

The Discovery of a Putative Allosteric Site in the SARS-CoV-2 Spike Protein Using an Integrated Structural/Dynamic Approach

Luisa Di Paola,* Hamid Hadi-Alijanvand, Xingyu Song, Guang Hu,* and Alessandro Giuliani

Cite This: <https://dx.doi.org/10.1021/acs.jproteome.0c00273>

Read Online

ACCESS |

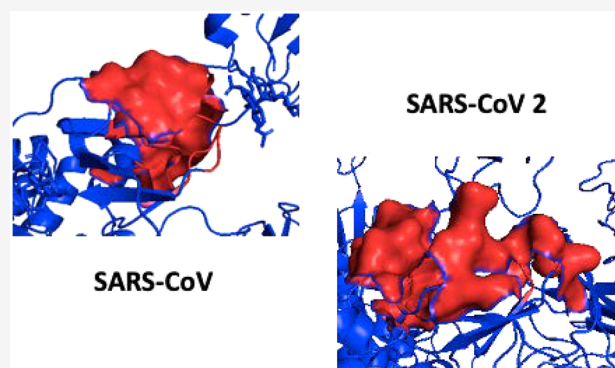
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: SARS-CoV-2 has caused the largest pandemic of the twenty-first century (COVID-19), threatening the life and economy of all countries in the world. The identification of novel therapies and vaccines that can mitigate or control this global health threat is among the most important challenges facing biomedical sciences. To construct a long-term strategy to fight both SARS-CoV-2 and other possible future threats from coronaviruses, it is critical to understand the molecular mechanisms underlying the virus action. The viral entry and associated infectivity stems from the formation of the SARS-CoV-2 spike protein complex with angiotensin-converting enzyme 2 (ACE2). The detection of putative allosteric sites on the viral spike protein molecule can be used to elucidate the molecular pathways that can be targeted with allosteric drugs to weaken the spike-ACE2 interaction and, thus, reduce viral infectivity. In this study, we present the results of the application of different computational methods aimed at detecting allosteric sites on the SARS-CoV-2 spike protein. The adopted tools consisted of the protein contact networks (PCNs), SEPAS (Affinity by Flexibility), and perturbation response scanning (PRS) based on elastic network modes. All of these methods were applied to the ACE2 complex with both the SARS-CoV2 and SARS-CoV spike proteins. All of the adopted analyses converged toward a specific region (allosteric modulation region [AMR]), present in both complexes and predicted to act as an allosteric site modulating the binding of the spike protein with ACE2. Preliminary results on hepcidin (a molecule with strong structural and sequence with AMR) indicated an inhibitory effect on the binding affinity of the spike protein toward the ACE2 protein.

KEYWORDS: COVID-19, spike protein, ACE2 binding, protein contact networks, allosteric drugs, binding patch softness, elastic networks modeling, drug discovery



INTRODUCTION

The recent pandemic caused by SARS-CoV-2 (COVID-19) has posed the most challenging global health and sanitation concern of the twenty-first century, with 6,663,304 confirmed cases of infected and 392,802 confirmed fatalities (World Health Organization, updated on June 6th, 2020¹). Moreover, this pandemic has caused a widespread lockdown, affecting more than 1 in 2 individuals in the world.

The virus responsible for COVID-19, SARS-CoV-2, emerged in late 2019, in the region of Wuhan, China.² SARS-CoV-2 belongs to the family *Coronaviridae*, which is characterized by large, enveloped, positive-stranded RNA viruses that can potentially cause gastrointestinal, nervous system, and respiratory distress.³ This family also includes SARS-CoV and MERS-CoV,⁴ which, despite several similarities, present some interesting differences from SARS-CoV2.^{4,5}

SARS-CoV-2, MERS-CoV, and SARS-CoV use an envelope protein, termed the spike (S) protein, to gain access to host cells.^{4,6} The S protein displays the highest degree of genetic variability in the virus genome.⁶ Additionally, the SARS-CoV-2 S

protein exhibits a greater affinity for the ACE2 receptor with respect to analogs in SARS-CoV and MERS-CoV.^{7,8} The SARS-CoV-2 S protein shares a 96.2% sequence similarity with the bat SARS-like RaTG13 S protein, indicating the natural origin of the virus.^{5,9}

Due to its crucial role in the very first stage of viral infection, several studies^{7,10–14} point to the S protein as an elective target to develop coronaviruses therapies and vaccines (the S protein is also the main target of neutralizing antibodies^{15–17}).

The S protein belongs to the viral fusion class I proteins,⁷ has a trimeric, crown-like shape, protrudes from the virus envelope, and targets diverse host cell receptors in different species (see Table 1 in ref 4). Both SARS and SARS-like (MERS and SARS-

Special Issue: Proteomics in Pandemic Disease

Received: April 22, 2020

Published: June 17, 2020

CoV-2) coronaviruses target the angiotensin-converting-enzyme 2 (ACE2) present in pneumocytes and enterocytes in the respiratory system and other susceptible cells (e.g., intestinal mucosal, renal tubular epithelial, cerebral neurons, and immune cells).¹⁸ The S-ACE2 interaction occurs at the very first stage of viral entry to initiate viral fusion with the host cell.^{19,20} The hallmark of SARS-CoV-2 is that its S protein is optimized for ACE2 binding.⁹

The S protein comprises two subunits, termed S1 and S2, which are cleaved during the first stage of viral infection. The S1 subunit is responsible for the interaction of the S protein with the host receptor (ACE2), while S2 promotes viral fusion.²¹ The receptor binding domain (RBD) resides between residues 318 and 513²² in SARS-CoV-2 and 318 and 510 in SARS-CoV.¹³

The development of therapeutics targeting the S-ACE2 complex requires a comprehensive study of this complex interface and its associated dynamics.²³ Thus, investigating the allosteric nature of the S-ACE2 complex will help to understand and quantify the extent by which the S protein binds to ACE2, and therefore the molecular basis of virus infectivity.

The identification of allosteric modulation regions (AMRs) allows for the design and testing of allosteric drugs (i.e., drugs targeting a different site from the complex interface) with enhanced bioactivity with respect to orthosteric drugs targeting the protein–protein interaction interface,²⁴ which are often tightly packed and scarcely accessible. The field of allosteric drugs is a rapidly developing field in drug discovery and may play a crucial role in the therapeutic strategy for SARS-CoV-2.^{25–28}

In this study, we adopted a multifaceted computational approach to identify promising “druggable” allosteric sites of viral S protein. We first applied the method of protein contact networks (PCNs) to the S-ACE2 complex in SARS-CoV-2 and SARS-CoV. In previous studies, we adopted this method to investigate both protein–protein interactions^{29–31} and the allosteric nature of binding.^{32,33}

In this work, the application of PCNs clearly highlighted a region in the SARS-CoV and SARS-CoV-2 S protein, which acts as the putative *locus* for allosteric modulation (AMR). AMR was further investigated in terms of its relevance in the stability of the complex and ability to transmit modulating signals to the binding site by means of independent computational approaches.

Specifically, we adopted a monomer-based approach³⁴ which can predict the affinity of the S protein to its ligand. In this step, the ensembles for the ground state of the structures have been produced using the anisotropic network model (ANM) approach.³⁵ Moreover, we adopted an unsupervised blind procedure to predict possible interaction sites on S proteins and their affinities for tentative partners using a softness-based prediction of intersubunits affinity (SEPAS). In addition, the dynamical difference in RBD between the S proteins of the two virus strains suggested that they may have different binding and allosteric properties.

We also provided biophysical evidence based on the elastic network modeling (ENM) approach, combined with perturbation-response scanning (PRS)³⁶ that AMRs in both viruses acted as a mediator of intermolecular allostery between the S protein and ACE2. “The three methods converged in allosteric character for residues in AMR in both complexes. This allowed us to state that residues in AMR for SARS-CoV-2 AMR are more susceptible to allosteric drug targeting than for SARS-CoV.

A recent study³⁷ suggested that the many residues in the AMR may be one of the most efficient epitopes for antibody recognition.

Preliminary docking studies individuated some molecules that bind to AMR (hepcidin), which were independently shown to exhibit strong sequence similarities with the AMR.³⁸

MATERIALS AND METHODS

The analyzed structures consist of the following: complex SARS-spike glycoprotein-human ACE2 complex (stabilized variant, all ACE2-bound particles, PDB code 6CS2,³⁹ termed the SARS-CoV S/ACE2 complex) and the SARS-CoV-2 analogous spike glycoprotein-human ACE2 complex⁴⁰.

Protein Contact Network Methods

Purposed software was used to transform the full structural information in the PDB files into a protein contact network (PCN): the network nodes are the amino acid residues represented by α -carbons. Links between nodes (residues) exist if the mutual distance of the residues (centered on α -carbons) were in the range between 4 and 8 Å, thus including only significant (<8 Å) noncovalent (>4 Å) bonds, while discarding “obliged” contacts due to proximity along the sequence.

The adjacency matrix **A** mathematically represented the PCN in terms of undirected, unweighted network and is defined as

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

The topological role of nodes (residues) addressed the functional role at the corresponding residues, based on the value of network descriptors. The method is widely discussed elsewhere.⁴¹

The basic network descriptor was the node degree k_i , defined as

$$k_i = \sum_j \mathbf{A}_{ij} \quad (2)$$

To detect allosteric sites and, more generally, the functional regions activating upon binding, we adopted a method based on network spectral clustering.^{42,43} Spectral clustering was based on the Laplacian matrix spectral decomposition, where the Laplacian matrix **L** is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \quad (3)$$

where **D** is the degree matrix (i.e., a diagonal matrix whose diagonal is the degree vector).

The spectral clustering was based on the eigenvalue decomposition of Laplacian **L**. The sign of the Fiedler vector (the eigenvector corresponding to the second minor eigenvalue) was used for binary partition (each network was parted into two clusters, which in turn could be parted into four, and so on), as outlined in refs 43 and 44. We previously demonstrated that network clusters (group of residues) correspond to functional regions in the protein.⁴³

Once network nodes (again, protein residues) were parted into the given number of clusters (an independent user-defined parameter for the clustering software), it was possible to define the participation coefficient P

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

where k_{si} is the degree of the i th residue computed in the cluster it belongs (i.e., accounting only for links with nodes pertaining to the same cluster).

This network descriptor has been demonstrated to represent the functional role of residues in binding and stability.^{33,41,43,45} Residues with a high P value are responsible for the communication between clusters (functional regions) and are thus addressed via the role in allosteric communication.⁴⁶ In this study, the participation coefficient maps were projected onto ribbon protein structures (as heat maps), to highlight activating hotspots in the spike-ACE2 complex. Participation coefficient maps are visualized by PyMol (<https://pymol.org/2/>).

We characterized the spike protein/ACE2 interface by means of descriptors stemming from the structural information and the network analysis described above:

- the interchain degree k_i^{IC} was defined as the node degree, but was only computed over contacts between nodes (residues) belonging to different chains

The interface roughness Q/R^{29} was defined for each chain participating in an interface, Q was the number of chain residues in the interface and R , the sequence range of chain residues in the interface.

- the interface amino acid range²⁹ IAR = R/N , where N was the total number of residues in the chain
- the interface energy matrix, \mathbf{E} , defined as

$$\mathbf{E} = \mathbf{E}_{ij} = \begin{cases} e_{ij} = \frac{1}{d_{ij}} & \text{if } 4 \text{ \AA} < d_{ij} < 8 \text{ \AA} \\ & \text{and the residues belong to different chains} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

the interface energy E_{INT} is the sum of e_{ij} and $\langle E_{\text{INT}} \rangle$ is the value averaged over the whole number of residues at the interface.

Elastic Network Modeling of ANM and GNM

Elastic network models (ENMs) provide efficient methods for investigating the intrinsic dynamics and allosteric communication pathways in proteins. We adopted two elastic network models: (1) anisotropic network model (ANM); and (2) Gaussian network model (GNM).

In ANM,⁴⁷ the interaction potential for a protein of N residues was:

$$V_{\text{ANM}} = \frac{\gamma}{2} \sum_{ij}^N \left(|R_{ij}| - |R_{ij}^0| \right)^2 \quad (6)$$

In ANM, the $3N \times 3N$ Hessian matrix, \mathbf{H} , determined the systems dynamics.

\mathbf{H} generic element was

$$\mathbf{H}_{ij} = \begin{bmatrix} \frac{\partial^2 V}{\partial X_i \partial X_j} & \frac{\partial^2 V}{\partial X_i \partial Y_j} & \frac{\partial^2 V}{\partial X_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Y_i \partial X_j} & \frac{\partial^2 V}{\partial Y_i \partial Y_j} & \frac{\partial^2 V}{\partial Y_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Z_i \partial X_j} & \frac{\partial^2 V}{\partial Z_i \partial Y_j} & \frac{\partial^2 V}{\partial Z_i \partial Z_j} \end{bmatrix} \quad (7)$$

where X_i , Y_i , and Z_i represented the Cartesian components of the i th residue, V was the potential energy of the system. We selected $r_c = 13 \text{ \AA}$.

ANMs provide information about the amplitudes, as well as the direction of residue fluctuations. The similarity between two ANM modes, u_k and v_l , evaluated for proteins with two different conformations, can be quantified in terms of the inner product of their eigenvectors:

$$O(u_k, v_l) = u_k \cdot v_l \quad (8)$$

GNM³⁶ was based on the description of the protein structure as a network of C_α connected by springs of uniform force constant γ if they were located within cutoff distance r_c . In GNM, the interaction potential for a protein of N residues was

$$V_{\text{GNM}} = -\frac{\gamma}{2} \left[\sum_{i=1}^{N-1} \sum_{j=i+1}^N (R_{ij} - R_{ij}^0) \cdot (R_{ij} - R_{ij}^0) \Gamma_{ij} \right] \quad (9)$$

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

$$\Gamma_{ij} = \begin{cases} -1 & i \neq j, R_{ij} \leq r_c \\ 0 & i \neq j, R_{ij} > r_c \\ -\sum_{i:i \neq j} \Gamma_{ij} & i = j \end{cases} \quad (10)$$

Thus, the square fluctuations were

$$\begin{aligned} \langle (\Delta R_i)^2 \rangle &= (3kT/\gamma) \cdot [\Gamma^{-1}]_{ii} \text{ and } \langle \Delta R_i \cdot \Delta R_j \rangle \\ &= (3kT/\gamma) \cdot [\Gamma^{-1}]_{ij} \end{aligned} \quad (11)$$

The normal modes were extracted by eigenvalue decomposition of the matrix $\Gamma = U\Lambda U^T$, with U being the orthogonal matrix whose k th column u_k was the k th mode eigenvector. Λ was the diagonal matrix of eigenvalues λ_k .

The perturbation response scanning (PRS) approach was based on the linear response theory and allowed evaluation of residue displacements in response to external forces.⁴⁸ Our study performed a PRS analysis based on GNM by constructing the inverse Laplacian matrix, Γ^{-1} .

The N -dimensional vector ΔR of node displacements in response to the application of a perturbation (a N -dimensional force vector \mathbf{F}) obeys Hooke's law ($\mathbf{F} = \mathbf{H} \cdot \Delta R$).

The purpose of PRS was to exert a force of a given magnitude on the network, one residue at a time, and F for other residues not being perturbed, equal to zero. Thus, the force exerted on residue i was expressed as

$$F^{(i)} = (F_x^{(i)}, F_y^{(i)} F_x^{(i)})^T \quad (12)$$

and the resulting response was

$$\Delta R^{(i)} = H^{-1} F^{(i)} \quad (13)$$

$\Delta R^{(i)}$ was an N -dimensional vector that described the deformation of all the residues in response to $F^{(i)}$. A metric for the response of residue k was the magnitude $\langle \|\Delta R_k^{(i)}\|^2 \rangle$ of the k th block of $\Delta R^{(i)}$ averaged over multiple $F^{(i)}$, expressed as the ik th element of the $N \times N$ PRS matrix, S_{PRS} . The elements of S_{PRS} referred to the unit (or uniform) perturbing force. The response to unit deformation at each perturbation site was obtained by dividing each row by its diagonal value

$$\bar{S}_{\text{PRS}} = \begin{pmatrix} 1/d_1 & & 0 \\ & \ddots & \\ 0 & & 1/d_N \end{pmatrix} S_{\text{PRS}} \quad (14)$$

The average effect of the perturbed effector site i on all other residues was computed by averaging over all sensors (receivers) residues j and can be expressed as $\langle (\Delta R^i)^2 \rangle_{\text{sensor}}$. The effector profile $\langle (\Delta R^i)^2 \rangle_{\text{effector}}$ described the average effect that the local perturbation in the effector site i had on all other residues. The maxima along the effector and sensor profiles corresponded to the functional mobile residues that underwent allosteric structural change.

Because the crystallized structure presented one snapshot of the protein dynamic life at the bottom of a potential well, we created a trajectory of protein dynamics in its ground state by setting the RMSD cutoff to 5 Å using the ANM^{36,47} approach implemented in Prody.⁴⁹ To predict the most possible protein binding patches (PBPs) on spike proteins, we utilized ISPRED to predict the interface residues.⁵⁰ The ISPRED-predicted interface residues acted as a seed to create a possible PBP based on the procedure as previously described.³⁴ The predictions of monomer-based affinities based on the mechanical softness of PBPs were performed by utilizing SEPAS ver. 1.³⁴ We introduced ALA mutations into the Covid-spike using FoldX suite.⁵¹

Additional Computational Methods

We computed the general interface properties of the spike/ACE2 complex using PISA software,⁵² specifically the interface area, S_{INT} , and the energy gain upon the interface formation, ΔG_{INT} .

RESULTS AND DISCUSSION

Protein Contact Network Results

Figure 1 reports the cluster partition (2 clusters) of the SARS-CoV S/ACE2 complex (Figure 1B). The chains are also shown

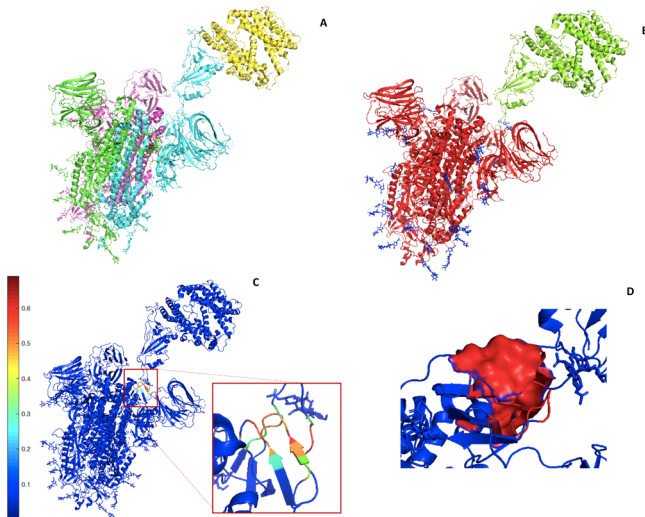


Figure 1. SARS-CoV S/ACE2 complex (PDB code 6cs2³⁹). (A) Partition into two clusters (green cluster 1, comprising the ectodomain of ACE2 and the fusion peptide, in red cluster 2, the remaining part of the spike protein); (B) chains in the complex: in green chain A, in cyan chain B, carrying the fusion peptide, in magenta chain C, in yellow the ACE2 ectodomain.

to highlight which chains participate in the two clusters (Figure 1A).

The interface between the fusion peptide and the ACE2 ectodomain was so strong (in terms of number of contacts between viral and host proteins) that the clustering partition algorithm recognized the fusion+ACE2 region as a single cluster, even though it comprised sequences belonging to different chains.

The map of the participation coefficient projected onto the ribbon structure of the SARS-CoV/ACE2 complex (Figure 1C) shows an active region ($P > 0$) in the junction between the fusion peptide and the trimeric bulk phase of the spike protein. Active residues are shown more in greater detail in red in Figure 1D. As previously demonstrated, the participation coefficient describes the attitude of residues toward participating in the intercommunication between clusters; thus, it was a putative score of the involvement of the residues in allosteric communication. Therefore, this region was a good candidate to intervene in the allosteric regulation of complex formation and there was value in investigating allosteric drugs that target this portion of the protein.

Figure 2 refers to the analogous analysis for the SARS-CoV2 S/ACE2 complex.

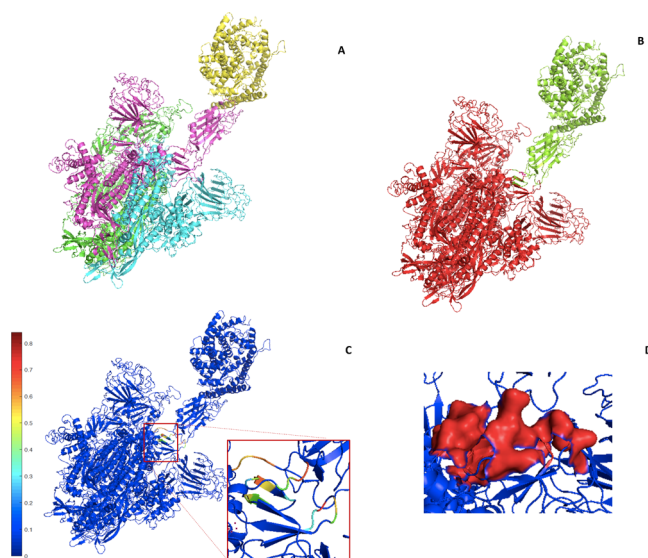


Figure 2. SARS-CoV2 S/ACE2 complex: (A) Partition into two clusters (green cluster 1, comprising the ectodomain of ACE2 and the fusion peptide, in red cluster 2, the remaining part of the spike protein); (B) chains in the complex: in green the chain A, in cyan the chain B, in magenta the chain C, carrying the fusion peptide, in yellow the ACE2 ectodomain.

A similar active region appeared at the junction of the fusion peptide to the body of the spike protein; however, the active region in the SARS CoV S/ACE2 was more compact than that of SARS-CoV2 S/ACE2. Active regions in both complexes were characterized by two β -sheets and unfolded traits, and could thus be targeted to peptides due to flexibility of the region and absence of a pocket to bind small organic molecules.⁵³

Table 1 reports the values of the participation coefficient P in residues with $P > 0$, all of which were located in the chain carrying the fusion peptides of the SARS-CoV S/ACE2 (chain B, in cyan in Figure 1A) and SARS-CoV2 S/ACE2 (chain C, in magenta in Figure 1A) complexes.

Table 1. Values of the Participation Coefficient P for the SARS-CoV S/ACE2 and nCoV2019 S/ACE2 Complexes (Only Positive Values Are Reported, with Corresponding Position, in Italic Values Larger than 0.6)

SARS CoV S/ACE2		SARS-CoV2 S/ACE2	
aa	P	aa	P
312D	0.31	320Q	0.61
313V	0.56	321P	0.56
314 V	0.69	322T	0.64
315R	0.47	323E	0.67
316F	0.23	531N	0.75
516S	0.31	532L	0.75
517T	0.64	533V	0.44
518D	0.64	534 K	0.31
519L	0.56	537C	0.15
520I	0.36	538V	0.44
527F	0.19	539N	0.70
528N	0.51	540F	0.51
529N	0.61	542F	0.21
530N	0.64	546T	0.31
532L	0.36	547G	0.36
563R	0.21	548T	0.51
564D	0.27	549G	0.44
565P	0.61	576R	0.21
566 K	0.56	577D	0.21
567T	0.56	578P	0.84
568S	0.64	580T	0.56
-	-	581L	0.36

The SARS CoV S/ACE2 and SARS-CoV2 S/ACE2 complexes accounted for 21 and 22 active residues, respectively. The average value was similar for the two distributions (0.47 for SARS CoV S/ACE2, 0.48 for SARS-CoV2 S/ACE2). However, the absolute values of higher P residues (>0.6) were higher for COVID19 S/ACE2, indicative of a corresponding larger reactivity of the active region.

Characterization of the S-ACE2 Interface

Table 2 reports the properties of interface S/ACE2 in the two complexes. The interface area and the absolute value of energy of

Table 2. Properties of the Spike Protein/ACE2 Interface

	SARS-CoV SA/ACE2	SARS-CoV2 S/ACE2
IAR_{SP}	0.01	0.008
IAR_{ACE}	0.62	0.56
$(Q/R)_{SP}$	0.45	0.37
$(Q/R)_{ACE}$	0.05	0.04
E_{INT}	5.37	4.35
$\langle E_{INT} \rangle$	0.17	0.17
$n_{IC} = Q_{SP} + Q_{ACE}$	31	26
$S_{INT}, \text{Å}^2$	962.2	739.9
$\Delta G_{INT}, \text{kcal/mol}$	-7.9	-7.4
$(\frac{\Delta G_{INT}}{\Delta S_{INT}}, \text{cal}/(\text{mol Å}^2))$	8.21	10.00

the SARS CoV S/ACE2 complex were higher than the corresponding area for the SARS-CoV-2 S/ACE2 complex; however, the specific values for the unit area and single residues were higher for SARS-CoV-2 S/ACE2, indicating a more efficient yet less stable interface.

This implied an optimized complex formation with the ACE2 of SARS-CoV-2, as observed in ref 9. This helped to explain the

more rapid kinetics of infection, mediated by a faster complex formation with main receptor, ACE2.

The RBDs for the two proteins are reported as a wide peptide, which comprised approximately 200 residues for both structures; however, a more limited core region, termed the receptor binding motif (RBM), was claimed to actively participate in the complex formation of the spike protein with its receptor, ACE2. We independently identified the residues in the spike protein for SARS-CoV and SARS-CoV2 using the method of PCNs. The position and the value of the interchain degree are reported in Table 3 for both complexes. The two residues in the two structures with the highest value of interface degree were 488G in SARS-CoV and 502G in SARS-CoV-2, respectively.

Table 3. Residues in the Interface S/ACE2, with Corresponding Interface Degree

SARS-CoV S/ACE2		SARS-CoV2 S/ACE2	
Position	k_{IC}	Position	k_{IC}
462P	3	475A	2
463D	2	476G	1
464G	1	487N	2
473N	3	489Y	1
475Y	1	493Q	1
479N	3	500T	4
482G	1	501N	4
486T	4	502G	8
487T	4	503V	3
488G	6	504G	1
489I	4	505Y	2
490G	1	-	-
491Y	2	-	-

These two residues could be considered the hotspots of the PPI interface in the S/ACE2 complex; both are glycine residues, likely due to the high steric adaptability of the small glycine residues in the most tightly packed region of the S-ACE2 interface. This confirmed the high density of the interface and its extremely scarce accessibility to small molecules targeting the region, which also suggested the need to interfere with the S/ACE2 interface through allosteric regulation.

Spike Binding Affinity and the Allosteric Modulation Region (AMR)

In the previous sections, the AMRs of the S/ACE2 complexes have been identified for SARS-CoV-2 and SARS-CoV using the PCNs approach (Table 1). The AMR is accessible for binding, which is at odds with the PBD.

One critical property of the protein complexes was the intersubunit affinity. Therefore, besides the S/ACE2 stability, we predicted the affinity between the AMR and incoming-designed molecule (peptide molecules), which will be designed to affect the function of AMR.

This was a crucial step that will provide an independent proof-of-concept of both the allosteric features of AMR (by definition, an allosteric site must “sense” the microenvironment) and the druggability of the predicted AMR.

We reviewed the need for complete 3D knowledge of the S/ACE2 complex by means of a recently introduced method to predict the intersubunit affinity of protein complexes using protein binding patch (PBP) on a single subunit by computing the PBP mechanical softness.^{34,54}

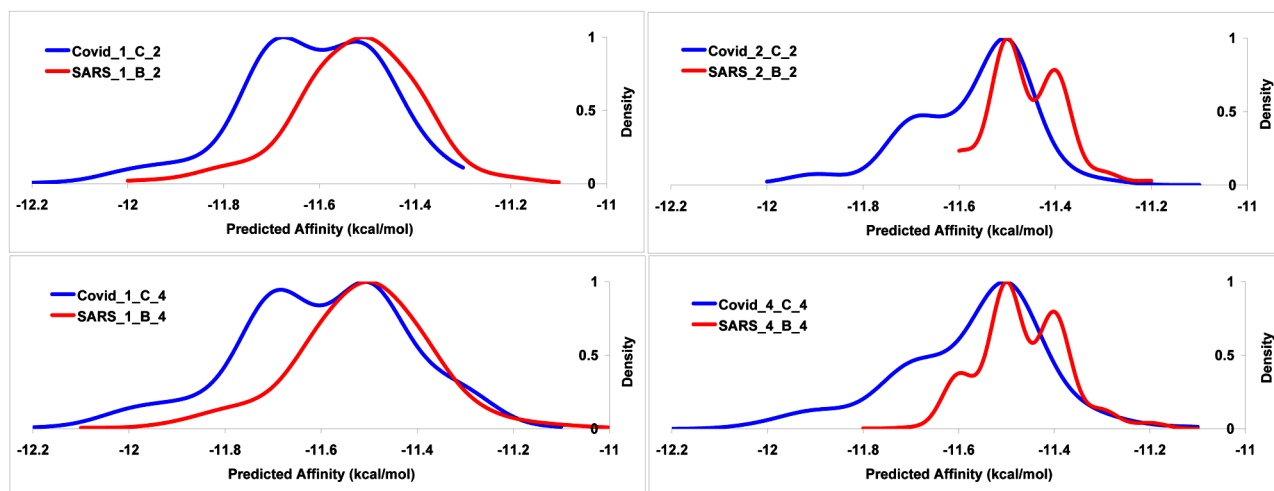


Figure 3. Predicted affinities of spikes to ACE2 are presented. The affinity of SARS-CoV 2 spike (blue curve) and SARS-CoV spike (red curve) to ACE2 is predicted using SARS-CoV 2/SARS_X_C/B_Y summarized the state of spike proteins in affinity prediction: X denotes number of chain in predictions. $X = 1$ means S1 protein in monomer form is considered, $X = 2$ means S1 + ACE2 complex, and $X = 4$ means whole spike complex + ACE2. Y notes the considered PBP in computations: $Y = 2$ means the PBP's residues of S1 in S1-ACE2 complex is used to predict the affinity and $Y = 4$ means the PBP's residues of S1 in whole spike proteins-ACE2 complex is used to predict the affinity between S1 and ACE2. Horizontal axis presents predicted affinity (kcal/mol). Normalized density denotes in Y-axis.

To alleviate the static presentation of the protein structure in a crystallized state, we created ensembles of protein complexes in their ground state using the ANM approach. In the first step, we predicted the affinity of the SARS-CoV-2 and SARS-CoV S proteins for ACE2.

We compared the predicted affinity between the S protein and ACE2 when the S protein was in its monomeric or trimeric state to obtain some hints about the effect of adjacent subunits on the affinity between spike and its ligand (Figure 3). Subunits (chains) are denoted as SA, SB, and SC.

To manage the potential states of the spike protein, we propose different forms of PBPs on the S protein:

- The SA subunit is alone and waiting for its partner (left column, Figure 3). We suppose that SA may interact with ACE2 using PBPs similar to those in the ensemble of S/ACE2 in its dimeric state (the first row, Figure 3) or by using PBPs similar to ones SA displays in the trimer spike and ACE2 ensemble (second row, Figure 3)
- Subunit SA finds ACE2 and binds it directly (right column, Figure 3). SA binds to ACE2 via the PBPs that it displays, either when no SB subunit is available (the first row, Figure 3) or in association with SB monomers (second row, Figure 3).

Our computations suggested that SARS-CoV-2 SA (dark blue curves, Figure 3) was more sensitive to the presence of the SB subunit than SARS-CoV SA (red curves, Figure 3). This result suggested approaches that potentially disrupted the SA-SB association may provide an efficient therapeutic route.

We observed a higher affinity between S and ACE2 in SARS-CoV-2 in SARS-CoV. This observation may explain the higher infectivity of SARS-CoV-2.

We are interested in predicting the affinities between possible PBPs on spike proteins and tentative protein partners. In this case, we did not know the PBPs on spike proteins because we did not have the resolved 3D structures of all possible protein complexes between the spike protein and other natural or synthesized partners. We predicted the seed interface residues on the spike surface using the ISPRED approach.⁵⁰ By

expanding the selection radius for defining the PBP region,³⁴ many possible PBPs were generated for both SARS-CoV-2 and SARS-CoV spike proteins. Using SEPAS, the affinities between the generated PBPs and unknown partners were predicted.

The affinities averaged over the whole are presented in Figure 4a and b (for SARS-CoV-2 and SARS-CoV spikes, respectively); the average affinities mapped on the surface of the spike proteins are reported in Supporting Information Figures 1 and 2. Some of the predicted regions with an affinity < -5 kcal/mol resided in the interface region of the spike subunits. There were some residues in the proposed PBPs that belonged to AMR (indicated as red triangles in Figure 4). This meant that AMR could act as a distinct PBP for some partners, suggesting that targeting AMR may be a good choice for drug design.

Thus, in considering AMR as distinct PBP, we were able to

1. Predict the affinities between AMR and unknown drug-like protein partners, simply by using a 3D structure of AMR on the spike subunit in different ensembles (monomer, case "C" in Figure 5a)
2. Account for the effect of ACE2 binding on AMR affinity for its possible partners. We assumed the spike monomer in association with ACE2 then predicted the affinity between AMR and its possible partners (dimer, case "DC" in Figure 5a)
3. Address the effect of the SB spike chains on the affinity of AMR in the SA spike chain for its partners, feeding SEPAS with the trimeric state of spikes (trimer, case "ABC" in Figure 5a)
4. Analyze the condition of a trimer spike and ACE2 in terms of a holo-complex (S/ACE2 complex, case "ABCD" in Figure 5a)

These results indicated that the AMR of SARS-CoV-2 spike exhibited a higher affinity for generic peptides than the AMR of SARS-CoV spike. Interestingly, the affinity of AMR for their partners was increased upon binding to ACE2, which suggested an interplay between the two distant regions, the ACE2-binding site of spike and AMR. This last result could be considered to be another proof-of-concept of the allosteric character of the AMR.

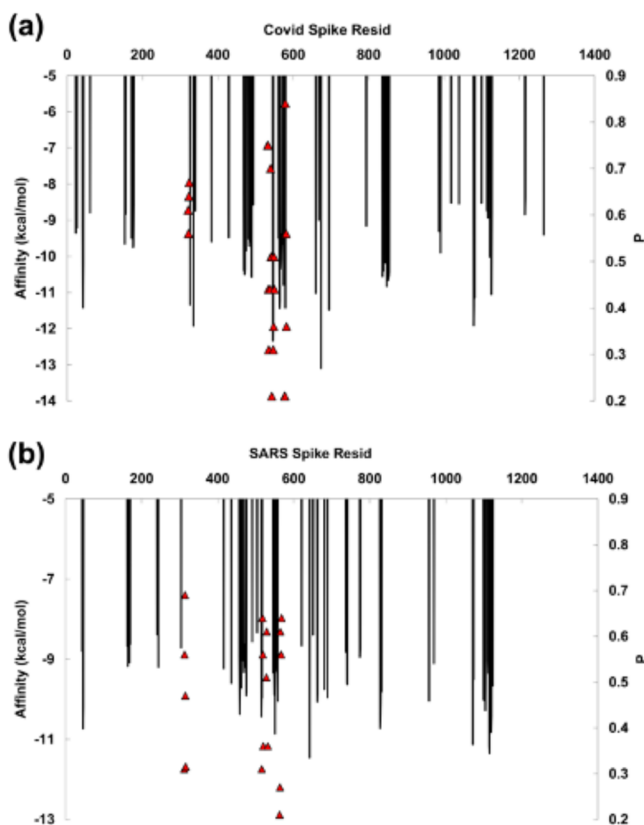


Figure 4. Affinity of possible PBPs on the S1 subunit of the spike to tentative proteins are predicted. Possible PBPs are defined for the S1 subunit. The affinity of the PBPs on SARS-CoV2 (a) and SARS-CoV (b) S1 protein for their possible partners is presented. Left vertical axis points to the averaged predicted affinity (kcal/mol) of the proposed PBPs. Right vertical axis denotes the P metric (Table 1) of the AMR participant residues. X-axis notes the residue of the PBP's central residue. Red triangles represent the position of AMR residues on the S1 protein sequence.

To further investigate the effect of the induced structural perturbation in AMR on the binding affinity of the spike protein to ACE2, we introduced two sets of ALA mutations in AMR. In one set, the AMR residues that play a major role in allosteric modulation ($P > 0.5$ in Table 1) were mutated to ALA, and in the other set, all residues of AMR on SARS-CoV-2 SA were mutated to ALA.

To simulate an SA perturbation, we used the ensemble of structures generated using the ANM approach after setting the RMSD cutoff to 8 Å for selecting normal mode states. Such global perturbation only distorts the ACE2-binding site of SA to less than 1.5 Å. In Figure 5b, we reported that wild-type SA exhibits a shoulder in its affinity density curve at a high affinity region. Our observations suggested that the structural perturbation of AMR due to the ALA mutation of native residues decreased the affinity of SA for ACE2. This was further confirmation of the allosteric properties of the AMR.

In order to explore AMR druggability, we started by recognizing the strict structural similarity of the AMR region with hepcidin, a crucial protein for iron regulation. Moreover, a recent preprint reported a distant (albeit relevant) sequence similarity between hepcidin and the SARS-CoV-2 spike protein.³⁸ It has been established that there was some proof of the similarity between the hydrophobicity spectra of ligand–

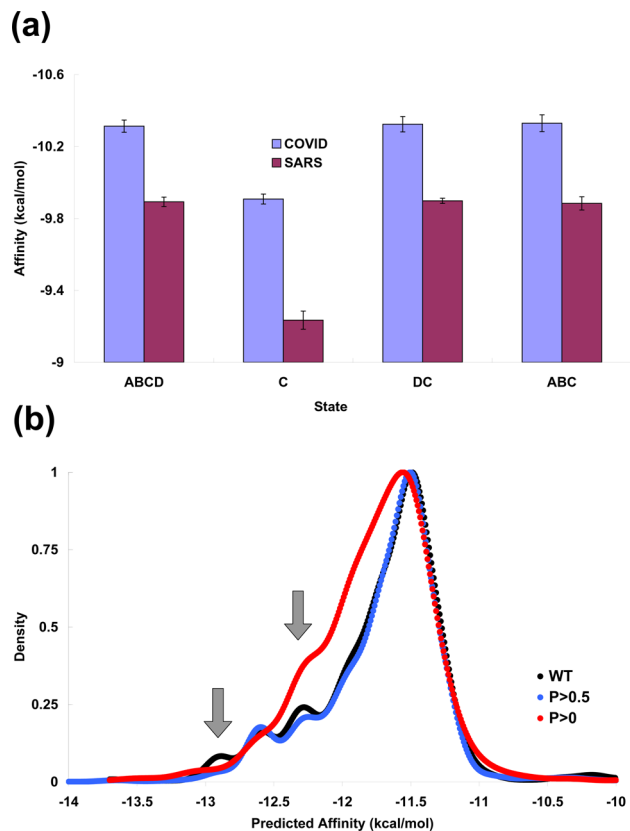


Figure 5. Affinity of AMR for tentative protein partners and the effect of AMR perturbation on ACE2 affinity are presented. (a) Affinity (kcal/mol) of AMR to possible partners are predicted when S1 is a monomer (state C), S1 is associated with ACE2 (state DC), S1 is accompanied by two S2 subunits (state ABC), and when the whole spike binds to ACE2 (state ABCD). (b) Affinity of S1 for ACE2 is predicted when AMR is intact (WT), residues of AMR with $P > 0.5$ are mutated to ALA ($P > 0.5$), or all AMR residues are mutated to ALA ($P > 0$). The affected region of affinity density curve is marked with vertical gray arrow. X-axis notes predicted affinity (kcal/mol) and vertical axis presents normalized density of affinities.

receptor pairs, which fostered the observed sequence similarity between hepcidin and the SARS-CoV-2 S-protein.^{55,56}

Moreover, hepcidin is strictly involved in interleukin 6 (IL-6) regulation,⁵⁷ given that IL-6 is the main actor associated with the “cytokine storm” that is linked to fatalities provoked by SARS-CoV-2 infection.⁵⁸ A preliminary docking analysis showed a perfect fit between hepcidin and the AMR region; however, this now must only be considered as a first step with which to prioritize different peptides as allosteric effectors.

Dynamical RBDs Comparison of SARS-CoV and SARS-CoV-2

In addition to the characterization of the S/ACE2 interface, a comparison of RBD dynamics between SARS-CoV and SARS-CoV-2 was also important to understand the functional mechanism of the S/ACE2 interaction. To this aim, the overlapping values between the slowest ANM modes (eigenvectors) were calculated to compare the intrinsic dynamics of RBDs in both S proteins (eq 8 in Materials and Methods). In Figure 6a, the dot products (overlap values) were plotted in the map for the RBDs in SARS-CoV versus SARS-CoV-2 for the first 10 ANM modes. In this map, the upper limit of 1 indicated a perfect overlap (red) correspondent to a perfectly shared dynamics, while 0 indicated no overlap (blue),

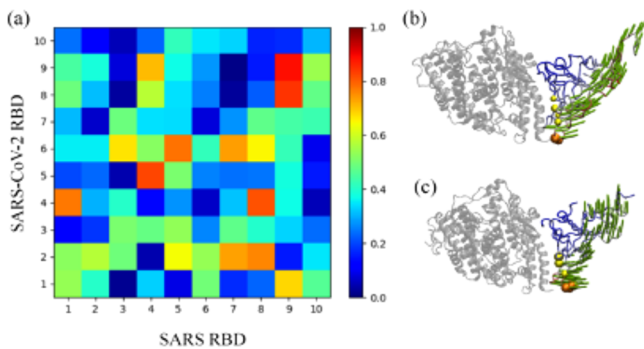


Figure 6. ANM results for RBDs. (a) Overlap map between the ten slowest ANM modes of SARS-CoV and SARS-CoV-2 RBDs, in which dark red and blue regions indicate strong and weak similar dynamics, respectively. The motions of the first ANM mode of RBDs in SARS-CoV (b) and SARS-CoV-2 (c). The color scale from red to blue changes from the highest to lowest flexibilities.

indicating that they showed independent dynamics. Although RBDs in the two complexes showed very high structural similarity with an RMSD of 0.68 Å, the overlap between the first 10 ANM modes of two RBDs exhibited some differences. Overlap values with ~ 0.6 for the first three modes indicated a weaker dynamical overlap, even at the lowest frequencies, which were those endowed with most important functional meaning.

To further understand the relative motions and dynamics of the interface structures, the ANM motions of the first mode for two RBDs are shown in Figure 6b and c. In ANM 1, most of the interfacial residues were very stable (blue), while the peripheral loop exhibited the highest fluctuation (red). The most relevant dynamical difference between the two systems referred to this loop motion. For SARS-CoV RBD, the highest fluctuations

point to residues 463D and 465 K, while the SARS-CoV-2 RBD displayed higher flexibilities at 476G to 479P (orange beads). This loop was found to be located at the C terminal of the $\alpha 1$ helix of the RBD, interacting with ACE2 through van der Waals forces, hereinafter referred to as an “interfacial C loop”. We suppose that the difference in the dynamics of such an interfacial C loop may be caused by the amino acid mutations in their respective interfaces with ACE2 (i.e., 442Y to 455L, 443L to F, 460F to 473Y, and 479N to 493Q) from SARS-CoV-RBD to SARS-CoV-2-RBD.⁵⁹ Thus, we further mapped these interfacial mutations onto the structures (yellow beads). Thus, the lower fluctuations of these mutations in SARS-CoV-2-RBD may account for the higher affinity of such interface.

Accordingly, despite the overall structural similarity, RBDs in the two S proteins showed different dynamics, especially for the S-ACE2 interfaces. These differences in dynamics may also indicate that these two S-ACE2 interfaces achieved their activities, including substrate binding and allosteric regulation.

PRS Analysis Identifies the Long-Range Allosteric Potential of the AMR

We performed perturbation response scanning (PRS) based on GNM to quantify the allosteric properties of AMRs and identify the key residues for which perturbation provokes structural dynamical changes at a long distance. First, the PRS map provided information regarding the sensitivity and effectiveness of a given residue in transmitting signals. The column of the matrix corresponding to a single residue (e.g., for r313V) described how well-connected that residue was within the AMR region and its likely relevance in transmitting allosteric signals to distal parts of the S-ACE2 complex.

Here, we adapted this method to probe the allosteric response of the AMR region upon perturbation on residues with $P > 0.5$ from the PCN prediction (Table 1).

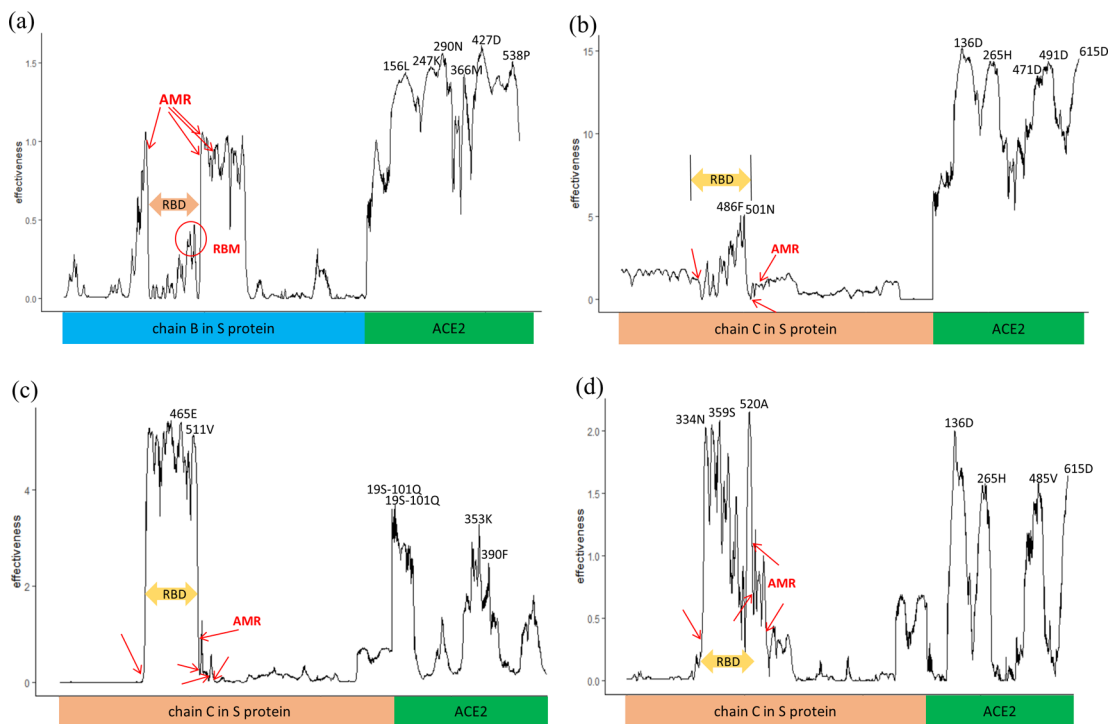


Figure 7. Allosteric effects obtained upon unit perturbation on the AMR residues, obtained from the PRS analysis. Effectiveness/influence profiles with respect to the linker residue (A) 568S in SARS-CoV AMR, and (B) 320Q, 531N, and 580T in SARS-CoV-2 AMR. Peaks indicate the most influential residues, while some important regions are highlighted.

For the S/ACE2 complex in SARS-CoV, the dynamical changes of the other residues in AMR (red arrows in Figure 7a) have been caught due to their close distance in space. In addition, all AMR residues have allosteric effects on the interfacial C loop with high fluctuation. Because the effectiveness of the three residues with $P > 0.5$ was almost identical, we only reported in Figure 7a for the case of the perturbation of 568S that had the highest P value (0.64). It is appreciated that three residues (313V, 529N, and 568S) displayed long-range allosteric effects on some particular regions in ACE2 through the S/ACE interface. The residues relative to the AMR in SARS-CoV-2 could be classified into three types according to their allosteric character.

The first class included most residues, including 320Q, 321P, 322T, 323E, 539N, 548T, and 578P. Upon perturbation, these residues exerted a moderate dynamical effect on other AMR residues, and a relatively weak but sensible intramolecular allosteric effect on residues of the S/ACE2 interface, especially on the highly flexible C loop. The perturbation of these AMR residues largely affected the dynamics of ACE2 through stronger long-range allosteric communications.

The perturbation of the second type of (531N and 532L) residues provokes a large intermolecular allosteric effect in the RBD, including some peaks at the interfacial C loop (Figure 7c). In addition, the perturbation weakened the intermolecular communications from AMR to ACE2.

The last type only included 580T, which exhibited similar dynamical effects on three regions (AMR, RBD, and ACE2) via long-range communications.

Accordingly, our PRS analysis provided important insight into AMR allostery. Among other findings, 3 residues in SARS-CoV AMR and 10 residues in SARS-CoV-2 AMR were predicted to have wide-ranging allosteric effects on the ACE2 protein. This result in conjunction with the statistically significant higher effectiveness of the SARS-CoV-2 AMR perturbation (SI Figure 3), suggested that its allosteric character was stronger than that of AMR in SARS-CoV. Notably, in both complexes, the interfacial C loop formed a bridge to transmit the signal from AMR to ACE2. The reported RBD-ACE2 interfacial mutations from SARS-CoV to SARS-CoV-2 were close to the interfacial C loop. Therefore, we propose that the different allosteric properties of AMR residues were due to mutations in the S-ACE2 interface.

In Figure S3, the box plot shows that AMR in SARS-CoV-2 had a statistically higher effectiveness than that of SARS-CoV (under the p -value of 7.4×10^{-5} by Wilcoxon signed-ranked test).

CONCLUSIONS

The application of PCNs, SEPAS (Affinity by Flexibility) and PRS based on Elastic Network modes to two complexes of the spike protein with its receptor, ACE2, generated highly consistent results. The integrated computational approach disclosed active zones in the SARS-CoV and SARS-CoV2 complexes with the ACE2 ectodomain. These active zones, as demonstrated in previous studies^{41,60,61} and in the present research by the convergence among different approaches, have an extremely high probability of acting as allosteric sites, which can modulate spike/ACE2 protein binding.

These regions are worthy of further investigation from the perspective of drug or vaccine development as indicated by a preliminary docking simulation. From a methodological perspective, our results highlighted the need to identify

convergence of different simulation methods arising from independent theoretical premises, in order to obtain a reliable picture of the associated biomolecular systems.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00273>.

Figure S1: Affinity for the SARS-CoV S/ACE2 complex.
Figure S2: Affinity for the SARS-CoV2 S/ACE2 complex.
Figure S3: Wilcoxon test for effectiveness (PDF)

AUTHOR INFORMATION

Corresponding Authors

Luisa Di Paola – Unit of Chemical-Physics Fundamentals in Chemical Engineering, Department of Engineering, Università Campus Bio-Medico di Roma, 00128 Rome, Italy; orcid.org/0000-0001-5329-8689; Email: l.dipaola@unicampus.it

Guang Hu – Center for Systems Biology, Department of Bioinformatics, School of Biology and Basic Medical Sciences, Soochow University, Suzhou 215123, China; orcid.org/0000-0002-8754-1541; Email: huguang@suda.edu.cn

Authors

Hamid Hadi-Alijanvand – Department of Biological Sciences, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran

Xingyu Song – Center for Systems Biology, Department of Bioinformatics, School of Biology and Basic Medical Sciences, Soochow University, Suzhou 215123, China

Alessandro Giuliani – Environmental and Health Department, Istituto Superiore di Sanità, 00161 Rome, Italy

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00273>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank the encouraging support and the fruitful discussion from Prof. C. Serangeli. The project was supported by National Natural Science Foundation of China (31872723) and the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions.

REFERENCES

- (1) Romanov, B. K. Coronavirus disease COVID-2019 (In Russ.). *Safety and Risk of Pharmacotherapy* **2020**, *8* (1), 3–8.
- (2) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727–733.
- (3) Chen, Y.; Liu, Q.; Guo, D. Emerging Coronaviruses: Genome Structure, Replication, and Pathogenesis. *J. Med. Virol.* **2020**, *92*, 418–423.
- (4) Belouzard, S.; Millet, J. K.; Licitra, B. N.; Whittaker, G. R. Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. *Viruses* **2012**, *4* (6), 1011–1033.
- (5) Benvenuto, D.; Giovanetti, M.; Ciccozzi, A.; Spoto, S.; Angeletti, S.; Ciccozzi, M. The 2019-new Coronavirus Epidemic: Evidence for Virus Evolution. *J. Med. Virol.* **2020**, *92*, 455.

- (6) Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* **2016**, *3* (1), 237–261.
- (7) Tai, W.; He, L.; Zhang, X.; Pu, J.; Voronin, D.; Jiang, S.; Zhou, Y.; Du, L. Characterization of the Receptor-Binding Domain (RBD) of 2019 Novel Coronavirus: Implication for Development of RBD Protein as a Viral Attachment Inhibitor and Vaccine. *Cell. Mol. Immunol.* **2020**, *17*, 613.
- (8) 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 (80-.)* **2020**, *1*.
- (9) Andersen, K. G.; Rambaut, A.; Lipkin, W. I.; Holmes, E. C.; Garry, R. F. The Proximal Origin of SARS-CoV-2. *Virological* **2020**, 1–7, DOI: 10.2106/JBJS.F.00094.
- (10) Wang, Q.; Wong, G.; Lu, G.; Yan, J.; Gao, G. F. MERS-CoV Spike Protein: Targets for Vaccines and Therapeutics. *Antiviral Res.* **2016**, *133*, 165–177.
- (11) Mustafa, S.; Balkhy, H.; Gabere, M. Peptide-Protein Interaction Studies of Antimicrobial Peptides Targeting Middle East Respiratory Syndrome Coronavirus Spike Protein: An In Silico Approach. *Adv. Bioinf.* **2019**, *2019*, 1.
- (12) Zhou, Y.; Yang, Y.; Huang, J.; Jiang, S.; Du, L. Advances in MERS-CoV Vaccines and Therapeutics Based on the Receptor-Binding Domain. *Viruses* **2019**, *11* (1), 1–18.
- (13) Du, L.; He, Y.; Zhou, Y.; Liu, S.; Zheng, B. J.; Jiang, S. The Spike Protein of SARS-CoV - A Target for Vaccine and Therapeutic Development. *Nat. Rev. Microbiol.* **2009**, *7* (3), 226–236.
- (14) Jiang, S.; Lu, L.; Liu, Q.; Xu, W.; Du, L. Receptor-Binding Domains of Spike Proteins of Emerging or Re-Emerging Viruses as Targets for Development of Antiviral Vaccines. *Emerging Microbes Infect.* **2012**, *1* (March), 1–8.
- (15) He, Y.; Zhou, Y.; Liu, S.; Kou, Z.; Li, W.; Farzan, M.; Jiang, S. Receptor-Binding Domain of SARS-CoV Spike Protein Induces Highly Potent Neutralizing Antibodies: Implication for Developing Subunit Vaccine. *Biochem. Biophys. Res. Commun.* **2004**, *324* (2), 773–781.
- (16) Ahmed, S. F.; Quadeer, A. A.; McKay, M. R. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* **2020**, *12* (3), 254.
- (17) Grifoni, A.; Sidney, J.; Zhang, Y.; Scheuermann, R. H.; Peters, B.; Sette, A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* **2020**, 1–10.
- (18) Gu, J.; Korteweg, C. Pathology and Pathogenesis of Severe Acute Respiratory Syndrome. *Am. J. Pathol.* **2007**, *170*, 1136–1147.
- (19) Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* **2016**, *3* (1), 237–261.
- (20) Li, W.; Moore, M. J.; Vasliava, N.; Sui, J.; Wong, S. K.; Berne, M. A.; Somasundaran, M.; Sullivan, J. L.; Luzuriaga, K.; Greeneugh, T. C.; Choe, H.; Farzan, M. Angiotensin-Converting Enzyme 2 Is a Functional Receptor for the SARS Coronavirus. *Nature* **2003**, *426* (6965), 450–454.
- (21) Letko, M.; Marzi, A.; Munster, V. Functional Assessment of Cell Entry and Receptor Usage for SARS-CoV-2 and Other Lineage B Betacoronaviruses. *Nat. Microbiol.* **2020**, *5* (4), 562–569.
- (22) Liu, Z.; Xiao, X.; Wei, X.; Li, J.; Yang, J.; Tan, H.; Zhu, J.; Zhang, Q.; Wu, J.; Liu, L. Composition and Divergence of Coronavirus Spike Proteins and Host ACE2 Receptors Predict Potential Intermediate Hosts of SARS-CoV-2. *J. Med. Virol.* **2020**, 1–7.
- (23) 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*, 1–9.
- (24) Wagner, J. R.; Lee, C. T.; Durrant, J. D.; Malmstrom, R. D.; Feher, V. A.; Amaro, R. E. Emerging Computational Methods for the Rational Discovery of Allosteric Drugs. *Chem. Rev.* **2016**, *116* (11), 6370–6390.
- (25) Peracchi, A.; Mozzarelli, A. Exploring and Exploiting Allostery: Models, Evolution, and Drug Targeting. *Biochim. Biophys. Acta, Proteomics* **2011**, *1814* (8), 922–933.
- (26) Csermely, P.; Nussinov, R.; Szilágyi, A. From Allosteric Drugs to Allo-Network Drugs: State of the Art and Trends of Design, Synthesis and Computational Methods. *Curr. Top. Med. Chem.* **2013**, *13* (1), 2–4.
- (27) Nussinov, R.; Tsai, C.-J. The Different Ways through Which Specificity Works in Orthosteric and Allosteric Drugs. *Curr. Drug Metab.* **2012**, *18* (9), 1311–1316.
- (28) Amamuddy, O. S.; Veldman, W.; Manyumwa, C.; Khairallah, A.; Agajanian, S.; Oluyemi, O.; Verkhivker, G. M.; Bishop, Ö. T. Integrated Computational Approaches and Tools for Allosteric Drug Discovery. *Int. J. Mol. Sci.* **2020**, *21* (3), 1 DOI: 10.3390/ijms21030847.
- (29) Di Paola, L.; Mei, G.; Di Venere, A.; Giuliani, A. Exploring the Stability of Dimers through Protein Structure Topology. *Curr. Protein Pept. Sci.* **2015**, *17* (1), 30–36.
- (30) Minicozzi, V.; Di Venere, A.; Nicolai, E.; Giuliani, A.; Caccuri, A. M.; Di Paola, L.; Mei, G. Non-Symmetrical Structural Behavior of a Symmetric Protein: The Case of Homo-Trimeric TRAF2 (Tumor Necrosis Factor-Receptor Associated Factor 2). *J. Biomol. Struct. Dyn.* **2020**, *1*.
- (31) Platania, C. B.; Paola, L. D.; Leggio, G. M.; Romano, G. L.; Drago, F.; Salomone, S.; Bucolo, C. Molecular Features of Interaction between VEGFA and Anti-Angiogenic Drugs Used in Retinal Diseases: A Computational Approach. *Front. Pharmacol.* **2015**, *6*, 1 DOI: 10.3389/fphar.2015.00248.
- (32) Cimini, S.; Di Paola, L.; Giuliani, A.; Ridolfi, A.; De Gara, L. GH32 Family Activity: A Topological Approach through Protein Contact Networks. *Plant Mol. Biol.* **2016**, *92* (4–5), 401–410.
- (33) Di Paola, L.; Giuliani, A. Protein Contact Network Topology: A Natural Language for Allostery. *Curr. Opin. Struct. Biol.* **2015**, *31*, 43–48.
- (34) Hadi-Alijanvand, H. Complex Stability Is Encoded in Binding Patch Softness: A Monomer-Based Approach to Predict Inter-Subunit Affinity of Protein Dimers. *J. Proteome Res.* **2020**, *19* (1), 409–423.
- (35) Gur, M.; Zomot, E.; Bahar, I. Global Motions Exhibited by Proteins in Micro- to Milliseconds Simulations Concur with Anisotropic Network Model Predictions. *J. Chem. Phys.* **2013**, *139* (12), 121912.
- (36) Yang, L.; Bahar, I.; Radar, A.; Chennubhotia, C. The Gaussian Network Model. In *Normal Mode Analysis*; Chapman and Hall/CRC: 2005; pp 41–64.
- (37) Pinto, D.; Park, Y.; Beltramello, M.; Walls, A. C.; Tortorici, M. A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; De Marco, A.; Peter, A.; Guarino, B.; Spreafico, R.; Camerini, E.; Case, J. B.; Chen, R. E.; Havenar-daughton, C.; Snell, G.; Telenti, A.; Virgin, H. W.; Lanzavecchia, A.; Diamond, M. S.; Fink, K.; Veesler, D.; Corti, D. Cross-Neutralization of SARS-CoV-2 by a Human Monoclonal SARS-CoV Antibody. *Nature* **2020**, *1* DOI: 10.1038/s41586-020-2349-y.
- (38) Ehsani, S. *Distant Sequence Similarity between Hepcidin and the Novel Coronavirus Spike Glycoprotein: A Potential Hint at the Possibility of Local Iron Dysregulation in COVID-19*; 2020.
- (39) Kirchdoerfer, R. N.; Wang, N.; Pallesen, J.; Wrapp, D.; Turner, H. L.; Cottrell, C. A.; Corbett, K. S.; Graham, B. S.; McLellan, J. S.; Ward, A. B.; Benvenuto, D.; Giovannetti, M.; Ciccozzi, A.; Spoto, S.; Angeletti, S.; Ciccozzi, M.; Jahn, R.; Lang, T.; Südhof, T. C.; Li, F.; Li, W.; Farzan, M.; Harrison, S. C.; Ahn, H.; Weaver, M.; Lyon, D.; Eunyong Choi, R. N.; Roger, B.; Fillingim, P.; Nomaguchi, M.; Fujita, M.; Miyazaki, Y.; Adachi, A.; Shulla, A.; Gallagher, T.; Walls, A. C.; Xiong, X.; Park, Y. J.; Tortorici, M. A.; Snijder, J.; Quispe, J.; Camerini, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; Zamboni, M.; Rey, F. A.; Corti, D.; Veesler, D. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **2018**, *176* (47), 1864–1868.
- (40) Zhang, Y.; Zheng, N.; Hao, P.; Cao, Y.; Zhong, Y. A Molecular Docking Model of SARS-CoV S1 Protein in Complex with Its Receptor, Human ACE2. *Comput. Biol. Chem.* **2005**, *29* (3), 254–257.
- (41) De Ruvo, M.; Giuliani, A.; Paci, P.; Santoni, D.; Di Paola, L. Shedding Light on Protein-Ligand Binding by Graph Theory: The

Topological Nature of Allostery. *Biophys. Chem.* **2012**, *165–166*, 21–29.

(42) Cimini, S.; Di Paola, L.; Giuliani, A.; Ridolfi, A.; De Gara, L. GH32 Family Activity: A Topological Approach through Protein Contact Networks. *Plant Mol. Biol.* **2016**, *92* (4–5), 401–410.

(43) Tasdighian, S.; Di Paola, L.; De Ruvo, M.; Paci, P.; Santoni, D.; Palumbo, P.; Mei, G.; Di Venere, A.; Giuliani, A. Modules Identification in Protein Structures: The Topological and Geometrical Solutions. *J. Chem. Inf. Model.* **2014**, *54* (1), 159–168.

(44) Vishveshwara, S.; Brinda, K. V.; Kannan, N. Protein Structure: Insights from Graph Theory. *J. Theor. Comput. Chem.* **2002**, *1* (1), 187–211.

(45) Giuliani, A.; Di Paola, L.; Setola, R. Proteins as Networks: A Mesoscopic Approach Using Haemoglobin Molecule as Case Study. *Curr. Proteomics* **2009**, *6* (4), 235–245.

(46) Tasdighian, S.; Di Paola, L.; De Ruvo, M.; Paci, P.; Santoni, D.; Palumbo, P.; Mei, G.; Di Venere, A.; Giuliani, A. Modules Identification in Protein Structures: The Topological and Geometrical Solutions. *J. Chem. Inf. Model.* **2014**, *54* (1), 159–168.

(47) Atilgan, A. R.; Durell, S. R.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model. *Biophys. J.* **2001**, *80* (1), 505–515.

(48) 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* (3), 933–943.

(49) Bakan, A.; Meireles, L. M.; Bahar, I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **2011**, *27* (11), 1575–1577.

(50) Savojardo, C.; Fariselli, P.; Martelli, P. L.; Casadio, R. ISPRED4: Interaction Sites PREDiction in Protein Structures with a Refining Grammar Model. *Bioinformatics* **2017**, *33* (11), 1656–1663.

(51) Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX Web Server: An Online Force Field. *Nucleic Acids Res.* **2005**, *33*, 1 DOI: [10.1093/nar/gki387](https://doi.org/10.1093/nar/gki387).

(52) Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372* (3), 774–797.

(53) Higuero, A. P.; Jubbe, H.; Blundell, T. L. Protein-Protein Interactions as Druggable Targets: Recent Technological Advances. *Curr. Opin. Pharmacol.* **2013**, *13* (5), 791–796.

(54) Hadi-Alijanvand, H. Soft Regions of Protein Surface Are Potent for Stable Dimer Formation. *J. Biomol. Struct. Dyn.* **2019**, 1–12.

(55) Mandell, A. J.; Owens, M. J.; Selz, K. A.; Neal Morgan, W.; Shlesinger, M. F.; Nemeroff, C. B. Mode Matches in Hydrophobic Free Energy Eigenfunctions Predict Peptide-Protein Interactions. *Biopolymers* **1998**, 1 DOI: [10.1002/\(sici\)1097-0282\(199808\)46:2<89::aid-bip4>3.3.co;2-d](https://doi.org/10.1002/(sici)1097-0282(199808)46:2<89::aid-bip4>3.3.co;2-d).

(56) Giuliani, A.; Benigni, R.; Zbilut, J. P.; Webber, C. L.; Sirabella, P.; Colosimo, A. Nonlinear Signal Analysis Methods in the Elucidation of Protein Sequence-Structure Relationships. *Chem. Rev.* **2002**, 1 DOI: [10.1021/cr0101499](https://doi.org/10.1021/cr0101499).

(57) Wrighting, D. M.; Andrews, N. C. Interleukin-6 Induces Hepcidin Expression through STAT3. *Blood* **2006**, *108*, 3204.

(58) Mehta, P.; McAuley, D. F.; Brown, M.; Sanchez, E.; Tattersall, R. S.; Manson, J. J. COVID-19: Consider Cytokine Storm Syndromes and Immunosuppression. *Lancet* **2020**, *395*, 1033.

(59) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural Basis for the Recognition of the SARS-CoV-2 by Full-Length Human ACE2. *Science* **2020**, *2762*, 1–10.

(60) Di Paola, L.; Platania, C. B. M.; Oliva, G.; Setola, R.; Pascucci, F.; Giuliani, A. Characterization of Protein-Protein Interfaces through a Protein Contact Network Approach. *Front. Bioeng. Biotechnol.* **2015**, 3.

(61) Cimini, S.; Di Paola, L.; Giuliani, A.; Ridolfi, A.; De Gara, L. GH32 Family Activity: A Topological Approach through Protein Contact Networks. *Plant Mol. Biol.* **2016**, *92* (4–5), 401–410.