

Some mechanistic underpinnings of molecular adaptations of SARS-COV-2 spike protein by integrating candidate adaptive polymorphisms with protein dynamics

Nicholas James Ose¹, Paul Campitelli¹, Tushar Modi¹, I Can Kazan¹, Sudhir Kumar^{2,3,4*}, Sefika Banu Ozkan^{1*}

¹Department of Physics and Center for Biological Physics, Arizona State University, Tempe, United States; ²Institute for Genomics and Evolutionary Medicine, Temple University, Philadelphia, United States; ³Department of Biology, Temple University, Philadelphia, United States; ⁴Center for Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia

*For correspondence: s.kumar@temple.edu (SK); banu.ozkan@asu.edu (SBO)

Competing interest: The authors declare that no competing interests exist.

Funding: See page 16

Sent for Review
29 August 2023

Preprint posted
15 September 2023

Reviewed preprint posted
06 November 2023

Reviewed preprint revised
03 April 2024

Version of Record published
07 May 2024

Reviewing Editor: Donald Hamelberg, Georgia State University, United States

© Copyright Ose et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract We integrate evolutionary predictions based on the neutral theory of molecular evolution with protein dynamics to generate mechanistic insight into the molecular adaptations of the SARS-COV-2 spike (S) protein. With this approach, we first identified candidate adaptive polymorphisms (CAPs) of the SARS-CoV-2 S protein and assessed the impact of these CAPs through dynamics analysis. Not only have we found that CAPs frequently overlap with well-known functional sites, but also, using several different dynamics-based metrics, we reveal the critical allosteric interplay between SARS-CoV-2 CAPs and the S protein binding sites with the human ACE2 (hACE2) protein. CAPs interact far differently with the hACE2 binding site residues in the open conformation of the S protein compared to the closed form. In particular, the CAP sites control the dynamics of binding residues in the open state, suggesting an allosteric control of hACE2 binding. We also explored the characteristic mutations of different SARS-CoV-2 strains to find dynamic hallmarks and potential effects of future mutations. Our analyses reveal that Delta strain-specific variants have non-additive (i.e., epistatic) interactions with CAP sites, whereas the less pathogenic Omicron strains have mostly additive mutations. Finally, our dynamics-based analysis suggests that the novel mutations observed in the Omicron strain epistatically interact with the CAP sites to help escape antibody binding.

eLife assessment

This **important** study investigates various variants of the SARS-COV-2 spike protein using established computational methods, complemented by experimental validation efforts. The evidence, bolstered by an evolutionary approach and protein dynamics, is **solid**. Placing this research in the broader context of the field could further enrich the article. It will interest biophysicists focused on allostery and protein evolution.

Introduction

Since 2019, the evolution of SARS-CoV-2 in humans has been characterized by the spread of mutations, many notably found within the spike (S) glycoprotein. The S protein is directly related to the human immune response to COVID-19 and, as such, has been one of the most studied and targeted

proteins in SARS-CoV-2 research (*Shang et al., 2020; Harvey et al., 2021; Jackson et al., 2022; Carabelli et al., 2023; Markov et al., 2023*). Subsequently, research into the biophysical properties and mutational patterns associated with S protein evolution not only remains critical to understanding the pandemic but also emerges as a useful system to understand the mechanics of molecular adaptation within viruses.

For successful infection of a human host, the S glycoprotein of SARS-CoV-2 binds to the human ACE2 (hACE2) receptor through its receptor-binding domain (RBD). Evidence indicates that fine-tuning S protein interactions with hACE2 significantly affects viral reproduction (*Rehman et al., 2020; Saputri et al., 2020; Rochman et al., 2021*). Previous evolutionary studies show a complex network of interactions among mutated residues (*Changeux and Edelman, 2005; Doshi et al., 2016; O'Rourke et al., 2016; Mishra and Jernigan, 2018*). Therefore, there has been a vast effort to uncover which mutations are important steps of adaptation for the S protein (*Cagliani et al., 2020; Damas et al., 2020; Singh and Yi, 2021; Kistler et al., 2022; Maher et al., 2022; Neher, 2022*). In particular, a significant aspect of many such studies was a focus on understanding adaptive mutations of SARS-CoV-2 that contributed to the leap to human hosts (*Cagliani et al., 2020; Damas et al., 2020; Singh and Yi, 2021; Starr et al., 2022c*). This is because SARS-CoV-2 has continuously mutated since its early detection (*Kistler et al., 2022*), causing the emergence of CDC-designated 'variants of concern' (VOCs) that may be driven by an accelerated substitution rate (*Tay et al., 2022*).

Predicting how new mutations impact the biophysical properties of the S protein remains a challenge, let alone explaining their complex interactions with one another and how they might affect hACE2 binding; many factors affect hACE2 interactions. Binding affinity with hACE2 can be enhanced directly through stronger receptor interactions or mediated through changes in RBD opening (*Teruel et al., 2021b; Zhang et al., 2021; Díaz-Salinas et al., 2022*). The RBD exhibits both 'closed' and 'open' conformational states. In the closed state, the RBD is shielded from receptor binding. In the open state, the RBD is accessible for hACE2 binding (*Kirchdoerfer et al., 2016; Gur et al., 2020; Henderson et al., 2020; Hoffmann et al., 2020*). While some mutations may affect the transition between these states (*Henderson et al., 2020; Yurkovetskiy et al., 2020; Gobeil et al., 2021a; Sztain et al., 2021; Zhang et al., 2021; Shoemark et al., 2022*), other mutations may allosterically regulate RBD openings through Furin cleavage site (residue ID range: 681–695) interactions to regulate hACE2 binding (*Deng et al., 2021; Gobeil et al., 2021b; Khan et al., 2021; Laiton-Donato et al., 2021*).

Moreover, as new mutations accumulate, culminating in the emergence of a new VOC, these mutations must occur on varied sequence backgrounds containing neutral, nearly-neutral, and adaptive mutations. While many studies have explored the impacts of individual mutations, VOCs result in a substantial difference in protein function compared to their individual effects (*Moulana et al., 2022; Starr et al., 2022c; Moulana et al., 2023; Witte et al., 2023*). Here we integrate an evolutionary approach with protein dynamics analysis to address the fundamental mechanisms of mutations dictating VOCs and the impact of their epistatic interaction on the function of the S protein. Many earlier studies have combined phylogeny and evolutionary theory to identify adaptive mutations and analyze how the viral sequence has changed over time (*Frost et al., 2018; Boni et al., 2020; Cagliani et al., 2020; Damas et al., 2020; Tang et al., 2020*). Similarly, we first use a well-established evolutionary probability (EP) approach (*Liu et al., 2016*) that utilizes phylogenetic trees in combination with the neutral theory of molecular evolution to determine candidate adaptive polymorphisms (CAPs) using the early Wuhan sequence as a variant. CAPs are substitutions in SARS-CoV-2 that are rarely observed in other closely related sequences (*Figure 1A*), which implies a degree of functional importance and makes them candidates for adaptation (*Liu et al., 2016*). Adaptation in this case means a virus which can successfully infect human hosts. As CAPs are unexpected polymorphisms under neutral theory, their existence implies a non-neutral effect. This can come in the form of functional changes (*Liu et al., 2016*) or compensation for functional changes (*Ose et al., 2022b*). Therefore, we suspect that these CAPs may be partially responsible for the functional change allowing the infection of human hosts. In support of this method, we find an overlap between our list of sites containing CAPs and putative adaptive sites identified by others (*Cagliani et al., 2020; Singh and Yi, 2021; Kistler et al., 2022; Starr et al., 2022b*). Second, we use a suite of computational tools to analyze how CAPs that arose in the early and late phases of the COVID-19 pandemic modulate the dynamics of the S protein. We also explore the complex interactions between these sets of CAPs to gain mechanistic

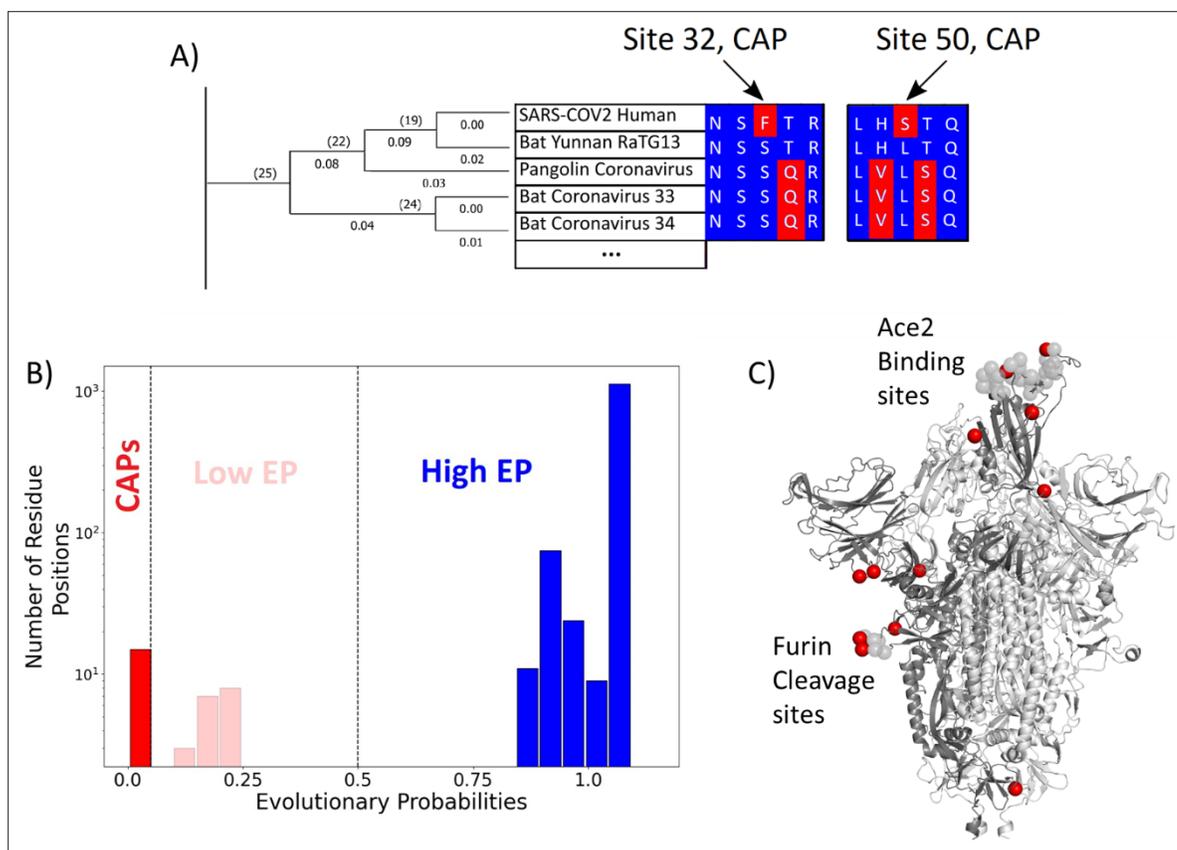


Figure 1. Evolutionary probability (EP) in SARS-CoV-2. **(A)** The EPs of each amino acid in the S protein sequence are calculated by taking the multiple sequence alignment of the S proteins through their evolutionary tree and using Bayesian inferences to determine the likelihood of finding a particular residue at a particular location within a given sequence. Simply, if the residue is found at a location 'x' in closely related sequences, it will have a higher EP at location 'x' in the target sequence. Residues with an EP < 0.05 in the target sequence are candidate adaptive polymorphisms (CAPs) (red). CAPs are found at sites 32, 50, 218, 346, 372, 478, 486, 498, 519, 604, 681, 682, 683, 684, and 1125. **(B)** The distribution of EP scores of the wild-type residues in the S protein. Here, lower EP scores are shown in red, and higher EP scores in blue. While the vast majority of the wild-type (reference) protein consists of high EP residues, a few residues have low EP. **(C)** The CAPs are also highlighted as red spheres in the open configuration of the S protein, with the open chain in a darker shade. We observe that a majority of the CAP positions reside at the receptor-binding domain (RBD) and the Furin cleavage site (676-689; [Wrobel et al., 2020](#)) shown as transparent light gray spheres.

insight into the behavior of molecular adaptation involving the S protein. In particular, we focused on how mutations modulate protein dynamics as we and others have previously found that rather than changing a protein's structure solely, mutations modulate conformational dynamics, leading to changes in biophysical properties such as stability, flexibility, and allosteric dynamic coupling, any of which may affect protein binding ([Swint-Kruse et al., 1998](#); [Keskin et al., 2000](#); [Bhabha et al., 2013](#); [Nussinov and Tsai, 2013](#); [Campbell et al., 2016](#); [Ma and Nussinov, 2016](#); [Saavedra et al., 2018](#); [Kuzmanic et al., 2020](#)).

With this evolutionary-dynamics unified approach, we aim to answer the following questions about VOCs: Are all the mutations in VOCs adaptive in nature? Are they coupled to one another or provide some measure of biophysical, dynamical, or mechanical epistasis? While many of these mutations are found within the RBD, numerous others are located distal to this region; hence, we aim to investigate the functional roles of distal mutations, particularly from a protein dynamics perspective. Our integrated analysis revealed that protein dynamics play a significant role in the evolution of the S protein. The flexibility of sites within the S protein shows a strong, direct correlation with substitution rate, and newly evolving CAPs are mostly additive mutations that modulate the dynamics of the hACE2 binding site. Yet other CAPs, 346R, 486F, and 498Q, show highly epistatic (i.e., non-additive) modulation of the hACE2 binding site and provide immune escape benefits.

Results and discussion

Candidate adaptive mutations in the S protein are recognized via EPs

SARS-CoV-2 is part of a family of coronaviruses, many of which infect mainly animals and are less capable of infecting humans (*Dicken et al., 2021*). Therefore, to identify the most likely mutations responsible for the infection of human hosts (i.e., putative adaptive mutations for humans), we estimated the (neutral) EP scores of mutations found within the S protein (*Liu et al., 2016*). EP scores of the amino acid variants of the S protein were obtained using a maximum likelihood phylogeny (*Kumar et al., 2018*) built from 19 orthologous coronavirus sequences. Sequences were selected by examining available non-human sequences with a sequence identity of 70% or above to the human SARS CoV-2's S protein sequence. This cutoff allows for divergence over evolutionary history such that each amino acid position had ample time to experience purifying selection, whilst limiting ourselves to closely related coronaviruses. (*Figure 1A*). The likelihood of finding a particular amino acid in the sequence is then determined using a Bayesian framework, with calculations carried out by MEGA X software (*Kumar et al., 2018*). As apparent in the name, EP scores obtained for the amino acids in the sequence provide information regarding the likelihood of finding them at their position, given the history of the sequence. Amino acid residues receiving low EP scores (<0.05) at a position are less likely to be found in a given position within the sequence because they are non-neutral. Generally, positions with low EP amino acids are far less common than those containing mutations with high EP, a trend also realized in the CoV-2 S protein (*Figure 1B*).

Of particular interest is an observed evolutionary change where an amino acid with high EP is replaced by an amino acid residue with low EP. While amino acids with low EP should be harmful or deleterious to viral fitness due to functional disruption or change, fixation of a low EP amino acid at a position suggests an underlying mechanism for natural selection to operate. These fixed, low EP mutations are called CAPs as they are predicted to alter protein function, and adaptive pressures may drive their prevalence (*Kumar and Patel, 2018*). Indeed, there is an overlap between these CAPs and the mutations suggested by other methods to be adaptive for the S protein (*Cagliani et al., 2020; Singh and Yi, 2021; Kistler et al., 2022; Starr et al., 2022b*). Within these studies, sites 478, 486, 498, and 681 have been implicated in SARS-CoV-2 evolution, leaving the remaining 11 CAPs as undiscovered candidate sites for adaptation.

Interestingly, most of the CAP residues are at functionally critical sites, including the RBD and the Furin cleavage site (*Figure 1C*). As mentioned earlier, the RBD plays a key role in initiating the infection of a healthy cell by binding with the host organism's ACE2 protein. Before ACE2 binding, one chain of the homotrimer comprising the S protein must open to expose the RBD (*Kirchdoerfer et al., 2016; Henderson et al., 2020; Hoffmann et al., 2020; Sztain et al., 2021*). The Furin cleavage site plays a key role in the opening process as the binding of host cell protease Furin aids in the cleavage of the S protein into two domains: S1 and S2 (*Wrobel et al., 2020: 13*). Another host cell protease, TMPRSS2, facilitates viral attachment to the surface of target cells upon binding either to sites Arg815/Ser816, or Arg685/Ser686 which overlap with the Furin cleavage site 676–689, further emphasizing the importance of this area (; *Fraser et al., 2022*). Similar cleavage sites have been found in related coronaviruses, including HKU1 and Middle East respiratory syndrome coronavirus (MERS-CoV), which infects humans (*Chan et al., 2008: 1; Millet and Whittaker, 2014; Millet and Whittaker, 2015*), and the acquisition of similar cleavage sites is associated with increased pathogenicity in other viruses such as the influenza virus (*Steinhauer, 1999*). Interestingly, however, CAPs do not display such an overwhelming tendency to occur at well-known critical sites within human proteins studied with similar methods (*Ose et al., 2022b*), yet mutations at those sites are associated with disease, indicating their critical role in inducing functional change. Therefore, the identified CAPs in the S protein, which are signs of recent evolution, can provide mechanistic insights regarding the molecular adaptation of the virus. In particular, we aimed to analyze how these CAP positions in the S protein modulate the interaction with hACE2 using our protein dynamics-based analysis (*Gerek and Ozkan, 2011; Nevin Gerek et al., 2013; Larrimore et al., 2017; Kumar et al., 2015*).

Asymmetry in communications among the network of interactions in spike describes how CAPs regulate the dynamics of the S protein

A mutation at a given amino acid position inevitably not only alters local interactions, but this change cascades through the residue–residue interaction network, which gives rise to a variation in native

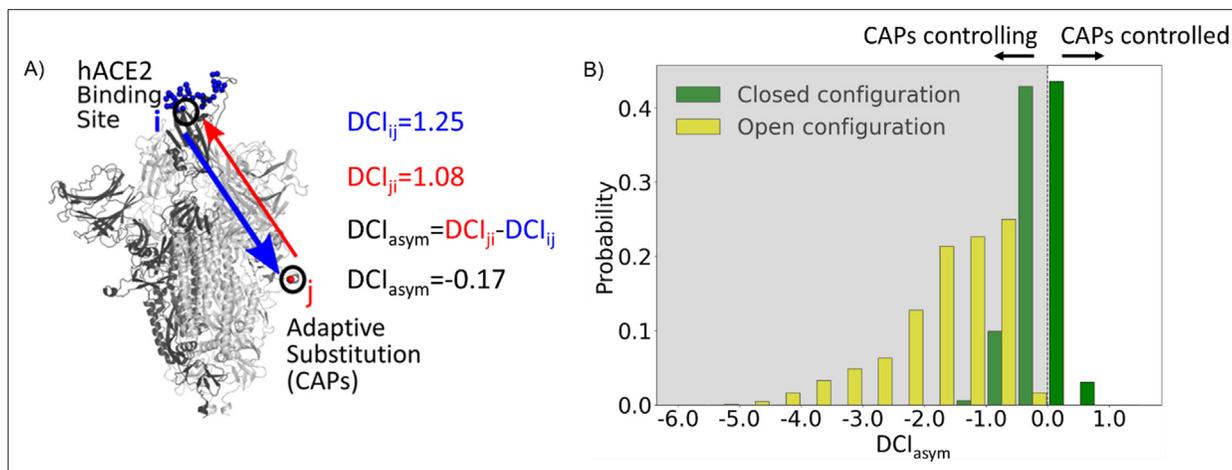


Figure 2. Candidate adaptive polymorphisms (CAPs) control binding sites across chains. **(A)** Schematic representation of DCI_{asym} . **(B)** DCI_{asym} of CAP residue positions with the binding interface of receptor-binding domain (RBD) in the open chain. Residues in the closed chains with a low evolutionary probability (EP) amino acid in the reference sequence dominate the binding site interface of RBD in the open chain. There is a significant difference between the asymmetry profiles of the closed ($M = -0.06$, $SD = 0.33$) and open ($M = -1.68$, $SD = 0.89$) conformations ($p < 0.001$).

ensemble dynamics to modulate function (Dror et al., 2012; Labbadia and Morimoto, 2015; Sekhar and Kay, 2019; Campitelli et al., 2020). Many groups have already examined the conformational dynamics of the S protein using normal mode analysis to explore mutation sites and interactions with different receptors (Zhou et al., 2020; Majumder et al., 2021; Teruel et al., 2021a; Verkhivker, 2022). However, our study will focus mainly on the role of CAPs, of which F486 and Q498 have already been identified through perturbation response scanning (PRS) as potential allosteric sites (Verkhivker, 2022). Thus, we analyze the internal dynamics of the system to understand the functional role of CAPs in S proteins. This analysis allows us to gain a mechanistic understanding of the relationship between CAP mutations and biophysical outcomes (Teruel et al., 2021b). First, we implement the dynamic coupling index (DCI) approach to study long-distance coupling between the CAPs and the hACE2 binding sites emerging from the 3D network of interactions across the S protein system. DCI calculation combines PRS and linear response theory (LRT) to capture the strength of a displacement response for position *i* upon perturbation of position *j*, relative to the average fluctuation response of position *i* to all other positions in the protein. It represents the strength of dynamic coupling between positions *i* and *j* upon perturbation to *j* (Larrimore et al., 2017; Kumar et al., 2015).

Further, asymmetry can be captured in the DCI values as dynamic coupling is not necessarily symmetric due to an anisotropic interaction network. That is, each amino acid has a set of positions to which it is highly coupled, and this anisotropy in connections gives rise to unique differences in coupling between a given *i, j* pair of amino acids which do not have direct interactions. By calculating the coupling of the hACE2 binding interface in the RBD with respect to the CAP residue positions and vice versa, we can generate DCI_{asym} (Figure 2A) as the difference between the normalized displacement response of position *j* upon a perturbation to position *i* (DCI_{ij}) and the normalized displacement response of position *i* upon a perturbation to position *j* (DCI_{ji}) (see 'Methods'). If the DCI_{asym} values significantly differ from zero, it shows asymmetry in coupling and presents a cause-effect relationship between the *i, j* pair in terms of force/signal propagation. This metric has been used previously in a variety of systems to analyze the unique behavior of positions within a protein and a given position's propensity to effect biophysical changes upon mutation, particularly at long distances (Modi and Ozkan, 2018; Campitelli and Ozkan, 2020; Kolbaba-Kartchner et al., 2021; Ose et al., 2022a; Kazan et al., 2023; Campitelli et al., 2021).

Recent work from our group has shown an enhancement in cross-chain communication within the main protease of SARS COV-2 compared to SARS COV-1 (Campitelli et al., 2022). Furthermore, previous studies have shown that allosteric inter-chain communication is important to S protein function (Zhou et al., 2020; Spinello et al., 2021; Tan et al., 2022; Xue et al., 2022). In support of these findings, we observe through DCI_{asym} that when the S protein is in its pre-fusion conformation with one chain open, the CAPs in the closed chains have negative coupling asymmetry with respect to the

hACE2 binding site interface in the RBD-open chain. This indicates an allosteric control where the hACE2 binding site is dominated by the dynamics of the CAPs in closed chains (**Figure 2B**, yellow bars). As this open-state RBD is critical for the viral infection of host cells (**Kirchdoerfer et al., 2016**), our results suggest that this type of closed-to-open cross-chain interaction is important for viral proliferation. Our prior studies on DCI_{asym} show a similar trend in lactose inhibitor (Lacl), a protein with a functional role in gene expression through binding DNA. The allosteric mutations (i.e., mutations on the sites that are far from the DNA binding sites) that alter DNA binding affinity not only exhibited unique asymmetry profiles with the DNA binding sites of Lacl, but also regulated the dynamics of these binding sites (**Campitelli et al., 2021**).

Similarly, it is possible that mutations to such residue positions within the S protein can be used to regulate the dynamics of the hACE2 bindings sites of the open RBD state. We, therefore, propose that the residue positions with CAP substitutions hold the potential for mutations in the spike sequence which can alter the opening and closing dynamics of the RBD. This hypothesis is further supported by many mutations already observed at these residue positions which alter the infection rate (**Brister et al., 2015**). Interestingly, residues responsible for extremely low asymmetry values (<-4) lie overwhelmingly in the region 476–486. These same residues were suggested to stabilize S protein dynamics and prime it for host Furin proteolysis (**Raghuvamsi et al., 2021**).

Moreover, as a control, we performed the same analysis on the S protein with the RBDs of all chains in the closed configuration. In this case, we observed that the DCI_{asym} of the CAPs residue positions with respect to the hACE2 interface in the other chains yields a largely symmetric distribution about 0 (**Figure 2B**, green bars). This verifies that the asymmetry in the coupling of CAPs with the exposed binding site interface in pre-fusion configuration results from one of the RBDs opening up and further suggests the allosteric role played by CAPs in locking the S protein in the RBD open state.

Dynamics analysis shows that rigid sites tend to be more highly conserved than flexible sites

CAPs represent important S protein amino acid changes between related coronaviruses across multiple species and the Wuhan-Hu-1 reference sequence (MN908947). Since SARS-CoV-2 first spread to

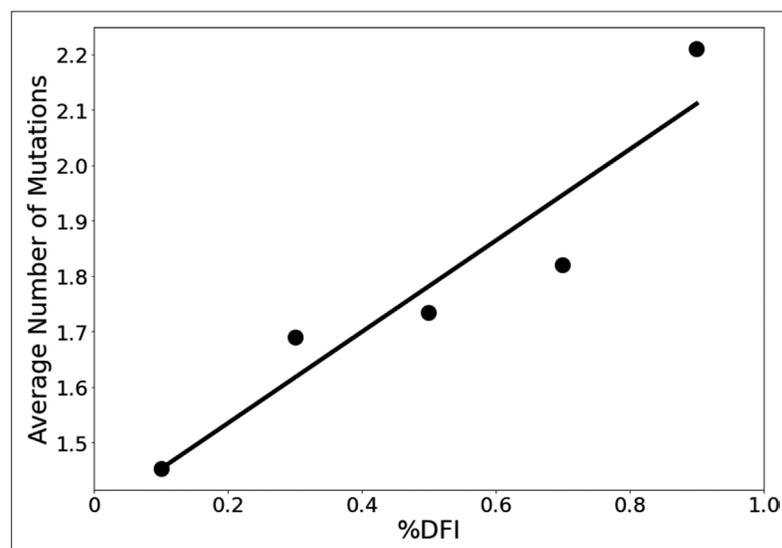


Figure 3. The average number of variants observed among residues of different flexibility. Residues were sorted into one of five bins based on flexibility. Then, the average number of variants for residues within that bin was calculated. Here, the number of variants is defined as the number of different amino acid varieties found at that site. Mutational data was calculated across approximately 24,000 SARS-CoV-2 S protein sequences from the NCBI Datasets Project (**Brister et al., 2015**). Residue flexibility, as reported here via %DFI, was computed using PDB id 6vsb from the Protein DataBank (**Berman et al., 2000**). More rigid residues tend to have fewer variants ($r = 0.94$).

The online version of this article includes the following source data for figure 3:

Source data 1. A comparison of the amino acid sequences for different variant spike proteins used in this article.

humans, it has continued to mutate and evolve rapidly, particularly regarding the S protein (Amicone et al., 2022; Liu et al., 2022; Tay et al., 2022). Just as mutations leading to the Wuhan strain caused an increase in binding affinity to hACE2, continued evolution in human hosts has resulted in further altered binding affinities as well as different phenotypic outcomes for those infected (Ali et al., 2021; Barton et al., 2021; Ozono et al., 2021).

We explore whether protein dynamics has played a role in the selection of mutational sites during the evolution of the S protein since 2019. Our previous work has indicated that the rate of evolution per positional site exhibits a positive correlation with positional flexibility; generally, positions that exhibit higher flexibility are also sites that experience a higher number of amino acid substitutions (Liu and Bahar, 2012; Maguid et al., 2008; Maguid et al., 2006; Mikulska-Ruminska et al., 2019; Nevin Gerek et al., 2013). To confirm these findings for the evolution of the S protein using the sequenced variants of infected humans, we examine the flexibility of the S protein, an analysis conducted in other studies which resulted in several important findings (Nguyen et al., 2020; Socher et al., 2021; Teruel et al., 2021b; Pipitò et al., 2022; Verkhiyker, 2022; Abduljalil et al., 2023). For example, Teruel et al., 2021a used their Elastic Network Contact Model to find how certain highly observed mutations make the open state more rigid and the closed state more flexible. For our own flexibility analysis, we measure the site-specific amino acid flexibility using the dynamic flexibility index (DFI). Using the same mathematical foundation as DCI, DFI evaluates each position's displacement response to random force perturbations at other locations in the protein (Gerek and Ozkan, 2011; Nevin Gerek et al., 2013), and it can be considered a measure of a given position's ability to explore its local conformational space. We found that the COVID-19 S protein shows the expected high correlation between the occurrence of mutations and site flexibility (Figure 3) when we compare %DFI (DFI ranked by percentile) to the average number of variants per position found within a given %DFI bin. Previous studies have indicated that rigid residues are critical for functional dynamics, thus more likely to impact function if mutated and, generally, can lead to a loss of function and thus more conserved (Kim et al., 2015; Butler et al., 2018; Modi et al., 2021b; Modi et al., 2021a; Kazan et al., 2022; Ose et al., 2022a; Stevens et al., 2022; Kumar et al., 2015). This analysis also agrees with these previous studies and highlights the power of negative selection, in line with the neutral theory of molecular evolution, stating that deleterious mutations (i.e., those on the rigid positions) should be eliminated and therefore not observed (Kimura, 1983).

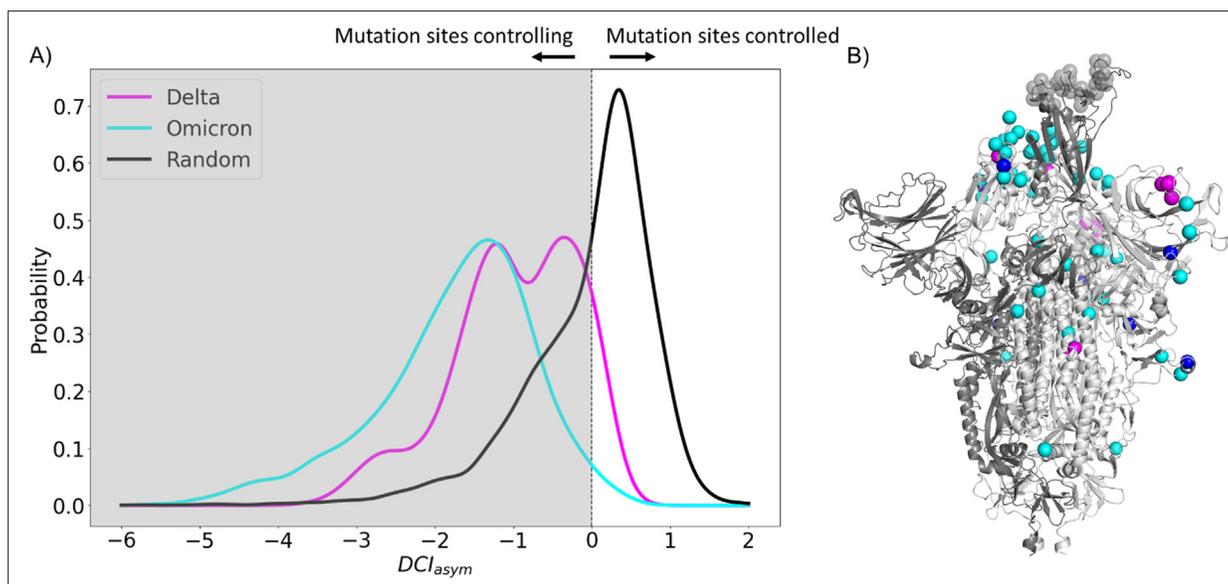


Figure 4. Mutations in variants of concern (VOCs) present different asymmetry profiles. **(A)** DCI_{asym} with low evolutionary probability (EP) characteristic mutation sites of Delta or Omicron strains in the closed chains and the binding interface of receptor-binding domain (RBD) in the open chain. Delta displays a second peak closer to zero, suggesting that Delta mutation sites ($M = -0.98$, $SD = 0.80$) have less allosteric control over the hACE2 binding sites than Omicron mutation sites ($M = -1.74$, $SD = 1.00$) ($p < 0.001$). However, both sets of sites have far more control over hACE2 binding sites than expected, based on a random control group ($M = 0.03$, $SD = 0.85$) ($p < 0.001$). **(B)** S protein structure showing binding interface sites (transparent gray), Delta mutation sites (magenta), Omicron mutation sites (cyan), and sites mutated in both Omicron and Delta (blue).

Continued mutations within human hosts have resulted in a multitude of variants. Indeed, by fitting various molecular clock models to genome sequence data, VOC emergence is punctuated by an episodic period of rapid evolution, with a substitution rate of up to fourfold greater than the background substitution rate (Kumar et al., 2021; Tay et al., 2022). With such an aggressive evolutionary rate, we are finding VOCs to consist of a number of different characteristic mutations, almost all of which are CAPs. Sequence differences between the various VOCs used in this article can be found in **Figure 3—source data 1**.

To explore the dynamic effects of the evolution of the spike in humans, we examine asymmetry with these new potentially adaptive sites, namely the low EP (CAP) characteristic mutation sites observed in the Delta variant, the widely dominant variant from December 2021 to January 2022 (Thye et al., 2021), and the Omicron variant, a highly transmissible variant whose lineages have remained dominant since January 2022 (Kim et al., 2021; **Figure 4**). This analysis revealed a mechanism similar to that for the CAPs in the reference protein (**Figure 2**) as the open-chain binding interface is also allosterically controlled by these potentially new adaptive sites. Regarding this, we see that the asymmetry is much more pronounced in observed mutations of Omicron variants, suggesting that these new mutations have a stronger power in controlling the dynamics of open-chain hACE2 binding interface compared to those observed in Delta variants. We surmise that the difference in virulence and infection rates between Omicron and Delta (Earnest et al., 2022; Bager et al., 2021; Sheikh et al., 2021; Twohig et al., 2022; Houhamdi et al., 2022; Menni et al., 2022) might be due to these specific CAPs within each variant and their differences in allosterically controlling the dynamics of open RBD binding sites as also observed in the DCI_{asym} analysis of the Wuhan variant in **Figure 2**.

Experimental results motivate the use of EpiScore within the SARS-CoV-2 S protein

The fact that the identified CAPs in the reference protein and the more recently evolved CAPs of Delta and Omicron variants both show a high degree of control over the functional sites begs the question: what is the complex interaction between these previous and new CAP sites? Motivated by this concept, we explore the interplay of mutational pairs to understand the effects of the specific amino acid backgrounds associated with these two predominant variants. Some CAP sites in Delta and Omicron have already been considered adaptive (Kemp et al., 2021; Kistler et al., 2022; Maher et al., 2022; Neher, 2022).

It is well understood that the impact of even a single mutation to a protein sequence can sometimes dramatically alter the biophysical behavior of the system. However, the mechanistic impact of point mutations can only be fully understood when the sequence background upon which it is made is accounted for. This means that, in the case of strains with multiple mutations, the interplay between mutated positions will ultimately impact a protein as an aggregate behavior, where the presence of previous mutations may strongly (or weakly) influence some mutations. This concept of non-additivity is known as epistasis. In fact, studies of evolutionary pathways of mutations have suggested that a majority of the mutations have a second or a higher order epistasis among them (Bershtein et al., 2006). Nature exploits this higher order complex relationship between the mutations to evolve their function.

To computationally capture and interpret the pairwise effects of mutations, we have developed an in-house computational tool called EpiScore (**Figure 5A**). Here, we evaluate how a given position pair ij may affect other critical positions k of the protein. EpiScore is the relative coupling strength to a position k when positions i and j are perturbed *simultaneously* compared to the average dynamic coupling strength of i to k and j to k . EpiScore has previously been used successfully to capture overarching trends in GB1 deep mutational scan data as well as specific instances of the development of antibiotic resistance in various enzymatic systems (Campitelli and Ozkan, 2020). An EpiScore of 1 indicates perfect coupling additivity, and deviations from this value represent non-additive behavior between position pairs and functionally important sites. Prior EpiScore work has shown a difference in EpiScore between the sites of compensatory and non-compensatory mutations, where both yield distributions with peaks around 1, but non-compensatory mutations show higher deviation in their EpiScore distribution (Ose et al., 2022b).

Many studies have confirmed epistasis between residues within the S protein (Moulana et al., 2022; Starr et al., 2022a; Moulana et al., 2023; Witte et al., 2023). These epistatic residues can have various

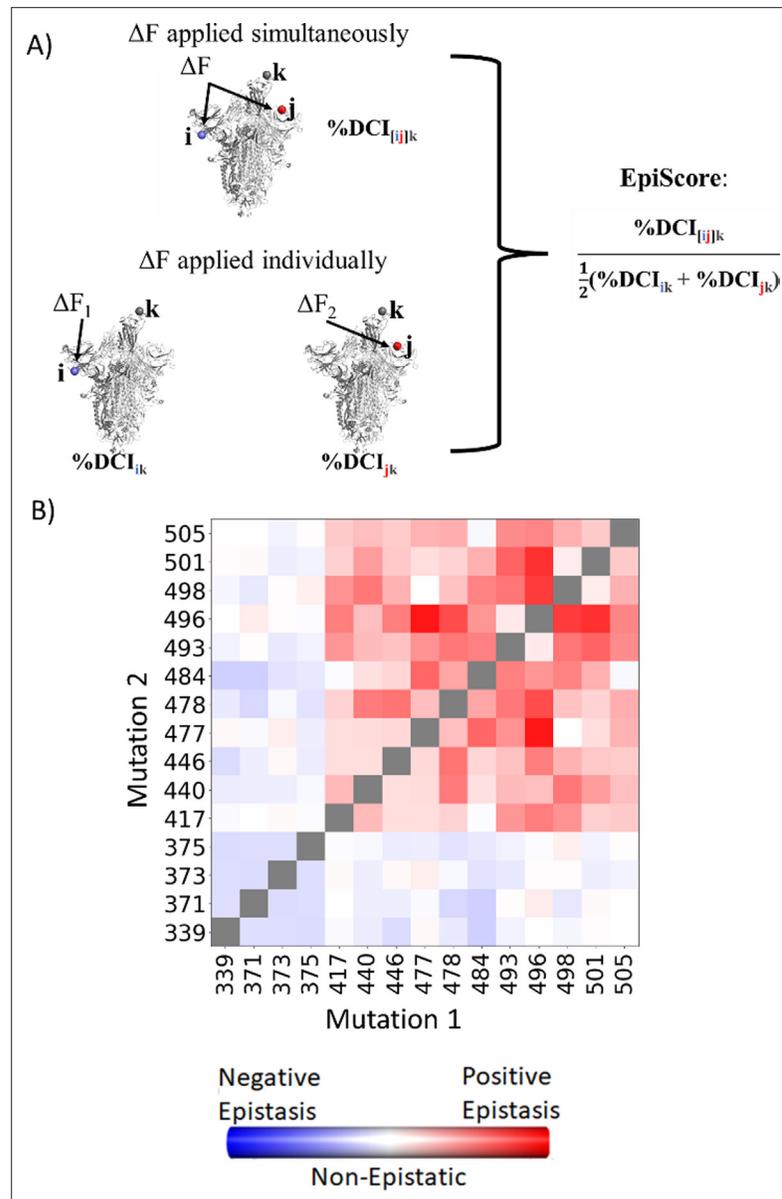


Figure 5. EpiScore provides a measurement of epistasis. **(A)** Schematic representation of cross-chain EpiScore, describing i, j , in chains B and C, respectively, and its impact in receptor-binding domain (RBD) binding position k in the open RBD conformer chain A. **(B)** Colors indicate EpiScore values for given mutation pairs, averaged over hACE2 binding sites. Cross-chain residue pairs in the upper right tend to be highly epistatic, similar to pairwise second-order interaction coefficients from *Moulana et al., 2022*.

effects on hACE2 or antibody binding. Within SARS-CoV-2, here we calculate the EpiScore (**Figure 5B**) of a set of mutation pairs used by *Moulana et al., 2022* and compare our results to quantified epistatic effects determined by the experimental hACE2 binding affinity of ‘first-order’ single mutation variants compared to ‘second-order’ mutation pairs. Our EpiScore results and the experimentally determined epistasis both captured highly epistatic behavior among residues 493, 496, 498, 501, and 505, as well as a lack of epistatic behavior for residues 339, 371, 373, and 375; however, EpiScore generally showed higher epistasis values than experiment for residues 417, 440, 446, 477, 478, and 484.

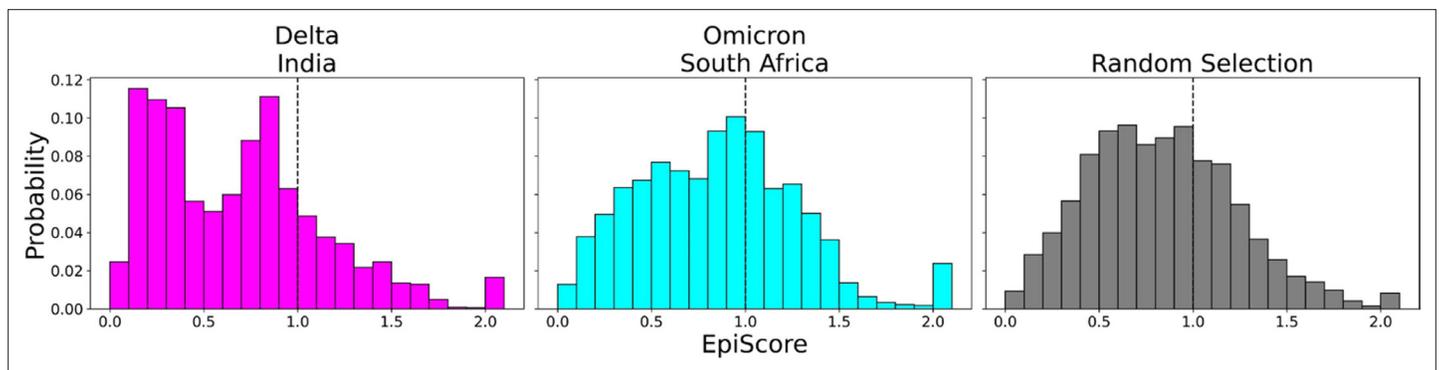


Figure 6. EpiScores of variants of concern (VOC) mutations. Here, i = low evolutionary probability (EP) Delta mutation sites (magenta), low EP Omicron mutation sites (cyan), j = low EP sites in the Wuhan variant, and k = the binding interface of the open chain. EpiScores using sites of either variant (Delta: $M = 0.70$, $SD = 0.50$, Omicron $M = 0.86$, $SD = 0.46$) are significantly different ($p < 0.001$) from a set of EpiScores using random sites ($M = 0.84$, $SD = 0.41$), but the distribution for Delta variants differs much more from the other two. EpiScores for other variants can be found in **Figure 6—figure supplement 1**.

The online version of this article includes the following figure supplement(s) for figure 6:

Figure supplement 1. EpiScores for more variants of concern (VOCs).

Figure supplement 2. EpiScores in the N-terminal domain (NTD).

EpiScore highlights the epistatic relationship between the recent adaptive mutations in VOCs and the CAPs of the Wuhan reference

Seeking further to understand the role of epistasis within S protein variants, we explored the possibility of epistatic relationships between the CAPs of the Wuhan variant and the new CAPs in VOCs. Thus, we computed the EpiScore of these CAP positions in the closed RBDs (i.e., chains B and C) with respect to functional hACE2 binding interface sites of the open RBD chain (chain A) (**Figure 5B**) and obtained EpiScore distributions.

To contrast these variants, Omicron (**Figure 6**, cyan) shows a high proportion of additive mutations compared to the Delta variant (**Figure 6**, magenta), with a peak centered on 1. The comparatively more pathogenic Delta variant exhibited many non-additive mutations with EpiScores below one. This again suggests that the cross-communication between the open and closed chain of the S protein is important for regulating the function. Four out of seven low EP Delta mutation sites used in this analysis often resulted in EpiScores below 1. Each of those is found in the N-terminal domain (NTD) on or near the N3 loop and is implicated in antibody escape in recent studies (*Chi et al., 2020; Weisblum et al., 2020; Harvey et al., 2021; Klinakis et al., 2021; Cantoni et al., 2022*). The low EpiScores of NTD mutations suggest that they dampen the control of Wuhan variant CAPs over the hACE2 binding sites in addition to their effects on antibody binding. It is possible that what the Delta variant gained in transmission rate also came with being more harmfully pathogenic due in part to negatively epistatic interactions. In contrast, the mutations leading to the development of the Omicron strain were additive with respect to Wuhan variant CAPs, possibly leading to a lower pathogenicity and higher effective immune escape, resulting in an overall higher transmission rate. Another possible explanation for the higher transmission rate of Omicron comes from a different normal mode analysis study, which found that despite reduced binding affinity with hACE2, Omicron showed increased occupancy of the open state compared to the closed state, which would increase chances of interaction with hACE2 (*Teruel et al., 2021b*). It is worth noting that other variants contain NTD mutations which result in low EpiScores; however, the proportion of these mutations within the set is considerably less than in Delta (**Figure 6—figure supplement 2**).

One of the more notable features of generated EpiScore distributions is the presence of a tail of values upward of 2.0, indicating highly epistatic behavior. Interestingly, these tails are largely due to three different CAPs: 346R, 486F, and 498Q. Those residues are nearby one another within the protein structure and have been reported to play a role in antibody binding, either being known antibody binding sites (346R and 486F) or having received very high antibody accessibility scores (498Q) (*Harvey et al., 2021; Raghuvamsi et al., 2021*). These observed high EpiScore values also support

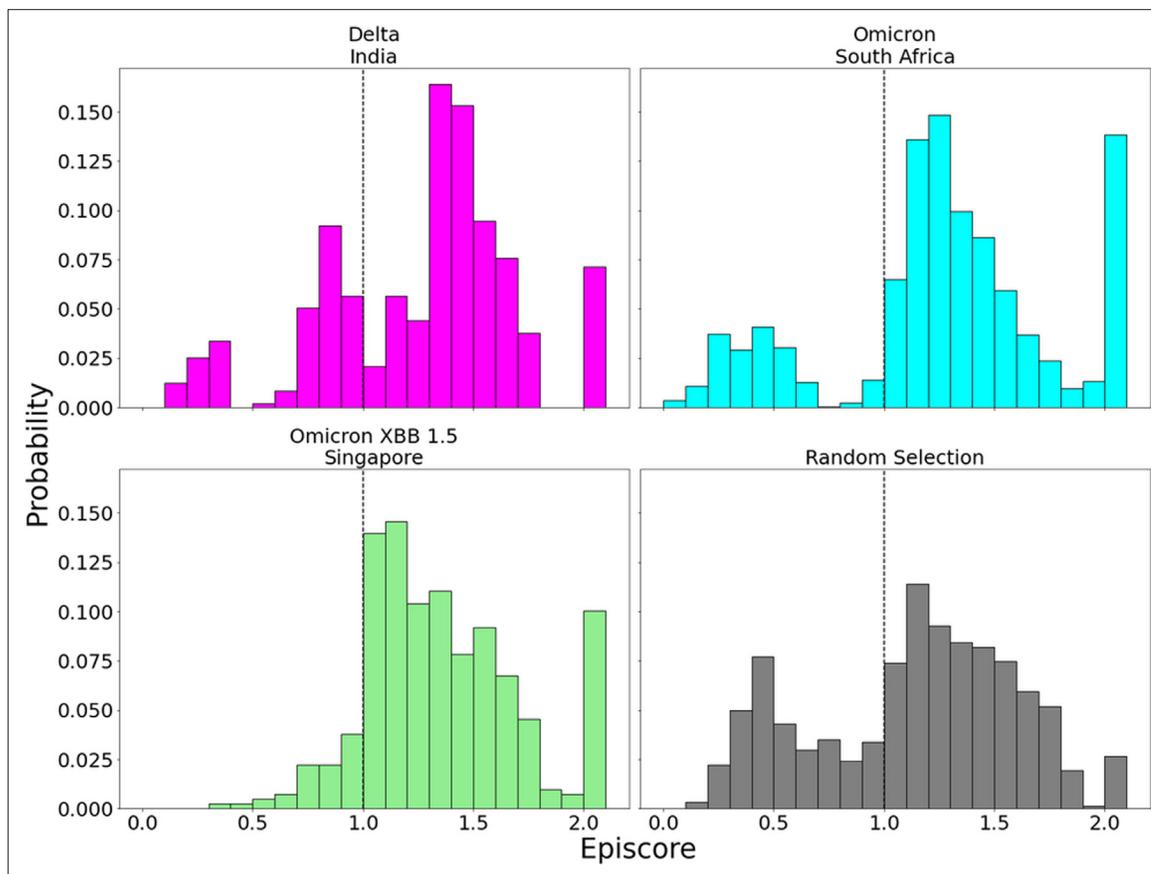


Figure 7. EpiScores with candidate adaptive polymorphism (CAP) site 486. Here, i = low evolutionary probability (EP) Delta mutation sites (magenta), low EP Omicron mutation sites (cyan), and a random selection of sites (gray), j = site 486, and k = the binding interface of the open chain. CAP and hACE2 and antibody binding site 486 displays epistasis with almost all XBB 1.5 variant sites at almost every hACE2 binding site ($M = 1.40$, $SD = 0.46$) and presents a significantly different profile from other variant sites ($p < 0.01$). EpiScores involving 486 for other variants can be found in **Figure 7—figure supplement 1**.

The online version of this article includes the following figure supplement(s) for figure 7:

Figure supplement 1. EpiScores for more variants of concern (VOCs) with candidate adaptive polymorphism (CAP) site 486.

other studies indicating the epistatic interactions between these CAPs and the mutations of the VOCs within the S protein are crucial for maintaining binding affinity of hACE2 whilst evading immunity (**Hong et al., 2022; Starr et al., 2022b**).

Inspection of EpiScores of Delta and Omicron potentially adaptive mutation sites with only CAP site 486F (a binding site for both hACE2 and antibodies) (**Huang et al., 2020; Ali et al., 2021; Harvey et al., 2021; Raghuvamsi et al., 2021**) shows highly epistatic interactions at other hACE2 binding sites (**Figure 6**). However, within a recent and rapidly spreading subvariant of Omicron, XBB 1.5, we see a mutation of S to P, a rare double nucleotide mutation, at site 486 (preceding Omicron variants included mutation F486S from the original CAP). This new variant has unprecedented immune escape capabilities, resisting neutralizing antibodies almost entirely (**Qu et al., 2023**). EpiScores of other XBB 1.5-specific mutation sites with 486P are almost entirely greater than 1, showing an even higher degree of epistasis with the binding sites of RBD (**Figure 7**). These results present a threefold importance for the S486P mutation: not only does this residue alter antibody (i.e., immune escape) and hACE2 binding by directly modifying a binding site, but it may also be responsible for modifying hACE2 binding via epistatic cooperation with other co-occurring mutations.

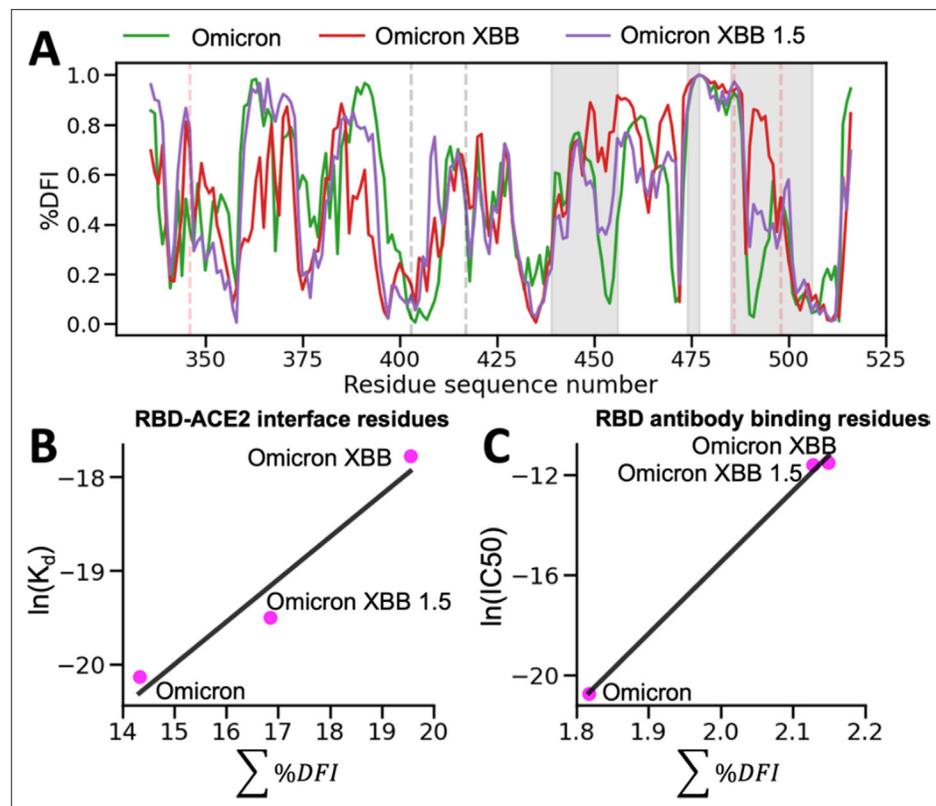


Figure 8. The %DFI calculations for variants Omicron, XBB, and XBB 1.5. **(A)** %DFI profile of the variants are plotted in the same panel. The gray shaded areas and dashed lines indicate the ACE2 binding regions, whereas the red dashed lines show the antibody binding residues. **(B)** The sum of %DFI values of RBD-ACE2 interface residues. The trend of total %DFI with the log of K_d values overlaps with the one seen with the experiments ($r = 0.97$). **(C)** The receptor-binding domain (RBD) antibody binding residues are used to calculate the sum of %DFI. The ranking captured with the total %DFI agrees with the log of IC50 values from the experiments.

The online version of this article includes the following figure supplement(s) for figure 8:

Figure supplement 1. Non-additive interactions influence variant behavior.

Change in flexibility of RBD binding site correlates with experimental binding affinities for Omicron and Omicron XBB variants

Experimental studies have tracked hACE2 binding for different variants since the virus first spread (Ali *et al.*, 2021; Barton *et al.*, 2021; Ozono *et al.*, 2021; Wu *et al.*, 2022). Within the Omicron variant, for example, characteristic mutations on the RBD are shown to increase the overall binding affinity of the virus to the ACE2 receptor, which is suspected to allow it to spread more easily (Kim *et al.*, 2021). Furthermore, the new Omicron XBB and Omicron XBB 1.5 variants contain additional mutations in the RBD and antibody binding residues, which may further impact their dynamics and interactions with the host.

To gain deeper insights into the impact of dynamics on the binding affinity of hACE2 and antibodies with the recent Omicron XBB variants, we conducted molecular dynamics (MD) simulations. By analyzing the resulting trajectories, we investigated how these mutations influence the flexibility and rigidity of the RBD and antibody binding residues, consequently affecting their binding affinity and potential for immune evasion (Figure 8). To understand the overall flexibility changes, we measured the sum of DFI of the ACE2 binding residues, as well as the sum of DFI of the antibody binding residues, calculated from the MD trajectories and compared them with experimental viral binding (dissociation constants) and immunity evasion antibody IC50 values (Yue *et al.*, 2023).

This investigation elucidated the impact of mutations in the RBD and antibody binding residues on the binding affinity of the S protein and immune evasion by modulating their flexibility and rigidity (Figure 8A). The Omicron XBB variant exhibits heightened flexibility in hACE2 and antibody

binding residues, reducing infectivity and enhancing immune evasion. Conversely, the Omicron XBB 1.5 variant induces distinct dynamics in these regions, rendering the RBD-ACE2 interface more rigid while increasing flexibility in antibody binding residues. These effects indicate that Omicron XBB 1.5 retains its antibody escape capabilities while regaining ACE2 binding affinity comparable to previous Omicron variants, in accordance with experimental findings (Yue *et al.*, 2023). These findings suggest that mutations in the RBD and antibody binding residues can have complex effects on the dynamics of the protein and, ultimately, on the virus's ability to infect and evade the host immune system through an alteration of binding site dynamics. However, we do note that the effect of mutations will heavily depend on the genetic background in which they are mutated. The effects of a mutation on a Delta variant protein may differ entirely from the effects of a mutation on an Omicron variant protein (Figure 8—figure supplement 1).

Conclusion

We analyzed the evolutionary trajectory of the CoV-2 S protein in humans to understand the dynamic and epistatic interactions of the mutations defining specific VOCs. We first obtain the phylogenetic tree of the COV-2 S protein and identify the sites of certain recent mutations known as CAPs. CAPs are considered adaptive because mutations rarely tolerated in closely related sequences have suddenly become fixed, implying a degree of functional importance or evolution (Liu *et al.*, 2016). In addition, our earlier work has shown that CAPs can also be compensatory; multiple CAPs may dynamically compensate for one another, changing the dynamic landscape and allowing for different mutations (Ose *et al.*, 2022b). We then explored the mechanistic insights and epistatic relationship between the observed mutations in different VOCs and CAP sites, and, particularly, the relationship between CAP sites and the functionally critical RBD using our dynamic coupling analysis (Kumar *et al.*, 2015).

We find a mechanistic pattern in the S protein evolution that is common amongst previously studied systems, where allosteric sites exert control over the dynamics of the binding sites, and mutations of these allosteric sites modulate function. Coupling our analysis with evolutionary theory showed that many of these allosteric sites regulating function of the S protein may have been subjected to adaptive evolution as observed in mutations in VOCs. Our dynamics analysis also provides a mechanistic insight where the Omicron-defining sites have greater control over the binding sites than the Delta variant and are dynamically additive with the functional advantage of CAPs, thus, the greatest infectivity may not be a coincidence.

Specifically, we find that the interactions between CAP sites and VOC-defining mutations show fingerprints of non-additive dynamics within the Delta variant. In contrast, mutations leading to the Omicron variant are largely additive, driving critical dynamical behavior closer to the patterns observed within the wild-type. These interactions may also drive observed behavior similar between the reference and Omicron strains yet differ in the delta strain, such as the severity of infection as evidenced by hospitalization rates (Houhamdi *et al.*, 2022; Menni *et al.*, 2022). It has also been shown that the Omicron variant has a lower binding affinity with hACE2 than previous variants (Wu *et al.*, 2022), which may contribute to its low pathogenicity.

Long-ranged interactions between different sites within a given protein are critically important for protein function (Peters and Lively, 1999; Bershtein *et al.*, 2006; Collins *et al.*, 2006; Ekeberg *et al.*, 2013; Levy *et al.*, 2017; Harrigan *et al.*, 2018; Otten *et al.*, 2018; Rojas Echenique *et al.*, 2019; Shimagaki and Weigt, 2019; de la Paz *et al.*, 2020; Rizzato *et al.*, 2020; Yang *et al.*, 2020; Bisardi *et al.*, 2022) and for the CoV-2 S protein in particular (Zeng *et al.*, 2020; Castiglione *et al.*, 2021; Dong *et al.*, 2021; Garvin *et al.*, 2021; Nielsen *et al.*, 2022; Ramarao-Milne *et al.*, 2022; Rochman *et al.*, 2022; Rodriguez-Rivas *et al.*, 2022). By showing dynamic differences between the interactions of CAPs, which have likely played a major role in allowing the virus to infect human hosts, the binding site, and the characteristic mutations of dominant Delta and Omicron strains, we see a 'fine-tuning' of protein behavior. As variants continue to evolve, Omicron subvariants are of growing concern due in large part to further increased immune evasion (Callaway, 2022; Wang *et al.*, 2022a; Wang *et al.*, 2022b), and we observe that the new mutations observed in antibody binding sites yield more epistatic interaction with the CAPs. In addition to supporting previous dynamic research on the S protein, this analysis provides the insight that CAP sites are of continued importance to protein function and should be given special attention when considering the impact of future mutations.

Methods

Dynamic flexibility and dynamic coupling

The DFI utilizes a PRS technique that combines the elastic network model (ENM) and LRT (**Gerek and Ozkan, 2011; Nevin Gerek et al., 2013**). In ENM, the protein is considered as a network of beads at C α positions interacting with each other via a harmonic spring potential. Using LRT, $\Delta\mathbf{R}$ is calculated as the fluctuation response vector of residue j due to unit force's \mathbf{F} perturbation on residue i , averaged over multiple unit force directions to simulate an isotropic perturbation.

$$[\Delta R]_{3N \times 1} = [H]_{3N \times 3N}^{-1} [F]_{3N \times 1} \quad (1)$$

where \mathbf{H} is the Hessian, a $3N \times 3N$ matrix that can be constructed from 3-D atomic coordinate information and is composed of the second derivatives of the harmonic potential with respect to the components of the position's vectors of length $3N$. The Hessian inverse in this equation may be replaced with the covariance matrix \mathbf{G} obtained from MD simulations as follows:

$$[\Delta R]_{3N \times 1} = [G]_{3N \times 3N} [F]_{3N \times 1} \quad (2)$$

MD simulations were used to obtain the DFI profiles of Omicron, Omicron XBB, and Omicron XBB 1.5. In order to obtain DFI, each position in the structure was perturbed sequentially to generate a Perturbation Response Matrix \mathbf{A}

$$A_{N \times N} = \begin{bmatrix} |\Delta R^1|_1 & \cdots & |\Delta R^M|_1 \\ \vdots & \ddots & \vdots \\ |\Delta R^1|_N & \cdots & |\Delta R^M|_N \end{bmatrix} \quad (3)$$

where $|\Delta R^j|_i = \sqrt{\langle (\Delta R)^2 \rangle}$ is the magnitude of fluctuation response at position i due to perturbations at position j . The DFI value of position i is then treated as the displacement response of position i relative to the net displacement response of the entire protein, which is calculated by sequentially perturbing each position in the structure.

$$DFI_i = \frac{\sum_{j=1}^N |\Delta R^j|_i}{\sum_{i=1}^N \sum_{j=1}^N |\Delta R^j|_i} \quad (4)$$

It is also often useful to quantify position flexibility relative to the flexibility ranges unique to individual structures. To that end, DFI can be presented as a percentile rank, %DFI. All %DFI calculations present in this work used the DFI value of every residue of the full spike structure for ranking. The DFI parameter can be considered a measure of a given amino acid position's ability to explore its local conformational space.

Dynamic coupling index

Similar to DFI, the DCI (**Larrimore et al., 2017; Kumar et al., 2015**) also utilizes PRS with the ENM and LRT. DCI captures the strength of displacement response of a given position i upon perturbation to a single functionally important position (or subset of positions) j , relative to the average fluctuation response of position i when all of the positions within a structure are perturbed.

$$DCI_{ji} = \frac{\sum_j^{N_{functional}} |\Delta R^j|_i / N_{functional}}{\sum_{j=1}^N |\Delta R^j|_i / N} \quad (5)$$

When only positional pairs are concerned, this expression reduces to

$$DCI_{ji} = \frac{|\Delta R^j|_i}{\sum_{j=1}^N |\Delta R^j|_i / N} \quad (6)$$

As such, this parameter represents a measure of the dynamic coupling between i and j upon a perturbation to j . As with DFI, DCI_{ij} can also be presented as a percentile-ranked $\%DCI_{ij}$.

One of the most important aspects of DCI is that the entire network of interactions is explicitly included in subsequent calculations without the need for dimensionality reduction techniques. If one considers interactions such as communication directionality or dynamic coupling regulation between position pairs as inherent properties of an anisotropic interaction network, it is critical to include the interactions of the entire network to accurately model the effect one residue can have on another.

Here, we present two further extensions of DCI, which allow us to uniquely model coupling directionality and epistatic effects: DCI_{asym} and EpiScore, respectively. Interestingly, we can capture asymmetry between different residues within a protein through DCI, as a coupling in and of itself is asymmetric within an anisotropic network. That is, each amino acid has a set of positions to which it is highly coupled, and this anisotropy in connections gives rise to unique differences in coupling between a given ij pair of amino acids which do not have direct interactions (**Figure 2A**). DCI_{asym} , then, is simply DCI_{ij} (the normalized displacement response of position j upon a perturbation to position i) – DCI_{ji} (**Equation (7)**). Using DCI_{asym} , we can determine a cause–effect relationship between the ij pair in terms of force/signal propagation between these two positions.

$$DCI_{asym} = DCI_{ij} - DCI_{ji} \quad (7)$$

$$\%DCI_{asym} = \%DCI_{ij} - \%DCI_{ji} \quad (8)$$

where a positive DCI_{asym} value indicates communication from position i to position j .

EpiScore can identify or describe potential non-additivity in substitution behavior between residue pairs. This metric can capture the differences in a normalized perturbation response to a position k when a force is applied at two residues i and j simultaneously versus the average additive perturbation response when each residue i, j , is perturbed individually (**Figure 5A, Equation 9**).

$$EpiScore = \frac{\%DCI_{[ij]k}}{\frac{1}{2} (\%DCI_{ik} + \%DCI_{jk})} \quad (9)$$

EpiScore values <1 (>1) indicate that the additive perturbations of positions i and j generate a greater (lesser) response at position k than the effect of a simultaneous perturbation. This means that, when treated with a simultaneous perturbation at both sites i and j , the displacement response of k is lower (higher) than the average effect of individual perturbations to i and j , one at a time. As EpiScore is a linear scale, the further the value from 1, the greater the effect described above.

Molecular dynamics (MD)

The production simulations for the Omicron variants, including Omicron, Omicron XBB, and Omicron XBB 1.5, were performed with the AMBER software package. These variants, each characterized by specific mutations, were modeled based on the template PDB structure 6M0J. The initial protein configurations in the simulations were parameterized using the ff14SB force field (**Maier et al., 2015**). In order to create an appropriate solvation environment for the proteins, a solvation box was defined around them, maintaining a minimum separation distance of 16 Å from the protein to the box boundaries. This was accomplished by employing the explicit TIP3P water model (**Sun and Kollman, 1995**), with the addition of sodium and chloride ions to maintain overall charge neutrality.

The simulation procedure involves an initial energy minimization step, aimed at mitigating steric clashes and optimizing the system's energy. The steepest descent algorithm was applied, encompassing 11,000 steps. Subsequently, the system underwent a gradual temperature increase (heat up), up to 300 K, and was subjected to production simulations under a constant number of particles, pressure, and temperature ensemble (NPT).

During these production simulations, temperature was maintained at 300 K, with pressure regulation set at 1 bar. Temperature regulation was achieved through the utilization of the Langevin thermostat (**Hünenberger, 2005**) and Berendsen barostat (**Berendsen et al., 1984**), featuring a collision frequency of 1.0 picoseconds⁻¹. Hydrogen atom bond lengths were constrained using the SHAKE algorithm (**Pearlman et al., 1995**). The production trajectories were simulated for 1 μs each.

To ensure the reliability of the simulations and assess their convergence, a convergence criterion was employed. The achieved convergence was determined by monitoring the root mean square

deviation (RMSD) between the highest sampled conformation in consecutive time windows (**Sawle and Ghosh, 2016**). Specifically, convergence was defined as the point at which the RMSD between the highest sampled conformation in the last 300 ns window and the 300 ns window immediately preceding it dropped below 1 Å. Window sizes varying from 100 ns to 500 ns were employed to evaluate convergence, ensuring the robustness and stability of the obtained results.

To calculate DFI, covariance matrix data were computed over different time windows as discussed above. By default, utilizing the Hessian implies a restriction to a harmonic potential, assuming that the data are sampled from a Gaussian distribution. Ergodicity in both simulation time and initial structures sampled in each time interval ensures two key conditions: (i) consistency of potential energy across conformations sampled from the same distribution. (ii) The sampling of different initial conformations while computing covariance matrices at various time windows eliminates global motions and accurately captures equilibrium coordinates. Consequently, the final average DFI profiles are independent of time window size, resulting in consistent results across different time window sizes (e.g., 50 ns vs. 75 ns) and enabling the acquisition of statistically significant DFI values.

Statistical tests

Pearson correlation coefficients (r) were used to demonstrate linear relationships between continuous variables in **Figures 3 and 8**. Student's independent t-tests were performed to demonstrate a significant difference between distributions in **Figures 2, 4 and 6** (and **Figure 3—source data 1, Figure 6—figure supplements 1 and 2**), and **Figure 7** (and **Figure 7—figure supplement 1**), as demonstrated by the p-value.

Acknowledgements

Funding was provided to NJO, PC, TM, and ICK by the Gordon and Betty Moore Foundation (award number AWD00034439) and to SBO by the National Science Foundation (award numbers: 1715591 and 1901709) and the National Institutes of Health R01GM147635-01. SK acknowledges the National Science Foundation (GCR 1934848) and the National Institutes of Health (GM139540) grants.

Additional information

Funding

Funder	Grant reference number	Author
Gordon and Betty Moore Foundation	AWD00034439	Nicholas James Ose Paul Campitelli Tushar Modi I Can Kazan
National Science Foundation	1715591	Sefika Banu Ozkan
National Science Foundation	1901709	Sefika Banu Ozkan
National Science Foundation	GCR 1934848	Sudhir Kumar
National Institutes of Health	R01GM147635-01	Sefika Banu Ozkan
National Institutes of Health	GM139540	Sudhir Kumar

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Nicholas James Ose, Conceptualization, Data curation, Software, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing; Paul Campitelli, Conceptualization,

Software, Formal analysis, Methodology, Writing – original draft, Writing – review and editing; Tushar Modi, Conceptualization, Software, Formal analysis, Methodology, Writing – review and editing; I Can Kazan, Conceptualization, Software, Formal analysis, Investigation, Methodology, Writing – review and editing; Sudhir Kumar, Conceptualization, Data curation, Software, Supervision, Funding acquisition, Methodology, Writing – review and editing; Sefika Banu Ozkan, Conceptualization, Supervision, Visualization, Methodology, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs

Nicholas James Ose  <http://orcid.org/0000-0002-2194-5199>

Paul Campitelli  <http://orcid.org/0000-0001-5620-609X>

Tushar Modi  <http://orcid.org/0000-0001-9483-9170>

I Can Kazan  <https://orcid.org/0000-0003-2593-4179>

Sudhir Kumar  <http://orcid.org/0000-0002-9918-8212>

Sefika Banu Ozkan  <https://orcid.org/0000-0002-9351-3758>

Peer review material

Reviewer #2 (Public review): <https://doi.org/10.7554/eLife.92063.3.sa1>

Reviewer #3 (Public review): <https://doi.org/10.7554/eLife.92063.3.sa2>

Author response <https://doi.org/10.7554/eLife.92063.3.sa3>

Additional files

Supplementary files

- Supplementary file 1. DFI profile of the Omicron variant, as obtained from MD simulations.
- Supplementary file 2. DFI profile of the Omicron XBB variant, as obtained from MD simulations.
- Supplementary file 3. DFI profile of the Omicron XBB 1.5 variant, as obtained from MD simulations.
- Supplementary file 4. Mutation sites and EP values.
- Supplementary file 5. The alignment used to generate EP values.
- Supplementary file 6. Alignment of VOC sequences.
- MDAR checklist

Data availability

The code to perform DFI and DCI analysis is available at <https://github.com/SBOZKAN/DFI-DCI> (copy archived at [Ozkan, 2024](#)). Molecular Dynamics data are available at https://github.com/SBOZKAN/COV2SPIKE_MD. The mutation sites and EP values are contained in **Supplementary file 4**. The alignment used to generate EP values is also contained within the supporting information files as "EP_alignment.fas". Protein Databank ID number 6VXX ([Walls et al., 2020](#)) was used for closed conformation DCI calculations. 6VSB ([Wrapp et al., 2020](#)) was used for DFI calculations, EpiScore calculations, and open conformation DCI calculations. 6M0J ([Lan et al., 2020](#)) was used in molecular dynamics simulations of the RBD.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Ozkan SB	2024	COV2SPIKE_MD	https://github.com/SBOZKAN/COV2SPIKE_MD	GitHub, 91b9d69

Continued on next page

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C, Abiona O, Graham BS, McLellan JS	2020	Prefusion 2019-nCoV spike glycoprotein with a single receptor-binding domain up	https://www.rcsb.org/structure/6vsb	RCSB Protein Data Bank, 6VSB
Wang X, Lan J, Ge J, Yu J, Shan S	2020	Crystal structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2	https://www.rcsb.org/structure/6m0j	RCSB Protein Data Bank, 6MOJ
Walls AC, Park YJ, Tortorici MA, Wall A	2020	Structure of the SARS-CoV-2 spike glycoprotein (closed state)	https://www.rcsb.org/structure/6vxx	RCSB Protein Data Bank, 6VXX

References

- Abduljalil JM**, Elghareib AM, Samir A, Ezat AA, Elfiky AA. 2023. How helpful were molecular dynamics simulations in shaping our understanding of SARS-CoV-2 spike protein dynamics? *International Journal of Biological Macromolecules* **242**:125153. DOI: <https://doi.org/10.1016/j.ijbiomac.2023.125153>, PMID: [37268078](https://pubmed.ncbi.nlm.nih.gov/37268078/)
- Ali F**, Kasry A, Amin M. 2021. The new SARS-CoV-2 strain shows a stronger binding affinity to ACE2 due to N501Y mutant. *Medicine in Drug Discovery* **10**:100086. DOI: <https://doi.org/10.1016/j.medidd.2021.100086>, PMID: [33681755](https://pubmed.ncbi.nlm.nih.gov/33681755/)
- Amicone M**, Borges V, Alves MJ, Isidro J, Zé-Zé L, Duarte S, Vieira L, Guiomar R, Gomes JP, Gordo I. 2022. Mutation rate of SARS-CoV-2 and emergence of mutators during experimental evolution. *Evolution, Medicine, and Public Health* **10**:142–155. DOI: <https://doi.org/10.1093/emph/eoac010>, PMID: [35419205](https://pubmed.ncbi.nlm.nih.gov/35419205/)
- Bager P**, Wohlfahrt J, Rasmussen M, Albertsen M, Krause TG. 2021. Hospitalisation associated with SARS-CoV-2 delta variant in Denmark. *The Lancet. Infectious Diseases* **21**:1351. DOI: [https://doi.org/10.1016/S1473-3099\(21\)00580-6](https://doi.org/10.1016/S1473-3099(21)00580-6), PMID: [34487704](https://pubmed.ncbi.nlm.nih.gov/34487704/)
- Barton MI**, MacGowan SA, Kutuzov MA, Dushek O, Barton GJ, van der Merwe PA. 2021. Effects of common mutations in the SARS-CoV-2 Spike RBD and its ligand, the human ACE2 receptor on binding affinity and kinetics. *eLife* **10**:e70658. DOI: <https://doi.org/10.7554/eLife.70658>, PMID: [34435953](https://pubmed.ncbi.nlm.nih.gov/34435953/)
- Berendsen HJC**, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **81**:3684–3690. DOI: <https://doi.org/10.1063/1.448118>
- Berman HM**, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The protein data bank. *Nucleic Acids Research* **28**:235–242. DOI: <https://doi.org/10.1093/nar/28.1.235>, PMID: [10592235](https://pubmed.ncbi.nlm.nih.gov/10592235/)
- Bershtein S**, Segal M, Bekerman R, Tokuriki N, Tawfik DS. 2006. Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* **444**:929–932. DOI: <https://doi.org/10.1038/nature05385>, PMID: [17122770](https://pubmed.ncbi.nlm.nih.gov/17122770/)
- Bhabha G**, Ekiert DC, Jennewein M, Zmasek CM, Tuttle LM, Kroon G, Dyson HJ, Godzik A, Wilson IA, Wright PE. 2013. Divergent evolution of protein conformational dynamics in dihydrofolate reductase. *Nature Structural & Molecular Biology* **20**:1243–1249. DOI: <https://doi.org/10.1038/nsmb.2676>, PMID: [24077226](https://pubmed.ncbi.nlm.nih.gov/24077226/)
- Bisardi M**, Rodriguez-Rivas J, Zamponi F, Weigt M. 2022. Modeling sequence-space exploration and emergence of epistatic signals in protein evolution. *Molecular Biology and Evolution* **39**:msab321. DOI: <https://doi.org/10.1093/molbev/msab321>, PMID: [34751386](https://pubmed.ncbi.nlm.nih.gov/34751386/)
- Boni MF**, Lemey P, Jiang X, Lam TTY, Perry BW, Castoe TA, Rambaut A, Robertson DL. 2020. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nature Microbiology* **5**:1408–1417. DOI: <https://doi.org/10.1038/s41564-020-0771-4>, PMID: [32724171](https://pubmed.ncbi.nlm.nih.gov/32724171/)
- Brister JR**, Ako-adjei D, Bao Y, Blinkova O. 2015. NCBI viral genomes resource. *Nucleic Acids Research* **43**:D571–D577. DOI: <https://doi.org/10.1093/nar/gku1207>
- Butler BM**, Kazan IC, Kumar A, Ozkan SB, editor. 2018. Coevolving residues inform protein dynamics profiles and disease susceptibility of nSNVs. *PLOS Computational Biology* **14**:e1006626. DOI: <https://doi.org/10.1371/journal.pcbi.1006626>, PMID: [30496278](https://pubmed.ncbi.nlm.nih.gov/30496278/)
- Cagliani R**, Forni D, Clerici M, Sironi M. 2020. Computational Inference of Selection Underlying the Evolution of the Novel Coronavirus, Severe Acute Respiratory Syndrome Coronavirus 2. *Journal of Virology* **94**:e00411-20. DOI: <https://doi.org/10.1128/JVI.00411-20>, PMID: [32238584](https://pubmed.ncbi.nlm.nih.gov/32238584/)
- Callaway E**. 2022. COVID 'variant soup' is making winter surges hard to predict. *Nature* **611**:213–214. DOI: <https://doi.org/10.1038/d41586-022-03445-6>
- Campbell E**, Kaltenbach M, Correy GJ, Carr PD, Porebski BT, Livingstone EK, Afriat-Jurnou L, Buckle AM, Weik M, Hollfelder F, Tokuriki N, Jackson CJ. 2016. The role of protein dynamics in the evolution of new

- enzyme function. *Nature Chemical Biology* **12**:944–950. DOI: <https://doi.org/10.1038/nchembio.2175>, PMID: 27618189
- Campitelli P**, Modi T, Kumar S, Ozkan SB. 2020. The role of conformational dynamics and allostery in modulating protein evolution. *Annual Review of Biophysics* **49**:267–288. DOI: <https://doi.org/10.1146/annurev-biophys-052118-115517>, PMID: 32075411
- Campitelli P**, Ozkan SB. 2020. Allostery and epistasis: emergent properties of anisotropic networks. *Entropy* **22**:667. DOI: <https://doi.org/10.3390/e22060667>, PMID: 33286439
- Campitelli P**, Swint-Kruse L, Ozkan SB. 2021. Substitutions at nonconserved rheostat positions modulate function by rewiring long-range, dynamic interactions. *Molecular Biology and Evolution* **38**:201–214. DOI: <https://doi.org/10.1093/molbev/msaa202>, PMID: 32780837
- Campitelli P**, Lu J, Ozkan SB. 2022. Dynamic allostery highlights the evolutionary differences between the CoV-1 and CoV-2 main proteases. *Biophysical Journal* **121**:1483–1492. DOI: <https://doi.org/10.1016/j.bpj.2022.03.012>, PMID: 35300968
- Cantoni D**, Murray MJ, Kalemra MD, Dicken SJ, Stejskal L, Brown G, Lytras S, Coey JD, McKenna J, Bridgett S, Simpson D, Fairley D, Thorne LG, Reuschl AK, Forrest C, Ganeshalingham M, Muir L, Palor M, Jarvis L, Willett B, et al. 2022. Evolutionary remodelling of N-terminal domain loops fine-tunes SARS-CoV-2 spike. *EMBO Reports* **23**:e54322. DOI: <https://doi.org/10.15252/embr.202154322>, PMID: 35999696
- Carabelli AM**, Peacock TP, Thorne LG, Harvey WT, Hughes J, COVID-19 Genomics UK Consortium, Peacock SJ, Barclay WS, de Silva TI, Towers GJ, Robertson DL. 2023. SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nature Reviews. Microbiology* **21**:162–177. DOI: <https://doi.org/10.1038/s41579-022-00841-7>, PMID: 36653446
- Castiglione GM**, Zhou L, Xu Z, Neiman Z, Hung CF, Duh EJ, editor. 2021. Evolutionary pathways to SARS-CoV-2 resistance are opened and closed by epistasis acting on ACE2. *PLOS Biology* **19**:e3001510. DOI: <https://doi.org/10.1371/journal.pbio.3001510>, PMID: 34932561
- Chan CM**, Woo PCY, Lau SKP, Tse H, Chen HL, Li F, Zheng BJ, Chen L, Huang JD, Yuen KY. 2008. RETRACTED: spike protein, s, of human coronavirus HKU1: role in viral life cycle and application in antibody detection. *Experimental Biology and Medicine* **233**:1527–1536. DOI: <https://doi.org/10.3181/0806-RM-197>
- Changeux JP**, Edelstein SJ. 2005. Allosteric mechanisms of signal transduction. *Science* **308**:1424–1428. DOI: <https://doi.org/10.1126/science.1108595>, PMID: 15933191
- Chi X**, Yan R, Zhang J, Zhang G, Zhang Y, Hao M, Zhang Z, Fan P, Dong Y, Yang Y, Chen Z, Guo Y, Zhang J, Li Y, Song X, Chen Y, Xia L, Fu L, Hou L, Xu J, et al. 2020. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**:650–655. DOI: <https://doi.org/10.1126/science.abc6952>, PMID: 32571838
- Collins SR**, Schuldiner M, Krogan NJ, Weissman JS. 2006. A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biology* **7**:R63. DOI: <https://doi.org/10.1186/gb-2006-7-7-r63>
- Damas J**, Hughes GM, Keough KC, Painter CA, Persky NS, Corbo M, Hiller M, Koepfli K-P, Pfenning AR, Zhao H, Genereux DP, Swofford R, Pollard KS, Ryder OA, Nweeia MT, Lindblad-Toh K, Teeling EC, Karlsson EK, Lewin HA. 2020. Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *PNAS* **117**:22311–22322. DOI: <https://doi.org/10.1073/pnas.2010146117>, PMID: 32826334
- de la Paz JA**, Nartey CM, Yuvaraj M, Morcos F. 2020. Epistatic contributions promote the unification of incompatible models of neutral molecular evolution. *PNAS* **117**:5873–5882. DOI: <https://doi.org/10.1073/pnas.1913071117>, PMID: 32123092
- Deng X**, Garcia-Knight MA, Khalid MM, Servellita V, Wang C, Morris MK, Sotomayor-González A, Glasner DR, Reyes KR, Gliwa AS, Reddy NP, Martin CSS, Federman S, Cheng J, Balcerek J, Taylor J, Streithorst JA, Miller S, Kumar GR, Sreekumar B, et al. 2021. Transmission, Infectivity, and Antibody Neutralization of an Emerging SARS-CoV-2 Variant in California Carrying a L452R Spike Protein Mutation. medRxiv. DOI: <https://doi.org/10.1101/2021.03.07.21252647>
- Díaz-Salinas MA**, Li Q, Ejemel M, Yurkovetskiy L, Luban J, Shen K, Wang Y, Munro JB. 2022. Conformational dynamics and allosteric modulation of the SARS-CoV-2 spike. *eLife* **11**:e75433. DOI: <https://doi.org/10.7554/eLife.75433>, PMID: 35323111
- Dicken SJ**, Murray MJ, Thorne LG, Reuschl AK, Forrest C, Ganeshalingham M, Muir L, Kalemra MD, Palor M, McCoy LE, Jolly C, Towers GJ, Reeves MB, Grove J. 2021. Characterisation of B.1.1.7 and pangolin coronavirus spike provides insights on the evolutionary trajectory of SARS-CoV-2. *bioRxiv*. DOI: <https://doi.org/10.1101/2021.03.22.436468>, PMID: 33791702
- Dong A**, Zhao J, Griffin C, Wu R. 2021. The genomic physics of COVID-19 pathogenesis and spread. *Cells* **11**:80. DOI: <https://doi.org/10.3390/cells11010080>, PMID: 35011641
- Doshi U**, Holliday MJ, Eisenmesser EZ, Hamelberg D. 2016. Dynamical network of residue-residue contacts reveals coupled allosteric effects in recognition, catalysis, and mutation. *PNAS* **113**:4735–4740. DOI: <https://doi.org/10.1073/pnas.1523573113>, PMID: 27071107
- Dror RO**, Dirks RM, Grossman JP, Xu H, Shaw DE. 2012. Biomolecular simulation: A computational microscope for molecular biology. *Annual Review of Biophysics* **41**:429–452. DOI: <https://doi.org/10.1146/annurev-biophys-042910-155245>, PMID: 22577825
- Earnest R**, Uddin R, Matluk N, Renzette N, Turbett SE, Siddle KJ, Loreth C, Adams G, Tomkins-Tinch CH, Petrone ME, Rothman JE, Breban MI, Koch RT, Billig K, Fauver JR, Vogels CBF, Bilguvar K, De Kumar B, Landry ML, Peaper DR, et al. 2022. Comparative transmissibility of SARS-CoV-2 variants Delta and Alpha in

- New England, USA. *Cell Reports. Medicine* **3**:100583. DOI: <https://doi.org/10.1016/j.xcrm.2022.100583>, PMID: 35480627
- Ekeberg M**, Lövkvist C, Lan Y, Weigt M, Aurell E. 2013. Improved contact prediction in proteins: Using pseudolikelihoods to infer Potts models. *Physical Review E* **87**:012707. DOI: <https://doi.org/10.1103/PhysRevE.87.012707>
- Fraser BJ**, Beldar S, Seitova A, Hutchinson A, Mannar D, Li Y, Kwon D, Tan R, Wilson RP, Leopold K, Subramaniam S, Halabelian L, Arrowsmith CH, Bénard F. 2022. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nature Chemical Biology* **18**:963–971. DOI: <https://doi.org/10.1038/s41589-022-01059-7>, PMID: 35676539
- Frost SDW**, Magalis BR, Kosakovsky Pond SL. 2018. Neutral theory and rapidly evolving viral pathogens. *Molecular Biology and Evolution* **35**:1348–1354. DOI: <https://doi.org/10.1093/molbev/msy088>, PMID: 29688481
- Garvin MR**, Prates ET, Romero J, Cliff A, Machado Gazolla JGF, Pickholz M, Pavicic M, Jacobson D. 2021. Rapid Expansion of SARS-CoV-2 variants of concern is a result of adaptive epistasis. *bioRxiv*. DOI: <https://doi.org/10.1101/2021.08.03.454981>
- Gerek ZN**, Ozkan SB. 2011. Change in allosteric network affects binding affinities of PDZ domains. *PLOS Computational Biology* **7**:e1002154. DOI: <https://doi.org/10.1371/journal.pcbi.1002154>
- Gobeil SMC**, Janowska K, McDowell S, Mansouri K, Parks R, Manne K, Stalls V, Kopp MF, Henderson R, Edwards RJ, Haynes BF, Acharya P. 2021a. D614G mutation alters SARS-CoV-2 spike conformation and enhances protease cleavage at the S1/S2 junction. *Cell Reports* **34**:108630. DOI: <https://doi.org/10.1016/j.celrep.2020.108630>, PMID: 33417835
- Gobeil SM-C**, Janowska K, McDowell S, Mansouri K, Parks R, Stalls V, Kopp MF, Manne K, Li D, Wiehe K, Saunders KO, Edwards RJ, Korber B, Haynes BF, Henderson R, Acharya P. 2021b. Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity. *Science* **373**:eabi6226. DOI: <https://doi.org/10.1126/science.abi6226>, PMID: 34168071
- Gur M**, Taka E, Yilmaz SZ, Kilinc C, Aktas U, Golcuk M. 2020. Conformational transition of SARS-CoV-2 spike glycoprotein between its closed and open states. *The Journal of Chemical Physics* **153**:075101. DOI: <https://doi.org/10.1063/5.0011141>, PMID: 32828084
- Harrigan P**, Madhani HD, El-Samad H. 2018. Real-time genetic compensation defines the dynamic demands of feedback control. *Cell* **175**:877–886. DOI: <https://doi.org/10.1016/j.cell.2018.09.044>, PMID: 30340045
- Harvey WT**, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R, Rambaut A, COVID-19 Genomics UK (COG-UK) Consortium, Peacock SJ, Robertson DL. 2021. SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews. Microbiology* **19**:409–424. DOI: <https://doi.org/10.1038/s41579-021-00573-0>, PMID: 34075212
- Henderson R**, Edwards RJ, Mansouri K, Janowska K, Stalls V, Gobeil SMC, Kopp M, Li D, Parks R, Hsu AL, Borgnia MJ, Haynes BF, Acharya P. 2020. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nature Structural & Molecular Biology* **27**:925–933. DOI: <https://doi.org/10.1038/s41594-020-0479-4>, PMID: 32699321
- Hoffmann M**, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA, Drosten C, Pöhlmann S. 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**:271–280. DOI: <https://doi.org/10.1016/j.cell.2020.02.052>
- Hong Q**, Han W, Li J, Xu S, Wang Y, Xu C, Li Z, Wang Y, Zhang C, Huang Z, Cong Y. 2022. Molecular basis of receptor binding and antibody neutralization of Omicron. *Nature* **604**:546–552. DOI: <https://doi.org/10.1038/s41586-022-04581-9>, PMID: 35228716
- Houhamdi L**, Gautret P, Hoang VT, Fournier P, Colson P, Raoult D. 2022. Characteristics of the first 1119 SARS-CoV-2 Omicron variant cases. *Journal of Medical Virology* **94**:2290–2295. DOI: <https://doi.org/10.1002/jmv.27613>
- Huang Y**, Yang C, Xu X-F, Xu W, Liu S-W. 2020. Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19. *Acta Pharmacologica Sinica* **41**:1141–1149. DOI: <https://doi.org/10.1038/s41401-020-0485-4>, PMID: 32747721
- Hünenberger PH**. 2005. Thermostat Algorithms for molecular Dynamics simulations. Holm C, Kremer K (Eds). *Advanced Computer Simulation*. Springer. p. 105–149. DOI: <https://doi.org/10.1007/b99427>
- Jackson CB**, Farzan M, Chen B, Choe H. 2022. Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews. Molecular Cell Biology* **23**:3–20. DOI: <https://doi.org/10.1038/s41580-021-00418-x>, PMID: 34611326
- Kazan IC**, Sharma P, Rahman MI, Bobkov A, Fromme R, Ghirlanda G, Ozkan SB. 2022. Design of novel cyanovirin-N variants by modulation of binding dynamics through distal mutations. *eLife* **11**:eLife. DOI: <https://doi.org/10.7554/eLife.67474>, PMID: 36472898
- Kazan IC**, Mills JH, Ozkan SB. 2023. Allosteric regulatory control in dihydrofolate reductase is revealed by dynamic asymmetry. *Protein Science* **32**:e4700. DOI: <https://doi.org/10.1002/pro.4700>, PMID: 37313628
- Kemp SA**, Collier DA, Datir RP, Ferreira IATM, Gayed S, Jahun A, Hosmillo M, Rees-Spear C, Mlcochova P, Lumb IU, Roberts DJ, Chandra A, Temperton N, CITIID-NIHR BioResource COVID-19 Collaboration, COVID-19 Genomics UK (COG-UK) Consortium, Sharrocks K, Blane E, Modis Y, Leigh KE, Briggs JAG, et al. 2021. SARS-CoV-2 evolution during treatment of chronic infection. *Nature* **592**:277–282. DOI: <https://doi.org/10.1038/s41586-021-03291-y>, PMID: 33545711
- Keskin O**, Bahar I, Jernigan RL, Beutler JA, Shoemaker RH, Sausville EA, Covell DG. 2000. Characterization of anticancer agents by their growth inhibitory activity and relationships to mechanism of action and structure. *Anti-Cancer Drug Design* **15**:79–98 PMID: 10901296.

- Khan A**, Zia T, Suleman M, Khan T, Ali SS, Abbasi AA, Mohammad A, Wei D-Q. 2021. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. *Journal of Cellular Physiology* **236**:7045–7057. DOI: <https://doi.org/10.1002/jcp.30367>, PMID: 33755190
- Kim H**, Zou T, Modi C, Dörner K, Grunkemeyer TJ, Chen L, Fromme R, Matz MV, Ozkan SB, Wachter RM. 2015. A hinge migration mechanism unlocks the evolution of green-to-red photoconversion in GFP-like proteins. *Structure* **23**:34–43. DOI: <https://doi.org/10.1016/j.str.2014.11.011>, PMID: 25565105
- Kim S**, Nguyen TT, Taïtt AS, Jhun H, Park H-Y, Kim S-H, Kim Y-G, Song EY, Lee Y, Yum H, Shin K-C, Choi YK, Song C-S, Yeom SC, Kim B, Netea M, Kim S. 2021. SARS-CoV-2 omicron mutation is faster than the chase: multiple mutations on Spike/ACE2 interaction residues. *Immune Network* **21**:e38. DOI: <https://doi.org/10.4110/in.2021.21.e38>
- Kimura M**. 1983. The neutral theory of molecular evolution. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511623486>
- Kirchdoerfer RN**, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, Corbett KS, Graham BS, McLellan JS, Ward AB. 2016. Pre-fusion structure of a human coronavirus spike protein. *Nature* **531**:118–121. DOI: <https://doi.org/10.1038/nature17200>, PMID: 26935699
- Kistler KE**, Huddleston J, Bedford T. 2022. Rapid and parallel adaptive mutations in spike S1 drive clade success in SARS-CoV-2. *Cell Host & Microbe* **30**:545–555. DOI: <https://doi.org/10.1016/j.chom.2022.03.018>, PMID: 35364015
- Klinakis A**, Cournia Z, Rampias T. 2021. N-terminal domain mutations of the spike protein are structurally implicated in epitope recognition in emerging SARS-CoV-2 strains. *Computational and Structural Biotechnology Journal* **19**:5556–5567. DOI: <https://doi.org/10.1016/j.csbj.2021.10.004>, PMID: 34630935
- Kolbaba-Kartchner B**, Kazan IC, Mills JH, Ozkan SB. 2021. The role of rigid residues in modulating TEM-1 β -lactamase function and thermostability. *International Journal of Molecular Sciences* **22**:2895. DOI: <https://doi.org/10.3390/ijms22062895>, PMID: 33809335
- Kumar A**, Glembo TJ, Ozkan SB. 2015. The role of conformational dynamics and allostery in the disease development of human ferritin. *Biophysical Journal* **109**:1273–1281. DOI: <https://doi.org/10.1016/j.bpj.2015.06.060>, PMID: 26255589
- Kumar S**, Patel R. 2018. Neutral theory, disease mutations, and personal exomes. *Molecular Biology and Evolution* **35**:1297–1303. DOI: <https://doi.org/10.1093/molbev/msy085>, PMID: 29688514
- Kumar S**, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**:1547–1549. DOI: <https://doi.org/10.1093/molbev/msy096>, PMID: 29722887
- Kumar S**, Tao Q, Weaver S, Sanderford M, Caraballo-Ortiz MA, Sharma S, Pond SLK, Miura S. 2021. An evolutionary portrait of the progenitor SARS-CoV-2 and its dominant offshoots in COVID-19 pandemic. *Molecular Biology and Evolution* **38**:3046–3059. DOI: <https://doi.org/10.1093/molbev/msab118>
- Kuzmanic A**, Bowman GR, Juarez-Jimenez J, Michel J, Gervasio FL. 2020. Investigating cryptic binding sites by molecular dynamics simulations. *Accounts of Chemical Research* **53**:654–661. DOI: <https://doi.org/10.1021/acs.accounts.9b00613>, PMID: 32134250
- Labbadia J**, Morimoto RI. 2015. The biology of proteostasis in aging and disease. *Annual Review of Biochemistry* **84**:435–464. DOI: <https://doi.org/10.1146/annurev-biochem-060614-033955>, PMID: 25784053
- Laiton-Donato K**, Franco-Muñoz C, Álvarez-Díaz DA, Ruiz-Moreno HA, Usme-Ciro JA, Prada DA, Reales-González J, Corchuelo S, Herrera-Sepúlveda MT, Naizaque J, Santamaría G, Rivera J, Rojas P, Ortiz JH, Cardona A, Malo D, Prieto-Alvarado F, Gómez FR, Wiesner M, Martínez MLO, et al. 2021. Characterization of the emerging B.1.621 variant of interest of SARS-CoV-2. *Infection, Genetics and Evolution* **95**:105038. DOI: <https://doi.org/10.1016/j.meegid.2021.105038>, PMID: 34403832
- Lan J**, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X. 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**:215–220. DOI: <https://doi.org/10.1038/s41586-020-2180-5>, PMID: 32225176
- Larrimore KE**, Kazan IC, Kannan L, Kendle RP, Jamal T, Barcus M, Bolia A, Brimijoin S, Zhan C-G, Ozkan SB, Mor TS. 2017. Plant-expressed cocaine hydrolase variants of butyrylcholinesterase exhibit altered allosteric effects of cholinesterase activity and increased inhibitor sensitivity. *Scientific Reports* **7**:10419. DOI: <https://doi.org/10.1038/s41598-017-10571-z>, PMID: 28874829
- Levy RM**, Haldane A, Flynn WF. 2017. Potts Hamiltonian models of protein co-variation, free energy landscapes, and evolutionary fitness. *Current Opinion in Structural Biology* **43**:55–62. DOI: <https://doi.org/10.1016/j.sbi.2016.11.004>, PMID: 27870991
- Liu Y**, Bahar I. 2012. Sequence evolution correlates with structural dynamics. *Molecular Biology and Evolution* **29**:2253–2263. DOI: <https://doi.org/10.1093/molbev/mss097>, PMID: 22427707
- Liu L**, Tamura K, Sanderford M, Gray VE, Kumar S. 2016. A molecular evolutionary reference for the human variome. *Molecular Biology and Evolution* **33**:245–254. DOI: <https://doi.org/10.1093/molbev/msv198>, PMID: 26464126
- Liu X**, Liu X, Zhou J, Dong Y, Jiang W, Jiang W. 2022. Rampant C-to-U deamination accounts for the intrinsically high mutation rate in SARS-CoV-2 spike gene. *RNA* **28**:917–926. DOI: <https://doi.org/10.1261/rna.079160.122>
- Ma B**, Nussinov R. 2016. Conformational footprints. *Nature Chemical Biology* **12**:890–891. DOI: <https://doi.org/10.1038/nchembio.2212>
- Maguid S**, Fernández-Alberti S, Parisi G, Echave J. 2006. Evolutionary conservation of protein backbone flexibility. *Journal of Molecular Evolution* **63**:448–457. DOI: <https://doi.org/10.1007/s00239-005-0209-x>, PMID: 17021932

- Maguid S**, Fernandez-Alberti S, Echave J. 2008. Evolutionary conservation of protein vibrational dynamics. *Gene* **422**:7–13. DOI: <https://doi.org/10.1016/j.gene.2008.06.002>, PMID: 18577430
- Maher MC**, Bartha I, Weaver S, di Iulio J, Ferri E, Soriaga L, Lempp FA, Hie BL, Bryson B, Berger B, Robertson DL, Snell G, Corti D, Virgin HW, Kosakovsky Pond SL, Telenti A. 2022. Predicting the mutational drivers of future SARS-CoV-2 variants of concern. *Science Translational Medicine* **14**:eabk3445. DOI: <https://doi.org/10.1126/scitranslmed.abk3445>
- Maier JA**, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation* **11**:3696–3713. DOI: <https://doi.org/10.1021/acs.jctc.5b00255>
- Majumder S**, Chaudhuri D, Datta J, Giri K. 2021. Exploring the intrinsic dynamics of SARS-CoV-2, SARS-CoV and MERS-CoV spike glycoprotein through normal mode analysis using anisotropic network model. *Journal of Molecular Graphics & Modelling* **102**:107778. DOI: <https://doi.org/10.1016/j.jmgm.2020.107778>, PMID: 33099199
- Markov PV**, Ghafari M, Beer M, Lythgoe K, Simmonds P, Stilianakis NI, Katzourakis A. 2023. The evolution of SARS-CoV-2. *Nature Reviews Microbiology* **21**:361–379. DOI: <https://doi.org/10.1038/s41579-023-00878-2>
- Menni C**, Valdes AM, Polidori L, Antonelli M, Penamakuri S, Nogal A, Louca P, May A, Figueiredo JC, Hu C, Molteni E, Canas L, Österdahl MF, Modat M, Sudre CH, Fox B, Hammers A, Wolf J, Capdevila J, Chan AT, et al. 2022. Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of omicron and delta variant dominance: a prospective observational study from the ZOE COVID Study. *The Lancet* **399**:1618–1624. DOI: [https://doi.org/10.1016/S0140-6736\(22\)00327-0](https://doi.org/10.1016/S0140-6736(22)00327-0)
- Mikulska-Ruminska K**, Shrivastava I, Krieger J, Zhang S, Li H, Bayir H, Wenzel SE, VanDemark AP, Kagan VE, Bahar I. 2019. Characterization of differential dynamics, specificity, and allostery of lipoygenase family members. *Journal of Chemical Information and Modeling* **59**:2496–2508. DOI: <https://doi.org/10.1021/acs.jcim.9b00006>, PMID: 30762363
- Millet JK**, Whittaker GR. 2014. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *PNAS* **111**:15214–15219. DOI: <https://doi.org/10.1073/pnas.1407087111>, PMID: 25288733
- Millet JK**, Whittaker GR. 2015. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research* **202**:120–134. DOI: <https://doi.org/10.1016/j.virusres.2014.11.021>, PMID: 25445340
- Mishra SK**, Jernigan RL. 2018. Protein dynamic communities from elastic network models align closely to the communities defined by molecular dynamics. *PLOS ONE* **13**:e0199225. DOI: <https://doi.org/10.1371/journal.pone.0199225>, PMID: 29924847
- Modi T**, Ozkan SB. 2018. Mutations utilize dynamic allostery to confer resistance in TEM-1 β -lactamase. *International Journal of Molecular Sciences* **19**:3808. DOI: <https://doi.org/10.3390/ijms19123808>, PMID: 30501088
- Modi T**, Campitelli P, Kazan IC, Ozkan SB. 2021a. Protein folding stability and binding interactions through the lens of evolution: a dynamical perspective. *Current Opinion in Structural Biology* **66**:207–215. DOI: <https://doi.org/10.1016/j.sbi.2020.11.007>, PMID: 33388636
- Modi T**, Risso VA, Martinez-Rodriguez S, Gavira JA, Mebrat MD, Van Horn WD, Sanchez-Ruiz JM, Banu Ozkan S. 2021b. Hinge-shift mechanism as a protein design principle for the evolution of β -lactamases from substrate promiscuity to specificity. *Nature Communications* **12**:1852. DOI: <https://doi.org/10.1038/s41467-021-22089-0>, PMID: 33767175
- Moulana A**, Dupic T, Phillips AM, Chang J, Nieves S, Roffler AA, Greaney AJ, Starr TN, Bloom JD, Desai MM. 2022. Compensatory epistasis maintains ACE2 affinity in SARS-CoV-2 Omicron BA.1. *Nature Communications* **13**:7011. DOI: <https://doi.org/10.1038/s41467-022-34506-z>, PMID: 36384919
- Moulana A**, Dupic T, Phillips AM, Chang J, Roffler AA, Greaney AJ, Starr TN, Bloom JD, Desai MM. 2023. The landscape of antibody binding affinity in SARS-CoV-2 Omicron BA.1 evolution. *eLife* **12**:e83442. DOI: <https://doi.org/10.7554/eLife.83442>, PMID: 36803543
- Neher RA**. 2022. Contributions of adaptation and purifying selection to SARS-CoV-2 evolution. *Virus Evolution* **8**:veac113. DOI: <https://doi.org/10.1093/ve/veac113>, PMID: 37593203
- Nevin Gerek Z**, Kumar S, Banu Ozkan S. 2013. Structural dynamics flexibility informs function and evolution at a proteome scale. *Evolutionary Applications* **6**:423–433. DOI: <https://doi.org/10.1111/eva.12052>, PMID: 23745135
- Nguyen HL**, Lan PD, Thai NQ, Nissley DA, O'Brien EP, Li MS. 2020. Does SARS-CoV-2 bind to human ACE2 More strongly than does SARS-CoV? *The Journal of Physical Chemistry. B* **124**:7336–7347. DOI: <https://doi.org/10.1021/acs.jpcc.0c04511>, PMID: 32790406
- Nielsen BF**, Li Y, Sneppen K, Simonsen L, Viboud C, Levin SA, Grenfell BT. 2022. Immune heterogeneity and epistasis explain punctuated evolution of SARS-CoV-2. *bioRxiv*. DOI: <https://doi.org/10.1101/2022.07.27.22278129>
- Nussinov R**, Tsai CJ. 2013. Allostery in disease and in drug discovery. *Cell* **153**:293–305. DOI: <https://doi.org/10.1016/j.cell.2013.03.034>, PMID: 23582321
- O'Rourke KF**, Gorman SD, Boehr DD. 2016. Biophysical and computational methods to analyze amino acid interaction networks in proteins. *Computational and Structural Biotechnology Journal* **14**:245–251. DOI: <https://doi.org/10.1016/j.csbj.2016.06.002>

- Ose NJ**, Butler BM, Kumar A, Kazan IC, Sanderford M, Kumar S, Ozkan SB. 2022a. Dynamic coupling of residues within proteins as a mechanistic foundation of many enigmatic pathogenic missense variants. *PLOS Computational Biology* **18**:e1010006. DOI: <https://doi.org/10.1371/journal.pcbi.1010006>, PMID: 35389981
- Ose NJ**, Campitelli P, Patel RP, Ozkan SB, Kumar S. 2022b. Protein dynamics provide mechanistic insights about the epistatic relationships among highly observed potentially adaptive missense variants. *Biophysical Journal* **121**:456a. DOI: <https://doi.org/10.1016/j.bpj.2021.11.488>
- Otten R**, Liu L, Kenner LR, Clarkson MW, Mavor D, Tawfik DS, Kern D, Fraser JS. 2018. Rescue of conformational dynamics in enzyme catalysis by directed evolution. *Nature Communications* **9**:1314. DOI: <https://doi.org/10.1038/s41467-018-03562-9>
- Ozkan SB**. 2024. Dfi-DCI. swh:1:rev:7a3a54bbe03e8356036bdb164ad1b2c01c103cc5. Software Heritage. <https://archive.softwareheritage.org/swh:1:dir:9f38b3c516e31bcfe8b796f7de4082eec6ecba4c;origin=https://github.com/SBOZKAN/DFI-DCI;visit=swh:1:snp:d723c9bc100d4a98dab72447e9b39707e2e33f9b;anchor=swh:1:rev:7a3a54bbe03e8356036bdb164ad1b2c01c103cc5>
- Ozono S**, Zhang Y, Ode H, Sano K, Tan TS, Imai K, Miyoshi K, Kishigami S, Ueno T, Iwatani Y, Suzuki T, Tokunaga K. 2021. SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity. *Nature Communications* **12**:848. DOI: <https://doi.org/10.1038/s41467-021-21118-2>, PMID: 33558493
- Pearlman DA**, Case DA, Caldwell JW, Ross WS, Cheatham TE, DeBolt S, Ferguson D, Seibel G, Kollman P. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications* **91**:1–41. DOI: [https://doi.org/10.1016/0010-4655\(95\)00041-D](https://doi.org/10.1016/0010-4655(95)00041-D)
- Peters AD**, Lively CM. 1999. The red queen and fluctuating epistasis: a population genetic analysis of antagonistic coevolution. *The American Naturalist* **154**:393–405. DOI: <https://doi.org/10.1086/303247>, PMID: 10523486
- Pipitò L**, Rujan R-M, Reynolds CA, Deganutti G. 2022. Molecular dynamics studies reveal structural and functional features of the SARS-CoV-2 spike protein. *BioEssays* **44**:e2200060. DOI: <https://doi.org/10.1002/bies.202200060>, PMID: 35843871
- Qu P**, Faraone JN, Evans JP, Zheng YM, Carlin C, Anghelina M, Stevens P, Fernandez S, Jones D, Panchal A, Saif LJ, Oltz EM, Xu K, Gumina RJ, Liu SL. 2023. Extraordinary evasion of neutralizing antibody response by omicron XBB.1.5, CH.1.1 and CA.3.1 variants. *bioRxiv*. DOI: <https://doi.org/10.1101/2023.01.16.524244>, PMID: 36711991
- Raghuvamsi PV**, Tulsian NK, Samsudin F, Qian X, Purushotorman K, Yue G, Kozma MM, Hwa WY, Lescar J, Bond PJ, MacAry PA, Anand GS. 2021. SARS-CoV-2 S protein:ACE2 interaction reveals novel allosteric targets. *eLife* **10**:e63646. DOI: <https://doi.org/10.7554/eLife.63646>, PMID: 33554856
- Ramarao-Milne P**, Jain Y, Sng LMF, Hosking B, Lee C, Bayat A, Kuiper M, Wilson LOW, Twine NA, Bauer DC. 2022. Data-driven platform for identifying variants of interest in COVID-19 virus. *Computational and Structural Biotechnology Journal* **20**:2942–2950. DOI: <https://doi.org/10.1016/j.csbj.2022.06.005>, PMID: 35677774
- Rehman SU**, Shafique L, Ihsan A, Liu Q. 2020. Evolutionary trajectory for the emergence of novel coronavirus SARS-CoV-2. *Pathogens* **9**:240. DOI: <https://doi.org/10.3390/pathogens9030240>, PMID: 32210130
- Rizzato F**, Coucke A, de Leonardis E, Barton JP, Tubiana J, Monasson R, Cocco S. 2020. Inference of compressed Potts graphical models. *Physical Review. E* **101**:012309. DOI: <https://doi.org/10.1103/PhysRevE.101.012309>, PMID: 32069678
- Rochman ND**, Wolf YI, Faure G, Mutz P, Zhang F, Koonin EV. 2021. Ongoing global and regional adaptive evolution of SARS-CoV-2. *PNAS* **118**:e2104241118. DOI: <https://doi.org/10.1073/pnas.2104241118>
- Rochman ND**, Faure G, Wolf YI, Freddolino PL, Zhang F, Koonin EV. 2022. Epistasis at the SARS-CoV-2 receptor-binding domain interface and the propitiously boring implications for vaccine escape. *mBio* **13**:e0013522. DOI: <https://doi.org/10.1128/mbio.00135-22>, PMID: 35289643
- Rodriguez-Rivas J**, Croce G, Muscat M, Weigt M. 2022. Epistatic models predict mutable sites in SARS-CoV-2 proteins and epitopes. *PNAS* **119**:e2113118119. DOI: <https://doi.org/10.1073/pnas.2113118119>, PMID: 35022216
- Rojas Echenique JI**, Kryazhinskiy S, Nguyen Ba AN, Desai MM, editor. 2019. Modular epistasis and the compensatory evolution of gene deletion mutants. *PLOS Genetics* **15**:e1007958. DOI: <https://doi.org/10.1371/journal.pgen.1007958>, PMID: 30768593
- Saavedra HG**, Wrabl JO, Anderson JA, Li J, Hilser VJ. 2018. Dynamic allostery can drive cold adaptation in enzymes. *Nature* **558**:324–328. DOI: <https://doi.org/10.1038/s41586-018-0183-2>
- Saputri DS**, Li S, van Eerden FJ, Rozewicki J, Xu Z, Ismanto HS, Davila A, Teraguchi S, Katoh K, Standley DM. 2020. Flexible, functional, and familiar: characteristics of SARS-CoV-2 Spike protein evolution. *Frontiers in Microbiology* **11**:2112. DOI: <https://doi.org/10.3389/fmicb.2020.02112>, PMID: 33042039
- Sawle L**, Ghosh K. 2016. Convergence of molecular dynamics simulation of protein native states: feasibility vs self-consistency dilemma. *Journal of Chemical Theory and Computation* **12**:861–869. DOI: <https://doi.org/10.1021/acs.jctc.5b00999>
- Sekhar A**, Kay LE. 2019. An NMR view of protein dynamics in health and disease. *Annual Review of Biophysics* **48**:297–319. DOI: <https://doi.org/10.1146/annurev-biophys-052118-115647>, PMID: 30901260
- Shang J**, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, Li F. 2020. Cell entry mechanisms of SARS-CoV-2. *PNAS* **117**:11727–11734. DOI: <https://doi.org/10.1073/pnas.2003138117>
- Sheikh A**, McMenamin J, Taylor B, Robertson C, Public Health Scotland and the EAVE II Collaborators. 2021. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *The Lancet* **397**:2461–2462. DOI: [https://doi.org/10.1016/S0140-6736\(21\)01358-1](https://doi.org/10.1016/S0140-6736(21)01358-1), PMID: 34139198

- Shimagaki K**, Weigt M. 2019. Selection of sequence motifs and generative Hopfield-Potts models for protein families. *Physical Review E* **100**:032128. DOI: <https://doi.org/10.1103/PhysRevE.100.032128>
- Shoemark DK**, Oliveira ASF, Davidson AD, Berger I, Schaffitzel C, Mulholland AJ. 2022. Molecular dynamics of spike variants in the locked conformation: RBD interfaces, fatty acid binding and furin cleavage sites. *bioRxiv*. DOI: <https://doi.org/10.1101/2022.05.06.490927>
- Singh D**, Yi SV. 2021. On the origin and evolution of SARS-CoV-2. *Experimental & Molecular Medicine* **53**:537–547. DOI: <https://doi.org/10.1038/s12276-021-00604-z>
- Socher E**, Conrad M, Heger L, Paulsen F, Sticht H, Zunke F, Arnold P. 2021. Mutations in the B.1.1.7 SARS-CoV-2 spike protein reduce receptor-binding affinity and induce a flexible link to the fusion peptide. *Biomedicines* **9**:525. DOI: <https://doi.org/10.3390/biomedicines9050525>, PMID: 34066729
- Spinello A**, Saltalamacchia A, Borišek J, Magistrato A. 2021. Allosteric cross-talk among spike's receptor-binding domain mutations of the SARS-CoV-2 South African variant triggers an effective hijacking of human cell receptor. *The Journal of Physical Chemistry Letters* **12**:5987–5993. DOI: <https://doi.org/10.1021/acs.jpcllett.1c01415>
- Starr TN**, Greaney AJ, Hannon WW, Loes AN, Hauser K, Dillen JR, Ferri E, Farrell AG, Dadonaite B, McCallum M, Matreyek KA, Corti D, Veessler D, Snell G, Bloom JD. 2022a. Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. *Science* **377**:420–424. DOI: <https://doi.org/10.1126/science.abo7896>, PMID: 35762884
- Starr TN**, Greaney AJ, Stewart CM, Walls AC, Hannon WW, Veessler D, Bloom JD. 2022b. Deep mutational scans for ACE2 binding, RBD expression, and antibody escape in the SARS-CoV-2 Omicron BA.1 and BA.2 receptor-binding domains. *PLOS Pathogens* **18**:e1010951. DOI: <https://doi.org/10.1371/journal.ppat.1010951>, PMID: 36399443
- Starr TN**, Zepeda SK, Walls AC, Greaney AJ, Alkhovsky S, Veessler D, Bloom JD. 2022c. ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. *Nature* **603**:913–918. DOI: <https://doi.org/10.1038/s41586-022-04464-z>
- Steinhauer DA**. 1999. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. *Virology* **258**:1–20. DOI: <https://doi.org/10.1006/viro.1999.9716>, PMID: 10329563
- Stevens AO**, Kazan IC, Ozkan B, He Y. 2022. Investigating the allosteric response of the PICK1 PDZ domain to different ligands with all-atom simulations. *Protein Science* **31**:e4474. DOI: <https://doi.org/10.1002/pro.4474>, PMID: 36251217
- Sun Y**, Kollman PA. 1995. Hydrophobic solvation of methane and nonbond parameters of the TIP3P water model. *Journal of Computational Chemistry* **16**:1164–1169. DOI: <https://doi.org/10.1002/jcc.540160910>
- Swint-Kruse L**, Matthews KS, Smith PE, Pettitt BM. 1998. Comparison of simulated and experimentally determined dynamics for a variant of the LacI DNA-binding domain, Nlac-P. *Biophysical Journal* **74**:413–421. DOI: [https://doi.org/10.1016/s0006-3495\(98\)77798-7](https://doi.org/10.1016/s0006-3495(98)77798-7), PMID: 9449341
- Sztain T**, Ahn S-H, Bogetti AT, Casalino L, Goldsmith JA, Seitz E, McCool RS, Kearns FL, Acosta-Reyes F, Maji S, Mashayekhi G, McCammon JA, Ourmazd A, Frank J, McLellan JS, Chong LT, Amaro RE. 2021. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nature Chemistry* **13**:963–968. DOI: <https://doi.org/10.1038/s41557-021-00758-3>
- Tan ZW**, Tee WV, Samsudin F, Guarnera E, Bond PJ, Berezovsky IN. 2022. Allosteric perspective on the mutability and druggability of the SARS-CoV-2 Spike protein. *Structure* **30**:590–607. DOI: <https://doi.org/10.1016/j.str.2021.12.011>
- Tang X**, Wu C, Li X, Song Y, Yao X, Wu X, Duan Y, Zhang H, Wang Y, Qian Z, Cui J, Lu J. 2020. On the origin and continuing evolution of SARS-CoV-2. *National Science Review* **7**:1012–1023. DOI: <https://doi.org/10.1093/nsr/nwaa036>
- Tay JH**, Porter AF, Wirth W, Duchene S. 2022. The emergence of SARS-CoV-2 variants of concern is driven by acceleration of the substitution rate. *Molecular Biology and Evolution* **39**:msac013. DOI: <https://doi.org/10.1093/molbev/msac013>, PMID: 35038741
- Teruel N**, Crown M, Bashton M, Najmanovich R. 2021a. Computational analysis of the effect of SARS-CoV-2 variant omicron spike protein mutations on dynamics, ACE2 binding and propensity for immune escape. *bioRxiv*. DOI: <https://doi.org/10.1101/2021.12.14.472622>
- Teruel N**, Mailhot O, Najmanovich RJ. 2021b. Modelling conformational state dynamics and its role on infection for SARS-CoV-2 Spike protein variants. *PLOS Computational Biology* **17**:e1009286. DOI: <https://doi.org/10.1371/journal.pcbi.1009286>
- Thye AYK**, Law JWF, Pusparajah P, Letchumanan V, Chan KG, Lee LH. 2021. Emerging SARS-CoV-2 Variants of Concern (VOCs): an impending global crisis. *Biomedicines* **9**:1303. DOI: <https://doi.org/10.3390/biomedicines9101303>
- Twohig KA**, Nyberg T, Zaidi A, Thelwall S, Sinnathamby MA, Aliabadi S, Seaman SR, Harris RJ, Hope R, Lopez-Bernal J, Gallagher E, Charlett A, De Angelis D, Presanis AM, Dabrera G, COVID-19 Genomics UK (COG-UK) consortium. 2022. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *The Lancet Infectious Diseases* **22**:35–42. DOI: [https://doi.org/10.1016/S1473-3099\(21\)00475-8](https://doi.org/10.1016/S1473-3099(21)00475-8), PMID: 34461056
- Verkhivker G**. 2022. Allosteric determinants of the SARS-CoV-2 Spike protein binding with nanobodies: examining mechanisms of mutational escape and sensitivity of the omicron variant. *International Journal of Molecular Sciences* **23**:2172. DOI: <https://doi.org/10.3390/ijms23042172>, PMID: 35216287

- Walls AC**, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. 2020. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**:281–292. DOI: <https://doi.org/10.1016/j.cell.2020.02.058>, PMID: 32155444
- Wang Q**, Guo Y, Iketani S, Nair MS, Li Z, Mohri H, Wang M, Yu J, Bowen AD, Chang JY, Shah JG, Nguyen N, Chen Z, Meyers K, Yin MT, Sobieszczyk ME, Sheng Z, Huang Y, Liu L, Ho DD. 2022a. Antibody evasion by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4 and BA.5. *Nature* **608**:603–608. DOI: <https://doi.org/10.1038/s41586-022-05053-w>, PMID: 35790190
- Wang Q**, Iketani S, Li Z, Liu L, Guo Y, Huang Y, Bowen AD, Liu M, Wang M, Yu J, Valdez R, Luring AS, Sheng Z, Wang HH, Gordon A, Liu L, Ho DD. 2022b. Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *bioRxiv*. DOI: <https://doi.org/10.1101/2022.11.23.517532>
- Weisblum Y**, Schmidt F, Zhang F, DaSilva J, Poston D, Lorenzi JC, Muecksch F, Rutkowska M, Hoffmann H-H, Michailidis E, Gaebler C, Agudelo M, Cho A, Wang Z, Gazumyan A, Cipolla M, Luchsinger L, Hillyer CD, Caskey M, Robbiani DF, et al. 2020. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife* **9**:e61312. DOI: <https://doi.org/10.7554/eLife.61312>, PMID: 33112236
- Witte L**, Baharani VA, Schmidt F, Wang Z, Cho A, Raspe R, Guzman-Cardozo C, Muecksch F, Canis M, Park DJ, Gaebler C, Caskey M, Nussenzweig MC, Hatziioannou T, Bieniasz PD. 2023. Epistasis lowers the genetic barrier to SARS-CoV-2 neutralizing antibody escape. *Nature Communications* **14**:302. DOI: <https://doi.org/10.1038/s41467-023-35927-0>, PMID: 36653360
- Wrapp D**, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**:1260–1263. DOI: <https://doi.org/10.1126/science.abb2507>, PMID: 32075877
- Wrobel AG**, Benton DJ, Xu P, Roustan C, Martin SR, Rosenthal PB, Skehel JJ, Gamblin SJ. 2020. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nature Structural & Molecular Biology* **27**:763–767. DOI: <https://doi.org/10.1038/s41594-020-0468-7>, PMID: 32647346
- Wu L**, Zhou L, Mo M, Liu T, Wu C, Gong C, Lu K, Gong L, Zhu W, Xu Z. 2022. SARS-CoV-2 Omicron RBD shows weaker binding affinity than the currently dominant Delta variant to human ACE2. *Signal Transduction and Targeted Therapy* **7**:8. DOI: <https://doi.org/10.1038/s41392-021-00863-2>, PMID: 34987150
- Xue Q**, Liu X, Pan W, Zhang A, Fu J, Jiang G. 2022. Computational insights into the allosteric effect and dynamic structural features of the SARS-CoV-2 spike protein. *Chemistry* **28**:e202104215. DOI: <https://doi.org/10.1002/chem.202104215>, PMID: 34962015
- Yang QE**, MacLean C, Papkou A, Pritchard M, Powell L, Thomas D, Andrey DO, Li M, Spiller B, Yang W, Walsh TR. 2020. Compensatory mutations modulate the competitiveness and dynamics of plasmid-mediated colistin resistance in *Escherichia coli* clones. *The ISME Journal* **14**:861–865. DOI: <https://doi.org/10.1038/s41396-019-0578-6>, PMID: 31896787
- Yue C**, Song W, Wang L, Jian F, Chen X, Gao F, Shen Z, Wang Y, Wang X, Cao Y. 2023. Enhanced transmissibility of XBB.1.5 is contributed by both strong ACE2 binding and antibody evasion. *bioRxiv*. DOI: <https://doi.org/10.1101/2023.01.03.522427>
- Yurkovetskiy L**, Wang X, Pascal KE, Tomkins-Tinch C, Nyalile TP, Wang Y, Baum A, Diehl WE, Dauphin A, Carbone C, Veinotte K, Egri SB, Schaffner SF, Lemieux JE, Munro JB, Rafique A, Barve A, Sabeti PC, Kyratsous CA, Dudkina NV, et al. 2020. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell* **183**:739–751. DOI: <https://doi.org/10.1016/j.cell.2020.09.032>, PMID: 32991842
- Zeng HL**, Dichio V, Rodríguez Horta E, Thorell K, Aurell E. 2020. Global analysis of more than 50,000 SARS-CoV-2 genomes reveals epistasis between eight viral genes. *PNAS* **117**:31519–31526. DOI: <https://doi.org/10.1073/pnas.2012331117>, PMID: 33203681
- Zhang J**, Cai Y, Xiao T, Lu J, Peng H, Sterling SM, Walsh RM Jr, Rits-Volloch S, Zhu H, Woosley AN, Yang W, Sliz P, Chen B. 2021. Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **372**:525–530. DOI: <https://doi.org/10.1126/science.abb2303>, PMID: 33727252
- Zhou T**, Tsybovsky Y, Gorman J, Rapp M, Cerutti G, Chuang GY, Katsamba PS, Sampson JM, Schön A, Bimela J, Boyington JC, Nazzari A, Olia AS, Shi W, Sastry M, Stephens T, Stuckey J, Teng IT, Wang P, Wang S, et al. 2020. Cryo-EM structures of SARS-CoV-2 spike without and with ACE2 reveal a pH-dependent switch to mediate endosomal positioning of receptor-binding domains. *Cell Host & Microbe* **28**:867–879. DOI: <https://doi.org/10.1016/j.chom.2020.11.004>, PMID: 33271067