cryo-EM structure of the SARS-CoV-2 trimeric spike in the prefusion conformation reconstructed an asymmetrical trimer with a single RBD in the optimal "up" conformation in which SARS-CoV-2 trimer binds ∼10 times stronger to the ACE2 host cell receptor as compared with SARS-CoV protein.[26] The cryo-EM structures of the full-length human ACE2 in the presence of the neutral amino acid transporter B^0AT1 with or without the RBD of the SARS-CoV-2 unveiled a complex structural arrangement with the concurrent binding of two S trimers to an ACE2 homodimer.[27] The two cryo-EM structures of the full-length S protein in the prefusion and postfusion conformations revealed the structural basis of the spontaneous transition to the postfusion state, discovering for the first time an ordered conformation of the ∼25-residue FPPR segment that is disordered in other structures and demonstrating that RBDs of the prefusion trimer can be locked in the closed form, arguably to allow long-range allosteric couplings to emerge between the stabilized RBD and FPPR regions.[28]

It was conjectured that allosteric cross-talk between the RBD and the FPPR regions, which may be mediated by the CTD1 segment, can regulate allosteric functional transformations between closed and host receptor-accessible open RBD forms, as the presence of a structured FPPR can restrict movements and lock down the RBD segment, whereas a disordered FPPR segment may release constraints on the RBD movements and promote the transition to the open "up" form.[28] A series of structural and biophysical studies employed protein engineering approaches to generate a prefusion-stabilized SARS-CoV-2 S protein by targeting regions that dramatically change between the prefusion and postfusion conformations as well as regions involved in the transitions between the open and closed prefusion states.[29] Through the design of disulfide bonds, proline mutations to curtail the mobility of flexible loops, and the introduction of hydrophobic modifications to fill empty voids in the prefusion structure, this study discovered a SARS-CoV-2 S variant with several proline substitutions that displayed ∼10 times higher expression than the original construct.[29] The cryo-EM structure of this engineered mutant retained the topology and molecular details of the prefusion trimer conformation.[29] By using structure-based vector analysis of spike glycoprotein structures, two soluble ectodomain constructs were designed for the SARS-CoV-2 S protein, in which the mobile RBD segment is either is locked in the "down" position or adopts "up" state conformations more readily than the wild-type S protein.[30] Through targeted mutagenesis design, negative-stain electron microscopy (NSEM), and cryo-EM structure determination, this study revealed a remarkable variability in the S trimer and the differential stabilization of the mutant ectodomain constructs, showing that the conformational equilibrium of the SARS-CoV-2 S trimer can be manipulated using modifications of contact regions between the RBD and S2 (S383C/D985C) and the RBD and NTD (D398L/S514L/E516L) and the interfaces between the SD1 and S2 (N866I/A570L).[30]

This pioneering study demonstrated that whereas the overall topological arrangements of the virus spike trimers are conserved, subtle differences in the interdomain contacts could play a major role in modulating the surface antigenicity and the equilibrium partitioning between the closed and open states. The cryo-EM analysis of the engineered SARS-CoV-2 S ectodomain trimer with several intermolecular disulfide bonds demonstrated that additional stabilization factors can effectively lock the trimer in the closed state without altering its architecture, enhancing the SARS-CoV-2 S resistance to proteolysis.[31] Another illuminating biophysical investigation combined targeted design and mutagenesis with structure determination of thermostable SARS-CoV-2 spike trimers, showing the subtle heterogeneity of the SARS-CoV-2 landscape concealed by the conserved trimer topology in which disulfide-bonded S-protein trimers using pairs (S383, D385) and (G413, V987) can be trapped in the structurally distinct locked and closed states of the prefusion form with only ∼20% of the population sampling one open state.[32] In the locked form, the S1 domains undergo a movement in which RBDs move closer to the three-fold axis and form a tightly packed trimer. At the same time, it was observed that for SARS-CoV-2 variants with proline substitutions K986P and V987C, ∼20% of the trimers are in the locked or closed state, and ∼80% of molecules adopt the open "up" trimer form. This study suggested that D427 and D428 in the RBD interact with the positively charged K986, which electrostatically stabilizes the closed form. The loss of these electrostatic contacts could act as a potential regulatory switch of the spike dynamic equilibrium by promoting a large allosteric change and shifting the population toward the open trimer.[32] The cryo-EM structures of the postfusion SARS-CoV S2 trimer showed dramatic conformational changes of the SARS-CoV S2 machinery during membrane fusion, where the HR1 and the CH helical segments can form a long continuous stem helix pointing toward the target membrane in the postfusion trimer.[33]

The shielding of receptor binding sites by glycans is a common feature of viral glycoproteins, as observed on SARS-CoV proteins, where glycosylation can camouflage immunogenic protein epitopes.[34,35] Site-specific glycan mass spectrometric analysis of the SARS-CoV-2 spike further enabled mapping of the glycan-processing states,[36] revealing the structural positions of 22 glycans on the SARS-CoV-2 S protein and highlighting vulnerabilities in the RBD regions on the exposed surfaces of the trimer. The generated computational models of glycoforms of the SARS-CoV-2 spike protein and molecular dynamics (MD) simulations determined the extent to which glycan heterogeneity can affect the antigenicity of the S glycoprotein.[37] This computational study detailed at the molecular level that the protein surface is extensively shielded from antibody recognition by glycans, with the exception of the RBD regions that present points of vulnerabilities for the glycan shield.[37] The development and all-atom modeling of a fully glycosylated full-length SARS-CoV-2 S protein in a viral membrane were recently reported, representing a significant technical breakthrough in enabling rigorous simulations of the SARS-CoV-2 S trimer structures in a glycosylated environment (Figure S1).[38] MD simulations have been instrumental in predicting the shapes and motions of glycans for glycoproteins.[39−41] The unprecedented level of atomistic detail of the all-atom MD simulations of the full-length SARS-CoV-2 S glycoprotein with a complete glycosylation profile established that the dynamics of the glycan shield can be allosterically coupled to the conformational changes, dictating the heterogeneity of the conformational landscape and the response to the host receptor.[42] MD simulations of the SARS-CoV-2 spike glycoprotein identified the changes in the molecular properties due to conformational flexibility.[43] Network-based computational methods were developed to identify allosteric sites on the SARS-CoV-2 spike protein.[44] Several computational studies examined SARS-CoV-2 interactions with the ACE2 enzyme.[45−47] MD simulations also elucidated the determinants of the higher affinity of SARS-CoV-2 with ACE2, highlighting

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

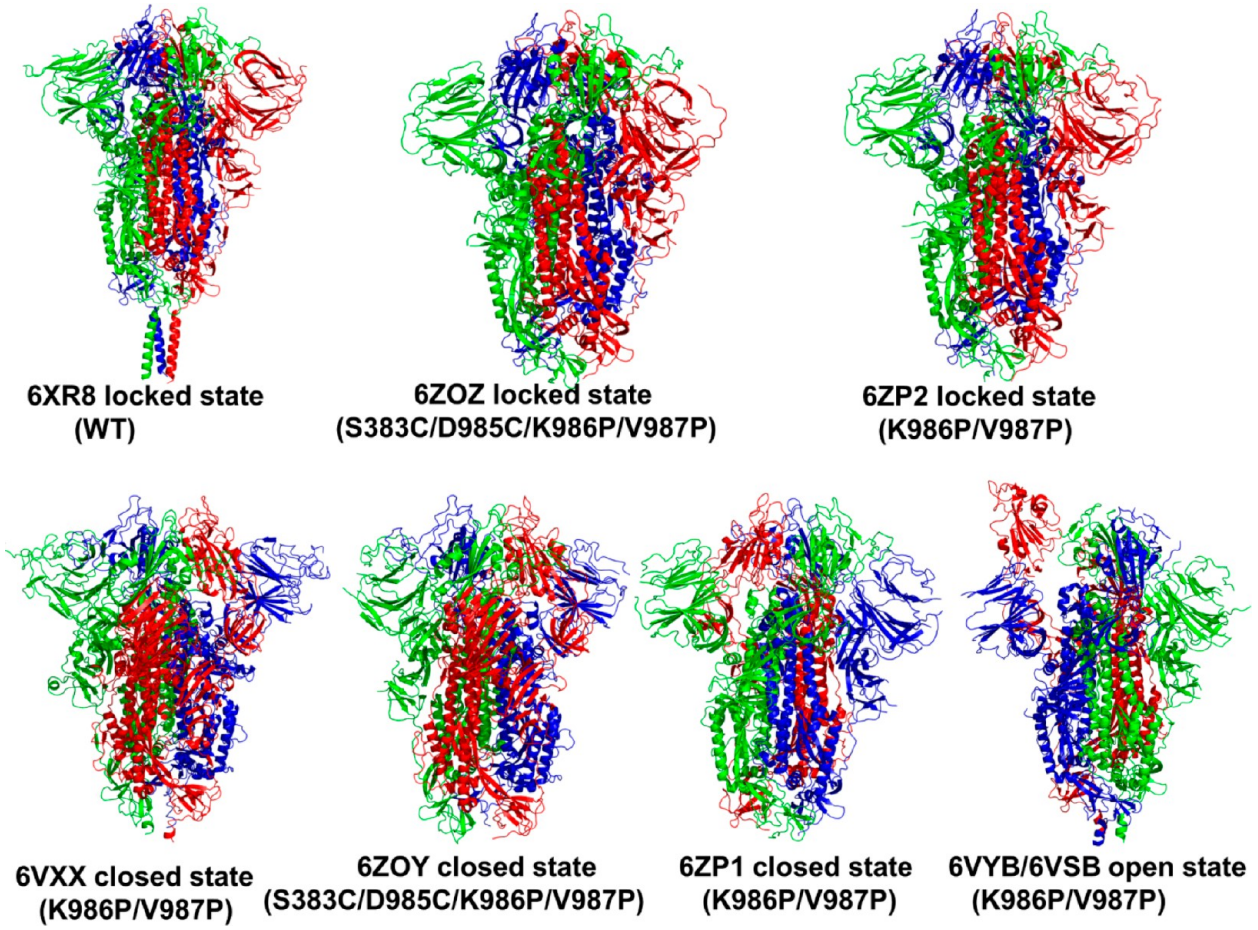


Figure 2. Cryo-EM structures of the SARS-CoV-2 spike (S) protein trimer in the locked and closed states of the prefusion form used in this study. Upper Panel: The following structures of the SARS-CoV-2 spike (S) protein trimer in the locked form were used in the simulations and modeling: the cryo-EM structure of the wild-type full-length S protein in the locked prefusion conformation (PDB ID 6XR8),[28] the cryo-EM structure of the disulfide-stabilized SARS-CoV-2 spike trimer (S383C, D985C, K986P, V987P, single R S1/S2 cleavage site) in the locked state (PDB ID 6ZOZ),[32] and the cryo-EM structure of the SARS-CoV-2 spike trimer in the locked state (K986P, V987P, single R S1/S2 cleavage site) (PDB ID 6ZP2).[32] Lower Panel: The following structures of the SARS-CoV-2 spike (S) protein trimer in the closed form were used for the simulations and modeling: the cryo-EM structure of the SARS-CoV-2 spike trimer in the closed state (K986P/V987P) (PDB ID 6VXX),[25] the cryo-EM structure of the disulfide-stabilized SARS-CoV-2 spike trimer (S383C, D985C, K986P, V987P, single R S1/S2 cleavage site) in the closed state (PDB ID 6ZOY),[32] and the cryo-EM structure of the SARS-CoV-2 spike trimer in the closed state (K986P, V987P, single R S1/S2 cleavage site) (PDB ID 6ZP1).[32] The following structures of the SARS-CoV-2 spike (S) protein trimer in the open form were subjected to simulations and modeling: the cryo-EM structure of the SARS-CoV-2 spike trimer ectodomain (K986P/V987P) in the open state (PDB ID 6VYB)[25] and the cryo-EM structure of the SARS-CoV-2 spike trimer (K986P/V987P) in the open state (PDB ID 6VSB).[26]

the role of the RBM motif[46] as well as role of the hydrophobic interactions and the elaborate hydrogen-bonding network in the SARS-CoV-2-RBD interface.[47]

Computational investigations have led to important methodological advances and an improved atomistic understanding of allosteric regulation in proteins.[48−51] The growing number of the SARS-CoV-2 S protein structures in different conformational states suggests that allosteric mechanisms may regulate functional transitions and the acquisition of host-accessible conformations for the SARS-CoV-2 S prefusion trimer. The current study introduces a hypothesis that the SARS-CoV-2 spike protein can function as an allosteric regulatory engine that fluctuates between dynamically distinct functional states and is controlled by allosteric switch centers that may determine the signal transmission and the regulation of virus entry. Coarse-grained (CG) simulations are employed together with perturbation-response scanning (PRS) of allosteric interactions and hierarchical network modeling to identify the salient signatures of allosteric interaction networks in the distinct conformational states of the SARS-CoV-2 spike trimers. The analysis of functional dynamics and the perturbation-response modeling identify the regulatory hotspots that drive transformations between locked, closed, and open SARS-CoV-2 spike forms. The results show that the stabilized locked form of the prefusion trimer is characterized by a broad allosteric network with key regulatory centers located in the RBD, FPPR, and HR regions that dictate the dynamic switching between conformational states of the SARS-CoV-2 spike trimer. This study offers a useful complementary perspective on the mechanisms underlying the conformational and dynamic changes in the SARS-CoV-2 spike trimer through the lens of functional dynamics and allosteric communications that enable us to identify the key residues and interactions responsible for the control of signal transmission and the regulation of virus entry to the host receptor.

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

MATERIALS AND METHODS

Sequence Conservation Analysis

Sequence conservation for SARS spike proteins was estimated using the ConSurf approach[52−56] by computing the residue-based conservation score profiles that measure evolutionary conservation. The low-score values are associated with the most conserved position in the protein. Multiple sequence alignment (MSA) was obtained using the MAFFT approach,[57] and homologues were obtained from UNIREF90.[58,59] MSA profiles were also obtained from the Pfam database of protein families (P59594, SPIKE_CVHSA, and CoV_S1_C, PF19209).[60−62]

Coarse-Grained Simulations and Elastic Network Models

CG models are computationally effective approaches that leverage a topology-based framework for the characterization of the protein structure and dynamics, enabling simulations of large systems over long time scales. In this study, the CABS CG model is used for simulations of the cryo-EM structures of the SARS-CoV-2 spike trimer in distinct conformational states (Table S1). This approach efficiently combines a CG model and Monte Carlo local moves, allowing the accurate reproduction of conformational ensembles of protein structures from all-atom MD simulations.[63−67] The sampling scheme of the CABS model employed for this study is based on replica-exchange simulations. The number of cycles during the CABS-CG simulations was set to 10 000, with the number of cycles between frames set to 100. Multiple CG simulations were performed for the distinct functional states of the SARS-CoV-2 S spike protein trimer (Figure 2). A total of 1 000 000 samples was generated for each system, and the total number of saved models in the trajectory used for the analysis was 10 000.

The structures of the SARS-CoV-2 spike proteins were obtained from the Protein Data Bank (PDB).[68,69] The recently reported fully glycosylated SARS-CoV-2 spike trimer atomistic models of the closed form (PDB ID 6VXX) and open forms (PDB ID 6VSB) (residues 1−1146)[38] were obtained from the CHARMM-GUI COVID-19 archive (http://www.charmm-gui.org/docs/archive/covid19) (Figure S1). The structure preparation process of other cryo-EM structures leveraged the developed full atomistic models for the closed and open states.[38] The missing residues were initially modeled based on these structures. Additionally, and for comparison, the missing loops in the cryo-EM structures were also reconstructed using template-based loop prediction approaches ModLoop,[70] ArchPRED server,[71] and DaReUS-Loop.[72]

The generated CABS-CG conformational ensembles were also subjected to all-atom reconstruction using the PULCHRA method[73] and the CG2AA tool.[74] The protein structures were then optimized using atomic-level energy minimization using the 3Drefine method.[75] The principal modes of protein motions are extracted from CG approaches and also using elastic network models (ENMs). In this work, the collective motions of the protein structures were also modeled using the Gaussian network model (GNM)[76,77] and the anisotropic network model (ANM) approaches[78] that are the most widely used ENM-based methods, computing the amplitudes of isotropic thermal motions and the directionality of anisotropic motions. The algorithmic details of this approach were discussed at length in related studies of molecular chaperones.[79] The essential mobility profiles were obtained using the DynOmics server[80] and the ANM server.[78]

Perturbation-Response Scanning

The PRS approach[81,82] estimates the residue response to external forces and was successfully used to locate allosteric hotspots and regulatory centers in various protein systems.[83−89] The PRS approach can also be efficiently combined with the dynamic analysis of residue interaction networks to identify and characterize allosteric hotspots and pathways of allosteric communications.[85−91] By monitoring the response to forces on the protein residues, the PRS approach can quantify allosteric couplings and link the directionality of the inserted force to the protein response in functional movements.[90,91] In the PRS approach, the $3N \times 3N$ Hessian matrix $\boldsymbol{H}$ of the second derivatives of the potential at the local minimum is computed. The $3N$-dimensional vector $\boldsymbol{\Delta R}$ of node displacements in response to the $3N$-dimensional perturbation force is evaluated according to Hooke's law $\boldsymbol{F} = \boldsymbol{H}^*\boldsymbol{\Delta R}$. A perturbation force is applied to each residue one at a time, and the response of the protein is estimated using the displacement vector $\Delta \boldsymbol{R}(\boldsymbol{i}) = \boldsymbol{H}^{-1}\boldsymbol{F}^{(i)}$. The deviations of each residue in the protein from the mean structure of the equilibrium ensemble and the covariance matrix $\Delta \boldsymbol{R}\Delta \boldsymbol{R}^T$ were calculated.

Each residue in the SARS-CoV-2 spike structures is perturbed one at a time by applying random forces distributed along a sphere of randomly selected directions. Using the residue displacements obtained from external force perturbations, the magnitude of the response of residue k is computed as $\langle |\Delta \boldsymbol{R}_k^{(i)}|^2 \rangle$ averaged over multiple perturbation forces $\boldsymbol{F}^i$, yielding the ikth element of the $N \times N$ PRS matrix. A measure of the response of residue k is the magnitude $\langle |\Delta \boldsymbol{R}_k^{(i)}|^2 \rangle$ of the kth block of $\Delta \boldsymbol{R}^i$ averaged over multiple $\mathbf{F}^{(i)}$, expressed as the ikth element of the $N \times N$ PRS matrix, $\mathbf{S}_{\mathrm{PRS}}$. The effector profile $\langle (\Delta \boldsymbol{R}^i)^2 \rangle_{\mathrm{effector}}$ represents the average effect of the perturbed effector site i on all other residues and is computed by averaging over all sensor residues j. Similarly, the jth column of the PRS matrix $\overline{\mathrm{S}}_{\mathrm{PRS}}$ represents the sensitivity profile of sensor residue j in response to perturbations of all residues, and its average is denoted as $\langle (\Delta \boldsymbol{R}^i)^2 \rangle_{\mathrm{sensor}}$.

Modeling and Community Analysis of Residue Interaction Networks

The residue interaction networks are represented as graphs with residues as network nodes and the inter-residue edges describing the inter-residue interactions.[92−94] The algorithmic details of constructing the residue interaction networks were presented and discussed in the previous studies.[95−97] The edges in the residue interaction networks are weighted by dynamic cross-correlations[94] and coevolutionary couplings.[97] The ensemble of shortest paths was determined by the Floyd−Warshall algorithm.[98] Network graph calculations were performed using the Python package NetworkX.[99] The betweenness of residue i is computed as the sum of the fraction of shortest paths between all pairs of residues that pass through residue i

$$C_b(n_i) = \sum_{j<k}^{N} \frac{g_{jk}(i)}{g_{jk}} \tag{1}$$

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 modified version of the Girvan−Newman algorithm[100−102] is used to identify local communities. The community centrality measure was computed on a hierarchical

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

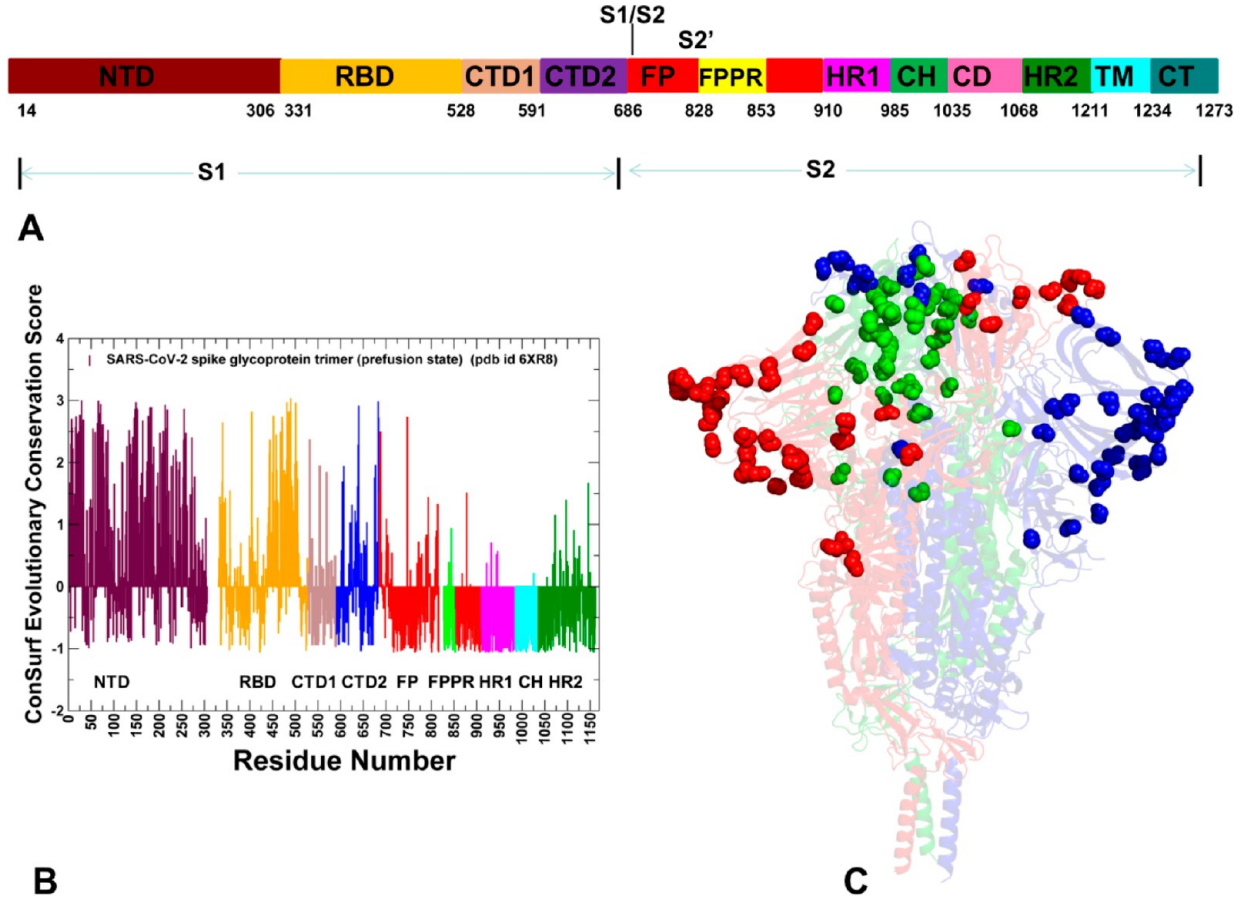


Figure 3. Sequence conservation analysis of the SARS-CoV-2 spike (S) glycoprotein. (A) Schematic representation of the domain organization and residue range for the full-length SARS-CoV-2 spike (S) protein. The subunits S1 and S2 include the N-terminal domain (NTD), receptor-binding domain (RBD), C-terminal domain 1 (CTD1), C-terminal domain 2 (CTD2), S1/S2 cleavage site (S1/S2), S2′ cleavage site (S2′), fusion peptide (FP), fusion peptide proximal region (FPPR), heptad repeat 1 (HR1), central helix region (CH), connector domain (CD), heptad repeat 2 (HR2), transmembrane anchor (TM), and cytoplasmic tail (CT). Annotation is as in Figure 1. (B) Normalized ConSurf conservation scores for the SARS-CoV-2 spike (S) glycoprotein projected onto the structure of the wild-type full-length SARS-CoV-2 spike (S) protein trimer in the locked closed ("RBD-down") prefusion conformation (PDB ID 6XR8).[28] The ConSurf profiles are shown as colored bars highlighting the conservation in different segments of the S protein. In particular, NTD residues are shown as maroon bars and RBD residues are shown as orange bars. The low negative ConSurf scores correspond to highly conserved sites, and high positive scores depict highly variable positions. (C) Structural mapping of variable positions onto the structure of the wild-type full-length SARS-CoV-2 spike (S) protein trimer (PDB ID 6XR8). The structure is shown as ribbons with reduced transparency, and the three protomers (A, B, and C) are colored green, red, and blue, respectively. The highly variable positions are shown as spheres colored according to the corresponding protomer (green, red, and blue). A high density of variable sites can be seen in the NTD and RBD regions.

meta-network whose nodes are the local communities using the following expression[103,104]

$$C_c(i) = \sum_{i \in j}^{N} \left(1 - \frac{1}{m} \sum_{i \in j \cap k}^{m} S(j, k)\right) \quad (2)$$

N is the number of communities to which node i belongs, and S is the Jaccard similarity coefficient between communities j and k. The network parameters were also evaluated using the ModuLand program in the Cytoscape package.[105−107]

RESULTS AND DISCUSSION

Sequence Analysis of the SARS-CoV-2 Spike Glycoproteins Reveals the Sharp Partition of Evolutionary Conservation between S1 and S2 Subunits

To understand the interplay between evolutionary, topological, and dynamic signatures of the SARS-CoV-2 spike glycoprotein trimer, the sequence conservation profiles are computed to identify regions of high conservation and variability (Figure 3). The sequence conservation ConSurf approach[52−56] was used to determine the residue-based conservation score profiles in which the low score is associated with the most conserved position in the protein. Consistent with previous studies,[108−110] this analysis showed that residues in the S1 domain are considerably less conserved than the S2 fusion domain. In particular, the NTD regions and the exposed RBM segment that forms a contact interface with the host receptor displayed a considerable variability (Figure 3A). In some contrast, the core region in SARS-CoV-2 RBD (residues 331−438) exhibited strong conservation, exemplified in this model by the low Consurf scores (Figure 3A). The RBM motif includes highly variable residues 472−486 that are not present in S proteins of coronavirus isolated in bats.[108] Among the variable residues are G482, Y484, F486, F490 N501, E471, E472, E468, and K444, which are consolidated near the flexible ridge loop subjected to structural changes upon the binding of SARS-CoV-2 RBD to the ACE2 receptor (Figure 3A). The structural studies of the SARS-

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

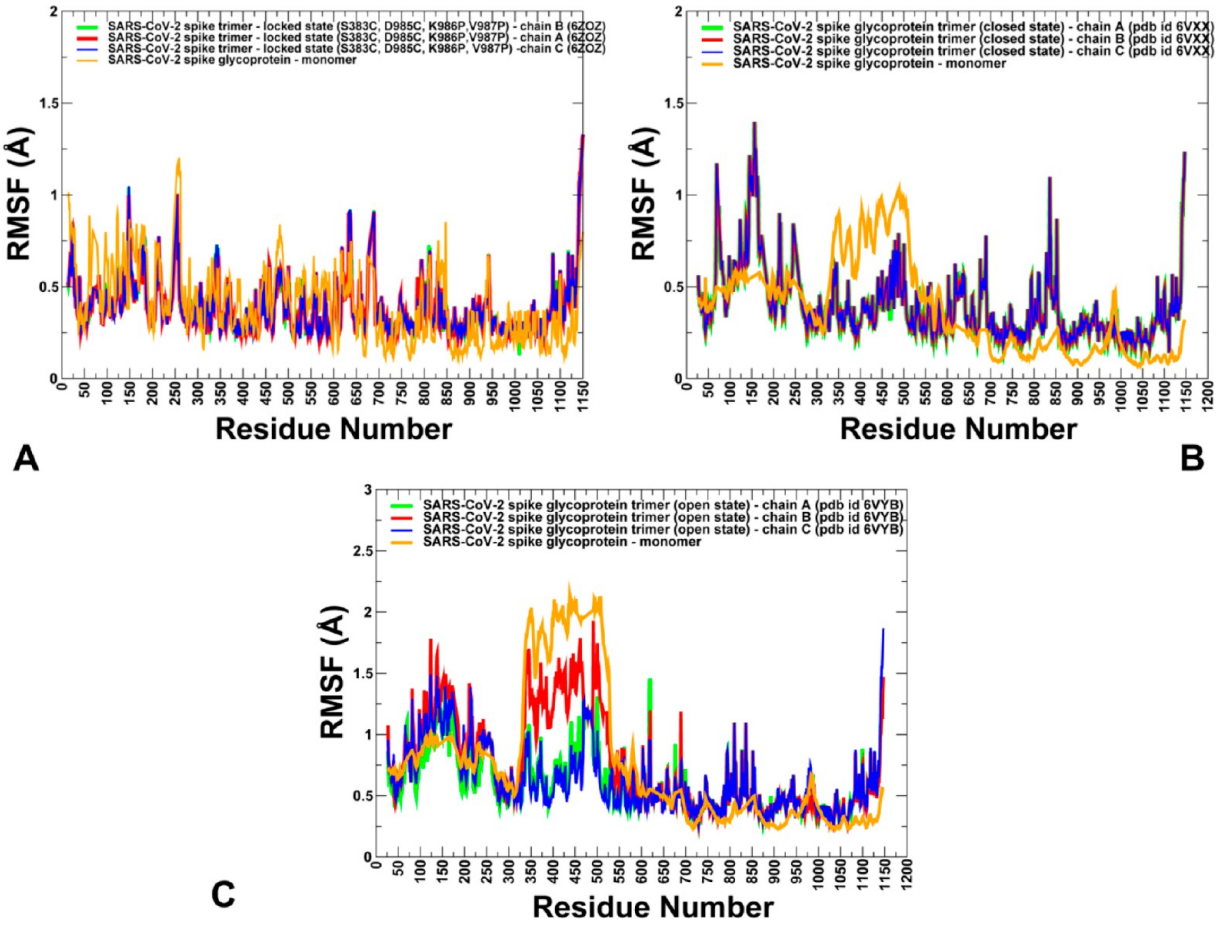


Figure 4. CABS-CG conformational dynamics of the SARS-CoV-2 spike (S) protein trimer in the locked (A), closed (B), and open states (C) of the prefusion form. (A) Root-mean-square fluctuation (RMSF) profiles obtained from CABS-CG simulations of the cryo-EM structure of the disulfide-stabilized SARS-CoV-2 S trimer (S383C, D985C, K986P, V987P, single R S1/S2 cleavage site) in the locked state (PDB ID 6ZOZ).[32] (B) RMSF profiles from simulations of the cryo-EM structure of the SARS-CoV-2 S trimer in the closed state (K986P/V987P) (PDB ID 6VXX).[25] The complete atomistic model of the full-length SARS-CoV-2 S protein developed in the study[38] was used in the simulations. (C) RMSF profiles from simulations of the cryo-EM structure of the SARS-CoV-2 S trimer ectodomain (K986P/V987P) in the open state (PDB ID 6VYB).[25] The profiles for protomer chains A, B, and C are shown as green, red, and blue lines, respectively. The conformational dynamics profile of the spike monomer is shown as an orange line.

CoV spike protein in the complex with ACE2 suggested that variable positions L472, N479, and T487 that are substituted for F486, Q493, and N501 in the SARS-CoV-2-RBD could be critical for binding selectivity.[20,21] Structural mapping of highly variable sites in the SARS-CoV-2 trimer highlighted the density of variable positions in the NTD and RBD regions (Figure 3A).

The furin cleavage site at the S1/S2 boundary is in a surface-exposed and disordered loop, and a small cluster of variable positions in this region is consistent with its functional role (Figure 3). Although the evolutionary analysis identified important variable RBM regions implicated in binding, a mix of conserved and variable positions in the RBD suggested that the interplay of stability and flexibility may enable dynamic adaptation of the spike protein to the host receptor and can be important for rendering the binding selectivity of the SARS-CoV-2 glycoprotein.

The S2 domain showed considerably higher conservation, which was pronounced for the FP region (812-SFIEDLFNK-VTLADAGF-829), the HR1 region (residues 910−985), the CH regions (residues 986−1035), and the CD residues (Figure 3). The high sequence conservation of the FP region is consistent with its functional role as a viral fusion motif that is composed of both small and bulky hydrophobic residues. The FPPR segment (residues 828−853) that is immediately downstream of the FP region also featured a high degree of conservation (Figure 3A). Interestingly, this functional segment was disordered in the SARS-CoV-2 ectodomain structures but adapted a well-defined helix-turn-helix motif in recently reported cryo-EM structures of the wild type and K986P/V987P mutants of the SARS-CoV-2 S trimer.[28]

Given the amino acid sequence conservation of the S2 subunit (Figure 3), it is possible that coronaviruses can operate through similar regulatory hotspots and regions to mediate virus−host membrane fusion in a unified manner. Interestingly, the cascade of tectonic conformational changes leading to the formation of the postfusion form implies a mechanism in which RBD binding to the host receptor and cleavage of the S2′ site can accelerate the release of the S1 subunit and refolding of helical regions in the remaining S2 trimer.[33,111] These dramatic rearrangements involve highly conserved HR1 and CH motifs that become exposed and vulnerable to the external intervention after dissociation of the S1 subunits.

To summarize, the striking partition of conservation and variability in the SARS-CoV-2 spike proteins between S1 and S2

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

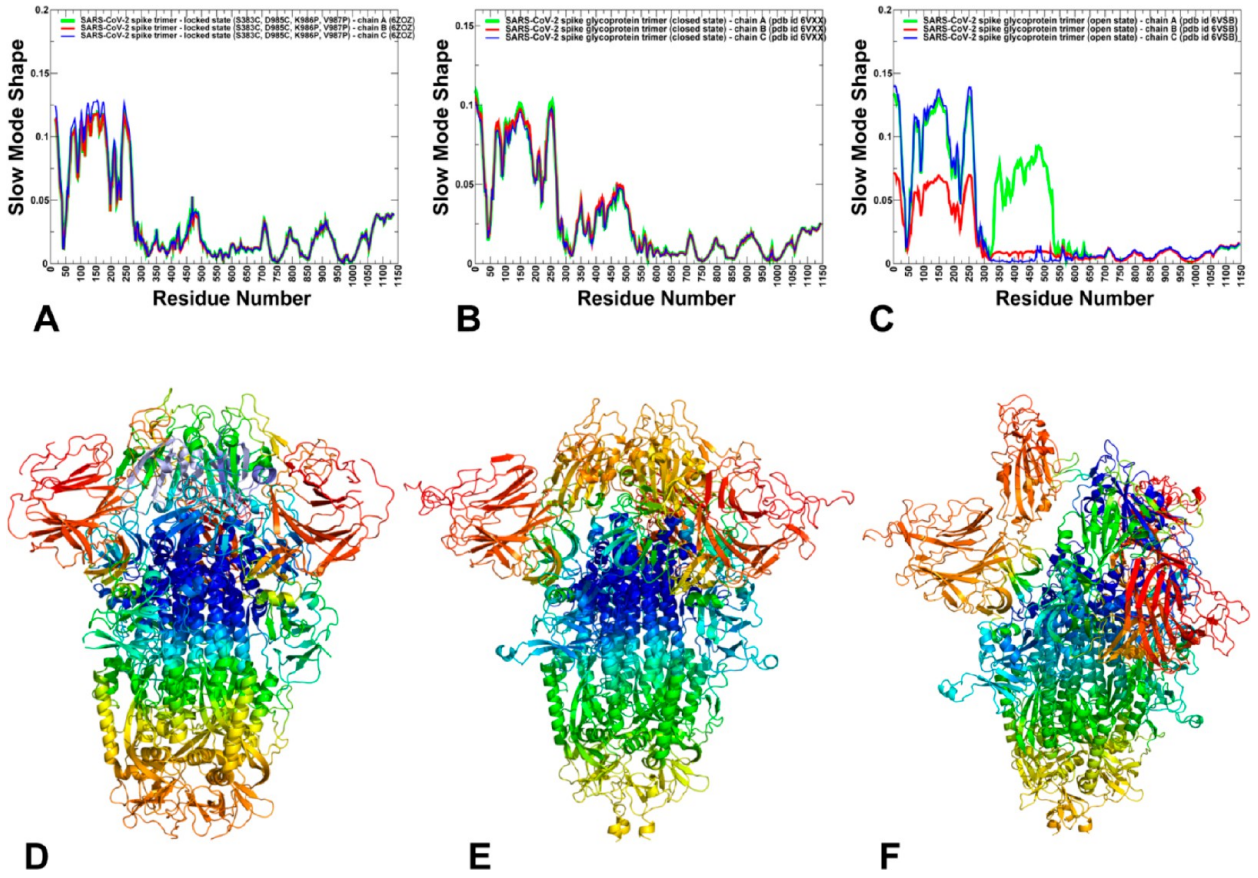


Figure 5. Functional dynamics and analysis of collective motions in the locked, closed, and open states of the SARS-CoV-2 spike trimer prefusion form. (A) Mean-square fluctuations averaged over the three lowest frequency modes for the locked state of the cryo-EM structure of the disulfide-stabilized SARS-CoV-2 S trimer (PDB ID 6ZOZ).[32] (B) Essential mobility profiles averaged over the three lowest frequency modes for the cryo-EM structure of the SARS-CoV-2 S trimer in the closed state (PDB ID 6VXX).[25] (C) Essential mobility profiles along the three slowest modes for the cryo-EM structure of the SARS-CoV-2 S trimer in the open state (PDB ID 6VSB).[25] The profiles for protomer chains A, B, and C are shown as green, red and blue lines, respectively. Structural maps of the essential mobility profiles for the locked state of the SARS-CoV-2 S prefusion trimer (D), closed state (E), and open state (F). The color gradient from blue to red indicates the decreasing structural rigidity and increasing flexibility. Notably, in the locked state of the SARS-CoV-2 spike trimer, the RBD residues are generally aligned with shallow minima positions, suggesting that large functional movements of RBD may be largely restricted and suppressed in the locked trimer form (Figure 5A). Several other local hinge points were precisely aligned with the RBD residues K386/L387 and D427/D428, and these hinge sites are unique for the locked form of the trimer (Figure 5A). The latter finding is of special interest and significance, as the biophysical studies showed that the stabilizing interactions of D427 and D428 with K986 are vital for the thermodynamic stabilization of the closed trimer and the formation of the rigid locked state.[32] The presented results indicated that these RBD residues could act as important functional hinge sites that regulate and restrict conformational transformations of RBD in the locked state.

subunits can be also embedded in the topology and drive the functional dynamics of the spike trimer in the prefusion form. In particular, this analysis suggested that conserved and structurally rigid elements of the S2 subunit can control and mediate dynamic changes on the landscape of the prefusion trimer, whereas highly variable NTD and RBD regions provide structural adaptation to external binding partners and host receptors.

Molecular Simulations of the SARS-CoV-2 Spike Trimers Characterize Salient Dynamic Signatures and the Redistribution of Mobility in Distinct Functional States

Ideally, all-atom MD simulations of the full-length SARS-CoV-2 S structures in different functional states with a complete glycosylation shield profile and a rigorous solvent description are required for the rigorous assessment of atomistic fluctuations and interaction details. The recent pioneering computational studies of the SARS-CoV-2 S protein with a complete and explicit glycosylation shield description[42] underscored the technical challenges of performing these simulations, which are time-consuming even with dedicated supercomputer resources. Accordingly, these simulations would require significant time and enormous resources for the execution of multiple SARS-CoV-2 S structures in fully glycosylated environments. The objective of this investigation is to provide a useful complementary analysis of topology-based dynamic signatures that are characteristic of the major functional states of the full-length SARS-CoV-2 S trimer by employing a large number of efficient CABS-CG simulations and subsequent atomistic reconstruction. Using these simulation approaches applied to full atomistic high-resolution structures of the SARS-CoV-2 S prefusion trimer in multiple functional states, a detailed comparative analysis of the conformational dynamics profiles was performed, and salient features of the dynamic conformational landscapes for the SARS-CoV-2 S trimer were analyzed (Figure 4). Whereas the dynamic fluctuations and variations of individual residues may be partly affected due to the absence of the explicit glycosylation shield, these molecular simulations can adequately reproduce global topological and dynamic patterns,

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

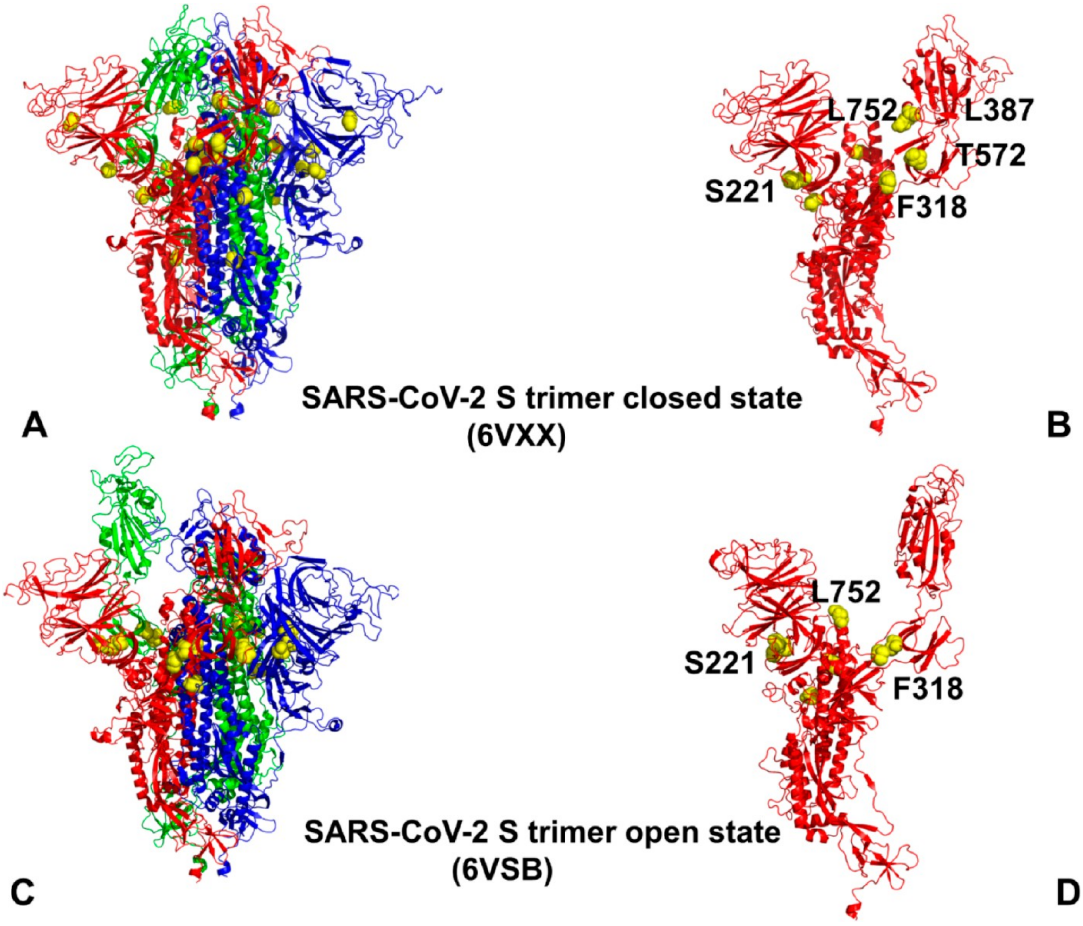


Figure 6. Structural maps of the hinge regions in the closed and open states of the SARS-CoV-2 spike trimer prefusion form. (A) Projection of hinge sites (shown as yellow spheres) onto the cryo-EM structure of the SARS-CoV-2 S trimer in the closed state (PDB ID 6VXX).[25] The SARS-CoV-2 S trimer is shown as ribbons with protomers A, B, and C in green, red, and blue, respectively. (B) Close-up map of hinge sites presenting a single protomer from the closed prefusion state. The protomer is shown as red ribbons, and hinge sites are highlighted with yellow spheres and annotated. (C) Projection of hinge sites (shown as yellow spheres) onto the cryo-EM structure of SARS-CoV-2 S trimer in the open state (PDB ID 6VSB).[25] (D) Close-up map of hinge sites presenting a single protomer undergoing the transition to the "up" prefusion state. The protomer is shown as red ribbons, and hinge sites are highlighted with yellow spheres and annotated.

allowing for a comparative analysis and enabling key differences between the dynamics of functional states to be pinpointed.

CABS-CG simulations of SARS-CoV-2 spike trimer in the locked, closed, and open forms of the perfusion state revealed important dynamic signatures of these states (Figure 4). Simulations of single monomers for each of the SARS-CoV-2 spike structures were also performed. A comparative analysis of the dynamic profiles generally showed very moderate thermal fluctuations of both the S1 and S2 subunits in the locked state (Figure 4A). Whereas the NTD regions and the furin cleavage site (residues 680−688) showed larger displacements, the mobility of the RBD residues was markedly suppressed in this form as compared with the more dynamic closed and open prefusion states (Figure 4A−C). Another important dynamic signature of the locked state is the thermal stabilization of the FP region and the structurally ordered FPPR motif (residues 828−853).

It was conjectured that CTD1 (residues 529−591) is a structural relay between RBD and FPPR (residues 828−853) that can sense the displacement on either side and communicate the signal from and to the fusion peptide. These segments of the S1 subunit are largely stable in the locked state of the S trimer (Figure 4A). This is in sharp contrast with the dynamic profiles of the closed state (Figure 4B) and open state (Figure 4C), where this segment becomes highly flexible and can be partially disordered. In addition, it was observed that both closed and open states featured significantly larger thermal fluctuations in the NTD and RBD regions, in particular, of the exposed RBM segment involved in the recognition of the host receptor. As expected, CG simulations of the open form of the trimer revealed the large fluctuations in the "up" monomer that become exposed for interactions with the host receptor (Figure 4C). Domains HR1, CH, and CD close to the viral transmembrane exhibited the least movement during MD simulations.

Of particular interest was a comparison with multiple simulations of the monomeric forms. Somewhat unexpectedly, the RBD fluctuations were also reduced in the monomer of the locked SARS-CoV-2 spike trimer (Figure 4A). Although this analysis is based on CABS-CG simulations that may underestimate the extent of the mobility of the isolated monomer, the comparison with other states clearly indicated that the monomeric form becomes highly flexible in the closed and open forms, with the RBD undergoing large conformational changes between the "down" and "up" forms (Figure 4). These simulations captured the fundamental topological features of the distinct conformational states, suggesting that despite the

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

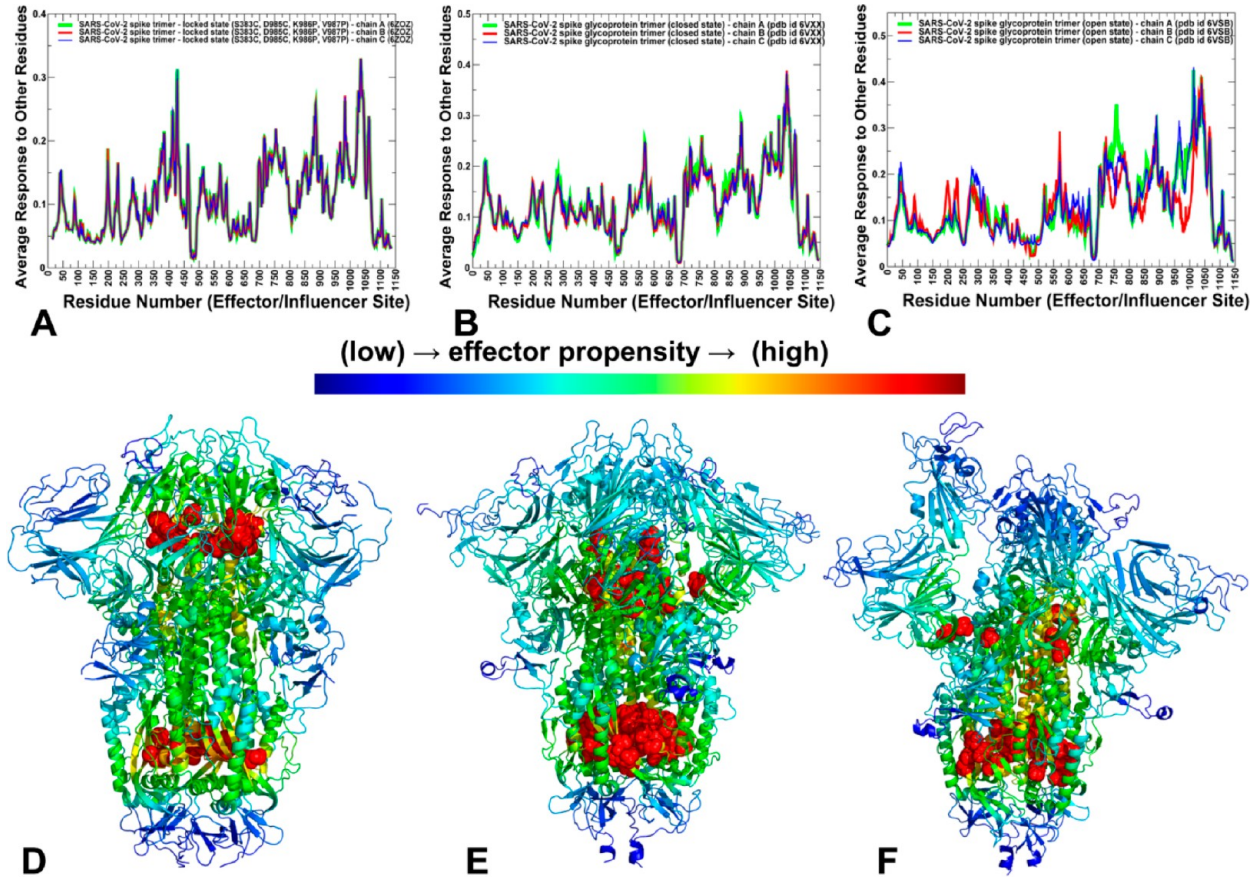


Figure 7. PRS effector profiles in the locked, closed, and open states of the SARS-CoV-2 spike trimer prefusion form. (A) PRS effector profile for the SARS-CoV-2 S trimer in the locked state (PDB ID 6ZOZ).[32] (B) Effector profile for the SARS-CoV-2 S trimer in the closed state (PDB ID 6VXX).[25] (C) Effector profiles for the SARS-CoV-2 S trimer in the open state (PDB ID 6VSB).[25] The profiles for protomer chains A, B, and C are shown as green, red, and blue lines, respectively. Structural maps of the PRS effector profiles for the locked state of the SARS-CoV-2 S prefusion trimer (D), closed state (E), and open state (F). The color gradient from blue to red indicates the increasing effector propensities. The clusters of residues with the high allosteric potential corresponding to the peaks of the effector profile are shown as red spheres.

structural similarities, the dynamics of the locked and closed states could be quite different.

Structural analysis suggested that the absence of an ordered FPPR segment in the closed states of the SARS-CoV-2 spike trimer may be linked to the higher mobility of the RBD and the emergence of the RBD-up conformation.[32] In fact, the "up" form of the RBD was not detected in the locked states of the trimer with the ordered FPPR segment. These simulations are consistent with this notion, showing that the dynamics of the FPPR and RBD regions may be allosterically coupled in all states of the SARS-CoV-2 spike trimer. The results also suggested that the RBD segment may be intrinsically predisposed to conformational selection between the "up" and "down" forms that can be modulated and stimulated (or suppressed) through the long-range couplings with the FP and FPPR motifs. Moreover, such redistribution of thermal fluctuations is more pronounced in SARS-CoV-2-RBD, which is coupled to a stronger stabilization of the entire binding interface. These findings imply that the binding mechanism of virus entry into the host receptor may exhibit signs of dynamically driven allostery, which is typically exemplified by the lack of structural changes between the unbound and bound forms, coupled to an exchange of conformational mobility between local protein regions.[99−102]

Functional Dynamics of the SARS-CoV-2 Spike Trimer Reveals the Migration of the Hinge Sites and the Release of Dynamic "Breaks" on the RBD Motions in the Closed State

To characterize collective motions and determine the distribution of hinge regions in the SARS-CoV-2 spike trimers, principal component analysis (PCA) of all-atom reconstructed trajectories was undertaken along with ENM analysis of the essential slow modes.[80] 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.[65] The analysis of collective motions revealed functionally important domain motions and hinge regions in the SARS-CoV-2 spike trimer structures (Figure 5). The simplified CG models used in the analysis did not affect the accuracy of the results and significantly reduced the computational cost so that the collective motions can be analyzed in great detail for all available SARS-CoV-2 trimer structures in the locked, closed, and open forms. By leveraging the efficiency of these methods, the essential profiles for the slowest modes were computed, and the functional dynamics profiles averaged over the first three major low-frequency modes were analyzed in detail (Figure 5). As expected within the realm of CG models, the overall shape of slow mode profiles was generally preserved in structurally similar locked and closed states but was largely altered in the open form of the trimer. Nonetheless, a detailed analysis of the hinge

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

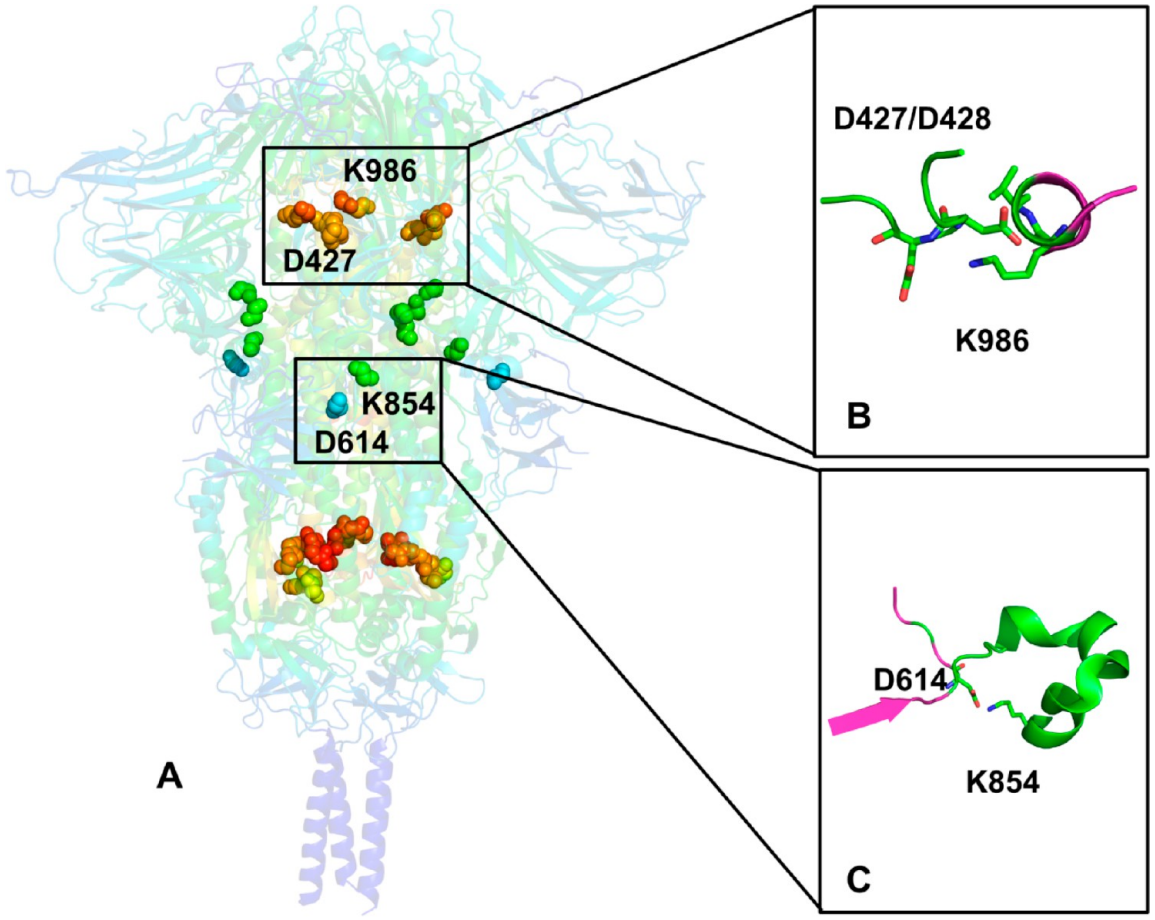


Figure 8. Detailed structural analysis of PRS effector profiles in the locked state of the SARS-CoV-2 spike trimer prefusion form. (A) Structural map of the PRS effector profiles for the wild-type full-length SARS-CoV-2 spike (S) protein trimer in the locked closed ("RBD-down") prefusion conformation (PDB ID 6XR8).[28] The color gradient from blue to red indicates the increasing effector propensities. The key effector centers and corresponding interaction clusters are shown as spheres colored according to the effector propensity. (B) Close-up of the interprotomer interaction cluster formed by effector hotspots D427, D428, and K986 residues. The residues are shown as sticks colored by atom type. (C) Close-up of the interprotomer interaction cluster formed by the effector center K854 with D614.

regions that correspond to local minima along the slow mode profiles revealed subtle important changes between the locked and closed states of the prefusion trimer form (Figure 5, Table S1).

The RBD regions of the S1 subunit are linked through two antiparallel hinge linkers to the NTD and CTD2 and allow the transition between closed and open conformations. Uniquely for the locked state, the hinge regions that can regulate the interdomain movements between RBD and NTD are fairly broadly distributed, forming shallow local minima aligned with a helical linker (residues 298−306) and extending through a flexible loop to another hinge site corresponding to a β-strand segment (residues 324-ESIV-327) that is connected to the RBD region (Figure 5A, Table S1). Interestingly, the recent protein engineering studies demonstrated that the dynamic equilibrium of the SARS-CoV-2 spike trimer can be manipulated using modifications of contact regions between the RBD and NTD through a triple D398L/S514L/E516L mutant[30] which can now be explained by the fact that D398 belongs to an important hinge site, and mutations in this regulatory switch position could alter the global equilibrium and distribution of states in the prefusion form of the trimer.

Together with the experimental evidence,[32] this suggested that D427/D428 and K986 may play the role of the specific regulatory switch of the spike dynamic equilibrium, where mutations or dynamic changes in these positions could trigger an allosteric conformational change and population shift between the closed and open trimers. In addition, the hinge sites in the locked state also corresponded to residues 570−572, 750−755, and 986−990 (Figure 5A). Similarly, modifications of hinge points A570 and T572 near interfaces between SD1 and S2 and SD1 and S2 domains can change the population of locked, closed, and open forms.[30] These positions were selected in the mutagenesis engineering studies to produce variants A570L, T572I, F855Y, and N856I of the SARS-CoV-2 S trimer that destabilized the closed state rather than stabilize the open state.[30]

Despite obvious structural similarities of the locked and closed trimer states, the essential mobility profile showed important changes reflecting a more flexible nature of the metastable closed form (Figure 5B). The most significant difference can be seen in the S1 subunit with only a single well-defined hinge site observed at the border of the NTD and RBD regions and aligned with the β-sheet motif (residues 315−320) (Figures 5B and 6A,B). Interestingly, the key hinge site near D427/D428 residues considerably weakened and corresponds to a very minor local minimum along the profile (Figure 5B), indicating that dynamic "brakes" on the RBD motions can be

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

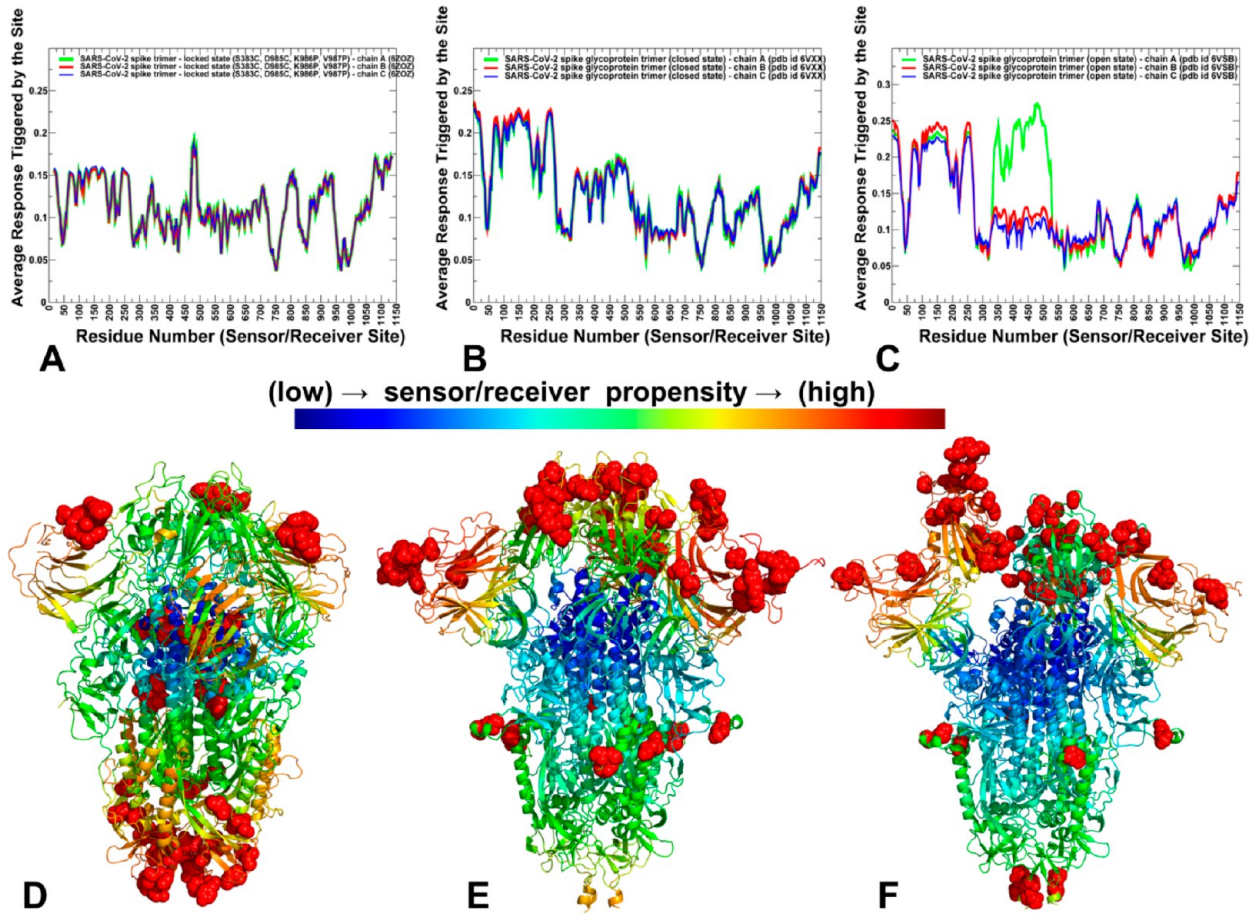


Figure 9. PRS sensor profiles in the locked, closed, and open states of the SARS-CoV-2 spike trimer prefusion form. (A) Sensor profile for the SARS-CoV-2 S trimer in the locked state (PDB ID 6ZOZ).[32] (B) Sensor profile for the SARS-CoV-2 S trimer in the closed state (PDB ID 6VXX).[25] (C) Sensor profile for the SARS-CoV-2 S trimer in the open state (PDB ID 6VSB).[25] The profiles for protomer chains A, B, and C are shown as green, red, and blue lines, respectively. Structural maps of the PRS sensor profiles for the locked state of the SARS-CoV-2 S prefusion trimer (D), closed state (E), and open state (F). The color gradient from blue to red indicates the increasing sensor (receiver) propensities. The clusters of residues with the high sensor potential corresponding to the peaks of the effector profile are shown as red spheres.

largely released in the closed state that becomes prone to large conformational transformations. Moreover, the RBD regions were mostly aligned with the local maxima of the slow mode profile, indicating that these regions in the closed form can experience functional movements between "down" and "up"' forms as the dynamic constraints on the RBD repositioning are removed in a more flexible closed state (Figure 5B). Other hinge sites were preserved in the closed trimer, showing the presence of local minima for residues 570−572 and 856−862 (Figure 5B). These positions may be involved in regulating functional motions near the interdomain regions and overlapped with positions used in protein design experiments (N866I/A570L, A570L/T572I/F855Y/N856I) to manipulate the dynamic equilibrium and promote a population shift of the SARS-CoV-2 spike protein.[30] Structural projection of the essential mobility profiles onto the SARS-CoV-2 S trimer structures in the locked state (Figure 5D) and closed state (Figure 5E) illustrated the migration of the immobilized regions away from the RBD during the transition to the dynamic closed form. The observed changes in the distribution of rigid and flexible regions in slow modes reflected the release of dynamic breaks on functional motions of the RBD in the closed state (Figure 5D,E). In the open form, the main hinge sites corresponded to residues 42−44 in the NTD and β-sheet motif (residues 315−320) (Figure 5C).

These regulatory positions of collective dynamics are associated with concerted functional movements of the NTD and RBD regions around the respective hinges. Notably, in the monomer undergoing a transition to the "up" conformation, the RBD regions are aligned with maxima along the profiles reflecting this large change (Figures 5C and 6C,D). The collective motions of the other two monomers corresponded to synchronous changes in the NTD, whereas the RBD regions remained in their locked closed positions (Figure 5C), suggesting that NTD adjustments may provide necessary room for the upward movement of the RBD in the single protomer of the trimer. Structural mapping of the slow mode profiles in the open state (Figure 5F) showed redistribution in the density of moving regions in which the RBD of the upward moving monomer cooperates with functional motions of the NTDs in the other two monomers. Overall, the results suggested that collective dynamics of the spike trimer may be driven by coordinated movements of the NTD and RBD regions around the more dynamically restricted S2 subunit.

The positions of 22 glycosylation sites of the S protein that constitute the glycan shield of the SARS-CoV-2 spike trimer[36] were aligned onto structural maps of functional motions in slow modes for the locked, closed, and open states (Figure S2). It is evident that the glycan shield can be especially effective at

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

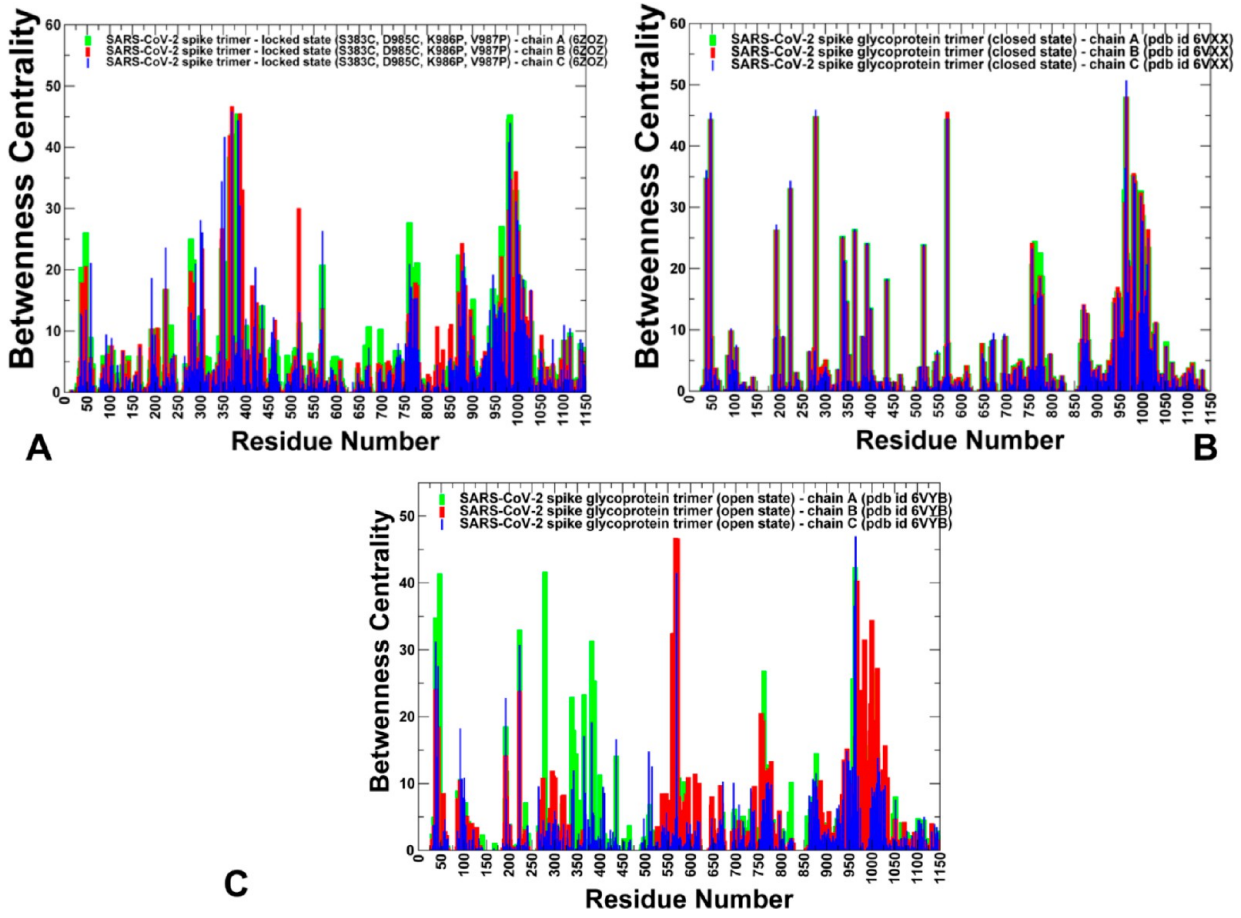


Figure 10. Residue-based betweenness centrality profiles in the locked, closed, and open states of the SARS-CoV-2 spike trimer prefusion form. The residue-based centrality values are computed by averaging the results over 1000 representative equilibrium samples of the SARS-CoV-2 spike trimer structure from CABS-CG simulations and atomistic reconstruction of trajectories. (A) Betweenness centrality profile for the SARS-CoV-2 S trimer in the locked state (PDB ID 6ZOZ).[32] (B) Betweenness centrality profile for the SARS-CoV-2 S trimer in the closed state (PDB ID 6VXX).[25] (C) Betweenness centrality profile for the SARS-CoV-2 S trimer in the open state (PDB ID 6VYB). The profiles for protomer chains A, B, and C are shown as green, red, and blue lines, respectively.

enhancing the protection of the stable locked state, as the majority of the moving regions tend to be localized in the NTD regions. In some contrast, in a more dynamic closed state of the prefusion S trimer, the glycan shield may display potential vulnerabilities in the RBDs, exposing the evolutionary variable RBD regions that undergo collective movements (Figure S2). As a result of the reduced protection, the more flexible closed state of the spike trimer could become susceptible to large-scale transformations and readily undergo allosteric structural changes to the open form, in which one RBD protomer is exposed for binding with the host-cell receptor.

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

Using the PRS method,[81−91] the allosteric effect of each residue in the protein structures in response to external perturbation was quantified and examined. PRS analysis produced the residue-based effector and sensor response profiles in different functional states of the SARS-CoV-2 S trimer (Figures 7 and 8). The effector profiles assess the propensities of a given residue to influence dynamic changes in other residues and are often applied to identify the regulatory hotspots of allosteric interactions as the local maxima along the profile (Figure 7), whereas the sensor/receiver profiles measure the ability of residues to serve as transmitters of allosteric perturbations and allow for the determination of the regions undergoing large structural changes during allosteric transformations. The peaks of the effector profiles may point to the allosteric sites that control the allosteric signal transmission and conformational transitions between the distinct functional states of the SARS-CoV-2 S protein. First, the effector profiles of the SARS-CoV-2 S homotrimer were analyzed in the locked closed form (Figure 7A), the closed form (Figure 7B), and the open receptor-accessible form (Figure 7C). In the locked form, the effector profiles featured three major peaks corresponding to the RBD residues 426-LPDDF-429, residues 1026−1040 near the homotrimer cavity, residues 883−893 and 986/987 (K986P/V987P), and also residues 850−855 from the FPPR region (residues 833−855) (Figure 7A). The peaks in the RBD region (D427/D428) and the FPPR motif appeared to be uniquely characteristic of the effector profile in the locked state (Figure 7A), as these peaks and the effector density in the RBD regions were considerably weakened in both the closed (Figure 7B) and open forms (Figure 7C, Table S1).

Interestingly, the effector peak in the RBD region is precisely aligned with the hinge residues D427 and D428 that form the interprotomer electrostatic interactions with K986. Although K986 and V987 are mutated to the neutral proline in the locked structures of the SARS-CoV-2 trimer, this interacting cluster

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

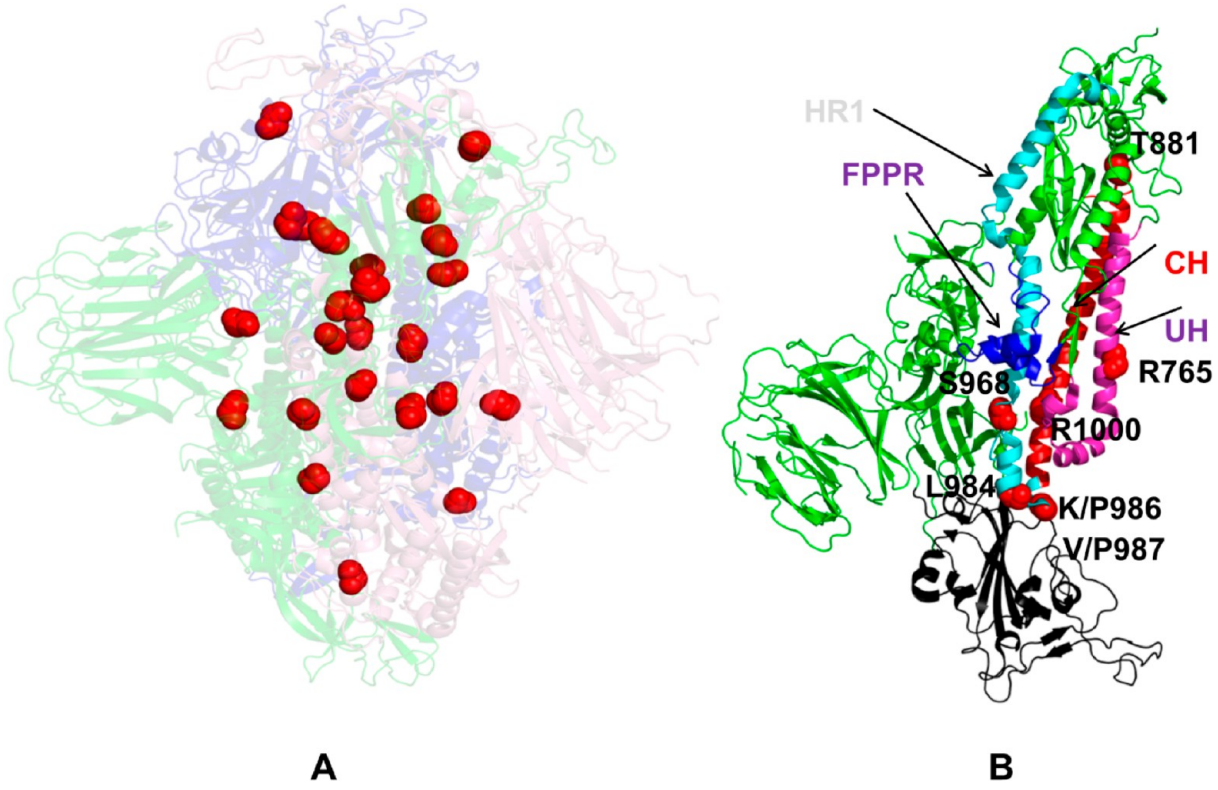


Figure 11. Structural map of high-centrality clusters for the locked state of the SARS-CoV-2 S prefusion trimer. (A) Projection of high-centrality clusters (shown in red spheres) onto the cryo-EM structure of SARS-CoV-2 spike trimer in the locked state (PDB ID 6ZOZ).[32] The S trimer structure is in ribbon representation with reduced transparency. The protomers A, B, and C are green, red, and blue, respectively. (B) Close-up of the S2 subunit with high-centrality sites (in red spheres) mapped in the HR1 (residues 910−985) and CH regions (residues 986−1035). The NTD is green, RBD is black, HR1 motif is cyan, CH is red, FPPR is blue, and upstream helices (UH) motif is magenta.

retains its structural integrity and is aligned with the unique allosteric hotspot in the RBD (Figure 7A). Another important signature of allosteric interactions in the locked trimer form is the emergence of the effector peak (residues 850−855) in the FPPR region, which is N-terminal to the FP and S2′ site. The key structural difference of the locked trimer form is the ordering of the FPPR segment into a well-defined motif, whereas in the majority of SARS-CoV-2 spike trimer structures of the closed and open forms, this region is poorly defined and highly flexible. In addition, the effector sites in FPPR are adjacent to K854, which forms a salt bridge with residue D614 that is known to be critical for the stability of the closed form (Table S1).

It may be argued that the regulatory center in the RBD (D427/D428) can communicate with the effector center in the FPPR region to cooperatively control allosteric couplings and interactions in the locked closed state of the SARS-CoV-2 spike trimer. Interestingly, some of the predicted effector sites are located in close proximity to residues that were engineered to produce a prefusion-stabilized SARS-CoV-2 spike protein that is more stable than the original construct.[29] For example, the effector residues 886−890 include S884 and A893 positions that were covalently linked by cysteine substitutions S884C/A893C and T791C/A879C, leading to the increased thermal stability of the closed form by improving the intermonomer packing.[29] Notably, some of these sites also coincided with the major hinge points of collective motions, suggesting that these positions may be involved in the regulation of allosteric interactions in the SARS-CoV-2 spike trimer (Table S1).

The effector profiles were also computed for other SARS-CoV-2 trimer structures in the locked state, including the full-length, wild-type SARS-CoV-2 spike trimer with the structurally determined FPPR region (Figure S3). The effector profiles were similar, highlighting the unique peak in the RBD region that is aligned with the D427/D2428 residues and indicating the critical role of this allosteric hotspot in signal transmission.

In the closed form of the SARS-CoV-2 spike trimer, the major effector peaks corresponded to residues 567−569 and 756−758 in the S2 CH repeat region, residues 886−890, as well as residues 1038−1040 in the CH/CD regions (Figure 7B, Figure S4). These characteristic peaks are shared in other closed trimer conformations (Figure S4), revealing an important redistribution of the effector sites and indicating the reduced allosteric potential of the RBD regions that in the closed form function mainly as sensors of allosteric signals. This redistribution of the effector allosteric centers and migration of the effector density is evidenced from structural mapping (Figure 7D,E). Notably, this dynamic switching of control points occurs in the absence of significant structural transformations between the locked and closed forms, suggesting that dynamically driven allostery[112−115] can govern the equilibrium between the locked and closed states.

In the open trimer conformation, the shape of the effector profile (Figure 7C) displayed the further migration of the effector density away from the RBD regions (Figure 7F) and exhibited density in the CTD1 region (residues 529−591) that is implicated as a structural relay between RBD and FPPR. Nonetheless, several clusters of conserved effector peaks were retained in the S2 subunit (residues 756−758, 887−891, and 1035−1045 in the S2 domain) (Figure 7C, Table S1). Structural mapping of effector hotspot clusters highlighted key findings, showing the evolution of the effector density in the trimer, from

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

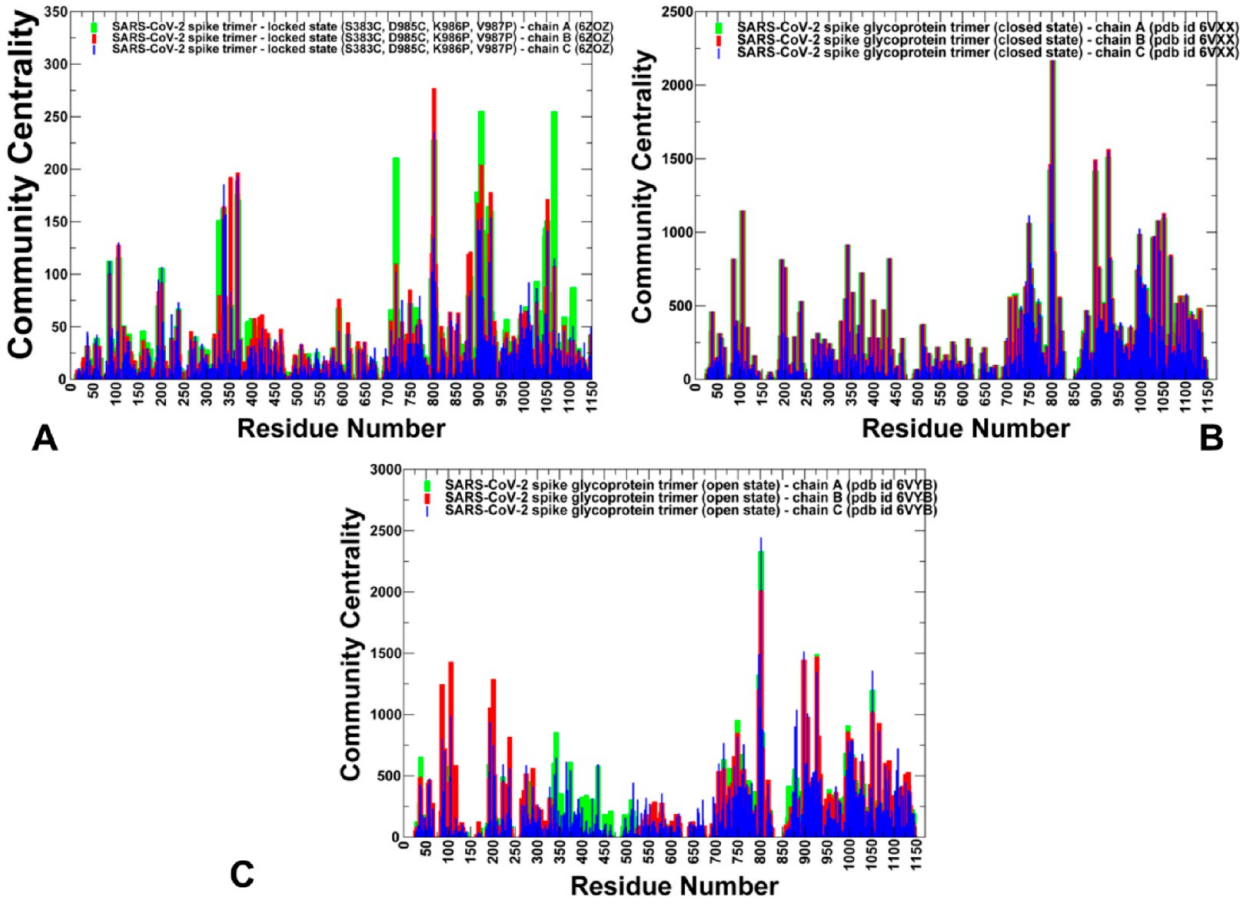


Figure 12. Community analysis and community centrality profiles in the locked, closed, and open states of the SARS-CoV-2 spike trimer prefusion form. The community centrality values are computed using community decomposition of the residue interaction networks and by averaging the results over 1000 representative equilibrium samples of the SARS-CoV-2 spike trimer. (A) Community centrality profile for the SARS-CoV-2 S trimer in the locked state (PDB ID 6ZOZ).[32] (B) Community centrality profile for the SARS-CoV-2 S trimer in the closed state (PDB ID 6VXX).[25] (C) Community centrality profile for the SARS-CoV-2 S trimer in the open state (PDB ID 6VYB). The profiles for protomer chains A, B, and C are shown as green, red, and blue lines, respectively.

high and broad density reaching the RBD regions in the locked closed state to a narrower density in the closed and open forms, in which allosteric couplings between FPPR and RBD regions are weakened (Figure 7D−F). These results are consistent with the notion that the locked conformation can represent an early intermediate state before opening of the RBD and binding to the host receptor.[32] The key finding of the PRS effector analysis is the emergence of a broad and dense allosteric network with multiple regulatory centers in the locked state and a cross-talk between the effector hotspots in the RBD and FPPR regions (Figure 7). Through these regulatory control points, allosteric signaling dictates the stable, down-regulated arrangement of the RBD regions. Interestingly, the allosteric interaction network may be weakened, become smaller, and diffuse in a more dynamic metastable closed form, where the effector centers regulate allosteric transitions. The allosteric couplings between effector hotspots from the conserved interdomain regions of the S2 subunit with FPPR (residues 850−855) and RBD regulatory centers (D427/D428) may be central to the coordination of allosteric interactions in the trimer, so that dynamic changes in these regions can determine the extent and directionality of the RBD motions. These results also provide an additional insight into the functional role of the intermonomer interactions of D427/D428 with K986 and K854 contacts with D614 to promote the stability of the closed state (Figure 8). The results are consistent with protein engineering studies,[29−32] suggesting that these predicted effector centers may function as regulatory switch points of the dynamic equilibrium in the SARS-CoV-2 spike trimer. The structural rigidity and the strategic position of the effector centers as mediating hubs in the interaction networks suggest that mutations or ligand-based targeting of these positions can affect the dynamic equilibrium and allosteric conformational transitions between the functional states of the SARS-CoV-2 spike protein, which may be relevant for the discovery of allosteric modulators.

Sensor Profiling of SARS-CoV-2 Spike Trimer Structures: Signal Transmitters of Allosteric Changes Are Localized in the RBD Regions and Expose the Vulnerabilities of Glycosylation Shielding

The computed sensor profiles quantified the propensity of residues to serve as carriers or transmitters of long-range allosteric interactions, revealing the key regions that participate in this function. The major clusters of the sensor density in the locked form of the trimer are aligned with a small cluster of RBM residues 478−491 that are involved in the recognition of the host receptor ACE2 (Figure 9A). According to the results, the allosteric control over RBM residues may be exerted through regulatory centers D427/D428 that are allosterically coupled to the FPPR effector sites. In the locked state, the RBD sensor residues maintain their down-regulated positions, in which the

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

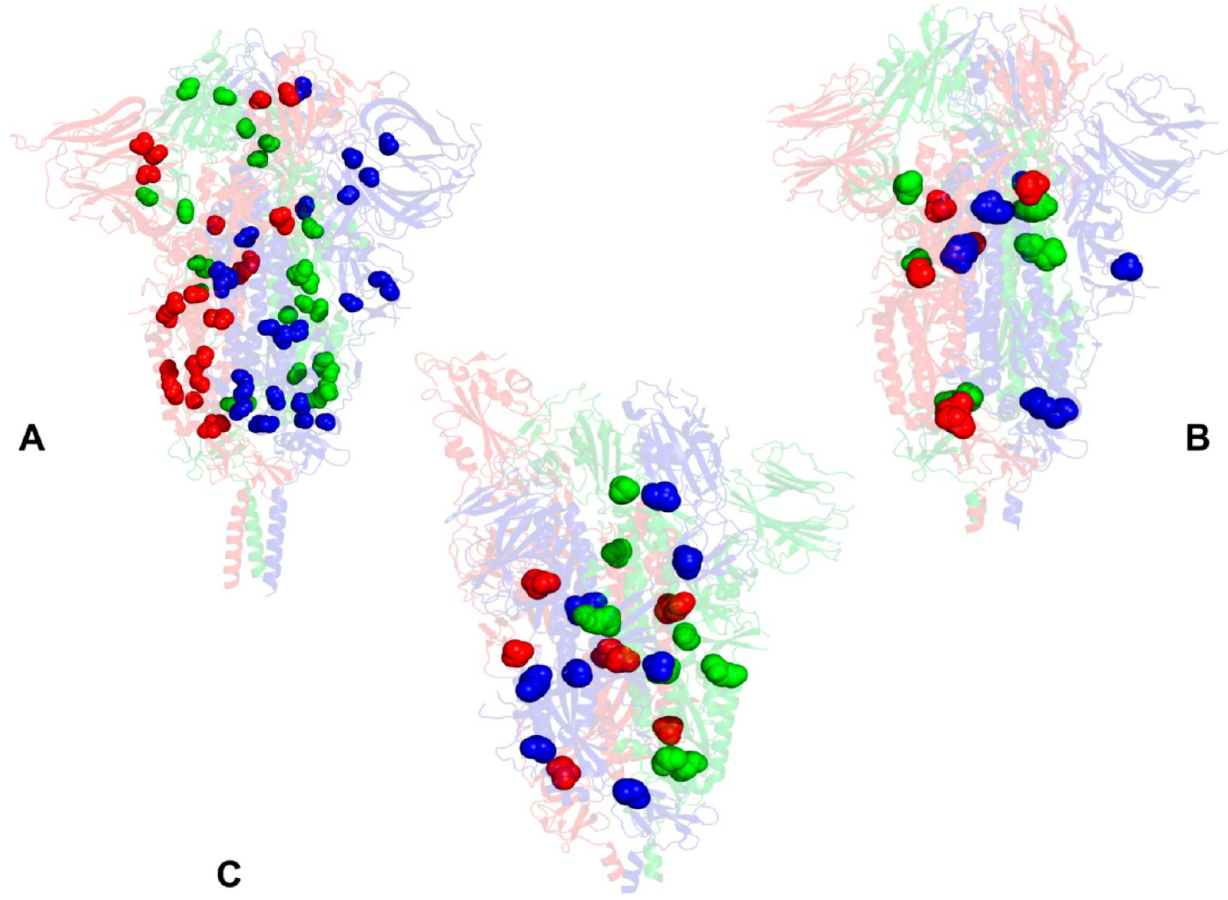


Figure 13. Structural map of community centrality clusters for the locked state of the SARS-CoV-2 S prefusion trimer. (A) Projection of community centrality clusters onto the cryo-EM structure of SARS-CoV-2 spike trimer in the locked state (PDB ID 6ZOZ).[32] (B) Mapping of community centrality clusters onto the SARS-CoV-2 S trimer in the closed state (PDB ID 6VXX).[25] (C) Projection of community centrality clusters on the SARS-CoV-2 S trimer in the open state (PDB ID 6VYB). The profiles for protomer chains A, B, and C are shown as green, red, and blue lines, respectively. The S trimer structure is in ribbon representation with reduced transparency. The community centrality clusters are depicted as spheres colored by their respective protomers.

RBD region is shielded from access to the host receptor (Figure 9A,D). In addition, the sensor sites are also consolidated near the FP motif (residues 812−829), with the peaks aligned with residues L821, L822, and F823 (L803, L804, and F805 in SARS-CoV) (Figure 9A, Table S1). Mutagenesis studies and infectivity assays showed the critical role of these residues in membrane fusion, as alanine mutants in these positions produced only limited fusion events in comparison with wild-type trypsin-treated cells.[116] Only small groups of RBM and FP residues form two clusters of sensor residues in the locked state of the trimer. Interestingly, these clusters of sensor residues are immediately proximal to their respective effector controllers (D427/D428 in the RBD and 850−855 residues of the FPPR).

It can be suggested that in the locked state these effectors could mediate the local mobility in the narrowly localized sensor regions, as allosteric signaling dictates only minor conformational changes of the rigid trimer form (Figure 9A,D). The sensor profiles were also computed for other SARS-CoV-2 trimer structures in the locked state, including the full-length, wild-type SARS-CoV-2 spike trimer with the structurally determined FPPR region (PDB ID 6XR8) (Figure S5). This comparison demonstrated the strong similarity of the sensor distributions in all locked states, where most sensor residues are localized in the flexible NTD regions. Hence, allosteric communications in the stable locked state may be primarily directed from the dominant effector region in the S2 subunit through the effector switch in the RBD (D427/D428) to the flexible NTD residues acting as primary receivers of the allosteric signal and enabling thermal breezing without altering the structural stability of the locked state.

A notable change can be observed in the closed form, where the major cluster of sensor sites is now localized in the flexible RBD regions (Figure 9B,E), and the second cluster in the RBD region becomes much broader and covers a wide range of exposed residues in the RBD region (residues 350−505). As a result of this expansion in the distribution of sensor sites, the entire RBD region becomes susceptible to large conformational changes that are governed by the effector centers in the conserved regions of the S2 subunit (Figure 9B, Table S1). In the closed form, the furin cleavage site (residues 682−685) also overlapped with the increased density of the sensor sites. Together, the effector and sensor sites can form a dynamic allosteric network that drives signal transmission in the prefusion trimer. The sensor profile of the open form of the trimer featured a different allocation of sensor sites in the monomers (Figure 9C,F). Most of the sensor sites in the monomer undergoing “up” transition are consolidated in the flexible RBD region (Figure 9C). The entire RBM motif in the “up” monomer becomes enriched with sensor sites transmitting the allosteric signal in the open trimer. At the same time, the

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

sensor profiles of the other two monomers revealed that the density localized mainly in the NTD regions.

The PRS analysis showed that sensor positions often overlap with evolutionary variable sites in the flexible RBD region that undergo mutations between SARS-CoV and SARS-CoV-2 glycoproteins. The partitioning between effector and sensor sites in the SARS-CoV-2 S trimer also reflects the evolutionary diversity elevated at the NTD and RBD regions and the prevalent conservation in the regions in the S2 domain, including the FP, FPPR, HR1, and the CH domains, which are subject to greater functional constraints. These findings are consistent with the properties of sensor sites in protein systems that are often tolerant of mutational changes and tend to be located in the mobile loops and exposed binding interfaces.[117] Computational studies of molecular chaperones conducted in my laboratory[89] and other groups[83−87] demonstrated that sensors are frequently found close to recognition sites and protein regions undergoing allosteric changes to adapt to binding and to accommodate diverse interactions.

Mechanistic details of the binding mechanism with the host receptors result from a complex interplay of thermodynamic forces determined by the intrinsic dynamics of the spike trimer coupled to the variability and adaptation of glycosylation shields.[42] To access the role of glycosylation sites in shielding sensor/receiver sites in the trimer, the results of the site-specific glycan analysis of the SARS-CoV-2 spike protein were utilized to project the determined 22 glycosylation sites on structural maps of allosteric sensor profiles of the locked closed, closed, and open forms (Figure S6). This analysis showed a considerable overlap of the glycosylation sites dispersed across both the S1 and S2 subunits, with the regions of increased sensor density in the locked trimer state (Figure S6). Strikingly, some of the important glycosylation sites N122, N149, N74, N165, N234, and N343 correspond precisely to the distinct and sharp peaks of the sensor profile in the locked state. The analysis indicates that N-linked glycans can shield the most sensor-sensitive regions of the SARS-CoV-2 in the NTD regions of the locked state (Figure S6). In the closed flexible state, an increasing density of sensor residues in the RBD regions that are not properly shielded may create vulnerabilities, leading to the reduced protection against conformational changes and ultimately promoting the transition to the open form. These findings are consistent with the pioneering studies by Amaro and colleagues,[42] who performed a series of unprecedented atomistic MD simulations of the SARS-CoV-2 spike trimers in a fully glycosylated environment, showing that the N-glycans can function to protect the glycoprotein but also to allow for a coordinated response to allosteric signaling and to promote the "down−up" RBD transition through observed shield vulnerabilities in the closed state.

Hierarchical Analysis of the Residue Interaction Networks: Betweenness Centrality and Community Centrality Profiling Unveil the Mediators of Allosteric Interactions

Using a graph-based representation of protein structures,[92−94] the residue interaction networks were constructed, in which the inter-residue edges were weighted using the residue cross-correlations obtained from simulations[94] and coevolutionary residue correlations.[97] Using the ensemble-averaged model of the residue interaction networks, residue betweenness (or residue centrality) is computed and is used to identify the key mediating centers of the allosteric interaction networks. The high-centrality positions in the interaction networks were aligned with many effector centers in the PRS profiles and have high allosteric propensities (Figure 10). The important result of this analysis is the emergence of several clusters of high-centrality sites in the locked prefusion state that mediate allosteric interactions in the global interaction network (Figure 10A). The first dense cluster of mediating residues is located in the RBD regions, with major centrality peaks corresponding to residues Y369, L387, and V382. The second cluster corresponds to residues in the CTD1 (residues 529−591), which is believed to function as an allosteric relay between the RBD and FPPR regions by communicating the signal from and to the FP. Another cluster was found near the HR1 motif and CH regions (Figures 10A and 11A). In the prefusion state, the HR1 (residues 910−985) and CH regions (residues 986−1035) form separate helices, and the HR1 helices and the CH helix are arranged in an antiparallel orientation. The high-centrality sites in the HR1 region are aligned with residues I980, R983, L984, K/P986, V/P987, and R1000 (Figure 11B). Interestingly, some of these sites act as regulatory switches of the dynamic equilibrium of the SARS-CoV-2 spike trimer (K986, V987), as proline mutations of these residues can dramatically reduce the fraction of the closed state and favor the open state of the trimer.[32] A high density of high-centrality sites in this important region of the S2 subunit is present and is shared in distinct conformational states of the SARS-CoV-2 spike protein (Figure 10A−C). Notably, the number of high-centrality peaks was reduced in the more dynamic closed form, particularly indicating the reduced density of global mediating centers in the RBD regions, which may be a consequence of the increased flexibility of the RBD regions (Figure 10B). These results are consistent with the PRS analysis in which the RBD regions in the dynamic closed form are mainly aligned with sensors of allosteric signals. In network terms, this implies that the allosteric interaction network in the closed form may be weakened and become more diffuse, reflecting the reduced stability and showing signs of the intermediate conformation preceding the transition to the open form. The distribution of the high-centrality sites was further altered in the open state, where the high density of mediating centers was detected near the RBD regions in the monomer undergoing "up" transition (Figure 10C).

Importantly, the cluster of high-centrality positions in HR1/CH regions retained its mediating role in the closed form of the trimer, suggesting that these regions play a key role during allosteric conformational changes from the locked to the open state. It is worth noting that global rearrangements in the HR1/CH regions harboring the proposed regulatory sites act as a nucleus during the formation of the postfusion state.[33] The important result of this analysis was a revelation that key regulatory sites that could dictate dynamic switching between conformational states of the SARS-CoV-2 spike trimer overlap with the high-centrality mediators of the interaction network that are also predicted to be effector centers with high allosteric potential (Figures 10 and 11). To further examine the network properties, the hierarchical multilayer community decomposition was performed, and the community centrality metric[118] of the residue interaction networks was computed for distinct functional states of the SARS-CoV-2 spike trimer (Figure 12). The community centrality profile revealed that network hubs that could bridge local modules in the locked state are located at the borders of the RBD regions (Figures 12A and 13A). The denser distribution of bridging centers in a stable locked state of the SARS-CoV-2 spike trimer (Figure 13A) is indicative of a broad and stable allosteric interaction network. The community

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

centrality hubs in the allosteric network are located near the experimentally known regulatory switches of dynamic equilibrium in the SARS-CoV-2 spike trimer. The structural mapping of the community centrality sites in the locked state (Figure 13A) highlighted a network of switch points that could enable allosteric couplings between spatially separated communities in the interaction network. The results also indicated that the number of local communities and community centrality sites can be reduced in the more dynamic intermediate closed state (Figure 13B) and open state of the prefusion trimer (Figure 13C).

Hence, the hierarchical community analysis confirmed a distinct organization of the interaction networks in functionally different locked and closed prefusion states. The regulatory sites that could dictate dynamic switching between the conformational states of the SARS-CoV-2 spike trimer overlap with the high-centrality sites and effector hotspots with high allosteric potential.

CONCLUSIONS

This study systematically examined the functional mechanisms of the SARS-CoV-2 spike prefusion trimer through the lens of allosteric regulation and signal transmission. Functional dynamics analysis and PRS characterized the hinges of conformational transitions and identified the regulatory hotspots that control the signal transmission and allosteric interactions in distinct states of the SARS-CoV-2 spike trimer. The results revealed the migration of the hinge sites during the transition from the stable locked to the more dynamic closed trimer state, leading to the release of dynamic breaks on the RBD motions. Through perturbation-based network analysis, it was demonstrated that the stabilized locked form of the prefusion trimer is characterized by a broad allosteric network with key regulatory centers located in the RBD, FPPR, and HR regions that dictate the dynamic switching between conformational states of the SARS-CoV-2 spike trimer. The PRS analysis identified key regulatory hotspots in the SARS-CoV-2 spike, showing that these allosteric centers are precisely aligned with known regulatory switches in the RBD and FPPR regions. The results are consistent with protein engineering studies and suggest that the predicted effector centers may function as regulatory switch points of the dynamic equilibrium in the SARS-CoV-2 spike trimer. The discovered effector centers can alter the distribution of functional states and conformational transitions between the closed and the host-accessible open states of the SARS-CoV-2 spike trimer, thus presenting opportunities for rational allosteric drug targeting of the SARS-CoV-2 spike proteins.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00654.

Figure S1. Atomistic models of fully glycosylated full-length SARS-CoV-2 spike trimer protein in the open and closed forms. Figure S2. Structural maps of the essential slow modes for the functional states of the of SARS-CoV-2 spike trimer. Figure S3. Effector profiles for other SARS-CoV-2 structures in the locked state. Figure S4. Effector profiles for other trimer structures in the closed state. Figure S5. Sensor profiles for SARS-CoV-2 trimer structures in the locked state. Figure S6. Structural mapping of glycosylation sites on allosteric sensor profiles of the functional SARS-CoV-2 states. Table S1. Summary of the studied structures and overview of the predicted allosteric regulatory regions (PDF)

AUTHOR INFORMATION

Corresponding Author

Gennady M. Verkhivker − *Graduate Program in Computational and Data Sciences, 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

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00654

Notes

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

REFERENCES

(1) Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K. S. M.; Lau, E. H. Y.; Wong, J. Y.; Xing, X.; Xiang, N.; Wu, Y.; Li, C.; Chen, Q.; Li, D.; Liu, T.; Zhao, J.; Liu, M.; Tu, W.; Chen, C.; Jin, L.; Yang, R.; Wang, Q.; Zhou, S.; Wang, R.; Liu, H.; Luo, Y.; Liu, Y.; Shao, G.; Li, H.; Tao, Z.; Yang, Y.; Deng, Z.; Liu, B.; Ma, Z.; Zhang, Y.; Shi, G.; Lam, T. T. Y.; Wu, J. T.; Gao, G. F.; Cowling, B. J.; Yang, B.; Leung, G. M.; Feng, Z. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199−1207.

(2) Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470−473.

(3) Yi, Y.; Lagniton, P. N. P.; Ye, S.; Li, E.; Xu, R. H. COVID-19: what has been learned and to be learned about the novel coronavirus disease. *Int. J. Biol. Sci.* **2020**, *16*, 1753−1766.

(4) Wu, A.; Peng, Y.; Huang, B.; Ding, X.; Wang, X.; Niu, P.; Meng, J.; Zhu, Z.; Zhang, Z.; Wang, J.; Sheng, J.; Quan, L.; Xia, Z.; Tan, W.; Cheng, G.; Jiang, T. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* **2020**, *27*, 325−328.

(5) 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.; Cheng, Z.; Yu, T.; Xia, J.; Wei, Y.; Wu, W.; Xie, X.; Yin, W.; Li, H.; Liu, M.; Xiao, Y.; Gao, H.; Guo, L.; Xie, J.; Wang, G.; Jiang, R.; Gao, Z.; Jin, Q.; Wang, J.; Cao, B. 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.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271−280.
(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.; Chen, H. D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R. D.; Liu, M. Q.; Chen, Y.; Shen, X. R.; Wang, X.; Zheng, X. S.; Zhao, K.; Chen, Q. J.; Deng, F.; Liu, L. L.; Yan, B.; Zhan, F. X.; Wang, Y. Y.; Xiao, G. F.; Shi, Z. L. 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.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng, Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J.; Wang, D.; Xu, W.; Holmes, E. C.; Gao, G. F.; Wu, G.; Chen, W.; Shi, W.; Tan, W. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, *395*, 565−574.
(13) Tortorici, M. A.; Veesler, D. Structural insights into coronavirus entry. *Adv. Virus Res.* **2019**, *105*, 93−116.
(14) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. 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) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215−220.
(21) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221−224.
(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.; Cameroni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; Zambon, M.; Rey, F. A.; Corti, D.; Veesler, D. 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.; Zhang, X.; Gao, G. F. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Commun.* **2017**, *8*, 15092.
(25) 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.
(26) 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.
(27) 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.
(28) Cai, Y.; Zhang, J.; Xiao, T.; Peng, H.; Sterling, S. M.; Walsh, R. M.; Rawson, S.; Rits-Volloch, S.; Chen, B. Distinct conformational states of SARS-CoV-2 spike protein. *Science* **2020**, *369* (6511), 1586−1592.
(29) Hsieh, C. L.; Goldsmith, J. A.; Schaub, J. M.; DiVenere, A. M.; Kuo, H. C.; Javanmardi, K.; Le, K. C.; Wrapp, D.; Lee, A. G.; Liu, Y.; Chou, C. W.; Byrne, P. O.; Hjorth, C. K.; Johnson, N. V.; Ludes-Meyers, J.; Nguyen, A. W.; Park, J.; Wang, N.; Amengor, D.; Lavinder, J. J.; Ippolito, G. C.; Maynard, J. A.; Finkelstein, I. J.; McLellan, J. S. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **2020**, *369*, 1501.
(30) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S. M. C.; Kopp, M.; Li, D.; Parks, R.; Hsu, A. L.; Borgnia, M. J.; Haynes, B. F.; Acharya, P. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* **2020**, DOI: 10.1038/s41594-020-0479-4.
(31) 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**, DOI: 10.1038/s41594-020-0483-8.
(32) Xiong, X.; Qu, K.; Ciazynska, K. A.; Hosmillo, M.; Carter, A. P.; Ebrahimi, S.; Ke, Z.; Scheres, S. H. W.; Bergamaschi, L.; Grice, G. L.; Zhang, Y.; Nathan, J. A.; Baker, S.; James, L. C.; Baxendale, H. E.; Goodfellow, I.; Doffinger, R.; Briggs, J. A. G. A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* **2020**, DOI: 10.1038/s41594-020-0478-5.
(33) 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.
(34) Yang, T.-J.; Chang, Y.-C.; Ko, T.-P.; Draczkowski, P.; Chien, Y.-C.; Chang, Y.-C.; Wu, K.-P.; Khoo, K.-H.; Chang, H.-W.; Hsu, S.-T. D. 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.
(35) Watanabe, Y.; Berndsen, Z. T.; Raghwani, J.; Seabright, G. E.; Allen, J. D.; Pybus, O. G.; McLellan, J. S.; Wilson, I. A.; Bowden, T. A.; Ward, A. B.; Crispin, M. Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat. Commun.* **2020**, *11*, 2688.
(36) 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*, eabb9983.
(37) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield: implications for immune recognition. *Sci. Rep.* **2020**, *10*, 14991.
(38) Woo, H.; Park, S. J.; Choi, Y. K.; Park, T.; Tanveer, M.; Cao, Y.; Kern, N. R.; Lee, J.; Yeom, M. S.; Croll, T. I.; Seok, C.; Im, W. Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane. *J. Phys. Chem. B* **2020**, *124*, 7128−7137.
(39) Jo, S.; Qi, Y.; Im, W. Preferred conformations of N-glycan core pentasaccharide in solution and in glycoproteins. *Glycobiology* **2015**, *26*, 19−29.

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

(40) 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.
(41) Amaro, R. E.; Li, W. W. Molecular-level simulation of pandemic influenza glycoproteins. *Methods Mol. Biol.* **2012**, *819*, 575−594.
(42) Casalino, L.; Gaieb, Z.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; Fadda, E.; Amaro, R. E. Shielding and Beyond: The Roles of Glycans in SARS-CoV-2 Spike Protein. *bioRxiv* **2020**, DOI: 10.1101/2020.06.11.146522.
(43) Kalathiya, U.; Padariya, M.; Mayordomo, M.; Lisowska, M.; Nicholson, J.; Singh, A.; Baginski, M.; Fahraeus, R.; Carragher, N.; Ball, K.; Haas, J.; Daniels, A.; Hupp, T. R.; Alfaro, J. A. Highly Conserved Homotrimer Cavity Formed by the SARS-CoV-2 Spike Glycoprotein: A Novel Binding Site. *J. Clin. Med.* **2020**, *9*, 1473.
(44) 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**, DOI: 10.1021/acs.jproteome.0c00273.
(45) Brielle, E. S.; Schneidman-Duhovny, D.; Linial, M. The SARS-CoV-2 Exerts a Distinctive Strategy for Interacting with the ACE2 Human Receptor. *Viruses* **2020**, *12*, 497.
(46) Spinello, A.; Saltalamacchia, A.; Magistrato, A. Is the Rigidity of SARS-CoV-2 Spike Receptor-Binding Motif the Hallmark for Its Enhanced Infectivity? Insights from All-Atom Simulations. *J. Phys. Chem. Lett.* **2020**, *11*, 4785−4790.
(47) 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.
(48) del Sol, A.; Tsai, C. J.; Ma, B.; Nussinov, R. The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure* **2009**, *17*, 1042−1050.
(49) Tsai, C. J.; Nussinov, R. A unified view of "how allostery works. *PLoS Comput. Biol.* **2014**, *10*, No. e1003394.
(50) Dokholyan, N. V. Controlling Allosteric Networks in Proteins. *Chem. Rev.* **2016**, *116*, 6463−6487.
(51) Wodak, S. J.; Paci, E.; Dokholyan, N. V.; Berezovsky, I. N.; Horovitz, A.; Li, J.; Hilser, V. J.; Bahar, I.; Karanicolas, J.; Stock, G.; Hamm, P.; Stote, R. H.; Eberhardt, J.; Chebaro, Y.; Dejaegere, A.; Cecchini, M.; Changeux, J. P.; Bolhuis, P. G.; Vreede, J.; Faccioli, P.; Orioli, S.; Ravasio, R.; Yan, L.; Brito, C.; Wyart, M.; Gkeka, P.; Rivalta, I.; Palermo, G.; McCammon, J. A.; Panecka-Hofman, J.; Wade, R. C.; Di Pizio, A.; Niv, M. Y.; Nussinov, R.; Tsai, C. J.; Jang, H.; Padhorny, D.; Kozakov, D.; McLeish, T. Allostery in Its Many Disguises: From Theory to Applications. *Structure* **2019**, *27*, 566−578.
(52) Goldenberg, O.; Erez, E.; Nimrod, G.; Ben-Tal, N. The ConSurf-DB: Pre-calculated evolutionary conservation profiles of protein structures. *Nucleic Acids Res.* **2009**, *37*, D323−D327.
(53) Glaser, F.; Rosenberg, Y.; Kessel, A.; Pupko, T.; Ben-Tal, N. The ConSurf-HSSP database: the mapping of evolutionary conservation among homologs onto PDB structures. *Proteins: Struct., Funct., Genet.* **2005**, *58*, 610−617.
(54) Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **2010**, *38*, W529−W533.
(55) Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **2016**, *44*, W344−W350.
(56) Ben Chorin, A.; Masrati, G.; Kessel, A.; Narunsky, A.; Sprinzak, J.; Lahav, S.; Ashkenazy, H.; Ben-Tal, N. ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. *Protein Sci.* **2020**, *29*, 258−267.
(57) Rozewicki, J.; Li, S.; Amada, K. M.; Standley, D. M.; Katoh, K. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* **2019**, *47*, W5−W10.
(58) Suzek, B. E.; Huang, H.; McGarvey, P.; Mazumder, R.; Wu, C. H. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **2007**, *23*, 1282−1288.
(59) Suzek, B. E.; Wang, Y.; Huang, H.; McGarvey, P. B.; Wu, C. H. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **2015**, *31*, 926−932.
(60) Finn, R. D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; Sonnhammer, E. L.; Tate, J.; Punta, M. Pfam: the protein families database. *Nucleic Acids Res.* **2014**, *42*, D222−D230.
(61) El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S. R.; Luciani, A.; Potter, S. C.; Qureshi, M.; Richardson, L. J.; Salazar, G. A.; Smart, A.; Sonnhammer, E. L. L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S. C. E.; Finn, R. D. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427−D432.
(62) Finn, R. D.; Miller, B. L.; Clements, J.; Bateman, A. iPfam: a database of protein family and domain interactions found in the Protein Data Bank. *Nucleic Acids Res.* **2014**, *42*, D364−D373.
(63) Kolinski, A. Protein modeling and structure prediction with a reduced representation. *Acta Biochim. Polym.* **2019**, *51*, 349−371.
(64) 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.
(65) 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.
(66) 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.
(67) 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.
(68) 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.
(69) Rose, P. W.; Prlic, A.; Altunkaya, A.; Bi, C.; Bradley, A. R.; Christie, C. H.; Costanzo, L. D.; Duarte, J. M.; Dutta, S.; Feng, Z.; Green, R. K.; Goodsell, D. S.; Hudson, B.; Kalro, T.; Lowe, R.; Peisach, E.; Randle, C.; Rose, A. S.; Shao, C.; Tao, Y. P.; Valasatava, Y.; Voigt, M.; Westbrook, J. D.; Woo, J.; Yang, H.; Young, J. Y.; Zardecki, C.; Berman, H. M.; Burley, S. K. The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res.* **2016**, *45*, D271−D281.
(70) Fiser, A.; Sali, A. ModLoop: Automated Modeling of Loops in Protein Structures. *Bioinformatics* **2003**, *19*, 2500−25001.
(71) Fernandez-Fuentes, N.; Zhai, J.; Fiser, A. ArchPRED: A Template Based Loop Structure Prediction Server. *Nucleic Acids Res.* **2006**, *34*, W173−W176.
(72) Karami, Y.; Guyon, F.; De Vries, S.; Tufféry, P. DaReUS-Loop: accurate loop modeling using fragments from remote or unrelated proteins. *Sci. Rep.* **2018**, *8*, 13673.
(73) Rotkiewicz, P.; Skolnick, J. Fast procedure for reconstruction of full-atom protein models from reduced representations. *J. Comput. Chem.* **2008**, *29*, 1460−1465.
(74) Lombardi, L. E.; Marti, M. A.; Capece, L. CG2AA: backmapping protein coarse-grained structures. *Bioinformatics* **2016**, *32*, 1235−1237.
(75) 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.
(76) 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.
(77) 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.
(78) Eyal, E.; Lum, G.; Bahar, I. The anisotropic network model web server at 2015 (ANM 2.0). *Bioinformatics* **2015**, *31*, 1487−1489.

(79) 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.
(80) 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.
(81) Atilgan, C.; Atilgan, A. R. Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLoS Comput. Biol.* **2009**, *5*, No. e1000544.
(82) Atilgan, C.; Gerek, Z. N.; Ozkan, S. B.; Atilgan, A. R. Manipulation of conformational change in proteins by single-residue perturbations. *Biophys. J.* **2010**, *99*, 933−943.
(83) General, I. J.; Liu, Y.; Blackburn, M. E.; Mao, W.; Gierasch, L. M.; Bahar, I. ATPase subdomain IA is a mediator of interdomain allostery in Hsp70 molecular chaperones. *PLoS Comput. Biol.* **2014**, *10*, No. e1003624.
(84) Dutta, A.; Krieger, J.; Lee, J. Y.; Garcia-Nafria, J.; Greger, I. H.; Bahar, I. Cooperative Dynamics of Intact AMPA and NMDA Glutamate Receptors: Similarities and Subfamily-Specific Differences. *Structure* **2015**, *23*, 1692−1704.
(85) Brown, D. K.; Penkler, D. L.; Sheik Amamuddy, O.; Ross, C.; Atilgan, A. R.; Atilgan, C.; Tastan Bishop, O. MD-TASK: a software suite for analyzing molecular dynamics trajectories. *Bioinformatics* **2017**, *33*, 2768−2771.
(86) Penkler, D.; Sensoy, O.; Atilgan, C.; Tastan Bishop, O. Perturbation-Response Scanning Reveals Key Residues for Allosteric Control in Hsp70. *J. Chem. Inf. Model.* **2017**, *57*, 1359−1374.
(87) Penkler, D. L.; Atilgan, C.; Tastan Bishop, O. Allosteric Modulation of Human Hsp90alpha Conformational Dynamics. *J. Chem. Inf. Model.* **2018**, *58*, 383−404.
(88) Penkler, D. L.; Tastan Bishop, Ö. Modulation of Human Hsp90α Conformational Dynamics by Allosteric Ligand Interaction at the C-Terminal Domain. *Sci. Rep.* **2019**, *9*, 1600.
(89) Stetz, G.; Tse, A.; Verkhivker, G. M. Dissecting Structure-Encoded Determinants of Allosteric Cross-Talk between Post-Translational Modification Sites in the Hsp90 Chaperones. *Sci. Rep.* **2018**, *8*, 19.
(90) Abdizadeh, H.; Atilgan, C. Predicting long term cooperativity and specific modulators of receptor interactions in human transferrin from dynamics within a single microstate. *Phys. Chem. Chem. Phys.* **2016**, *18*, 7916−7926.
(91) Jalalypour, F.; Sensoy, O.; Atilgan, C. Perturb-Scan-Pull: A Novel Method Facilitating Conformational Transitions in Proteins. *J. Chem. Theory Comput.* **2020**, *16*, 3825−3841.
(92) Brinda, K. V.; Vishveshwara, S. A Network Representation of Protein Structures: Implications for Protein Stability. *Biophys. J.* **2005**, *89*, 4159−4170.
(93) Vijayabaskar, M. S.; Vishveshwara, S. Interaction Energy Based Protein Structure Networks. *Biophys. J.* **2010**, *99*, 3704−3715.
(94) 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.
(95) Czemeres, J.; Buse, K.; Verkhivker, G. M. Atomistic Simulations and Network-Based Modeling of the Hsp90-Cdc37 Chaperone Binding with Cdk4 Client Protein: A Mechanism of Chaperoning Kinase Clients by Exploiting Weak Spots of Intrinsically Dynamic Kinase Domains. *PLoS One* **2017**, *12*, No. e0190267.
(96) 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.
(97) 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.
(98) Floyd, R. W. Algorithm 97: Shortest Path. *Commun. ACM* **1962**, *5*, 345.
(99) 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, CA; Varoquaux, G., Vaught, T., Millman, J., Eds.; 2008; pp 11−15.
(100) Girvan, M.; Newman, M. E. Community Structure in Social and Biological Networks. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 7821−7826.
(101) 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.
(102) 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.
(103) Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **2006**, *74*, 036104.
(104) Kalinka, A. T.; Tomancak, P. linkcomm: an R package for the generation, visualization, and analysis of link communities in networks of arbitrary size and type. *Bioinformatics* **2011**, *27*, 2011−2012.
(105) 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.
(106) 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.
(107) 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.
(108) 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.
(109) Shah, M.; Ahmad, B.; Choi, S.; Woo, H. G. Sequence variation of SARS-CoV-2 spike protein may facilitate stronger interaction with ACE2 promoting high infectivity. *Research Square* **2020**, DOI: 10.21203/rs.3.rs-16932/v1.
(110) Armijos-Jaramillo, V.; Yeager, J.; Muslin, C.; Perez-Castillo, Y. SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals undiscovered amino acids necessary for complex stability. *Evol. Appl.* **2020**, *13*, 2168.
(111) 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.
(112) Popovych, N.; Sun, S.; Ebright, R. H.; Kalodimos, C. G. Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **2006**, *13*, 831−838.
(113) Tzeng, S. R.; Kalodimos, C. G. Dynamic activation of an allosteric regulatory protein. *Nature* **2009**, *462*, 368−372.
(114) Huang, C.; Kalodimos, C. G. Structures of Large Protein Complexes Determined by Nuclear Magnetic Resonance Spectroscopy. *Annu. Rev. Biophys.* **2017**, *46*, 317−336.
(115) Jiang, Y.; Kalodimos, C. G. NMR Studies of Large Proteins. *J. Mol. Biol.* **2017**, *429*, 2667−2676.
(116) Madu, I. G.; Roth, S. L.; Belouzard, S.; Whittaker, G. R. Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide. *J. Virol.* **2009**, *83*, 7411−21.
(117) Ponzoni, L.; Bahar, I. Structural dynamics is a determinant of the functional significance of missense variants. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 4164−4169.

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608

(118) Stetz, G.; Astl, L.; Verkhivker, G. M. Exploring Mechanisms of Communication Switching in the Hsp90-Cdc37 Regulatory Complexes with Client Kinases through Allosteric Coupling of Phosphorylation Sites: Perturbation-Based Modeling and Hierarchical Community Analysis of Residue Interaction Networks. *J. Chem. Theory Comput.* **2020**, *16*, 4706−4725.

https://dx.doi.org/10.1021/acs.jproteome.0c00654
J. Proteome Res. 2020, 19, 4587−4608