Computational analysis of interior mutations of SARS-CoV-2 Spike protein suggest a balance of protein stability and S2: S1 separation propensity

Zhen-lu Li [a,b,*], Matthias Buck [b,c,*]

[a] School of Life Science, Tianjin University, Tianjin 300072, China
[b] Department of Physiology and Biophysics, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, OH 44106, USA
[c] Department of Pharmacology, Department of Neurosciences, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, OH 44106, USA

ABSTRACT

SARS-CoV-2 variants often include surface mutations in the Spike protein that are important for viruses to recognize host receptors and evade antibody neutralization. The Spike protein also has mutations in the interior of the protein likely to affect the Spike protein S1 – S2 subunit's separation propensity, the most important of which is the D614G mutation. Remarkably, the Omicron variant contains a large number of internal mutations at the S2: S1 interface, which have not been investigated yet. In this study, we examined the effects of such interfacial mutations on the S2: S1 and subunit domain interactions and on the subunit's dissociation process. We found that the interaction with S2 is mainly contributed by the three encapsulation domains, named INT, ED1 and ED2 of S1, which are sandwiched between the S1 RBD and *N*-terminal NTD domain. We found that D614 is the strongest contributor for the S2: S1 interaction which is greatly weakened by the D614G mutation. Surprisingly, we found that, mutations T547K, H655Y, N764K, N856K, N969K, L981F in the Omicron variant largely enhance the S2: ED1 interaction, partially compensating the loss of S2: ED2 interaction due to the D614G mutation. Lastly, these results, together with biological considerations, allow us to suggest that in addition to the binding strength of between the RBD and ACE2, the stability of the Spike protein and the propensity of Spike protein S2: S1 separation are critical factors which likely exist in a balance for a particular infectivity and pathogenicity of the virus.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creative-commons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The SARS-CoV-2 coronavirus has four variants so far defined by the WHO as variants of concern, i.e. Beta, Gamma, Delta and Omicron. Of these variants, Delta and Omicron have swept the globe due to their higher infectivity [1–3]. Mutations in the virus envelope Spike protein on the surface of the coronavirus are also closely related to the effectiveness of vaccines and to the antibodies generated. The Beta variant contains five mutations in the Spike protein – K417N, E484K, N501Y, D614G, A701V. The Gamma variant has K417T, E484K, N501Y, D614G, H655Y mutations. The Delta variant has four mutations as L452R, T478K, D614G, and P681R. The Omicron variant (BA. 1) contains as many as 34 Spike protein mutation sites (A67V, del69-70, T95I, del142–144, Y145D, del211, L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K, L981F). Of all these Spike protein mutations, amino acid substitution D614G is the only mutation common to these variants, suggesting, that by itself, it has a profound role on the structure-dynamics-function relationship of the Spike protein.

At present, for Spike protein variants, there is an extensive and growing literature focusing on the effect of mutations on the binding strength of the mutated Spike proteins to human cell surface receptor ACE2, as well as the effect of mutations on antibody neutralization effects. In general, studies have found that mutations reduce antibody neutralization and evolved for the virus to escape the immune system [3,4,5,6,30]. However, some of the reports are inconsistent with one another as to whether the mutations

* Corresponding authors at: Department of Physiology and Biophysics, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, OH 44106, USA.

E-mail addresses: zxl480@tju.edu.cn (Z.-l. Li), zxl480@tju.edu.cn (M. Buck).

enhance the binding strength of the Spike protein to ACE2 [7,8,11,26,31]. Intriguingly, some mutation sites are located outside the receptor-binding region (RBD, residues 318–541) of the Spike protein, such as D614, suggesting potent allosteric effects [28]. Importantly, ACE2 binding by the Spike protein is only possible when one or several RBD in the timer in an "up" state, "opening" the protein. In accord, studies have shown that the G614 spike variant adopts a more open "RBD up" conformation for ACE2 binding [12]. However, other experiments indicated that the D614G mutation has actually little effect on the binding affinity toward recombinant ACE2 [11,14].

The effect of Spike mutations is not limited to the binding between Spike protein and human receptors. The mutations potentially have various effects on different stages of the viral infection process, not just virus-host recognition, but also on virus endocytosis and transport, virus fusion, and virus assembly [13]. At the outset of our study we wanted to understand which Spike protein mutations are likely to significantly affect the internal interaction between spike protein S1 and S2 domains. The virus Spike protein is a trimeric complex, and the complex is processed by furin or the other enzymes and is cleaved into two sub-domains, the S1 domain (1–685) and S2 domain (686–1273) [17,18]. S2 dissociates from S1, refolds and inserts its fusion peptides into the host membrane. This is crucial for the virus-host membrane fusion process. The D614 is precisely at the interface between S1 and S2. In the case of the D614G mutation, a study by Chen's research team proposed that instead of altering the Spike: ACE2 binding affinity, D614G protects the Spike protein by decreasing the shedding of S2 from S1 before virus maturation [14]. However, another study found that D614G does not influence S2 shedding, but increases viral replication [15]. In another case, the P681R mutation in the Delta variant was found to facilitate cleavage of the spike protein and enhance the viral fusogenicity and pathogenicity [16].

By contrast to the prior variants, the Omicron variant contains a large number of mutations, many of which are also located at the S1: S2 interface [4]. These mutation sites include D614G, as mentioned, the T547K and H655Y mutations in the S1 domain and N764K, N856K, N969K, L981F mutations in the S2 domain. These mutations appeared first in the Omicron variant. Significantly, four neutral amino acids are mutated to charged lysine. At present, a comprehensive understanding of the impact of Spike protein mutations on the different aspects of viral invasion beyond host recognition is still absent. It is of great importance to investigate how the various mechanisms involving the Spike protein are affected by the mutations and how these effects alter virus infectivity.

In this study, we take advantage of computational methods to examine the potential influence of S1: S2 interface mutation on Spike protein stability and in terms of its dissociation. We study the exit mechanism of the S2 from the S1 domain, influenced by RBD opening/closure, the D614 mutation, as well as substantial S2: S1 interfacial mutation of the Omicron variant. We identify key S1: S2 interaction residues, and identify-three small S1 sub-domains that are responsible for encapsulating the S2 domain. We find that D614 has a significant role in maintaining the S2: S1 ED2 interaction. Interestingly the Omicron variants compensate the loss of such a S2: S1 interaction by largely enhancing the S2: S1 ED1 interaction. Lastly, we discuss the potent implication of S2: S1 interactions on the stability and separation propensity of the Spike protein.

2. Methods

2.1. Simulation systems

Our simulations compare the wild-type (wt) Spike protein, the D614G mutation, and the Omicron variant. The virus Spike protein

is a trimeric complex (unit A, B, C), and the complex is processed by the Furin protease and is cleaved into two sub-domains, the S1 domain (1–685) and S2 domain (686–1273). The three units of the S2 initially maintain a compact, oblong shaped complex (similar perhaps to three segments of an orange fruit sticking together), while the three subunits of S1 constitute a crown-like structure and form a shallow pocket cone on top of S2. Thus, S2 is partially encapsulated into the S1 pocket. The S1: S2 complex is mainly maintained by the interaction of amino acids in the encapsulation pocket. In addition, near the cleavage site outside the pocket, S2 residues 692–697 form a β-strand and this strand lies, spatially, in between two adjacent β-strands from S1 (including residues S1 652–656 and S1 669–674 respectively). The interior capped S2 region separates from S1 in a unique direction (Z direction here). By contrast, the exterior S2 β-strand region can dissociate spontaneously in any direction under the influence of thermal fluctuations or external forces (Fig. S1). In simulations, the application of a unidirectional external force in the required z-direction makes it hard to break the β–strand formed between S2 686–699 and S1 652–656 and S1 669–674. This β-sheet interaction would very significantly retard the complete dissociation of S2 and S1 on the time scale of the simulations, involving a separation distance of S2 as much as 14–20 nm [23]. Importantly, since all the mutations are at the S2: S1 interior interface and this β-sheet structure is the same in the wt Spike protein and its variants, its absence is not expected to influence the separation process when comparing the WT and variant proteins. Thus, in this study, we removed this small S2 β-hairpin, residues 686 to 699 from the simulation system, and we focused on the interaction of the S1: S2 interface within the encapsulation pocket. We selected S1 residues 1–685 and S2 residues 700–1146 to investigate the influence of S2: S1 interfacial mutations on the interactions between S1 and S2. Four kinds of simulation systems are listed in Table 1.

2.2. Initial structure modeling

The RBD of the Spike protein has two orientations relative to the rest of the Spike protein – RBD "up" or "open" and RBD "down" or "closed". For the wt Spike protein, a state with 2 subunits in the RBD "closed" state and 1 subunit in the RBD "open" state is popular.

The D614G mutation, on the other hand, significantly increases the population of the 2 subunit RBD "open" state [11]. For the Omicron variant, the current electron microscope results indicate the existence of both 2 RBD "open" and 1 RBD "open" states, but the relative ratio is unknown [10].

For the wt Spike protein, the structure was based on the available crystal structure (6vsb and 7a93). '1 RBD open' state of Spike protein was based on PDB entry 6vsb. In CHARMM-GUI, Woo et at constructed and provided the Spike protein structure as the full-length protein by adding missing loops and residues. We used the structures in this study for the simulations [19]. '2 RBD open' state of Spike protein was based on PDB entry 7a93 [12]. Several short missing loops/residues were added by referring to the '1 RBD open' structure. For the G614 variant and Omicron variant, they were all built with '2 RBD open' states. For the G614 variant, amino acid Aspartic acid 614 was mutated to Glycine 614. For the Omicron variant, we used PDB 7t9j as a template to build the structure [9]. However, in this structure only one RBD (in 'closed' state) is visible, and there are many missing densities in this electron microscopy structure, including for the other two RBDs. We firstly used Swiss-Model to add missing loops and residues (but not the missing RBDs). In PDB entry 7t9l, a recombinant Omicron RBD is bound to an ACE2 protein. We used this RBD structure and fused it to the Omicron Spike protein by referring (with superimposition) the steric position of the two "open state" RBDs in wt S protein

Table 1
The four sets of simulation systems used in this study.

Systems	Sequence	Variant	RBD state	PDB template(s)
System 1	(S1: 1–685 and	WT	2 'open'	7a93
System 2	S2: 700–1146 for all)	WT	1 'open'	6vsb
System 3		G614	2 'open'	7a93
System 4		Omicron	2 'open'	7t9j/7t9l

structure (PDB 7a93). Glycosylation sites of the Spike protein are distributed on its surface and have no direct influence on interior interaction. Therfore, for simplicity, glycosylation of the Spike protein was not considered (see also Fig. S2). For reference, relevant studies of the role of glycan in Spike protein are found elsewhere [27,32]. Added loops or mutated residues were relaxed with a 1 ns short simulation with the protein backbone fixed.

2.3. Steered molecular dynamics simulation

After the short relaxation simulation, we ran steered molecular dynamics (SMD) simulations in order to investigate the separation of S2 and S1 subdomains. We have carried out similar SMD simulation of S2: S1 separation in the presence and absence of the Spike protein (co-)receptor/host factor neuropilin-1 [23]. Recently, SMD simulations were also used for the investigation of RBD: ACE2 separation [6,33,34]. In the SMD simulation, the long axis of the S2: S1 trimer complex is aligned in the z direction. In the steered MD simulation, a potential bias is added to a simulation system:

$$U(t) = \frac{1}{2}k \left[vt - \left(\overrightarrow{R}(t) - \overrightarrow{R}_0 \right) \cdot \overrightarrow{n} \right]^2$$

where $\overrightarrow{R}(t)$ is the current and $\overrightarrow{R}_0$ is the initial center of mass of SMD atoms, $\overrightarrow{n}$ for pulling direction, k for the force constant and v for moving velocity of virtual reference coordination. In this study, S2 gradually detaches from S1 by pulling S2 in the z-direction. The time step is set to 1 fs. The pulling rate constant v is set to 0.0000001 nm/fs. The total simulation time is 60 ns, thus S2 has a total movement of 6.0 nm. Each set of simulations was repeated 8 times with different velocity starting seeds. As before [23] a force constant k of 0.25 kcal/(mol · nm^2) was used for the harmonic spring that generates the force to pull S2 away from S1. The S1 RBD (res. 321–534) of three Spike protein units were restrained and held during the pulling simulation, the same way as in our previous study [23] mimicking the situation when they are firmly bound by the ACE2 protein.

2.4. Simulation condition and parameters

We applied Generalized Born Implicit Solvent method to investigate the S2: S1 separation [22,23]. While explicit water models are generally more suitable for molecular dynamics simulations, here, we applied an implicit water model for the steered molecular dynamics for the following reasons: In order to pull S2 out of S1, we need to reserve a region of 6 nm in the z-direction (totally about 33 nm) in the simulation box, so that the S2 can be displaced in the Z direction. The number of atoms contained in this system would be very large, estimated around 860,000, which would have been beyond our computational resources. More importantly, the size of S2 is relatively large. Different from pulling a small molecule, the movement of a sizeable S2 subunit on one hand needs to push away a lot of water molecules, which would generate excessive pressure on the simulation systems in a size-limited simulation box. On the other hand, the movement of S2 leaves a large void region between S1 and S2 subunits. The water molecules will find it difficult to fill the void space left by S2 in the short simula-

tion time in a manner which would appear to be at equilibrium. Thus, we expect the movement of such waters would generate a significant level of noise. Therefore, we think that the use of an implicit water model is a reasonable choice for studying the process of studying S2: S1 separation. Last, but not least, we wanted to build on our earlier work [23] which used the same simulation protocol. In the implicit solvent simulation, a cut-off of 1.4 nm was used for calculating the interactions. Ion concentration was set as 150 mM NaCl. A cut-off of 1.2 nm was used for calculating the Born radius. The simulations were performed at 310 K. The CHAMRM36m force field [21] was used throughout and all simulations were performed with the NAMD/2.12 package [20].

3. Results

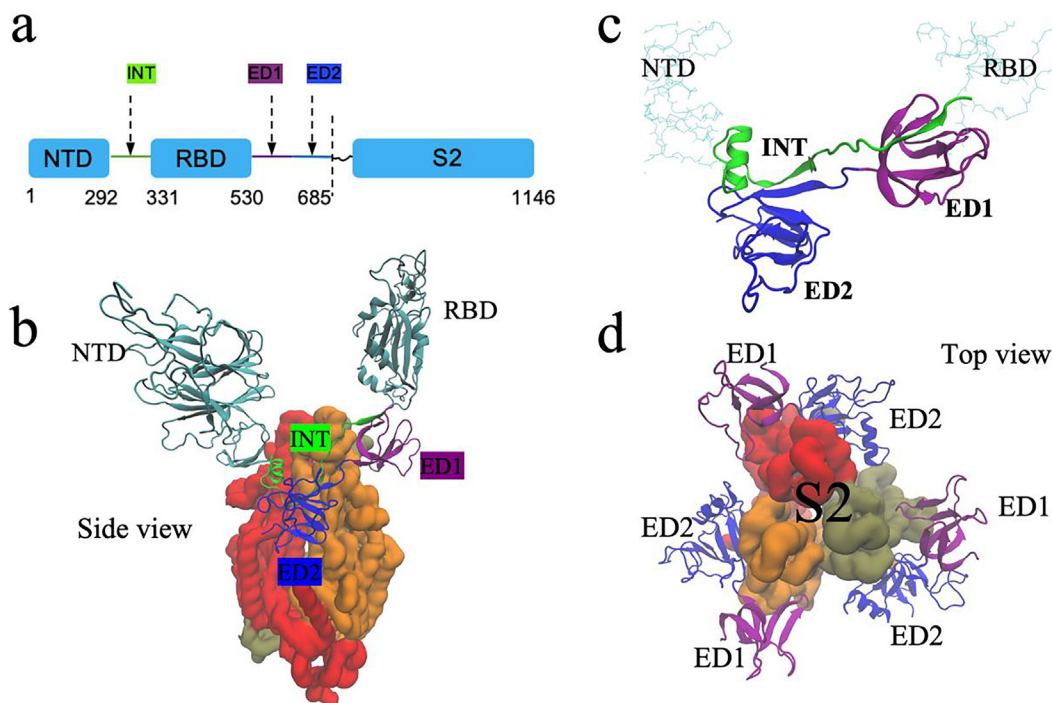### 3.1. Definition of critical domains for maintaining S2: S1 interface

SARS-CoV-2 Spike or S protein includes the S1 subunit and the S2 subunit. S1 subunit contains the NTD (*N*-terminal domain, residues 1–292) and the RBD (receptor binding domain, res. 331–530), which are well folded and spatially isolated (Fig. 1). Two linker sequences connect the NTD to RBD, and RBD to the *N*-terminus of S2 subunit respectively (Fig. 1a). As described below in this study, we found that the two linker regions sandwiched between NTD and the RBD domains are the critical regions that are responsible for maintaining the S1 and S2 binding interactions (Fig. 1b). However, these linker regions are largely understudied in previous reports. At present, the amino acid terminals of the linker sequence at both ends as well as the naming of the two linkers are inconsistent in different publications [e.g. [11,24,25,26]]. For the convenience of analysis of the S1: S2 interaction, and given the importance of these regions on encapsulating S2 subunits, here we divided the two linker regions (NTD-RBD linker and RBD-S2 linker) into 3 different subdomains (Fig. 1a and Table 2): the intermediary domain (INT, residues 293–330), encapsulation domain 1 (ED1, residues 531–590) and encapsulation domain 2 (ED2, residues 591–685).

Despite being distant in sequence, INT 310–318 and ED2 593–599 form two adjacent β-strands. Similarly, INT 324–329 and ED1 540–545 form two adjacent β-strands (Fig. 1c). It is worth noting that the interactions between the three subunits of S proteins are staggered, i.e., S2 of unit A interacts with S1 of unit B, S2 of unit B interacts with S1 of unit C, and S2 of unit C interacts with S1 of unit A (Fig. 1d).

3.2. Key S2: S1 interactions within the encapsulation pocket of the WT Spike protein

We first performed pulling simulations for the '2 RBD open' state. Under the action of the pulling force, S2 gradually dissociates from S1's encapsulation pocket (Fig. 2). In the beginning, the average interaction between S2 and S1 is the largest (Fig. 3a, the interaction energy was binned every 0.1 nm interval). After the native contacts between the S2: S1 amino acids are initially stretched while maintaining their interaction energy, the intensity of the interaction decreases rapidly (Fig. 3a). We further analyzed the

Fig. 1. Definition of structural subdomains for maintaining S2: S1 binding interactions. a) Structural subdomains. b) Complex highlighting the NTD, RBD, INT, ED2 and ED1 structural domain. c) Complex highlighting the INT, ED2 and ED1 structural domain. d) A top view (structure of b), 90° rotated around x-axis, of S2: S1 showing staggered interactions between S1 ED2/ED1 and S2.

Table 2
Definition of structural subdomains of S1 subunit of the SARS-Cov-2 Spike protein.

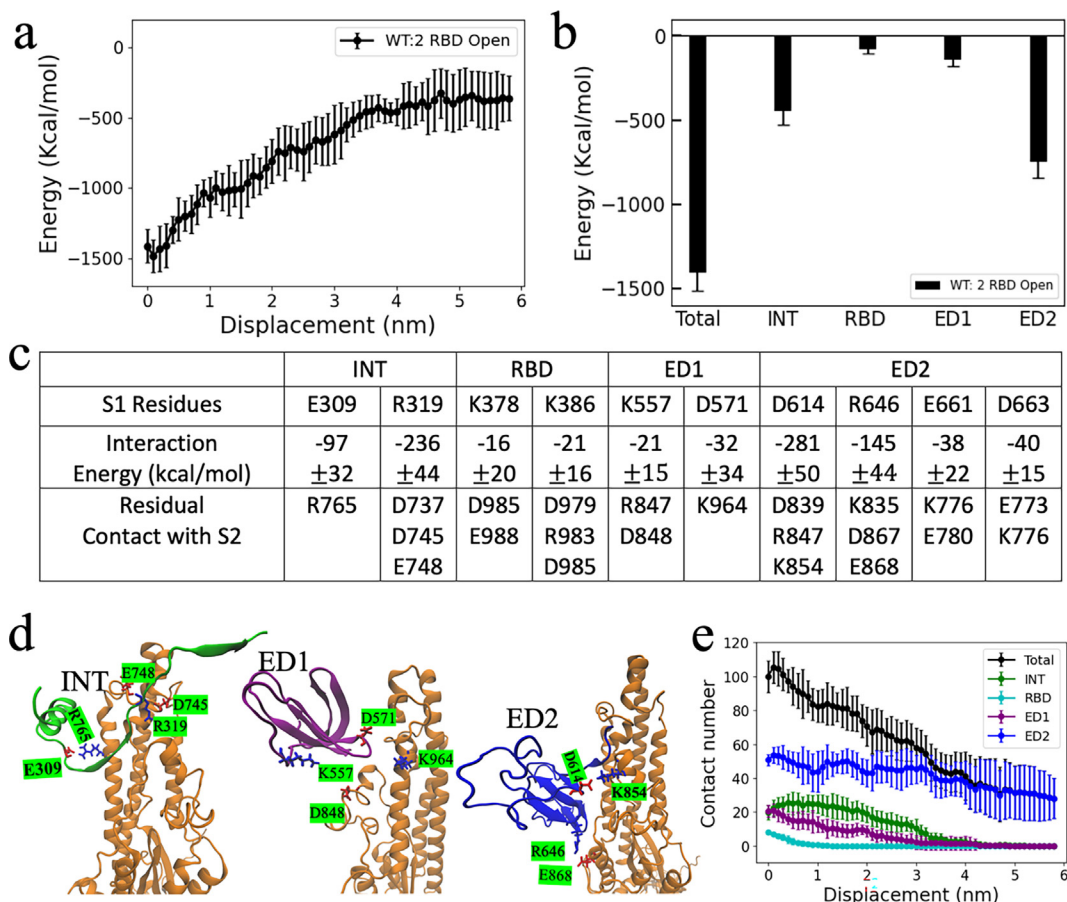| Structural subdomain (definition, this work)	Residues	Other names (in literatures)
NTD	1–292	–
INT	293–330	INT1, SD1
RBD	331–530	–
ED1	531–590	INT2, SD2, SD1
ED2	591–685	CTD, SD2
S2	686–1146	–

S2: S1 interaction at the subdomain- as well as residue level during the initial dissociation (0–0.5 nm displacement) process. By decomposing the interaction electrostatic energy into individual subdomains (Fig. 3b), we found that ED2 makes the highest contribution to the total interactions (53 %), INT contributes about 30 %,

while the ED1 contributes approx. 10 % of the total energy and the contribution of the RBDs is only about 5 % and involvement in the NTDs in the S1:S2 interaction is negligible. Looking at the atomistic level, we found that ED2 has the most amino acids which contribute to the S2: S1 interaction (Fig. 3c–e). In particular, ED2 D614 forms a strong interaction with K854. In addition, INT E309/R319 and ED2 R646 form strong interactions with S2 R765/D745/E748 and K835/D867/E868 respectively (Fig. 3c, d).

We calculated the number of residue-residue contacts between the S2 and S1 domains (Fig. 3e). We found that in S2′s exposure to S1, the interactions with subdomains INT, RBD, and ED1 is rapidly reduced, while S2′s exposure to S1 ED2 lasts longer. In this calculation, ED2 moves with S2, so the contact between the two lasts longer. This is because on one hand there is a stronger interaction between ED2 and S2, and on the other hand ED2 is localized initially further down on the S2 structure and can be thought of as
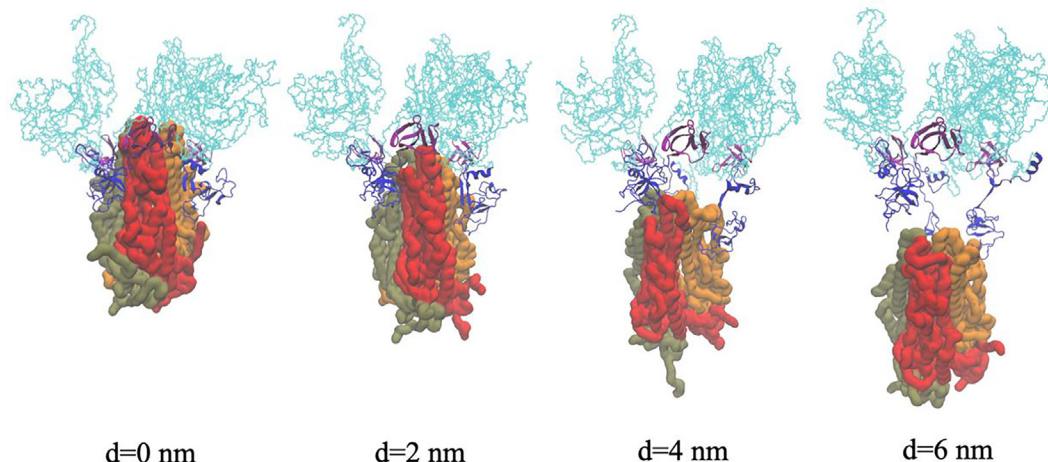
Fig. 2. A representative dissociation pathway of S2: S1. Note that the INT region is small and hidden from view.

	INT		RBD		ED1		ED2			
S1 Residues	E309	R319	K378	K386	K557	D571	D614	R646	E661	D663
Interaction Energy (kcal/mol)	-97 ±32	-236 ±44	-16 ±20	-21 ±16	-21 ±15	-32 ±34	-281 ±50	-145 ±44	-38 ±22	-40 ±15
Residual Contact with S2	R765	D737 D745 E748	D985 E988	D979 R983 D985	R847 D848	K964	D839 R847 K854	K835 D867 E868	K776 E780	E773 K776

Fig. 3. S2: S1 interactions and dissociation. a) S2: S1 interaction energy (direct pair interactions) as a function of separation distance of S2 from S1. b) Decomposition of S2: S1 interaction energy by individual subdomains when the separation distance between S2 and S1 is between 0 and 0.5 nm. c–d) Key S2: S1 electrostatic interactions. e) S2: S1 contact (number of residue-residue contact, with a cut-off of 4 Å) as a function of separation distance. All values represent the interactions among the three subunits in the Spike protein and are averaged over 8 repeat simulations.

being on the exit pathway of S2. Thus, it forms non-specific contact interactions during the dissociation process. Note that in none of the simulations, S1 and S2 are completely separated at 6 nm distance as the ED2 can the stretched considerably.

3.3. The opening and closure of RBD only slightly affects S2: S1 interactions

Any RBD, in the "closed" state, contacts and caps the top of S2, although these contacts are fewer and less strong than those of ED2. We compared the '1 RBD open' with the '2 RBD open' states (Fig. 4a). As the former has one more RBD capping the S2 region, RBDs should be able to establish more interactions with S2. However, in terms of the total interactions, we do not see a noticeable difference (Fig. 4b). At the RBD: S2 interface, amino acids K376 and K386 are major contributors for the RBD: S2 interaction, which interact with D985/E988 and D979 respectively (Fig. 4c). The decomposition of interactions shows that there are stronger interactions between the RBD and the S2 in '1 RBD open' state, but the increase of interaction is compensated by a decrease in interaction between ED2 and S2 (Fig. 4d). Overall, the RBD: S2 interaction contributes to the total energy, but the proportion is small.

### 3.4. The G614 variant reduces the separation barrier by reducing the S2: ED2 interaction

Examining the S2: S1 separation for the G614 variant, the energy barrier for S2: S1 separation is significantly reduced

(Fig. 5a, b). Comparing the interaction of D614/G614 with S2, we find that G614 has only a weak interaction with S2. It largely reduces the S2: ED2 interaction (Fig. 5c). Unlike aspartic acid, glycine is not able to form a strong interaction with adjacent positively charged amino acids. Thus, the original electrostatic interactions between D614 and positively charged residues are absent (Fig. 5d). The residue-residue contact between S2 and S1 are similar between 614G variant and WT protein (Fig. 5e).

S2 has much weaker interaction with S1 comparing G614 variant with WT protein (Fig. 5b). As discussed above, ED2 is at the exit pathway of S2. With the D614G mutation, the obvious decrease in the S2: ED2 interaction largely smoothens the separation process. Thus, the separation process of S2 and S1 becomes easier with the D614G mutation. Many studies emphasized a potent distal influence of D614G mutation on RBD: ACE2 binding. However, our study indicates that the D614G mutation, first of all, has significant local influence on the S2: S1 interaction.

3.5. The Omicron variant has a strongly enhanced S2: ED1 association

The Omicron variant has 34 mutations. Of these mutations, D614G, T547K, H655Y in S1 domain and S2 N764K, N856K, N969K, L981F are at the S2: S1 interface. These mutations create four lysine residues. Because the Spike protein contains three subunits, so these mutations will increase the number of positive charges by 12. By plotting the electrostatic potential energy of the S2 region, we found that the mutation drastically changes the surface electrostatic potential energy of S2 (Fig. 6a).
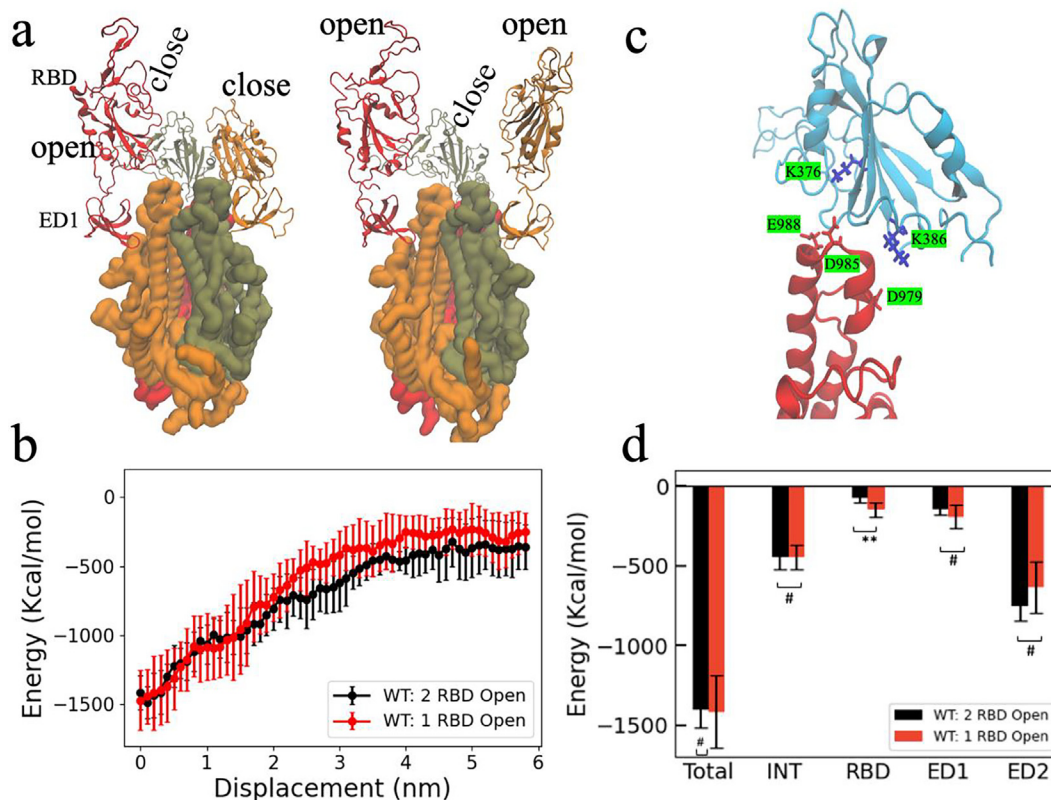
Fig. 4. Influence of RBD state (open or closed) on S2: S1 separation. a) Open and closed RBDs, comparing 1 RBD open (left) with 2 RBD open (right) systems. b) S2: S1 interaction energy as a function of separation distance of S2 from S1. c) Key S2: RBD electrostatic interaction sites. d) Decomposition of S2: S1 interaction by individual domains when the separation distance between S2 and S1 is between 0 and 0.5 nm. # indicates $p > 0.05$; * $p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

For the Omicron variant, the total S2: S1 interaction is almost the same in comparison to the S2: S1 interaction of WT (Fig. 6b), but the constituent of the total energy is dramatically different in comparison to the D614G variant. Amazingly, we found that the Omicron variant has significantly increased interactions between S2 and ED1 (Fig. 6c). This is because the T547K mutation establishes new interactions with S2 D745 and D979 (Fig. 6d, f). In the meantime, D571 and D568 establishes additional interactions with residue S2 856 which is also mutated to lysine (Fig. 6e, f). Altogether, these mutations enhance the association strength of S2 to S1, largely counteracting the weakening of S2 and S1 interactions caused by the D614G mutation (Fig. 5). The S2: ED2 interaction is reduced for the Omicron variant. This is similar to the G614 variant as the Omicron variant also has the D614G mutation. Due to the weakening of the S2: ED2 interaction, the S2: S1 dissociation is relatively smooth in comparison to wild type Spike protein (Fig. 6g).

To summarize, the G614 variant largely diminishes the S2: S1 interaction, but the Omicron variant has a small influence on the total S2: S1 interaction energy. However, it dramatically alters the balance of the S2: ED1 and S2: ED2 interaction. The Omicron variant has a large decrease in S2: ED2 interaction energy, but on the other hand, the interaction of S2: ED1 contacts are significantly enhanced.

4. Discussion

4.1. Amino acid mutations and the S2: S1 interaction

In this study, we investigated – surprisingly it appears for the first time – the detailed interactions between the S2 and S1 subdomains, INT, RBD, ED1 and ED2. The RBD has relatively weak interactions with S2, and its opening and closure relative to the rest of

S2, has less effect on the total S2: S1 interaction. We found that ED2 makes the biggest contribution to maintaining the S2: S1 interface for the WT virus. Of all the amino acids, D614 in ED2 is one of the key contributors to the S2: S1 interaction. Importantly, the D614G mutation dramatically reduces the S2: S1 interaction. Furthermore, in addition to D614G, the Omicron variant has six other mutations (T547K, H655Y, N764K, N856K, N969K, L981F) at the S2: S1 binding interface. These kinds of mutations are new compared to the previous widely distributed variants. Of these mutations, 4 amino acids are mutated to lysine, which greatly alters the electrostatic surface potential of the S2 protein and enhances the binding strength of S2 to S1 ED1. The enhancement of inter-domain and inter-subunit packing likely helps stabilize an open conformation of RBD, as proposed in the study by Ye et al [29]. However, a 614G variant itself already has a role in increasing the percentage of 'RBD open' conformation [11]. So, from this perspective, it has been unclear what made Omicron different from the other variants.

Glycosylation of the Spike protein is implicated for host recognition or immune evasion by shielding the protein surface from detection by antibodies. Importantly, the glycosylation sites are distributed all over the surface of the Spike protein (Fig. S2). Our research concerns the interior interactions between S2 and S1, but there are no glycosylation sites on the interior S2: S1 dissociation path. Therefore, we believe that glycosylation will not have a direct effect on the S2: S1 interaction and separation.

4.2. What make Omicron variant differs from 614G variant like Delta variant?

All the other circulating variants (such as Beta and Delta variants) only have the D614G mutation localized to the S2: S1
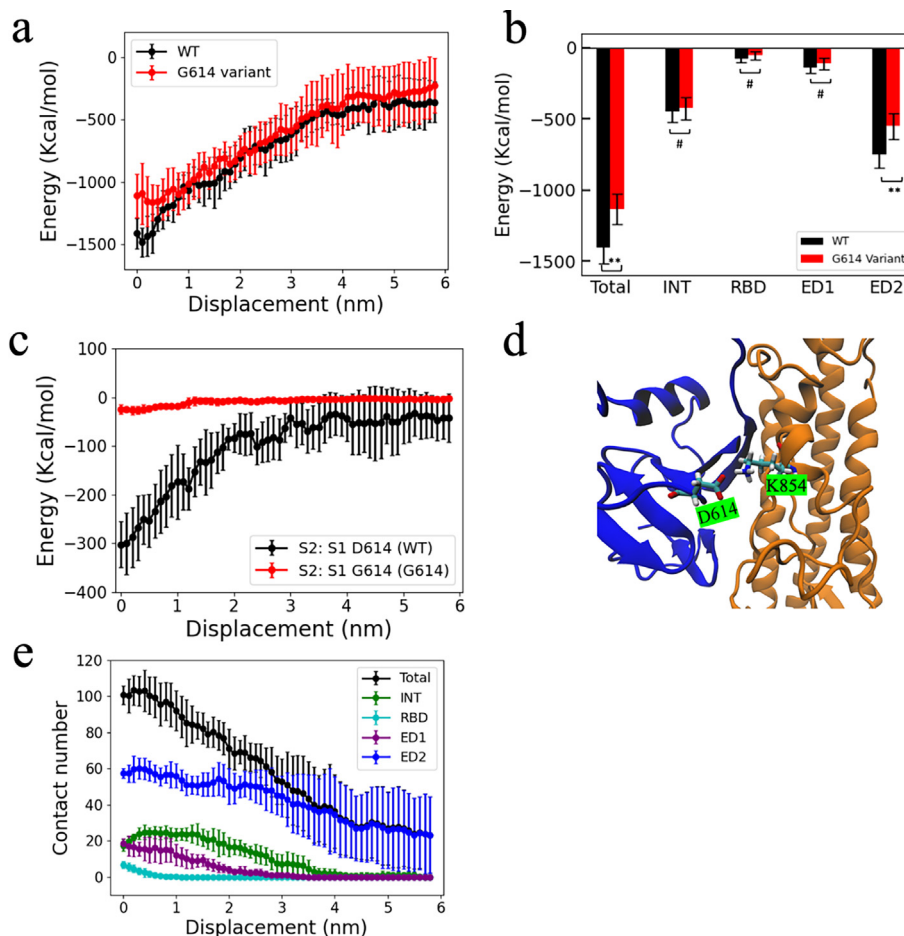
Fig. 5. The influence of the D614G mutation on S2: S1 interaction energy. a) S2: S1 interaction as a function of separation distance of S2 from S1. b) Decomposition of S2: S1 interaction by individual subdomains when the separation distance between S2 and S1 is between 0 and 0.5 nm. c) The interactions of S2: D614 and S2: G614 (i.e. S2 with only this one residue of S1) as a function of S2: S1 separation distance. d) Highlight of the D614: K854 interaction. e) Contact number as a function of separation distance.

interface. However, Omicron has multiple other mutations present at the interface between S2 and S1. The 614G variant dramatically weakens the S2: S1 interaction in comparison to the WT virus. However, for the Omicron variant the extent of decrease in S2: S1 interaction is not as significant as in the other 614G variants. By decomposing the interaction of S2 and S1, we found that similar to the other 614G containing variants, the S2: ED2 interaction is reduced, but the reduction in S2: ED2 interaction is largely compensated by an increase in interaction between S2: ED1. ED1 is next to the RBD and likely has a direct role in regulating RBD orientation. ED2 is distal from the RBD, but it forms more contacts with S2 and is at the exiting pathway of the S2 subunit.

4.3. S2:S1 interaction in biological context

The strength of the S2: S1 interaction is related to the stability of the Spike protein itself, but it is also closely related to the fission–fusion process (fission, i.e. separation of S2 and S1 and fusion of S2 –and the virus- with the host cell membrane). A strong S2: S1 interaction enhances protein stability, but is likely detrimental to protein dissociation. In contrast, a weak S2: S1 interaction would lead to a higher shedding effect of spike protein, which is harmful to virus assembly and maturation [14,15]. Ideally, the Spike protein should be stable enough before host receptor recognition, and easy to separate after binding with the host receptor. Therefore, a balance between Spike protein stability, S2: S1 separation propensity, and potentially S1: ACE2 binding affinity should be reached.

From a perspective of protein stability and protein separation propensity, we infer that the change in the relative strength of S2: ED1 and S2: ED2 interactions are likely biologically meaningful. The S2: ED1 interaction for the Spike protein of the Omicron variant is enhanced in comparison to the wild type Spike protein. However, the enhancement of the S2: ED1 interaction is largely compensated by the reduction of the S2: ED2 interaction (Fig. 7). Thus, the total interactions are of a similar strength. However, ED1 and ED2 are spatially distributed in different locations, and have different roles in maintaining protein stability and modulating protein dissociation. Specifically, ED1 locates in proximity to the RBD, and it has a pivotal role on gathering three loose RBDs on top of the S2 subunits. This prevents the RBD and the total S1 shedding from S2, especially when the RBDs are in the open conformation. On the other hand, ED2 is on the pathway of S2 dissociation, so the contact between S2 and ED2 exists throughout the dissociation process, thus the S2: ED2 interaction has a larger weight for modulating the dissociation of S2. Therefore, although the S2: ED1 interaction has been enhanced, the Omicron variant dissociates more easily compared to the WT protein as a result of the weakened S2: ED2 interaction.

So far, we still know very little about the effects of viral spike protein mutations on the entire process of virus infection. In this study, we investigated the potential role of S2: S1 interface mutations in affecting the S2: S1 interaction, and we suggest that a balance of S2: S1 stability and separation propensity should be an
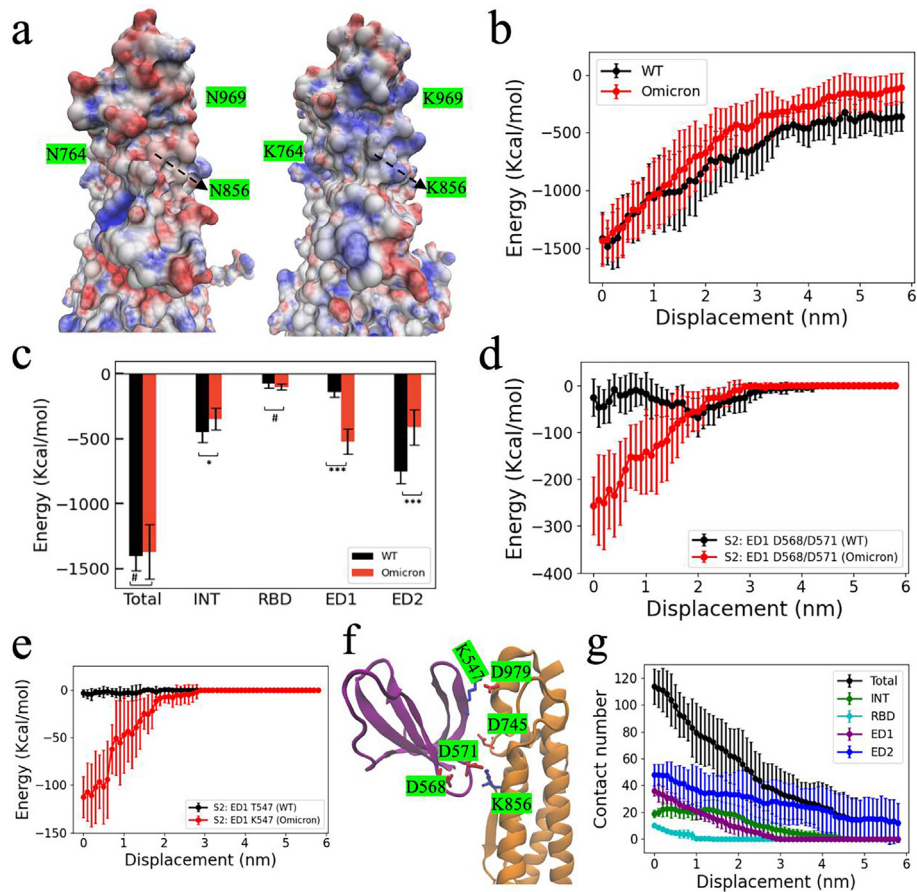
Fig. 6. The Omicron variant has a dramatically enhanced S2: ED1 interaction energy. a) Surface electrostatic potential. b) S2: S1 interaction as a function of separation distance of S2 from S1. c) Decomposition of S2: S1 interaction by individual subdomains when the separation distance between S2 and S1 is between 0 and 0.5 nm. d) Newly established interactions between S2 and ED1 K547, and e) between S2 and ED1 D568/D571. f) Interaction between S2 and S1 K547, and S2: D568/D571 interaction. g) Contact number as a function of separation distance.
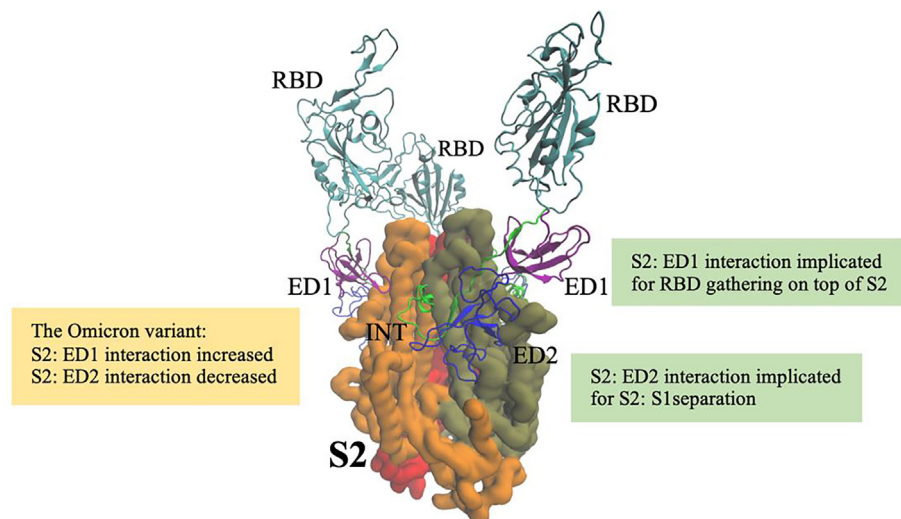
Fig. 7. Illustration of the role of balancing S2: ED1 and S2: ED2 interaction. For the Omicron variant, the S2: ED1 interaction is increased, while the S2: ED2 interaction is decreased.

important factor related to virus infectivity. Such a balance may be achieved in the context of protein structural variations, by balancing the relative binding strength of S2 with the different S1 subdomain ED1 and ED2.

CRediT authorship contribution statement

Zhen-lu Li: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Writing – review & editing. **Mat-**

thias Buck: Funding acquisition, Data curation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by a NIH R01 grant from the National Eye Institute R01EY029169 and previous grants from NIGMS (R01GM073071 and R01GM092851) to the Buck lab. The simulations were run at the local computing resource in the core facility for Advanced Research Computing at Case Western Reserve University.

Author contribution

Z.L. and M.B designed the studies, interpreted the data and wrote the manuscript. Z.L. performed the sequence analysis and computational modeling.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.10.044.

References

[1] Liu Y, Liu J, Plante KS, Plante JA, Xie X, Zhang X, et al. The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. Nature 2022 Feb;602(7896):294–9.
[2] Ulrich L, Halwe NJ, Taddeo A, Ebert N, Schön J, Devisme C, et al. Enhanced fitness of SARS-CoV-2 variant of concern Alpha but not Beta. Nature 2022 Feb;602(7896):307–13.
[3] Zhou B, Thao TTN, Hoffmann D, Taddeo A, Ebert N, Labroussaa et al. SARS-CoV-2 spike D614G change enhances replication and transmission. Nature. 2021 Apr;592(7852):122-127. 3.2 Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Sheffield COVID-19 Genomics Group, McDanal C, Perez LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC, Saphire EO, Montefiori DC. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell. 2020 Aug 20;182(4):812-827.e19.
[4] Dejnirattisai W, Huo J, Zhou D, Zahradník J, Supasa P, Liu C, et al. SARS-CoV-2 Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses. Cell 2022;185(3). 467-484.e15.
[5] Cele S, Jackson L, Khoury DS, Khan K, Moyo-Gwete T, Tegally H, et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. Nature 2022 Feb;602(7898):654–6.
[6] Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature 2022 Feb;602(7898):657–63.
[7] Han P, Li L, Liu S, Wang Q, Zhang D, Xu Z, et al. Receptor binding and complex structures of human ACE2 to spike RBD from omicron and delta SARS-CoV-2. Cell 2022;185(4). 630-640.e10.
[8] Zhang L, Jackson CB, Mou H, Ojha A, Peng H, Quinlan BD, et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. Nat Commun 2020 Nov 26;11(1):6013.
[9] Mannar D, Saville JW, Zhu X, Srivastava SS, Berezuk AM, Tuttle KS, et al. SARS-CoV-2 Omicron variant: Antibody evasion and cryo-EM structure of spike protein-ACE2 complex. Science. 2022 Feb 18;375(6582):760–4.
[10] McCallum M, Czudnochowski N, Rosen LE, Zepeda SK, Bowen JE, Walls AC, et al. Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement. Science 2022 Feb 25;375(6583):864–8.
[11] Yurkovetskiy L, Wang X, Pascal KE, Tomkins-Tinch C, Nyalile TP, Wang Y, et al. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. Cell. 2020 Oct 29;183(3):739-751.e8.
[12] Benton DJ, Wrobel AG, Roustan C, Borg A, Xu P, Martin SR, Rosenthal PB, et al.. The effect of the D614G substitution on the structure of the spike glycoprotein of SARS-CoV-2. Proc Natl Acad Sci U S A. 2021 Mar 2;118(9):e2022586118.
[13] Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. Proc Natl Acad Sci U S A 2020 May 26;117(21):11727–34.
[14] Zhang J, Cai Y, Xiao T, Lu J, Peng H, Sterling SM, et al. Structural impact on SARS-CoV-2 spike protein by D614G substitution. Science 2021 Apr 30;372(6541):525–30.
[15] Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, et al. Spike mutation D614G alters SARS-CoV-2 fitness. Nature 2021 Apr;592(7852):116–21.
[16] Saito A, Irie T, Suzuki R, Maemura T, Nasser H, Uriu K, et al. Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R mutation. Nature 2022 Feb;602(7896):300–6.
[17] Johnson BA, Xie X, Bailey AL, Kalveram B, Lokugamage KG, Muruato A, et al. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. Nature 2021 Mar;591(7849):293–9.
[18] Hoffmann M, Kleine-Weber H, Pöhlmann S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. Mol Cell 2020 May 21;78(4):779–784.e5.
[19] Woo H, Park SJ, Choi YK, Park T, Tanveer M, Cao Y, et al. Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane. J Phys Chem B 2020 Aug 20;124(33):7128–37.
[20] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem 2005 Dec;26(16):1781–802.
[21] Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods 2017 Jan;14(1):71–3.
[22] Onufriev AV, Case DA. Generalized Born Implicit Solvent Models for Biomolecules. Annu Rev Biophys 2019 May;6(48):275–96.
[23] Li ZL, Buck M. Neuropilin-1 assists SARS-CoV-2 infection by stimulating the separation of Spike protein S1 and S2. Biophys J 2021 Jul 20;120(14):2828–37.
[24] Adhikari P, Li N, Shin M, Steinmetz NF, Twarock R, Podgornik R, et al. Intra- and intermolecular atomic-scale interactions in the receptor binding domain of SARS-CoV-2 spike protein: implication for ACE2 receptor binding. PCCP 2020 Sep 7;22(33):18272–83.
[25] Edwards RJ, Mansouri K, Stalls V, Manne K, Watts B, Parks R, et al. Cold sensitivity of the SARS-CoV-2 spike ectodomain. Nat Struct Mol Biol 2021 Feb;28(2):128–31.
[26] Gobeil SM, Janowska K, McDowell S, Mansouri K, Parks R, Stalls V, et al. Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity. Science 2021;373(6555). eabi6226.
[27] Sztain T, Ahn SH, Bogetti AT, Casalino L, Goldsmith JA, Seitz E, et al. A glycan gate controls opening of the SARS-CoV-2 spike protein. Nat Chem 2021 Oct;13(10):963–8.
[28] Ray D, Le L, Andricioaei I. Distant residues modulate conformational opening in SARS-CoV-2 spike protein. Proc Natl Acad Sci U S A. 2021 Oct 26;118(43):e2100943118.
[29] Ye G, Liu B, Li F. Cryo-EM structure of a SARS-CoV-2 omicron spike protein ectodomain. Nat Commun 2022 Mar 3;13(1):1214.
[30] Pramanick I, Sengupta N, Mishra S, Pandey S, Girish N, Das A, et al. Conformational flexibility and structural variability of SARS-CoV2 S protein. Structure 2021 Aug 5;29(8):834–845.e5.
[31] Dokainish HM, Re S, Mori T, Kobayashi C, Jung J, Sugita Y. The inherent flexibility of receptor binding domains in SARS-CoV-2 spike protein. Elife 2022 Mar;24(11):e75720.
[32] Mori T, Jung J, Kobayashi C, Dokainish HM, Re S, Sugita Y. Elucidation of interactions regulating conformational stability and dynamics of SARS-CoV-2 S-protein. Biophys J 2021 Mar 16;120(6):1060–71.
[33] Kim S, Liu Y, Ziarnik M, Cao Y, Zhang XF, Im W. Binding of Human ACE2 and RBD of Omicron Enhanced by Unique Interaction Patterns Among SARS-CoV-2 Variants of Concern. bioRxiv [Preprint]. 2022 Jan 25:2022.01.24.477633.
[34] Cao W, Dong C, Kim S, Hou D, Tai W, Du L, et al. Biomechanical characterization of SARS-CoV-2 spike RBD and human ACE2 protein-protein interaction. Biophys J 2021 Mar 16;120(6):1011–9.