





Article

Conformational and Stability Analysis of SARS-CoV-2 Spike Protein Variants by Molecular Simulation

Gustavo E. Olivos-Ramirez ¹, Luis F. Cofas-Vargas ¹, Tobias Madl ² and Adolfo B. Poma ^{1,*}

¹ Department of Biosystems and Soft Matter, Institute of Fundamental Technological Research, Polish Academy of Sciences, ul. Pawińskiego 5B, 02-106 Warsaw, Poland; golivos@ippt.pan.pl (G.E.O.-R.); fcofas@ippt.pan.pl (L.F.C.-V.)

² Division of Medical Chemistry, Otto Loewi Research Center for Vascular Biology, Immunology and Inflammation, Medical University of Graz, Neue Stiftingtalstraße 6, 8010 Graz, Austria; tobias.madl@medunigraz.at

* Correspondence: apoma@ippt.pan.pl

Abstract: We performed a comprehensive structural analysis of the conformational space of several spike (S) protein variants using molecular dynamics (MD) simulations. Specifically, we examined four well-known variants (Delta, BA.1, XBB.1.5, and JN.1) alongside the wild-type (WT) form of SARS-CoV-2. The conformational states of each variant were characterized by analyzing their distributions within a selected space of collective variables (CVs), such as inter-domain distances between the receptor-binding domain (RBD) and the N-terminal domain (NTD). Our primary focus was to identify conformational states relevant to potential structural transitions and to determine the set of native contacts (NCs) that stabilize these conformations. The results reveal that genetically more distant variants, such as XBB.1.5, BA.1, and JN.1, tend to adopt more compact conformational states compared to the WT. Additionally, these variants exhibit novel NC profiles, characterized by an increased number of specific contacts distributed among ionic, polar, and nonpolar residues. We further analyzed the impact of specific mutations, including T478K, N500Y, and Y504H. These mutations not only enhance interactions with the human host receptor but also alter inter-chain stability by introducing additional NCs compared to the WT. Consequently, these mutations may influence the accessibility of certain protein regions to neutralizing antibodies. Overall, these findings contribute to a deeper understanding of the structural and functional variations among S protein variants.

Keywords: molecular dynamics; conformational space; native contact map; probability states; collective variables; protein stability; SARS-CoV-2



Academic Editor: Alessandra Borsetti

Received: 13 February 2025

Revised: 8 March 2025

Accepted: 11 March 2025

Published: 12 March 2025

Citation: Olivos-Ramirez, G.E.; Cofas-Vargas, L.F.; Madl, T.; Poma, A.B. Conformational and Stability Analysis of SARS-CoV-2 Spike Protein Variants by Molecular Simulation. *Pathogens* **2025**, *14*, 274. <https://doi.org/10.3390/pathogens14030274>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Analysis of the conformational space of proteins is important for understanding their functional properties. To date, critical aspects of the conformational space in proteins ranging from 10 to 80 amino acids in length have been revealed through long-scale molecular dynamic (MD) simulation (on the microsecond to millisecond timescale) [1]. However, MD simulations face certain limitations in complex systems (>100 k atoms) such as the case of the SARS-CoV-2 spike (S) protein. These challenges include not only the required sampling time, but also the determination of well-converged specific features (e.g., distances, angles, etc.) to explain the most relevant events at the macroscopic level [2]. Therefore, a comprehensive and representative conformational sampling of all states is necessary. Only through exhaustive analysis will it be possible to understand the stability of proteins

and observe conformational changes that may rarely occur (e.g., slow transitions), yet are directly involved in key biological processes [3]. In this context, novel methods to efficiently analyze long trajectories are implemented; for instance, time-lagged independent component analysis (TICA) [4,5] is often used to deal with long trajectories and, as a result, capture the meta-states in proteins with greater accuracy. However, defining specific features to conduct this analysis can be challenging for single proteins and even more complicated for complex systems. Therefore, the careful selection of these features, also known as collective variables (CVs), is essential to improving predictions of the conformational space of a complex biomolecular system.

The SARS-CoV-2 S protein has been the subject of several biophysical studies in the context of the global COVID-19 pandemic [6–9]. This is a membrane protein involved in the molecular recognition of the host cell receptor [10]. It forms a homotrimer with two functional subunits: the S1 subunit is responsible for binding to the cell, and the S2 subunit mediates the fusion of the virus with the host cell membrane [11]. The S1 subunit contains the receptor-binding domain (RBD) that interacts with the human angiotensin-converting enzyme 2 (hACE2). This strong and stable binding initiates the virus' cell entry process. In addition, the S1 subunit presents an N-terminal domain (NTD) which facilitates surface cell adhesion and regulates the pre-fusion-to-postfusion transition [12]. This protein domain undergoes conformational transitions, particularly between the open (up) and closed (down) state [13,14]. During the COVID-19 pandemic, the SARS-CoV-2 S protein evolved and developed numerous mutations such as substitutions, deletions, and insertions, leading to the emergence of new variants with differing levels of virulence and enhanced transmissibility [15].

Mutations in both the RBD and NTD are critical to study, alongside the induced conformational states associated with them. The Delta variant was the first to become the most prevalent variant [16] in 2020. This variant presents 29 distinct mutations that distinguish it from the Wuhan variant, hereafter known as the wild type (WT) [17], including D614G, L452R, P681R, and T478K with a potential to increase the transmissibility of the virus [18,19]. By mid-November 2021, new variants emerged with an increased transmissibility, among them the BA.1 (Omicron) variant, which became the most prevalent variant worldwide. BA.1 carries an additional 30 mutations, 15 of which are located in the RBD, along with three deletions and three insertions (e.g., ins214EPE) [20]. Some of these mutations (e.g., Y505H, N786K, T95I, N211I, N856K, and V213R) provide a higher positive electrostatic surface potential that improves the affinity for the hACE2 receptor and promote the evasion of antibodies initially developed for the earlier variants [21]. By 2023, XBB.1.5 was identified as an Omicron subvariant resulting from the recombination of two BA.2 sub-lineages, with an important mutation, S486P in the RBD. This variant presented the highest levels of immune escape among previous Omicron sub-lineages [22]. The JN.1 variant emerged by the end of 2023, spreading fast in at least 41 countries [23]. JN.1 was designated a variant of concern due to its higher hACE2-binding affinity and its ability to evade the immune response acquired by vaccines [24]. This variant differs by 39 mutations in the S protein from the previous Omicron XBB.1.5, including the L455S mutation in the RBD, which is known to increase immune evasiveness [25].

Mutations in the S protein modify the behavior of the virus, but more importantly those occurring in the NTD and RBD can affect the protein–protein interaction that may lead to conformational changes. Experimentally, a decreased rate of antibody recognition was observed in the presence of mutations such as C480S/R, E484K/G/D, C488Y/S, and F490L/I/C in the RBD, and Y145D, K150E, and W152R in the NTD [26]. Moreover, the binding affinity and kinetics of the RBD/hACE2 complex in several variants was evaluated by surface plasmon resonance (SPR). The relevant study revealed that only one RBD mutation

(N501Y) can increase the binding affinity by ≈ 10 -fold whereas two mutations (K417N/T and E484K) can enhance binding and immune escape [27]. Similarly, single-molecule force spectroscopy by AFM elucidated the role of N501Y in the stability of the hACE2/RBD complex, generating new protein contacts with an energy gain of 10 kJ/mol [28]. The combination of relatively distant mutations in RBD, such as E484K and N501Y, can further increase the RBD/hACE2 interaction 12-fold [29]. The RBD-opening pathway was characterized by weighted ensemble MD, highlighting the role of N343 glycan and the formation of important inter-chain hydrogen bonds (HBs) between T415 of RBD_A and K986 of RBD_C, a salt bridge between R457 of RBD_A and D364 of RBD_B, and a salt bridge between K462 of RBD_A and D198 of the N-terminus (NTD_C) [30]. Such HB contacts are very frequent in MD trajectories, and some are considered highly frequent contacts with 70% presence in the MD simulation. MD simulation was used to capture the nanomechanical stability of the SARS-CoV-2 RBD protein, which was compared to SARS-CoV-1 from 2003 (previous outbreak). The MD study showed a gain of ≈ 50 pN in mechanical stability for the WT [8]. The effect of mechanical stability in the RBD-nanobody was studied by MD, considering several variants (e.g., Alpha, Delta, and XBB.1.5) in complex with the H11-H4 potent nano-antibody [31]. The computational study shows that earlier variants contributed to the mechanical stability of the complex while later variants (XBB.1.5) exhibited the lowest interface energy. In addition, it highlights the presence of some conserved contacts between the RBD-nanobody (e.g., Y489-Y104, Y489-L105, L482-Y104, and Q493-S103) across several variants [31].

An important aspect that has not been thoroughly investigated is the conformational space of the S protein variants and their assessment through native contact (NC) analysis. Long-equilibrium all-atom MD performed on the microsecond-to-millisecond timescale [32] has revealed that the folding mechanism of proteins engages several NCs. Here, protein stability is defined quantitatively as the persistence of a certain number of high-frequency NCs across the molecular trajectories for a given conformational state in equilibrium. Two limiting cases are expected: the native state maximizes the number of NCs, while the unfolded state minimizes them. Additionally, NC analysis provides essential information for understanding specific molecular interactions that stabilize a given protein conformation in single proteins and their complexes [31]. Furthermore, the distributions of certain CVs allow us to characterize structural differences in the S protein between SARS-CoV-2 variances. Due to experimental limitations, the identification of important NCs is typically approached by MD simulation, which provides atomic-resolution insights. Several computational methods have been developed to extract this information from protein structures, including a distance geometry with a cutoff parameter [33], Q-score [32], the shadow map [34], the Voronoi diagram [35], overlaps (OVs) of enlarged Van der Waals (VdW) spheres [36], contacts of structural units (CSUs) [37], and OVs combined with repulsive CSU (OV+rCSU) [38]; all of these differ in their protocols, but they extract this information from the protein structure. Generally, an NC is determined between non-adjacent residues with sequential distance, i.e., $|i - j| \geq 4$ and where the distance between heavy atoms ranges from 4 to 12 Å [38,39]. The use of NCs was initially introduced in Gō-like models [40] and can be currently found in coarse-grained (CG) models like the GōMartini and OLIVES approaches [41,42].

While most studies have only focused on studying the RBD-ACE2 (protein-receptor) interaction, there is limited knowledge about the intrinsic effect of mutations on the conformational states of the S protein. This is partly due to the complexity of the system and the high computational cost it requires. Furthermore, the flexibility of the trimeric S glycoprotein establishes a number of contacts unique to each macrostate, which remains unclear for most variants of concern. The presence of these contacts and their chemical nature can determine the stability of the structure and its conformational changes. Figure 1

shows two distinct states, each characterized by different NC numbers. Ultimately, understanding these differences may help explain the mechanisms of immune evasion driven by specific mutations. This study determines the role of NCs between the trimeric S protein protomers considering the following variants: WT, Delta, BA.1, JN.1, and XBB.1.5. Through a comprehensive selection of collective variables, we determined the predominant conformational states of the S protein and their corresponding NCs. In addition, we performed an analysis of the chemical character of the most particular states to better understand how certain variants can change the behavior of the S protein interface.

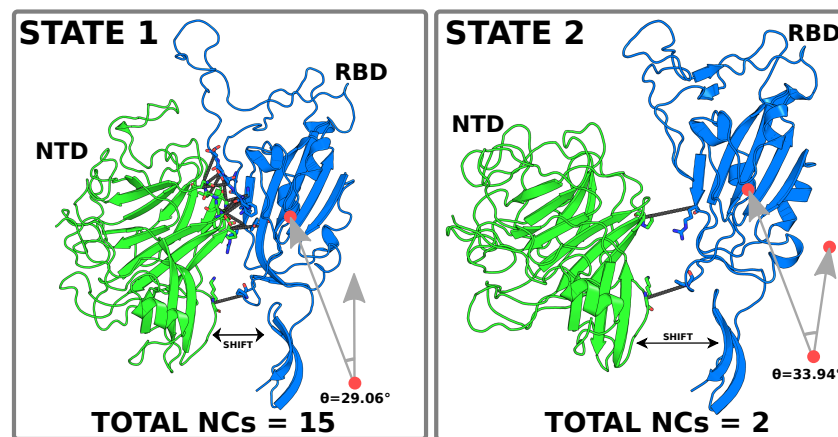


Figure 1. Representation of two states mediated by the formation of protein contacts (shown as solid lines) between the RBD and NTD of the S protein (JN.1 variant, Omicron). The left panel shows 15 NCs between the RBD (blue) and NTD (green), whereas the right panel shows the conformation with only two NCs. In the first case, both protein domains are closer together, while in the second, they are more separated due to the loss of most NCs. This leads to a greater angular fluctuation of the RBD regarding the HR1 domain. The NCs correspond to the set of high-frequency contacts (>60%) identified from the MD simulation.

2. Materials and Methods

2.1. All-Atom Molecular Dynamics Simulation

Trajectories of the WT S protein were obtained from [43,44]. The Amber ff14SB force field [45] and TIP3P water model [46] in a box with a 12 Å padding were used. The system was neutralized with 0.150 M NaCl. The system was minimized using the steepest descent algorithm for 2000 steps, followed by the conjugate gradient algorithm for 3000 steps. Heating was performed in two stages: NVT heating from 0 to 100 K over 50 ps and NPT heating from 100 to 300 K over 100 ps. During minimization and heating, harmonic restraints of $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ were applied to C-alpha atoms. These restraints were gradually reduced from 10 to $0.1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ during a short simulation of 6 ns at 300 K. Pressure was maintained at 1 atm using the Monte Carlo barostat, whereas temperature was controlled using the Langevin thermostat with a collision frequency of 1 ps^{-1} . The Monte Carlo barostat was chosen for its efficient pressure control in the NPT ensemble. Unlike the Berendsen barostat, which does not sample the correct ensemble, or the Parrinello–Rahman barostat (not implemented in AMBER), which can introduce artificial oscillations, the Monte Carlo barostat ensures more stable volume fluctuations. It is widely used in AMBER simulations and has proven effective for water and biomolecular systems [47]. Hydrogen mass repartition was used to enable a 4 fs integration time step [48]. Finally, production was performed for a time of 320 ns with 5 replicas (1.6 μs total). Frames were captured every 0.2 ns resulting in 1600 frames per replica.

Additionally, we performed independent MD simulations for four S protein variants: Delta (PDB ID: 7W92), BA.1 (PDB ID: 7XO5), JN.1 (PDB ID: 8Y5J), and XBB.1.5 (PDB ID:

8VKM), retrieved from the Protein Data Bank [49]. All variants were solved with one RBD in the open (up) state. BA.1 and XBB.1.5 structures were complexed with the hACE2 receptor. For consistency, the hACE2 receptor and all heteroatoms were removed during structure preparation. The Delta variant spanned amino acids 13–1147, BA.1 included residues 26–1146, JN.1 covered 26–1139, and XBB.1.5 ranged 14–1147. Missing residues within these ranges (detailed in Table S2) corresponded to loop regions which were remodeled with Modeller v10.5 [50]. Protonation states correspond to pH 7.4 and were assigned using the PDBFixer [51]. To avoid adding artificial charges, the first and last solved residues were capped with acetyl and N-methyl groups, respectively, to prevent unrealistic interactions and distorted dynamics.

All-atom MD simulations of the four variants were performed with the AMBER22 engine [52] and the FF19SB force field [53], using the pmemd.cuda module for high performance [54]. Each system was solvated in an octahedral box with 20 Å padding using the four-site OPC water model [55]. The systems were minimized using the steepest descent algorithm for 5000 steps. Afterwards, the system was equilibrated using NVT and NPT ensembles. Temperature equilibration was performed in four steps: 150, 200, 250, and 300 K, with each step lasting 200 ps. Harmonic restraints were applied to the heavy atoms of the proteins, with spring constants from 5.0, 4.0, 3.0, and 1 kcal mol⁻¹ Å⁻² used in each respective step to allow gradual relaxation. The systems were further equilibrated for 1 ns at 300 K and 1 bar under the NPT ensemble. Production simulations were performed under the NPT ensemble, with five 100 ns replicas for each system, resulting in 500 ns of simulation time per variant. Equilibration and production phases were performed using periodic boundary conditions. Long-range electrostatic interactions were treated with Ewald sums, utilizing a grid spacing of 1.0 Å. Direct electrostatic and Lennard–Jones interactions were both calculated with 9 Å. Hydrogen mass repartitioning with ParmEd [56] enabled a 4 fs integration time step [48]. Temperature control in the NVT ensemble was achieved using Langevin dynamics [57] with a collision frequency of 4.0 ps⁻¹, while pressure control in the NPT ensemble employed the Monte Carlo barostat [47] with a 2.0 ps relaxation time. Bond constraints for hydrogens were applied using the SHAKE algorithm [58].

2.2. Selection of Collective Variables

We established nine collective variables (CVs) to determine different conformational states in each system (see Figure 2). The first three CVs were defined by the distances between pairs of RBDs (Figure 2B). For this case, the distances were calculated considering the center of mass of each RBD without counting the most flexible region, the recognition binding motif (RBM, residues 412–480 in the WT model). The next three CVs were established by the distances between the center of mass of each NTD (Figure 2C). In this case, the whole domain sequence was considered to determine the center of mass. Finally, the last three CVs (Figure 2D) were composed of the angle (θ) formed between two vectors: \vec{v}_1 defined by the center of mass of one RBD and the center of mass of the heptad repeat 1 (HR1) domain of the three chains (residues 912–984) and \vec{v}_2 defined by the center of mass of the HR1 domain of the three chains and the center of the residues 949–969 of the three chains that are closely aligned to the Z-axis. Angles were determined with Equation (1). A detailed description of the residue set for each CV is given in Table 1.

$$\theta = \cos^{-1} \left(\frac{\vec{v}_1 \cdot \vec{v}_2}{|\vec{v}_1| |\vec{v}_2|} \right) \quad (1)$$

This equation, Equation (1), defines the angle (θ) between two vectors (\vec{v}_1 and \vec{v}_2) using the inverse cosine (arccos) function. The expression is derived from the dot product formula, where the numerator represents the scalar product of the two vectors, $\vec{v}_1 \cdot \vec{v}_2$, and the denominator normalizes this product by the magnitudes of the vectors, $|\vec{v}_1|$ and

$|\vec{v}_2|$. In our analysis, the magnitudes $|\vec{v}_1|$ and $|\vec{v}_2|$ are defined based on three spatially separated points, corresponding to the center of mass of selected residues in the RBD and HR1 (Table 1 and Figure 2D). As these vectors are always nonzero due to spatial separation, their magnitudes remain strictly positive. This ensures the validity of the equation within our framework, as it prevents division by zero or undefined operations.

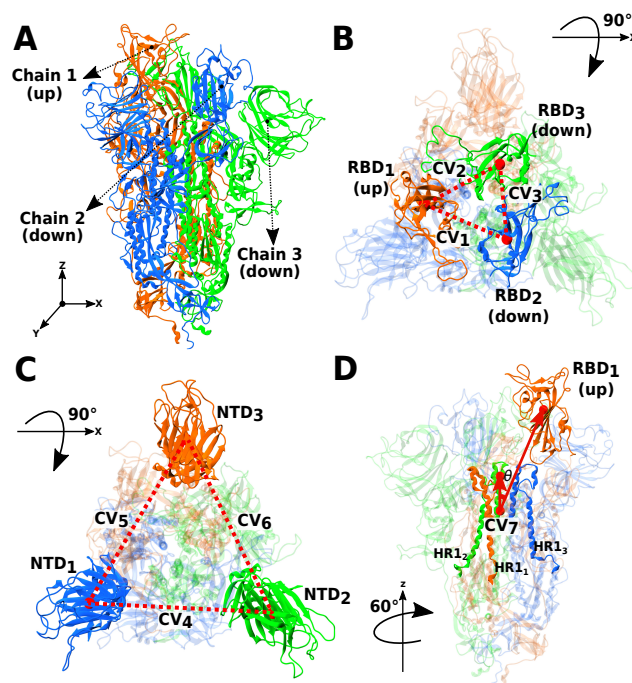


Figure 2. Description of the collective variables (CVs) used to characterize the conformational states of the S protein. Panel (A) illustrates the entire system, highlighting one chain with its receptor-binding domain (RBD) in the up state (orange). Panel (B) presents CV₁ to CV₃, representing the distances between two RBDs (top view). Panel (C) depicts CV₄ to CV₆, calculated based on the distances between two N-terminal domains (NTDs, top view). Panel (D) shows CV₇, defined as the angle (θ) formed between two vectors: one connecting an RBD and the center of mass of the three HR1 domains (side view). The same definition applies to CV₈ and CV₉.

Table 1. Definitions of the range of residues for the 9 collective variables (CVs) employed in this study. The type 1 CV involves the RBD, type 2 the NTD, and type 3 is defined between the RBD and center of mass (COM) of HR1.

CV	Measure	Index	Range of Residues Included in COM Calculations	
Type 1	d	1–3	RBD ₁ (up): I338-W442 to Q512-G532	RBD ₂ (down): I338-W442 to Q512-G532
Type 2	d	4–6	NTD ₁ : A31 to F311	NTD ₂ : A31 to F311
Type 3	angle (°)	7–9	$\vec{v}_1 =$ RBD ₁ (up): I338-G532 and HR1 ₁ –HR1 ₃ : L943-Q1015	$\vec{v}_2 =$ HR1 ₁ –HR1 ₃ : L943-Q1015 and HR1 ₁ –HR1 ₃ : S980-R1000

Note: d = distance in nm and (°) angle in degrees.

2.3. Probability Distribution of Conformational States

The probability distribution for each CV was calculated using a histogram for each variant to capture the conformational states of the S protein. A conformational state was considered formed if a peak in the probability distribution of a given CV. In addition, we performed Gaussian fitting as follows:

$$f(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2)$$

where μ represents the averages of the Gaussian distribution, σ is the standard deviation, and A is the amplitude. This equation remains nonzero, as the distribution $f(x)$ for a given CV is always positive and well-defined in our analysis. When the distributions exhibited two or three conformational states, double or triple Gaussian fitting was applied. The quality of each Gaussian fit was evaluated for statistical accuracy using the χ^2 value, calculated with the following formula:

$$\chi^2 = \sum_k \frac{(f_{\text{obs}}^k - f_{\text{fit}}^k)^2}{f_{\text{fit}}^k} \quad (3)$$

where f_{obs}^k represents the k -th observed value along a CV and f_{fit}^k is the corresponding fitted data according to Equation (2) in a given position. In our case, χ^2 always returns values larger than zero, ensuring that cases of over-fitting the data are not found.

2.4. Calculation of Native Contacts in MD Simulation

We used a combination of structural and chemical-based contact map (CM) determination known as the overlap (OV) and repulsive contacts of structural units (CSUs) approach [38,59]. The OV-CM method determines the presence of an NC between two non-sequential amino acid residues i and j , if the VdW radii associated with heavy atoms of those two residues show an overlap (i.e., $|j - i| \geq 4$). Meanwhile, the rCSU contact map evaluates the chemical properties of the atoms if they are close to each other, adding new contacts to the OV-CM. This approach considers the nature of hydrophobic, hydrophilic, aromatic, ionic, and repulsive (equal charge) residues to define an NC. The NCs were calculated for each state in the Gaussian distributions. Here, we evaluated the MD frames every 100 ps and the NCs were calculated taking all values within the centers of the distributions for each CV (i.e., $\mu \pm \frac{\sigma}{2}$ —see Equation (2))—capturing approximately 70% of the data under the curve. We determined the contacts between the RBDs in CV₁, CV₂, and CV₃. For CV₄ to CV₆, we determine the NCs between the RBDs and its nearest NTD. For CV₇ to CV₉, we determine the NCs between the RBDs and between an RBD and its nearest NTD. The types of interactions between the NCs of each conformational state were distinguished as the following types: ionic, polar, nonpolar, and nonspecific. Ionic interactions were established between NCs presenting one positively (K, R, or H) and one negatively charged (D or E) residue. While polar interactions were established between polar residues (S, T, Y, N, Q, or C) and for nonpolar interaction residues of the same type (G, A, V, L, I, M, F, W, or P), nonspecific interactions were defined between residues that did not present chemical affinity defined by the OV+rCSU CM. The raw data and scripts used in the analysis are located in an open database [60].

3. Results and Discussion

Here, we report the characterization of the conformational space of the S protein within the timescales of AA-MD, mapped using a set of nine CVs. Through the selected CV, we were able to identify distinct conformational states, which, in most cases, were well represented by a Gaussian fit (Table S1). The identification of well-defined states was based on the overlap of the distribution probabilities and the determination of NCs at the protein chain interface. The overall distinction between conformational states in the full-length S protein led to a one-to-one comparison between WT and the other SARS-CoV-2 S variants.

3.1. Elucidating the Conformational States Between Two RBDs

Figure 3A shows the distribution probability for CV₁, which represents the distance between the center of masses of RBD₁ and RBD₂ (see definition in Figure 2B and Table 1). Note that RBD₁ is in the open conformation (i.e., up state), while the others are in the close

conformation (i.e., down state). For WT and Delta variants, the CV_1 shows two Gaussian distributions whose centers are separated by ≈ 1 nm. This distance reflects the spatial flexibility between RBD_1 in the open state and the RBD_2 in the closed state. To verify the local nature of this interaction, we calculated the presence of NCs at the protein interface. For these two variants, no contacts were found, attributing the molecular flexibility to the associated thermal fluctuations common in MD simulations. In contrast, BA.1, XBB.1.5, and JN.1 variants present one single distribution and thus one single state in our MD simulations. Among these variants, JN.1 distribution occurs at the shortest distances and is well separated from the other variants by ≈ 1 nm. Notably, XBB.1.5 and JN.1 distributions are among the narrowest, with average σ values of 0.12 nm and 0.05 nm, respectively. In comparison, the WT exhibits $\sigma = 0.24$ nm at $CV_1 = 5.12$ nm and $\sigma = 0.11$ nm at $CV_1 = 4.25$ nm in WT (see Figures S1, S4, and S5).

On the other hand, the Gaussian distributions of CV_2 exhibited similar values, ranging from 3.90 nm to 4.60 nm across the WT and the Delta, BA.1, and XBB.1.5 variants. These variants showed a narrower distribution in their states and a smaller separation between their centers, which is ≤ 0.5 nm (Figure 3B). In this case, the distributions were based on distances between the RBD up (RBD_1) and the opposite RBD down (RBD_3) (see Figure 2A). We observed that WT, Delta, and XBB.1.5 variants display two states separated by the formation of specific NCs (see Tables S5, S6, and S8). The WT presented six and four nonspecific NCs that were formed between similar residues (S375, F486, T385, and F456) in both distribution states, respectively, with the center of distribution (μ) at 4.36 and 3.93 nm. Notably, the first state displays two extra specific NCs of nonpolar nature (Table S5). This difference of only two NCs of nonpolar nature suggests a smooth transition between states. In contrast, the NC profile differs greatly in the Delta variant, with 9 and 10 NCs for each state (Table S6). Unlike the WT, more NCs of both polar and nonpolar types are observed. In turn, the XBB.1.5 variant presented a state ($\mu = 3.90$ nm) defined by 24 NCs of ionic, polar, nonpolar, and nonspecific nature. The NCs are reduced to only 14 in the second state of XBB.1.5 ($\mu = 4.04$ nm), losing the two most relevant ionic contacts between D427-H505 and D428-H505 (Table S5). It is interesting to note that in XBB.1.5, the higher number of NCs causes larger stiffness between the RBDs, which is correlated with the shortest fluctuation in the distribution (Table S8). Meanwhile, BA.1 presented 1 nonpolar and 6 nonspecific contacts (Table S7), resulting in a single Gaussian distribution with a slightly greater distance compared to XBB.1.5, which supports the idea that few contacts facilitate conformational transitions. Unlike all variants, JN.1 shows a Gaussian distribution with CV_2 values spanning larger distances than the other variants, with a difference of >1 nm. In addition, 24 NCs were identified associated with its distribution (Table S9); however, 16 of these NCs are nonspecific, providing less energetic contribution, thus explaining the wide span.

For CV_3 (Figure 3C), we identified two distinct states in the WT and BA.1. Notably, independent states are formed despite the shorter distances between the two RBDs (down-down). The WT protein presented states with distances of $\mu = 3.42$ and 3.78 nm, which presented 2 and 8 NCs, respectively, most of them nonspecific (Table S5). The Delta variant only presented a single Gaussian distribution ($\mu = 3.44$ nm) with 5 NCs of nonspecific nature similar to WT, but the profile differs greatly (Table S6). In addition, the Delta distribution is within the range of WT, BA.1, and XBB.1.5. Meanwhile, BA.1 exhibited two distributions, separated only by three NCs in the first state and one NC in the second (Table S7). The XBB.1.5 and JN.1 variants only display one state with distances of $\mu = 3.46$ and 3.74 nm, respectively. In XBB.1.5, there are four NCs (one specific and three nonspecific); however, no NCs were observed in JN.1 (Tables S8 and S9). Since JN.1 did not exhibit NCs, its distribution displayed the largest fluctuation among all variants, with $\sigma = 0.20$ nm (see Figure 2). For this CV_3 , due to the inherent timescale in AA-MD, some states may be

separated by infrequent conformations, which could potentially merge with enhanced sampling techniques.

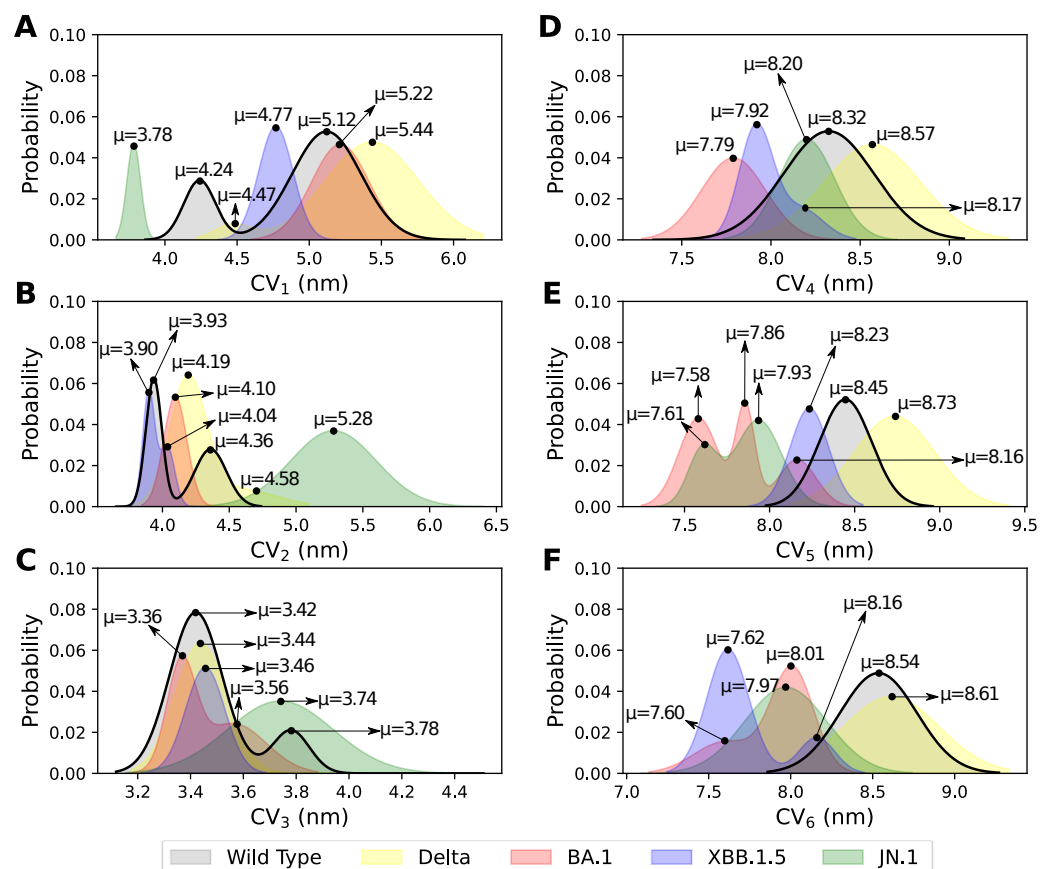


Figure 3. Gaussian distributions of the first 6 CVs based on distances between RBDs (A–C) and NTDs (D–F). All CVs are expressed in distances (nm). The distributions were colored for each variant: WT in gray, Delta in yellow, BA.1 in red, JN.1 in green, and XBB.1.5 in blue.

3.2. Conformational States and NCs Identified Between Two NTDs

The two NTDs in the S protein are more distantly separated compared to the two RBDs. Despite this, we observed distinct fluctuations in each variant. Therefore, we analyzed their distributions and NCs using the same approach. Notably, in these cases, the NCs were determined between each NTD and its nearest RBD.

CV₄ represents the distance between NTD₁ and NTD₂ (Figure 2C). Single distributions were observed for most variants, except for XBB.1.5, which exhibited two distributions (Figure 3D). The individual distributions were broader and well differentiated among the variants. In addition, these distributions exhibit few NCs as the NTD evaluated here is the closest to the RBD up conformation. For instance, WT ($\mu = 8.32$ nm) presented only four nonspecific NCs and BA.1 ($\mu = 7.79$ nm) presented two NCs (one nonspecific and one nonpolar) (Tables S10 and S12). No NC formation was observed in the Delta and JN.1 variants for CV₄ (Tables S11 and S14). In XBB.1.5, we identified one state ($\mu = 7.92$ nm) with 10 NCs, and the second state ($\mu = 8.17$ nm) with 12 NCs (Table S13), with similar proportions between specific and nonspecific NCs. Interestingly, the Delta and JN.1 variants have distributions close to WT, separated by approximately ± 0.2 nm. In contrast, BA.1 and XBB.1.5 are more separated in relation to the WT (≈ 0.5 nm) and presented the shortest distances (Figure 3D). This suggests a greater compactness within these domains (i.e., RBD and NTD) for these variants that are also more distant from WT in terms of genetic variability (see Figure S6).

In contrast, in CV₅, we observed Gaussian distributions that are more spread out regarding CV₄ (Figure 3E). In terms of NCs, we observed a higher number of contacts in variants that are distant. WT, Delta, and XBB.1.5 presented single distributions and these were close to each other, with centers separated by ≈ 0.3 nm. These three distributions were also rigid as the σ values ranged between 0.11 and 0.24 nm (see Figures S1, S2, and S5). In addition, the NCs of these three variants are mostly nonspecific, with R357 and P521 being relevant in these three profiles (Tables S10, S11, and S13). Conversely, BA.1 and JN.1 presented more than one Gaussian distribution, but with smaller distances compared to the previous ones. BA.1 presented up to three different distributions with NC profiles that varied significantly, with 21, 26, and 4 NCs for each state (Table S12). The two states with the shortest distance presented one ionic contact (R357-E167), which contributed to stronger binding between these domains. Additionally, a dominance of NCs involving positively charged residues was observed, including R355, R357, R466, H519, and K114. However, the third state only presented four nonspecific NCs. It is noteworthy that ionic contacts have been rarely observed at the S protein interface in previous variants (WT and Delta). In addition, the distribution of the JN.1 is within the range of BA.1, but with only two states, presenting 14 and 13 NCs, respectively, most of them nonspecific (Table S14).

The distances observed in CV₆ were smaller, suggesting a closer relationship between associated the NRD of the RBD down (Figure 3F). In this case, we observed that WT and Delta exhibited single distributions with their center at greater distances. The wide distribution of these variants is driven by the lack of contacts, only one in WT (P230-N394) and three in Delta (R357-P230, A520-K41, and P521-K41) (Tables S10 and S11). The BA.1, XBB.1.5, and JN.1 variants had smaller distributions similar to each other. BA.1 has two states with seven and six NCs, respectively. Notably, only one NC (Q116-I468) distinguished two states, but due to the AA-MD timescale, it may be required to extend the simulation or number of replicas to determine the statistical significance of the two-state model. In contrast, JN.1 exhibited a single distribution, dominated by ten NCs, one polar and nine nonspecific (Table S14). In the case of XBB.1.5, we found a greater separation and also a greater difference in the NC profile of each state. One of the states ($\mu = 7.62$ nm) is dominated by two polar and six nonspecific NCs; meanwhile, the second ($\mu = 8.16$ nm) is determined by the same two polar contacts and twelve other nonspecific NCs (Table S13). These results suggest that most states are determined by nonspecific contacts, a trend observed across the previously evaluated CVs. However, the presence of a few ionic, polar, or nonpolar NCs can contribute to the formation of distinct states. In the case of Omicron variants (e.g., BA.1 and XBB.1.5), these states are more compact compared to the WT. In the next section, we will introduce a different CV to analyze large conformational changes based on the angle between the RBD-HR1 subunits.

3.3. Conformational States and NCs Identified with Angle Between RBD-HR1 Subunits

The CV defined as an angle is a useful measure to evaluate fluctuations of large domains [61]. Here, we analyzed the distribution of angles formed between two specific vectors (Figure 2D), which reflects the flexibility of each RBD (Figure 2D). Likewise, for each case, we determined the occurrence of NCs between the RBDs and between one RBD and its nearest NTD unit.

CV₇ exhibited the largest angular fluctuation as it involves the RBD up (Figure 4A). Across all variants, no NCs were identified between the RBD up and its opposite RBD, consistent with CV₁ data indicating the absence of NCs. Contacts between the RBD and its closest NTD were mainly observed here. The WT and Delta variant display similar distributions, with two states separated by $\approx 6^\circ$. However, these states were clearly distinct, as Delta did not present NCs between the RBD and the NTD (Table S16). This is also in

agreement with CV_4 . Therefore, we can infer that Delta variability, particularly the T478K mutation in the RBD, induces a greater separation between these domains. The BA.1 variant is also close to WT in terms of Gaussian distribution (Figure 4A), but with a fluctuation of $\approx 3^\circ$. The NCs identified in BA.1 are large in number and mostly specific ones (Table S17) in the state close to the WT ($\mu = 32.34^\circ$). It is interesting to note the presence of two ionic contacts, R357-E167 and R466-E133, that provide better stability to those domains. In contrast, the second state ($\mu = 40.12^\circ$) lost all its NCs, retaining only one nonpolar NC, in agreement with its intrinsic flexibility. On the other hand, JN.1 and XBB.1.5 present distributions with lower angular values than the WT, which denotes increased rigidity. XBB.1.5 exhibits the most compact states due to the greater number of specific NCs between the domains. Here, we observe two states with similar NC profile and similar ratio between specific and nonspecific NCs (Figure 4A and Table S18). In contrast, JN.1 (Table S19) did not present NCs between these domains (i.e., NTD and RBD), similar to Delta. Therefore, its compactness is likely not determined by NCs but by structural changes in other domains.

Similarly, CV_8 , which measures the flexibility of RBD_2 (in down conformation), has a higher compactness in all variants denoted by the lower values compared to CV_7 , ranging from $\approx 24^\circ$ to 34° (Figure 4B). All variants except JN.1 exhibited a single state with a broad Gaussian distribution. The WT exhibits the highest flexibility, denoted by the σ value of 2.61 (see Figure S1). In contrast, the variants had a greater compactness, with lower σ values related to a significant presence of NCs (Tables S16–S19). Interestingly, with CV_8 , we observed NCs between both RBDs and between RBD and NTD units. We also noticed a large number of specific NCs between RBDs, the Delta variant with five specific NCs (Table S16), BA.1 with four specific NCs, among them an ionic D428-H505 arising from the Y505H mutation (Table S14), JN.1 with eight and thirteen specific NCs in its two states, respectively (Table S19), and XBB.1.5 with seven specific NCs of which two were ionic caused by the Y505H mutation similarly (Table S18). A consistent trend was observed: WT had fewer contacts, while the four variants displayed a higher number of NCs, most of which were specific (forming ionic, polar, and nonpolar interactions). Additionally, the JN.1 variant exhibited the highest number of contacts, with up to 36 NCs, most of which were nonspecific. This higher number of NCs indicates a marked difference for the RBD_2 of this particular variant, in agreement with the different distribution (Figure 4B).

Finally, CV_9 measures the flexibility of the RBD_3 (in down conformation). In this case, the WT is defined by two states, while all other variants displayed single distributions, ranging from ≈ 25 to 32° (Figure 4C). This RBD exhibited a fluctuation range similar to the opposite RBD, which was also in the down conformation ($\approx 20^\circ$ to 35° , Figure 4C). Therefore, we can infer that their dynamics in terms of protein stability are highly similar. In addition, the Gaussian distributions of all systems are closer to each other, with a maximum deviation of $\approx 4^\circ$ (Figure 4C). The states of WT are defined by mostly nonspecific NCs, six between RBDs and up to five between the RBD and NTD (Table S15). In Delta and BA.1 variants, we observed a similar number of contacts between the RBDs (four and one NCs, respectively) and NTDs (one and four NCs, respectively) (Tables S16 and S17). XBB.1.5 has NCs in both cases, between RBDs (three NCs) and between an RBD and NTD (nine NCs) (Table S18). The JN.1 variant displayed fewer NCs overall (Table S19), with no NCs between RBDs but ten NCs between RBD and NTD units. Additionally, we observed that JN.1 exhibited a markedly different profile between RBD_2 (CV_8) and RBD_3 , with numerous contacts around RBD_2 and none around RBD_3 .

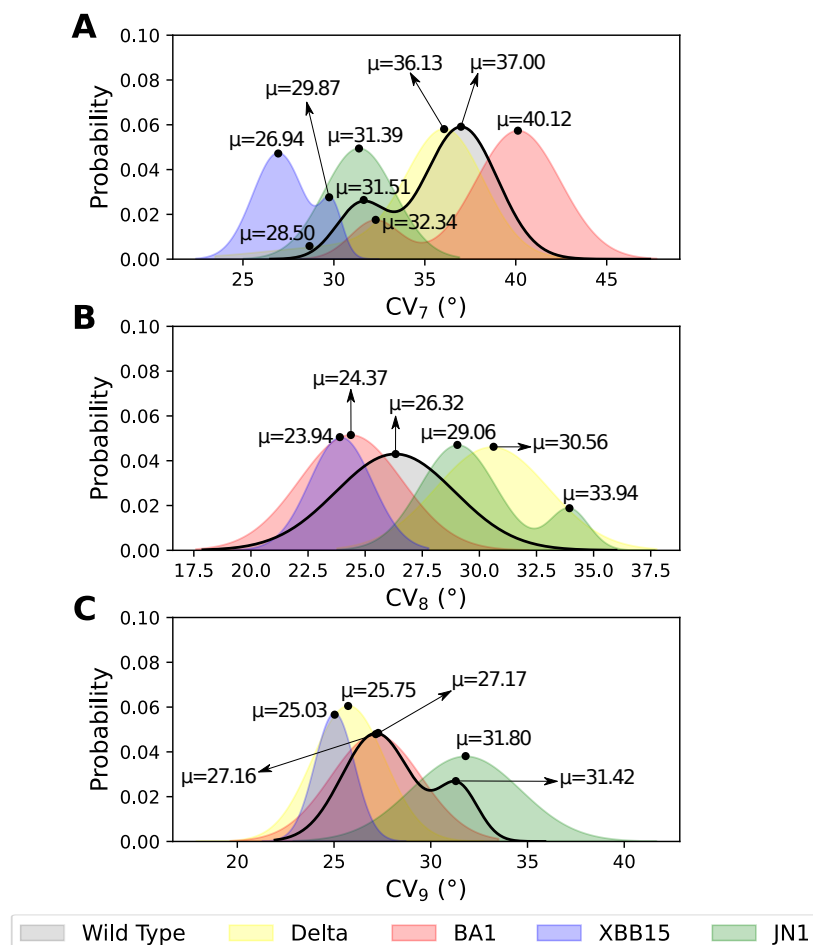


Figure 4. Gaussian distribution of the CVs defined by the angle formed between two specific vectors involving RBD₁ up (chain-A), RBD₂ down (chain-B), and RBD₃ down (chain-C). The angle is measured in degrees (°). The distributions follow the same color pattern for each SARS-CoV-2 variant.

3.4. Role of the Mutations on the Conformational States of the S Protein

We identified the formation of specific NCs in each state, considering the conserved and non-conserved residues. The NCs that occur in a conserved manner (with non-mutated residues) predominate between the RBD and NTD subunits, suggesting that surface stability between these two domains is maintained throughout the virus's evolution. This is consistent with the greater conservation observed near the stalk of the S protein [62]. In contrast, the NC profile for each state varied significantly between RBDs, with greater differences observed in variants with more distant sequences.

In the Delta variant, the mutation T478K drives the formation of three new NCs, as calculated with CV₂ (Figure 5A,B), which were not observed in the WT. Notably, T478K is located in the RBM region and is present in all the investigated variants (Figure S6). These new NCs are formed with conserved residues in the opposite chain, such as N370, S371, and A372 (Table S6). Furthermore, this group of new NCs occurs only in one state of CV₂, leading to a different state with greater distance between the RBDs. The T478K mutation has been listed as an important modification for altering the interface potential of the RBD by making it more positive. However, there are conflicting reports on whether this mutation enhances the specific interaction with the hACE2 receptor [63]. Additionally, no studies have reported that T478K improves H-bond or salt bridge formation with hACE2, but an improvement in electrostatic forces of 2.8 kcal/mol was reported in [64]. We point out that T478K does not induce the formation of new NCs in the following variants, and therefore, its role is limited to this variant.

In contrast, the BA.1 variant exhibited fewer NCs resulting from RBD mutations (Figure 5A,B). Among these, only the Y505H mutation generated a new specific interaction with D428. This mutation is particularly significant due to its ionic nature, which substantially enhances interactions between RBDs. The formation of this new NC was identified in both CV₂ and CV₈ (Tables S7 and S17). Previous studies have reported that YH505 affects the function of the S protein by decreasing its affinity for hACE2 [21,65,66]. In that sense, our results suggest that Y505H would stabilize the RBDs and make the interface more compact, at the cost of decreasing the affinity for hACE2. This observation aligns with reports indicating that this mutation increases the potential for antibody blockade [67]. Given this information, we propose that the BA.1 variant may exhibit a reduced accessible surface area at the protein interface, which could serve as an efficient mechanism to evade neutralizing antibodies.

In the case of XBB.1.5, we also identified new NCs attributed to mutations only among RBDs, in particular for CV₂, CV₃, CV₈, and CV₉ (Tables S8 and S18). The behavior of the RBDs in this variant is highly similar, as are the NCs attributed to specific mutations. This suggests that evaluating the same domain using different CV types (e.g., distance- and angle-based) may be beneficial for validating the impact of key mutations. In CV₂, three mutations (S373P, F486P, and Y505H) are associated with the formation of three NCs of specific nature (P373-P486, D427-H505, and D428-H505), while four mutations (S373P, K417N, Q498R, and N501Y) lead to five new nonspecific NCs (P373-N487, G381-N417, G413-R498, Q414-R498, and D427-Y501). CV₃ is less influenced by mutations as it presents few contacts; only D405N and S375F formed three new nonspecific NCs (N405-F374, N405-F375, H505-F375) (Table S8). In CV₈, we found three mutations (S373P, F486P, and Y505H) that lead to the formation of four new specific contacts (P373-P486, F374-P486, D427-H505, and D428-H505), two of them ionic. Likewise, mutations Q498R and N501Y give rise to four nonspecific NCs (G413-R498, G413 Y501, Q414 R498, and D427 Y501) for CV₈. For CV₉, we detected that the D405N mutation generates two new nonspecific NCs (N405-F374 and N405-F375). In agreement with the above, there are no NCs related to mutated residues in this variant (Table S18). Notably, in some cases, a single mutation can lead to the formation of multiple NCs, as seen with Y505H, which generates two new NCs.

The JN.1 was the variant with the highest number of mutated residues, resulting in a distinct NC profile. Specifically, with CV₈, we identified the highest number of NCs among two RBDs (RBD₁ and RBD₃), which are the closest to each other. Both of these states have new NCs attributed to mutations (Figure 5E,F). In the first state, the N500Y and Y504H mutations give rise to new NCs of polar and ionic nature, respectively. Other mutations of interest are N405, N417, S455, R497, and H504, although these only generate new nonspecific NCs. Notably, the R497 and H504 mutations increase the positive charge of the RBD surface, altering the domain's electrostatic potential, a mechanism that enhances antibody evasion, particularly in Omicron variants [68]. On the other hand, the second state has a greater number of mutations that generate new NCs. Among these, F485P, D405N, K417N, G455S, T483K, F485P, N500Y, and Y504H primarily interacted with conserved residues on the opposite chain (Table S19). This demonstrates that the S protein (Delta, BA.1, XBB.1.5, and finally JN.1) is moving towards a structure with a higher number of NCs, which makes the RBDs more compact and grants a high potential for immune evasion.

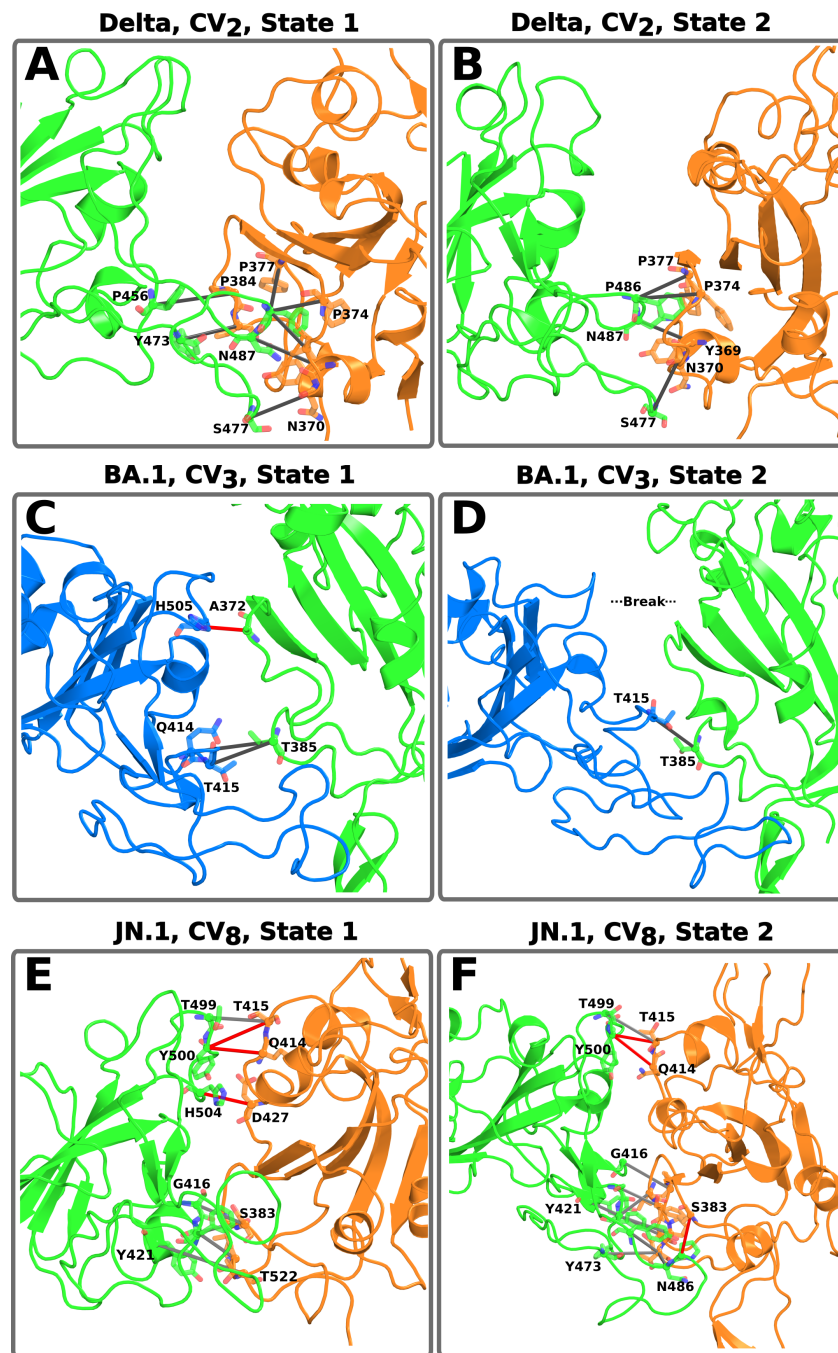


Figure 5. Formation of specific NC among the RBDs, mediated by conserved (gray dashes) and mutated (red dashes) residues. Panel (A,B) show the RBDs (CV₂) of the Delta variant. Panel (C) displays how mutation Y505H, associated with the BA.1 variant, leads to a new NCs in the distal part of the RBD, which were lost in the second state (D). Panel (E,F) demonstrate that in the JN.1 variant, multiple mutations (red dashes) lead to the formation of new NCs in both states, resulting in RBD compaction. RBD colors follow the same definitions from Figure 2.

4. Conclusions and Perspectives

Several MD studies have focused on determining the effect of mutations on RBD/hACE2 complex stability. In this study, using AA-MD simulation and contact map analysis, we reveal the key role of mutations in shaping different conformational states of the S protein and quantify the associated NC profiles. Through statistical analysis based on the distribution probabilities of distance- and angle-based CVs, we demonstrated that several variants exhibit a strong tendency to restrict RBD motion. This tendency is more

marked in Omicron variants (BA.1, XBB.1.5, and JN.1). Furthermore, we determined that mutations impact interactions between RBDs more than between the RBD and NTD. Additionally, our computational analysis highlights the importance of specific mutations in defining distinct conformational states in the S protein, beyond their well-established role in enhancing RBD/hACE2 interactions. In particular, T478K, N500Y, and Y504H are found to display this importance. Here, we have observed that these mutations also originate new NC pairs that determine different conformational states. Moreover, the NC profiles provide valuable insights into protein stability and facilitate comparative analyses of the conformational space across several SARS-CoV-2 S variants.

In addition, it is important to emphasize that the transitions between the up and down states are inherently stochastic and typically occur on longer timescales. Several studies have employed enhanced sampling techniques to characterize these transitions [14,30,69,70]. While these approaches (biased MD) extend accessible timescales, they may also introduce unphysical transition pathways. Therefore, to capture more realistic transition states, unbiased simulations are generally necessary. This is not only necessary to characterize the open and closed states, but also for the transition from prefusion to postfusion states and the effect of mutation for those cases. Although we performed extensive conformational sampling (500 ns) using unbiased MD of the full trimeric structure, this approach may still limit the observation of long-timescale conformational changes. However, our conformational sampling effectively captures fluctuations around the initial state (up conformation), allowing for a comparative analysis of the effect of mutations. Accessing timescales in the micro- or millisecond range is only possible with high-performance computational infrastructures, such as that realized in the aerosol virus simulation [71]. This limitation can be overcome in future studies through coarse-grained methods, which extend accessible timescales by one or two orders of magnitude. In this line, the characterized conformational states and NC profiles can further aid the development of effective CG models, enabling the exploration of larger length scales (e.g., virion level) and longer timescales (e.g., sub-millisecond), thereby bringing simulations closer to biological conditions.

Finally, our findings are consistent with both experimental [72–74] and computational studies [75]. As mentioned, the most recent variants (Omicron, particularly JN.1) tend to adopt a more compact conformational state in the interphase (near the RBD and NTD), accompanied by the formation of new NCs between mutated–mutated or mutated–conserved residues (e.g., T478K, N500Y, and Y504H). From this, we infer that the energy barrier for the transition between the open and closed states may be higher, which could affect the overall fitness of the virus. However, this structural alteration also confers an advantage in immune evasion, both for innate and acquired immunity (such as that induced by vaccination), as recently reported [76]. This involves reduced accessibility of critical epitopes, which could hinder recognition by neutralizing antibodies, contributing to immune escape. Although the stability of the RBDs is altered in the new variants, they remain functional for binding to host cell receptors, leading to increased viral infectivity. This reflects a balance between reduced hACE2 affinity and enhanced immune evasion, as suggested in other studies [77]. These changes in the dynamics of the protein underscore the adaptive mechanisms of the virus and suggest that variants with such structural modifications may exhibit both enhanced immune resistance and increased transmissibility.

Supplementary Materials: The following supporting information can be downloaded at the following link: <https://www.mdpi.com/article/10.3390/pathogens14030274/s1>; Figure S1: Gaussian distribution of the WT protein for all calculated CVs.; Figure S2: Gaussian distribution of the Delta variant for all calculated CVs.; Figure S3: Gaussian distribution of the BA.1 variant for all calculated CVs.; Figure S4: Gaussian distribution of the XBB.1.5 variant for all calculated CVs.; Figure S5: Gaussian distribution of the JN.1 variant for all calculated CVs.; Figure S6: Alignment of the primary sequences

of the WT S protein and its 4 variants.; Table S1: Values of each CV (calculated between RBDs), σ , and χ^2 averages of the different states obtained by the probability distribution.; Table S2: Values of each CV (calculated between NTDs), σ , and χ^2 averages of the different states obtained by the probability distribution.; Table S3: Values of each CV (calculated with angles of RBDs), σ , and χ^2 averages of the different states obtained by the probability distribution.; Table S4: Missing residues range remodeled into the chains of the S protein variants.; Table S5: High frequency NCs (freq > 60%) between the RBDs in the WT protein.; Table S6: High frequency NCs (freq > 60%) between the RBDs in the Delta variant.; Table S7: High frequency NCs (freq > 60%) between the RBDs in the BA.1 variant.; Table S8: High frequency NCs (freq > 60%) between the RBDs in the XBB.1.5 variant.; Table S9: High frequency NCs (freq > 60%) between the RBDs in the JN.1 variant.; Table S10: High frequency NCs (freq > 60%) between an NTD and RBD in the WT protein.; Table S11: High frequency NCs (freq > 60%) between an NTD and RBD in the Delta variant.; Table S12: High frequency NCs (freq > 60%) between an NTD and RBD in the BA.1 variant.; Table S13: High frequency NCs (freq > 60%) between an NTD and RBD in the XBB.1.5 variant.; Table S14: High frequency NCs (freq > 60%) between an NTD and RBD in the JN.1 variant.; Table S15: High frequency NCs (freq > 60%) for the CVs calculated with angles in the WT protein.; Table S16: High frequency NCs (freq > 60%) for the CVs calculated with angles in the Delta variant.; Table S17: High frequency NCs (freq > 60%) for the CVs calculated with angles in the BA.1 variant.; Table S18: High frequency NCs (freq > 60%) for the CVs calculated with angles in the XBB.1.5 variant.; Table S19: High frequency NCs (freq > 60%) for the CVs calculated with angles in the JN.1 variant.

Author Contributions: Conceptualization, A.B.P.; methodology, G.E.O.-R., L.F.C.-V. and T.M.; software, G.E.O.-R. and L.F.C.-V.; validation, G.E.O.-R. and L.F.C.-V.; formal analysis, G.E.O.-R., L.F.C.-V., T.M. and A.B.P.; investigation, G.E.O.-R.; resources, A.B.P.; data curation, G.E.O.-R. and L.F.C.-V.; writing—original draft preparation, G.E.O.-R.; writing—review and editing, L.F.C.-V., T.M. and A.B.P.; visualization, G.E.O.-R.; supervision, A.B.P.; project administration, A.B.P.; funding acquisition, A.B.P. All authors have read and agreed to the published version of the manuscript.

Funding: A.B.P. acknowledges financial support from the National Science Center, Poland, under grant 2022/45/B/NZ1/02519 and gratefully acknowledges the Polish high-performance computing infrastructure PLGrid (HPC Centers: ACK Cyfronet AGH) for providing computer facilities and support within computational grant nos. PLG/2023/016519 and PLG/2024/017332.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available at <https://doi.org/10.5281/zenodo.14793235>, accessed on 20 January 2025.

Acknowledgments: The authors thank Joseph L. Baker for carefully reading the manuscript and providing insightful comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lindorff-Larsen, K.; Piana, S.; Dror, R.O.; Shaw, D.E. How fast-folding proteins fold. *Science* **2011**, *334*, 517–520. [[CrossRef](#)] [[PubMed](#)]
2. Prašnikar, E.; Ljubič, M.; Perdih, A.; Borišek, J. Machine learning heralding a new development phase in molecular dynamics simulations. *Artif. Intell. Rev.* **2024**, *57*, 102. [[CrossRef](#)]
3. Ramanathan, A.; Agarwal, P.K. Computational identification of slow conformational fluctuations in proteins. *J. Phys. Chem. B* **2009**, *113*, 16669–16680. [[CrossRef](#)] [[PubMed](#)]
4. Prinz, J.H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J.D.; Schütte, C.; Noé, F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105. [[CrossRef](#)]
5. Wang, W.; Cao, S.; Zhu, L.; Huang, X. Constructing Markov State Models to elucidate the functional conformational changes of complex biomolecules. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1343. [[CrossRef](#)]

6. Henderson, R.; Edwards, R.J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S.; Kopp, M.; Li, D.; Parks, R.; Hsu, A.; et al. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 925–933. [[CrossRef](#)]
7. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.L.; Abiona, O.M.; Graham, B.S.; McLellan, J.S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263. [[CrossRef](#)]
8. Moreira, R.A.; Chwastyk, M.; Baker, J.L.; Guzman, H.V.; Poma, A.B. Quantitative determination of mechanical stability in the novel coronavirus spike protein. *Nanoscale* **2020**, *12*, 16409–16413. [[CrossRef](#)]
9. Ray, A.; Minh Tran, T.T.; Santos Natividade, R.d.; Moreira, R.A.; Simpson, J.D.; Mohammed, D.; Koehler, M.; L Petitjean, S.J.; Zhang, Q.; Bureau, F.; et al. Single-molecule investigation of the binding interface stability of SARS-CoV-2 variants with ACE2. *ACS Nanosci. Au* **2024**, *4*, 136–145. [[CrossRef](#)]
10. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444–1448. [[CrossRef](#)]
11. Walls, A.C.; Park, Y.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Veesler, D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **2020**, *181*, 281–292.e6. [[CrossRef](#)] [[PubMed](#)]
12. Klinakis, A.; Cournia, Z.; Rampias, T. N-terminal domain mutations of the spike protein are structurally implicated in epitope recognition in emerging SARS-CoV-2 strains. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 5556–5567. [[CrossRef](#)] [[PubMed](#)]
13. Ciccozzi, M.; Pascarella, S. Two sides of the same coin: The n-terminal and the receptor-binding domains of SARS-CoV-2 spike. *Future Virol.* **2023**, *18*, 75–78. [[CrossRef](#)]
14. Pang, Y.T.; Acharya, A.; Lynch, D.L.; Pavlova, Á.; Gumbart, J.C. SARS-CoV-2 spike opening dynamics and energetics reveal the individual roles of glycans and their collective impact. *Commun. Biol.* **2022**, *5*, 1170. [[CrossRef](#)]
15. Guruprasad, L.; Naresh, G.K.; Boggarapu, G. Taking stock of the mutations in human SARS-CoV-2 spike proteins: From early days to nearly the end of COVID-19 pandemic. *Curr. Res. Struct. Biol.* **2023**, *6*, 100107. [[CrossRef](#)] [[PubMed](#)]
16. McCrone, J.; Hill, V.; Bajaj, S.; Pena, R.E.; Lambert, B.; Inward, R.; Bhatt, S.; Volz, E.; Ruis, C.; Dellicour, S.; et al. Context-specific emergence and growth of the SARS-CoV-2 Delta variant. *Nature* **2022**, *610*, 154–160. [[CrossRef](#)]
17. Li, K.; Melnychuk, S.D.; Sandstrom, P.; Ji, H. Tracking the evolution of the SARS-CoV-2 delta variant of concern: Analysis of genetic diversity and selection across the whole viral genome. *Front. Microbiol.* **2023**, *14*, 1222301. [[CrossRef](#)]
18. Zhang, L.; Jackson, C.B.; Mou, H.; Ojha, A.; Peng, H.; Quinlan, B.D.; Rangarajan, E.S.; Pan, A.; Vanderheiden, A.; Suthar, M.S.; et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat. Commun.* **2020**, *11*, 6013. [[CrossRef](#)]
19. Dhawan, M.; Sharma, A.; Choudhary, P.; Thakur, N.; Rajkhowa, T.; Choudhary, O.P. Delta variant (B.1.617.2) of SARS-CoV-2: Mutations, impact, challenges and possible solutions. *Hum. Vaccines Immunother.* **2022**, *18*, 2068883. [[CrossRef](#)]
20. Shrestha, L.B.; Foster, C.S.P.; Rawlinson, W.D.; Tedla, N.; Bull, R.A. Evolution of the SARS-CoV-2 omicron variants BA.1 to BA.5: Implications for immune escape and transmission. *Rev. Med. Virol.* **2022**, *32*, e2381. [[CrossRef](#)]
21. Kumar, S.; Karuppanan, K.; Subramaniam, G. Omicron (BA.1) and sub-variants (BA.1.1, BA.2, and BA.3) of SARS-CoV-2 spike infectivity and pathogenicity: A comparative sequence and structural-based computational assessment. *J. Med. Virol.* **2022**, *94*, 4780–4791. [[CrossRef](#)] [[PubMed](#)]
22. Parums, D.V. Editorial: The xbb.1.5 ('kraken') subvariant of omicron SARS-CoV-2 and its rapid global spread. *Med. Sci. Monit.* **2023**, *29*, e939580-1. [[CrossRef](#)]
23. Satapathy, P.; Kumar, P.; Mehta, V.; Suresh, V.; Khare, A.; Rustagi, S.; Daulati, M.N.; Neyazi, M.; Najafi, E.; Neyazi, A. Global spread of COVID-19's JN.1 variant: Implications and public health responses. *New Microbes New Infect.* **2024**, *57*, 101225. [[CrossRef](#)]
24. Kaku, Y.; Okumura, K.; Padilla-Blanco, M.; Kosugi, Y.; Uriu, K.; Hinay, A.A.; Chen, L.; Plianachaisuk, A.; Kobiyama, K.; Ishii, K.J.; et al. Virological characteristics of the SARS-CoV-2 JN.1 variant. *Lancet Infect. Dis.* **2024**, *24*, e82. [[CrossRef](#)] [[PubMed](#)]
25. Lu, Y.; Ao, D.; He, X.; Wei, X. The rising SARS-CoV-2 JN.1 variant: Evolution, infectivity, immune escape, and response strategies. *MedComm* **2024**, *5*, e675. [[CrossRef](#)] [[PubMed](#)]
26. Haslwanter, D.; Dieterle, M.E.; Wec, A.Z.; O'Brien, C.M.; Sakharkar, M.; Florez, C.; Tong, K.; Rappazzo, C.G.; Lasso, G.; Vergnolle, O.; et al. A combination of receptor-binding domain and n-terminal domain neutralizing antibodies limits the generation of SARS-CoV-2 spike neutralization-escape mutants. *mBio* **2021**, *12*, e02473-21. [[CrossRef](#)]
27. Barton, M.I.; MacGowan, S.A.; Kutuzov, M.A.; Dushek, O.; Barton, G.J.; Merwe, P.A.v.d. Effects of common mutations in the SARS-CoV-2 spike RBD and its ligand, the human ace2 receptor on binding affinity and kinetics. *eLife* **2021**, *10*, e70658. [[CrossRef](#)]
28. Koehler, M.; Ray, A.; Moreira, R.A.; Juniku, B.; Poma, A.B.; Alsteens, D. Molecular insights into receptor binding energetics and neutralization of SARS-CoV-2 variants. *Nat. Commun.* **2021**, *12*, 6977. [[CrossRef](#)]
29. Laffeber, C.; Koning, K.d.; Kanaar, R.; Lebbink, J.H. Experimental evidence for enhanced receptor binding by rapidly spreading SARS-CoV-2 variants. *J. Mol. Biol.* **2021**, *433*, 167058. [[CrossRef](#)]
30. Sztain, T.; Ahn, S.; Bogetti, A.T.; Casalino, L.; Goldsmith, J.A.; Seitz, E.; McCool, R.S.; Kearns, F.L.; Acosta-Reyes, F.; Maji, S.; et al. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat. Chem.* **2021**, *13*, 963–968. [[CrossRef](#)]

31. Cofas-Vargas, L.; Moreira, R.; Poblete, S.; Chwastyk, M.; Poma, A. The Gōmartini approach: Revisiting the concept of contact maps and the modelling of protein complexes. *Acta Phys. Pol. A* **2024**, *145*, S9–S20. [[CrossRef](#)]
32. Best, R.B.; Hummer, G.; Eaton, W.A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17874–17879. [[CrossRef](#)] [[PubMed](#)]
33. Duarte, J.M.; Sathyapriya, R.; Stehr, H.; Filippis, I.; Lappé, M. Optimal contact definition for reconstruction of contact maps. *BMC Bioinform.* **2010**, *11*, 283. [[CrossRef](#)] [[PubMed](#)]
34. Noel, J.K.; Whitford, P.C.; Onuchic, J.N. The shadow map: A general contact definition for capturing the dynamics of biomolecular folding and function. *J. Phys. Chem. B* **2012**, *116*, 8692–8702. [[CrossRef](#)]
35. Poupon, A. Voronoi and voronoi-related tessellations in studies of protein structure and interaction. *Curr. Opin. Struct. Biol.* **2004**, *14*, 233–241. [[CrossRef](#)]
36. Tsai, J.; Taylor, R.; Chothia, C.; Gerstein, M. The packing density in proteins: Standard radii and volumes. *J. Mol. Biol.* **1999**, *290*, 253–266. [[CrossRef](#)]
37. Sobolev, V.M.; Sorokin, A.; Prilusky, J.; Abola, E.E.; Edelman, M. Automated analysis of interatomic contacts in proteins. *Bioinformatics* **1999**, *15*, 327–332. [[CrossRef](#)]
38. Wołek, K.; Gómez-Sicilia, Á.; Cieplak, M. Determination of contact maps in proteins: A combination of structural and chemical approaches. *J. Chem. Phys.* **2015**, *143*, 243105. [[CrossRef](#)]
39. Huang, E.S.; Subbiah, S.; Levitt, M. Recognizing native folds by the arrangement of hydrophobic and polar residues. *J. Mol. Biol.* **1995**, *252*, 709–720. [[CrossRef](#)]
40. Gō, N. Protein folding as a stochastic process. *J. Stat. Phys.* **1983**, *30*, 413–423. [[CrossRef](#)]
41. Pedersen, K.B.; Borges-Araújo, L.; Stange, A.D.; Souza, P.C.T.; Marrink, S.J.; Schiøtt, B. Olives: A gō-like model for stabilizing protein structure via hydrogen bonding native contacts in the Martini 3 coarse-grained force field. *J. Chem. Theory Comput.* **2024**, *20*, 8049–8070. [[CrossRef](#)] [[PubMed](#)]
42. Souza, P.C.T.; Borges-Araújo, L.; Brasnett, C.; Moreira, R.A.; Grünewald, F.; Park, P.; Wang, L.; Razmazma, H.; Borges-Araújo, A.C.; Cofas-Vargas, L.F.; et al. Gōmartini 3: From large conformational changes in proteins to environmental bias corrections. *BioRxiv* **2024**. [[CrossRef](#)]
43. Moreira, R.A.; Chwastyk, M.; Baker, J.L.; Guzman, H.V.; Poma, A.B. All atom simulations snapshots and contact maps analysis scripts for SARS-CoV-2002 and SARS-CoV-2 spike proteins with and without ACE2 enzyme. *Zenodo* **2020**. [[CrossRef](#)]
44. Moreira, R.A.; Guzman, H.V.; Boopathi, S.; Baker, J.L.; Poma, A.B. Characterization of structural and energetic differences between conformations of the SARS-CoV-2 spike protein. *Materials* **2020**, *13*, 5362. [[CrossRef](#)] [[PubMed](#)]
45. Maier, J.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.; Simmerling, C. Ff14sb: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [[CrossRef](#)]
46. Jorgensen, W.L. Quantum and statistical mechanical studies of liquids. 10. transferable intermolecular potential functions for water, alcohols, and ethers. application to liquid water. *J. Am. Chem. Soc.* **1981**, *103*, 335–340. [[CrossRef](#)]
47. Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B.O. Molecular dynamics simulations of water and biomolecules with a monte carlo constant pressure algorithm. *Chem. Phys. Lett.* **2004**, *384*, 288–294. [[CrossRef](#)]
48. Hopkins, C.W.; Grand, S.L.; Walker, R.C.; Roitberg, A.E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **2015**, *11*, 1864–1874. [[CrossRef](#)]
49. Berman, H.M. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)]
50. Šali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. [[CrossRef](#)]
51. Eastman, P.; Swails, J.M.; Chodera, J.D.; McGibbon, R.T.; Zhao, Y.; Beauchamp, K.A.; Wang, L.; Simmonett, A.C.; Harrigan, M.P.; Stern, C.; et al. Openmm 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **2017**, *13*, e1005659. [[CrossRef](#)] [[PubMed](#)]
52. Case, D.A.; Aktulga, H.M.; Belfon, K.; Cerutti, D.S.; Cisneros, G.A.; Cruzeiro, V.W.D.; Forouzes, N.; Giese, T.J.; Götz, A.W.; Gohlke, H.; et al. AmberTools. *J. Chem. Inf. Model.* **2023**, *63*, 6183–6191. [[CrossRef](#)] [[PubMed](#)]
53. Tian, C.; Kasavajhala, K.; Belfon, K.; Raguette, L.; Huang, H.; Miguez, A.N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; et al. Ff19sb: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J. Chem. Theory Comput.* **2019**, *16*, 528–552. [[CrossRef](#)]
54. Salomón-Ferrer, R.; Götz, A.W.; Poole, D.; Grand, S.L.; Walker, R.C. Routine microsecond molecular dynamics simulations with Amber on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888. [[CrossRef](#)]
55. Izadi, S.; Anandkrishnan, R.; Onufriev, A.V. Building water models: A different approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871. [[CrossRef](#)]
56. Swails, J. ParmEd. 2024. Available online: <https://github.com/ParmEd/ParmEd> (accessed on 12 December 2024).
57. Sindhikara, D.J.; Kim, S.; Voter, A.F.; Roitberg, A.E. Bad seeds sprout perilous dynamics: Stochastic thermostat induced trajectory synchronization in biomolecules. *J. Chem. Theory Comput.* **2009**, *5*, 1624–1631. [[CrossRef](#)]

58. Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [[CrossRef](#)]
59. Poma, A.B.; Thu, T.T.M.; Tri, L.T.M.; Nguyen, H.L.; Li, M.S. Nanomechanical stability of A β tetramers and fibril-like structures: Molecular dynamics simulations. *J. Phys. Chem. B* **2021**, *125*, 7628–7637. [[CrossRef](#)]
60. Olivos-Ramirez, G.E.; Poma, A.; Cofas-Vargas, L.F. Conformational dynamics and stability analysis of SARS-CoV-2 Spike glycoprotein. *Zenodo* **2025**. [[CrossRef](#)]
61. Arbeitman, C.R.; Rojas, P.A.; Ojeda-May, P.; Garcia, M.E. The SARS-CoV-2 spike protein is vulnerable to moderate electric fields. *Nat. Commun.* **2021**, *12*, 5407. [[CrossRef](#)]
62. Olukitibi, T.A.; Ao, Z.; Warner, B.M.; Unat, R.; Kobasa, D.; Yao, X. Significance of conserved regions in coronavirus spike protein for developing a novel vaccine against SARS-CoV-2 infection. *Vaccines* **2023**, *11*, 545. [[CrossRef](#)] [[PubMed](#)]
63. Pondé, R.A.d.A. Physicochemical effect of the N501Y, E484K/Q, K417N/T, L452R and T478K mutations on the SARS-CoV-2 spike protein RBD and its influence on agent fitness and on attributes developed by emerging variants of concern. *Virology* **2022**, *572*, 44–54. [[CrossRef](#)] [[PubMed](#)]
64. Goher, S.S.; Ali, F.; Amin, M. The delta variant mutations in the receptor binding domain of SARS-CoV-2 show enhanced electrostatic interactions with the ACE2. *Med. Drug Discov.* **2022**, *13*, 100114. [[CrossRef](#)]
65. Zaheer, M.; Ali, N.S.; Javed, H.; Munir, R.; Jamil, N. Uncovering the impact of SARS-CoV-2 spike protein variants on human receptors: A molecular dynamics docking and simulation approach. *J. Infect. Public Health* **2023**, *16*, 1544–1555. [[CrossRef](#)]
66. Cheng, Y.; Zheng, D.; Zhang, D.; Du, G.; Wang, Y.; Liu, W.; Liang, L.; Hu, J.; Luo, T. Molecular recognition of SARS-CoV-2 spike protein with three essential partners: Exploring possible immune escape mechanisms of viral mutants. *J. Mol. Model.* **2023**, *29*, 109. [[CrossRef](#)] [[PubMed](#)]
67. Wang, W.B.; Yang, B.; Lei, Z.H.; Zhang, X.F.; Li, J.; Li, S.S.; Dong, Z.; Liang, Y.; Li, Q.M.; Su, J.G. Identification of key mutations responsible for the enhancement of receptor-binding affinity and immune escape of SARS-CoV-2 omicron variant. *J. Mol. Graph. Model.* **2023**, *124*, 108540. [[CrossRef](#)]
68. Božič, A.; Podgornik, R. Changes in total charge on spike protein of SARS-CoV-2 in emerging lineages. *Bioinform. Adv.* **2024**, *4*, vbae053. [[CrossRef](#)]
69. Dokainish, H.; 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**, *11*, e75720. [[CrossRef](#)]
70. Wu, Y.; Qian, R.; Yang, Y.; Sheng, Y.; Li, W.; Wang, W. Activation pathways and free energy landscapes of the SARS-CoV-2 spike protein. *ACS Omega* **2021**, *6*, 23432–23441. [[CrossRef](#)]
71. Dommer, A.C.; Casalino, L.; Kearns, F.L.; Rosenfeld, M.A.; Wauer, N.; Ahn, S.; Russo, J.D.; Oliveira, A.S.F.; Morris, C.K.; Bogetti, A.T.; et al. #COVIDisairborne: AI-enabled multiscale computational microscopy of delta SARS-CoV-2 in a respiratory aerosol. *Int. J. High Perform. Comput. Appl.* **2022**, *37*, 28–44. [[CrossRef](#)]
72. Springer, D.N.; Camp, J.V.; Aberle, S.W.; Deutsch, J.; Lammel, O.; Weseslindtner, L.; Stiasny, K.; Aberle, J.H. Neutralization of SARS-CoV-2 omicron XBB.1.5 and JN.1 variants after COVID-19 booster-vaccination and infection. *J. Med. Virol.* **2024**, *96*, e29801. [[CrossRef](#)] [[PubMed](#)]
73. Yang, H.; Guo, H.; Wang, A.; Cao, L.; Fan, Q.; Jiang, J.; Wang, M.; Lin, L.; Ge, X.; Wang, H.; et al. Structural basis for the evolution and antibody evasion of SARS-CoV-2 BA.2.86 and JN.1 subvariants. *Nat. Commun.* **2024**, *15*, 7715. [[CrossRef](#)]
74. Tian, J.; Shang, B.; Zhang, J.; Guo, Y.; Li, M.; Hu, Y.; Bai, D.; She, J.; Han, Y.; Guo, P.; et al. T cell immune evasion by SARS-CoV-2 JN.1 escapees targeting two cytotoxic T cell epitope hotspots. *Nat. Immunol.* **2025**, *26*, 265–278. [[CrossRef](#)] [[PubMed](#)]
75. Sharma, A.; Maurya, S.; Tripathi, T.; Padhi, A.K. Integrated all-atom and coarse-grained simulations uncover structural, dynamics and energetic shifts in SARS-CoV-2 JN.1 and BA.2.86 variants. *Acta Trop.* **2024**, *260*, 107444. [[CrossRef](#)]
76. Paciello, I.; Maccari, G.; Pierleoni, G.; Perrone, F.; Realini, G.; Troisi, M.; Anichini, G.; Cusi, M.G.; Rappuoli, R.; Andreano, E. SARS-CoV-2 JN.1 variant evasion of IGHV3-53/3-66 B cell germlines. *Sci. Immunol.* **2024**, *9*, eadp9279. [[CrossRef](#)]
77. Raisinghani, N.; Alshahrani, M.; Gupta, G.; Verkhivker, G.M. AlphaFold2 modeling and molecular dynamics simulations of the conformational ensembles for the SARS-CoV-2 spike omicron JN.1, KP.2 and KP.3 variants: Mutational profiling of binding energetics reveals epistatic drivers of the ACE2 affinity and escape hotspots of antibody resistance. *Viruses* **2024**, *16*, 1458. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.