



Structural implications of SARS-CoV-2 Surface Glycoprotein N501Y mutation within receptor-binding domain [499-505] – computational analysis of the most frequent Asn501 polar uncharged amino acid mutations

Done Stojanov

To cite this article: Done Stojanov (2023) Structural implications of SARS-CoV-2 Surface Glycoprotein N501Y mutation within receptor-binding domain [499-505] – computational analysis of the most frequent Asn501 polar uncharged amino acid mutations, *Biotechnology & Biotechnological Equipment*, 37:1, 2206492, DOI: [10.1080/13102818.2023.2206492](https://doi.org/10.1080/13102818.2023.2206492)

To link to this article: <https://doi.org/10.1080/13102818.2023.2206492>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 04 May 2023.



[Submit your article to this journal](#)



Article views: 655



[View related articles](#)



[View Crossmark data](#)

Structural implications of SARS-CoV-2 Surface Glycoprotein N501Y mutation within receptor-binding domain [499-505] – computational analysis of the most frequent Asn501 polar uncharged amino acid mutations

Done Stojanov

Department of Computer Technologies and Intelligent Systems, Faculty of Computer Science, Goce Delčev University of Štip, Štip, North Macedonia

ABSTRACT

The aim of this study was to evaluate the impact of the most frequent Asn501 polar uncharged amino acid mutations upon important structural properties of SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) Surface Glycoprotein RBD – hACE2 (human angiotensin-converting enzyme 2) heterodimer. Mutations N501Y, N501T and N501S were considered and their impact upon complex solubility, secondary motifs formation and intermolecular hydrogen bonding interface was analyzed. Results and findings are reported based on 50ns run in Gromacs molecular dynamics simulation software. Special attention is paid on the biomechanical shifts in the receptor-binding domain (RBD) [499-505]: ProThrAsn(Tyr)GlyValGlyTyr, having substituted Asparagine to Tyrosine at position 501. The main findings indicate that the N501S mutation increases SARS-CoV-2S-protein RBD – hACE2 solubility over N501T, N501 (wild type): $SASA_{N501S} = 364.66 \pm 3.2842 \text{ nm}^2$, $SASA_{N501T} = 360 \pm 3.4156 \text{ nm}^2$, $SASA_{N501} = 359.56 \pm 3.6473 \text{ nm}^2$. The N501Y mutation shifts α -helix S-protein RBD [366-370]: SerValLeuTyrAsn into π -helix for $t > 38.5 \text{ ns}$. An S-protein RBD [503-505]: ValGlyTyr shift from 3_{10} -helix into a turn is observed due to the N501Y mutation in $t > 33 \text{ ns}$. An empirical proof for the presence of a Y501-binding pocket, based on RBD [499-505]: PTYGVGY C's RMSF peak formation is presented. There is enhanced electrostatic interaction between Tyr505 (RBD) phenolic -OH group and Glu37 (hACE2) side chain oxygen atoms due to the N501Y mutation. The N501Y mutation shifts the $\text{Gly502}_{S\text{-Protein}}\text{-N-H}\dots\text{O}=\text{CLys353}_{h\text{ACE2}}$ hydrogen bond into permanent polar contact; $E(\text{Gly502}_{S\text{-Protein}}\text{-N-H}\dots\text{O}=\text{CLys353}_{h\text{ACE2}})_{N501Y} = 4.816210332 \text{ kcal/mol}$; $E(\text{Gly502}_{S\text{-Protein}}\text{-N-H}\dots\text{O}=\text{CLys353}_{h\text{ACE2}})_{N501} = 3.909532232 \text{ kcal/mol}$.

Abbreviations: SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; S-protein/S-glycoprotein: surface/spike glycoprotein; RBD: receptor-binding domain; ACE2: Angiotensin-converting enzyme 2; hACE2: human Angiotensin-converting enzyme 2; VOC: variant of concern; VOI: variant of interest; PDB: Protein Data Bank; PME: Particle Mesh Ewald; NVT: Canonical Ensemble; NPT: isothermal–isobaric ensemble; SPC: simple point-charge; MD: molecular dynamics; RMSD: root-mean-square deviation; RMSF: root mean square fluctuations; SASA: solvent accessible surface area; DCLM: Double Lattice Method; DSSP: Dictionary of protein secondary structure; E: energy; h-bond: hydrogen bond; r_{DA} : donor-acceptor distance; $\angle HDA$: hydrogen-donor-acceptor angle; GISAD: Global Initiative on Sharing Avian Influenza Data.

ARTICLE HISTORY

Received 7 March 2023
Accepted 19 April 2023



KEYWORDS


SARS-CoV-2; N501Y; mutation; COVID-19; surface glycoprotein; hydrogen bonds

Introduction

Since the coronavirus disease (COVID-19) outbreak in 2019, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genome has undergone hundreds of mutations. Virus mutations arise causally during

replication, due to the lack of RNA polymerase proof-reading activity (1). Although most SARS-CoV-2 mutations have no significant impact upon the general course of COVID-19 (2), specific mutations may increase the infection severity and undermine current vaccination programs. The mutations occurring in

CONTACT Done Stojanov  done.stojanov@ugd.edu.mk  Faculty of Computer Science, Goce Delčev University of Štip, Krste Misirkov No.10-A P.O. Box 201, Štip 2000, North Macedonia.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/13102818.2023.2206492>.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

SARS-CoV-2 S-protein have attracted special attention, as the virus targets host cells by S-protein-to-hACE2 receptor interaction (3). It has been well-studied that the receptor-binding domain (RBD) of the S-glycoprotein serves as a binding interface to the human Angiotensin-converting enzyme 2 (hACE2) receptor, extracting kinetic and thermodynamic properties of the binding pocket (3). So far, SARS-CoV-2 has evolved into several VOCs (variants of concern) and VOIs (variants of interest) that may be seen as a direct consequence of the occurrence of specific mutations in the S-protein that resulted in increased infection severity and lucid immune escape, undermining current COVID-19 vaccine development efforts.

The alpha lineage, or B.1.1.7 variant, arose in the United Kingdom in September 2020 (4) and included 7 amino acid substitutions: N501Y, A570D, D614G, P681H, T716I, S982A, D1118H, and two deletions: HV69/70del and Y144del, in the S-protein (5). Substitutions D614G and N501Y were also found in the B.1.351 variant (beta lineage), in addition to: L18F, D80A, D215G, LAL242/244del, R246I, K417N, E484K, N501Y, D614G and A701V S protein mutations (5). The Gamma variant (P1) accumulated totally 12 S-protein substitutions: L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I and V1176F (5). The S-protein of the Delta variant (B.1.617.2) was characterized by 8 non-synonymous substitutions and 1 deletion: T19R, G142D, FR156/157del, R158G, L452R, T478K, D614G, P681R, D950N (6). The recent circulating variant of concern – Omicron (B.1.1.529) – has over 30 mutations in the S-protein (7), some of them: K417N, T478K, N501Y found in common to previous variants (8,9).

Several variants, such as: B.1.1.7, B.1.351 and P.1 share a common mutation N501Y located in the RBD of SARS-CoV-2 spike (S) protein, assumed to be of higher infectivity to humans (10,11). Although a study has discovered that the N501Y mutation diminishes the binding affinity to an animal ACE2 receptor (12), N501Y enhances viral binding to the human ACE2 receptor (hACE2) and confers moderate resistance to certain monoclonal antibodies (12). Liu et al. (10) regards the N501Y mutation as a hot spot for enhanced infection of the upper airway and transmission. Many studies have tried to explain the pathogenic relevance of the N501Y mutation on a biomechanical basis, but there is evident disagreement among the reported findings. Having a Y501 aromatic ring surrounded between residues hACE2^{K353} and hACE2^{Y41}, which leads to a better-formatted binding pocket (13), the interactions of Y501 to hACE2^{K353} and π - π stacking of the Y501-hACE2^{Y41} were provided as main arguments

explaining the N501Y increased viral fitness for replication (10–12, 14). On the other hand, it has been also found that Y501 increases the unbinding force towards the hACE2 receptor compared to N501 wild type (11). A recent study hypothesizes kinases (the epidermal growth factor receptor, EGFR) as possible factors that involve N501Y mutation binding through Y501 phosphorylation (15).

The aim of this study was to analyze the structural changes caused by the most frequent N501 to polar uncharged amino acids substitutions: N501Y, N501T and N501S in terms of hACE2–S-protein complex stability, solubility, secondary structure organization and intermolecular polar interactions. The change of polar contacts' affinities of 501-surrounding residues: P499, T500, G502, V503, G504, Y505, due to N501Y mutation, was also analyzed.

Materials and methods

Heterodimer – PDB file 6M0J: Crystal structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2 (<https://www.rcsb.org/structure/6m0j>), (16) was used as the wild type structure. The file was retrieved from Protein Data Bank (<https://www.rcsb.org/>). Zinc (Zn²⁺) metal cation bound to hACE2 characterized the receptor environment. Zn²⁺ ion contacted hACE2 residues His374/His378 and Glu375/Glu402; thus, promoting local stability. A total of five N-acetyl-beta-D-glucosamine (NAG) ligands were attached to the 6M0J hetero dimer. Four NAG glycans were attached to the hACE2 receptor, one NAG glycan bound to S-protein RBD. NAG glycans formed frequent contacts with Asparagine on the opposite side, such as: hACE2 Asn90, hACE2 Asn322 and RBD Asn343. No metal ion, no ligands were attached to the mutagenesis domain of interest in this study, S-protein RBD [499-505]: PTN(Y)GVGY. The wild-type complex with Asparagine at position 501 in RBD was mutated to Tyrosine(Y), Threonine(T) and Serine(S) using PyMol molecular graphics system (<https://pymol.org/2/>, ver. 2.5.4) mutagenesis object.

All four molecular systems – wild-type Asparagine501(N501) and mutants, Tyrosine501(Y501), Threonine501(T501) and Serine501(S501) – were prepared for molecular dynamics simulation in Gromacs (<https://www.gromacs.org/>, ver. 2020.1-Ubuntu-2020.1-1), (17). Topology files (.top) were compiled from .pdb files to have each molecule defined as object suitable for MD simulation, including standard properties, such as: atom types, charges, bonds, dihedrals; (command `pdb2gmx`). Charmm27 (18) all-atom force field was selected for simulation purposes. Long-range energy contributions were

calculated by the PME method in Fourier space (particle mesh Ewald), (19,20), with a real-space cutoff for PME calculations of 1 nm. Molecules were placed and centered in a cubic grid box at 1.0 nm minimum distance from the edge of solute-box, providing minimum 2.0 nm between periodic images of the protein.

As SARS-CoV-2 (virus)-to-host cell membrane fusion may occur in blood plasma (21) at physiological pH or near-physiological pH (22), and plasma is predominantly made of water (about 90%), water fits the simulation purpose as a solvent model. An SPC/E (simple point-charge/extended) water model under an SPC216 generic three-point solvent model (Brendsen, 1987) was used as a solvent. In total, 25 SPC/E water molecules were substituted with Na^+ ions to neutralize the native total negative charge in the system of $-25.000 e$. Refined geometry of structure, appropriate water molecules orientation, avoiding steric clashes, was provided through an energy minimization process applying the steepest descent minimization algorithm (23). Tuning maximum force $F_{\max} < 1000 \text{ kJmol}^{-1} \text{ nm}^{-1}$ required $1000 < \text{number of steps} < 2000$ energy optimization steps, for potential energy $E_{\text{pot}} < -10^5 \text{ kJmol}^{-1}$.

The systems were brought to desired temperature and pressure by NVT, followed by NPT-equilibration, with a coupling time of 2 ps. Then, 150-ps isothermal-isochoric NVT position-restrained equilibration at 310 K controlled by V-rescale thermostat (24) was performed. In this stage, pressure coupling was turned off, while particles initial velocities were assigned from Maxwell distribution at 310 K. To finalize the system equilibration, 150-ps isothermal-isobaric pressure coupling, regulated with Berendsen barostat (25) was performed. Referent coupling pressure of 1 bar and water isothermal compressibility of $4.45 \times 10^{-5} \text{ bar}^{-1}$ at NVT equilibrium of 310 K were used. Verlet (26) velocity algorithm was used to integrate Newton's motion equations. The equilibration process aimed to relax the complex – optimize the water molecules towards the heterodimer. The wild-type system (N501) and mutants (Y501, T501 and S501) were subjected to 50 ns md Gromacs simulation. Output trajectory, index files were used for the analysis.

Backbone dynamics and molecular dynamics convergence analysis

Molecular dynamics equilibrium (convergence) state and systems stability shifts were examined by the means of root mean square deviation (RMSD) analysis upon backbone atoms. By measuring the magnitude of atomic oscillations during the simulation ($t = 0 - 50 \text{ ns}$) relative to the initial conformation at $t = 0$, one can

analyze the impact of induced mutations upon general system dynamics/stability, identify periods of substantial conformational changes and inspect MD simulation convergence. Deviations produced during the course of simulation (27) are taken as stability estimate: flat-RMSD plot for relatively stable dynamics, spike-oscillating RMSD plot for destabilized behavior. The molecular dynamics convergence state was identified based on uninterrupted backbone atoms' oscillations ranging no more than 0.15 nm (1.5 Å), Appendix *RMSD_analysis.xlsx*.

Residues fluctuations analysis

Some mutations may disturb the dynamic behavior of a particular protein domains, which is evaluated in terms of the root mean square fluctuations (RMSF). RMSF per residue evaluates the average deviation of each residue relative to the starting, energetically minimized structure (28). Altered RMSFs as a consequence of mutation localized in the spike protein domain found in close proximity to the hACE2 receptor may indicate no-change, lesser or intensified virus-to-receptor interaction. To account for such behavior, molecular RMSF per residues were analyzed. S-protein motifs such that the RMSF per residue equals at least 0.03 nm (0.3 Å) were considered as localized conformational changes in the presence of a particular amino acid substitution. Special attention has focused on the S-protein RBD motif [499-505]: *ProThrAsn(Tyr)GlyValGlyTyr*, which in the presence of Tyrosine(Y) at position 501 instead of Asparagine(N) and coupled to the hACE2 receptor, exhibits major destabilization relative to the wild type (N501), Appendix *RMSF_analysis.xlsx*.

Solubility analysis

Amino acid substitutions at same points may significantly alter the contacts with surrounding solvent molecules, which may be seen as a consequence upon protein stability and protein conformational shifts. Molecular conformational changes were examined by means of solvent accessible surface area (SASA). The most significant conformational alterations are expected to happen due to hydrophilic-to-hydrophobic amino acid mutations and vice versa. Substituting wild-type Asparagine (N501) to also hydrophilic (29) Threonine (T501), Serine (S501) and dual-nature Tyrosine (Y501), primarily classified as a hydrophobic residue because of the aromatic ring in the side chain, being also capable of hydrogen bond formation due

Table 1. Average RMSD (nm)/SASA (nm²) per molecule of N501[Y, T, S] heterodimers.

Heterodimer		N501	N501Y	N501S	N501T
RMSD(nm)	Avg.	0.223	0.266	0.205	0.214
	S.D.	0.031	0.047	0.032	0.04
SASA(nm²)	Avg.	359.56	364.93	364.66	360
	S.D.	3.6473	3.3889	3.2842	3.4156

*Avg, average values; S.D., standard deviation.

to hydroxyl group (-OH) in its side chain, may also imply molecular conformational changes.

As per definition, SASA is the total surface area of a molecule that is accessible to the solvent. Amino acids mutations affect the overall SASA score (nm²) in one or another way, which plays an important role in understanding the structure-to-function protein properties. The Double Lattice Method-DCLM (30) is used to compute SASA of the wildtype and mutants, Appendix *SASA_analysis.xlsx*.

Secondary structure analysis

The impact of induced mutations upon secondary structure changes was monitored by DSSP (dictionary of protein secondary structure) algorithm (31). DSSP predicts secondary structures based on known hydrogen bond patterns found in common secondary motifs, such as: α -helix, β -sheet. The algorithm computes electrostatic interaction energy $E \left[\frac{kcal}{mol} \right]$ as a function of interatomic distances between main chain carbonyl and main chain amino group: $r_{NO}, r_{CH}, r_{OH}, r_{CN}$, by placing partial charges: $q_{C,O} = -0.42e$ (main chain carbonyl oxygen is considered to be universal acceptor), $q_{N,H} = +0.2e$ (main chain amino group is considered to be universal hydrogen bond donor). The presence of hydrogen bond is confirmed if E is at least 0.5 kcal/mol (32).

Hydrogen bonds dynamics analysis

Hydrogen bonding between molecules plays a crucial role in molecular recognition, as it provides directionality and specificity of the interaction (33). As hydrogen bonding dictates protein folding that is crucial for protein stability, selective macromolecular interactions are governed by specific hydrogen bonding dynamics. In this study, Module *gmx hbond* was used to analyze the impact of induced mutations upon SARS-CoV-2 S-protein RBD – hACE2 intermolecular hydrogen bonding dynamics and identify major changes in interactions due to induced amino acid substitutions. Several important aspects are reviewed when considering the impact of a particular intermolecular hydrogen bond, such as: the occupancy

(continuity) of the bond and h-bond energy (strength). As the strength of a hydrogen bond depends on the electronegativity of engaged atoms (34, pp. 597), r_{DA} (donor-acceptor distance), $\angle HDA$ (hydrogen-donor-acceptor angle) and environmental factors, such as pressure and temperature (35), hydrogen bonds can be classified as very strong, strong or weak following categorization provided in Andrews et al. (34, Table 1, pp. 597).

To get deeper insight into the conformational alterations due to N501Y S-protein mutation and understand the essentials of structural and biomechanical implications, specific aspects of the dynamics of the intermolecular hydrogen bonding interface between SARS-CoV-2 S-protein RBD and hACE2 receptor were analyzed in comparison to N501 wild type, which was used a referent structure. Intermolecular hydrogen bond contacts between S-protein RBD domain [499-505]: *ProThrAsn(Tyr)GlyValGlyTyr* and hACE2 receptor were analyzed, due to the fact that Asparagine(N) to Tyrosine(Y) transition at position 501 implies major RBD motif destabilization. The electrostatic impact of a specific hydrogen bond was estimated in terms of the occupancy of a h-bond and donor-acceptor distance (r_{DA}) for bonds formed between atoms of the same electronegativity in the mutant (N501Y) and the wild type (N501), Appendix *H_bonds_analysis.xlsx*. Since residues typically change their positions during the course of simulation, which leads to dynamic formation/breaking of hydrogen bonds, the criteria for presence of a h-bond at instance t will satisfy (36): **1**) $r_{DA} \leq 0.35nm$ and **2**) $\angle HDA \leq 30^\circ$. An adjusted version of the equation used in (31) – Eq. 1 was used to calculate the energy (kcal/mol) of the intermolecular main chain hydrogen bond of the highest occupancy (*Gly502_{S-Protein}*)N – H...O = C(*Lys353_{hACE2}*) based on charmm27 all atoms force field (37) parameters **a**) and **b**), (file: charmm27.ff/aminoacids.rtp), Appendix *Gly502_Lys353_H-bond_Energy.xlsx*:

- $q_{C,O}$ mainchain carbonyl oxygen partial charge = -0,51 e;
- $q_{N,H}$ mainchain amino hydrogen partial charge = +0,31 e;
- $r_{NO,avg}, r_{CH,avg}, r_{OH,avg}, r_{CN,avg}$ average interatomic

- distance measured in angstroms (\AA , $1 \text{\AA} = 10^{-1} \text{ nm}$) for instances where a h-bond is present;
- d. $f = 332$ is a dimensional factor;

$$E = fq_{C,O}q_{N,H} \left(\frac{1}{r_{NO,avg}} + \frac{1}{r_{CH,avg}} - \frac{1}{r_{OH,avg}} - \frac{1}{r_{CN,avg}} \right) \left(\frac{kcal}{mol} \right) Eq.(1)$$

Results

Of all variants under consideration in the study – N501Y, N501S, N501T and N501(wild type) – the S-protein N501Y mutation implied a major backbone shift, relative to the starting, energetically minimized state at $t = 0$ (Figure 1, Table 1). N501Y and N501T heterodimers exhibited highest backbone deviation relative to the average RMSD score: $N501Y_{St.Dev} = 0.047 \text{ nm}$, $N501T_{St.Dev} = 0.04 \text{ nm}$ (Table 1). The variants experienced peak backbone deviations at different time stamps during the simulation (Figure 1). The N501Y variant exhibited peak backbone oscillation of 0.49 nm at $t = 6.78 \text{ ns}$, N501S peak backbone oscillation of 0.34 nm at $t = 19.68 \text{ ns}$, N501T peak backbone oscillation of 0.42 nm at $t = 8.46 \text{ ns}$ and N501 peak backbone deviation of 0.35 nm at $t = 4.35 \text{ ns}$ (Figure 1). In terms of average backbone deviation, $t = [0-50] \text{ ns}$, only N501Y scored higher than N501 (Table 1). N501S/N501T average RMSD scored less than N501. Regardless of the difference in the magnitude of the backbone shifts, all four structures reached an equilibrium state at $t = 42.3 \text{ ns}$ (Figure 1). Throughout the phase $[42.3 - 50]$

ns backbone oscillations ranged no more than 0.15 nm , which indicates that an equilibrium point was reached (Figure 1). Reaching an equilibrium point confirms that the systems were well-prepared prior to the molecular dynamics simulation phase.

The Asparagine-to-Threonine transition at position 501 did not significantly affect the solvent accessibility surface area, and the scores in terms of complex solubility were very similar: $SASA_{N501T} = 360 \pm 3.4156 \text{ nm}^2$, $SASA_{N501} = 359.56 \pm 3.6473 \text{ nm}^2$ (Table 1 and Figure 2). Because Serine has a shorter side chain compared to Threonine, Serine exhibits higher polarity, expecting to contribute more favorably to protein solubility than Threonine. The obtained results support this suggestion: $SASA_{N501S} = 364.66 \pm 3.2842 \text{ nm}^2$ (Table 1 and Figure 2). Average SASA results scored almost the same for N501S and N501Y (Table 1). As can be observed in Figure 2, until 37.5 ns the N501S SASA score was even higher than N501Y, contributing more favorably to the complex solubility than any other hydrophilic amino acid (38).

Figure 3 shows the S-protein RBD C_{α} (alpha-carbon) root mean square fluctuation curves per residue for all four variants of interest (N501, N501Y, N501S and N501T), having coupled the SARS-CoV-2 S-protein RBD to the hACE2 receptor. The variants N501, N501S and N501T followed similar trends of RMSF curves, although a few domain exemptions were specific to one or another variant, such as S-protein RBD [367-373] *ValLeuTyrAsnSerAlaSer* where N501S and N501T variants' residues exhibited significantly lower C_{α} fluctuations

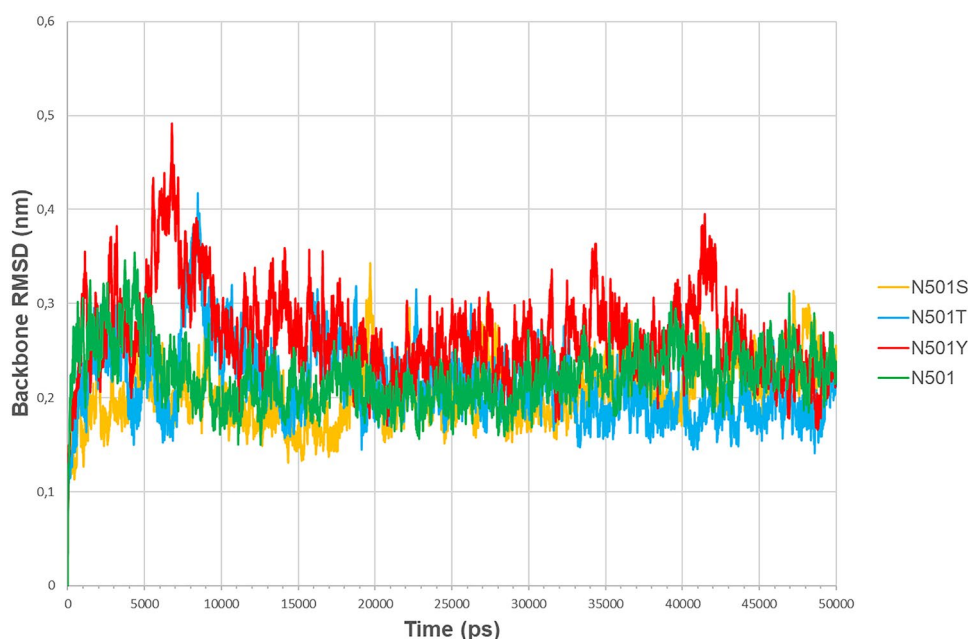
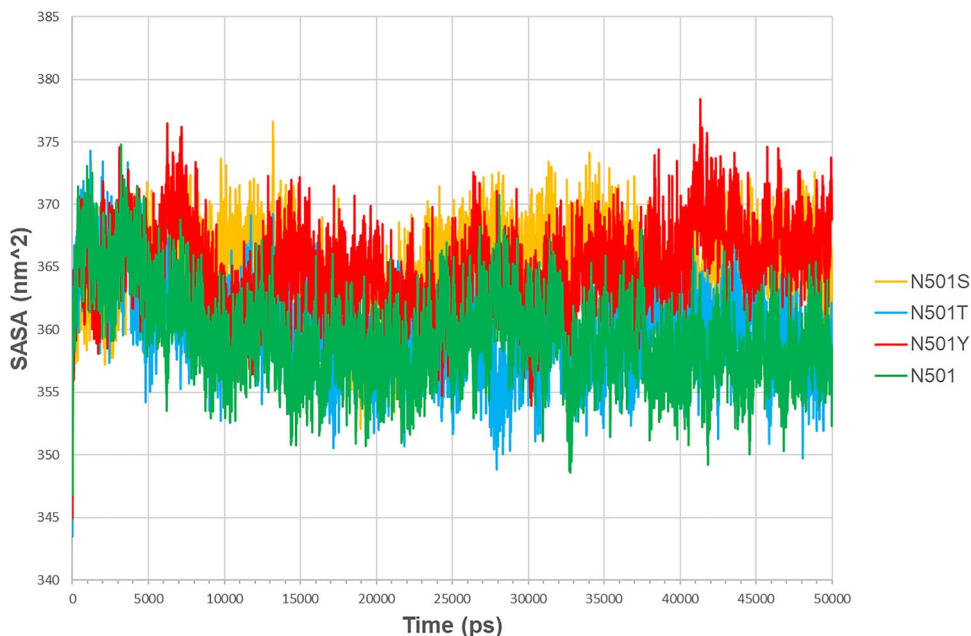
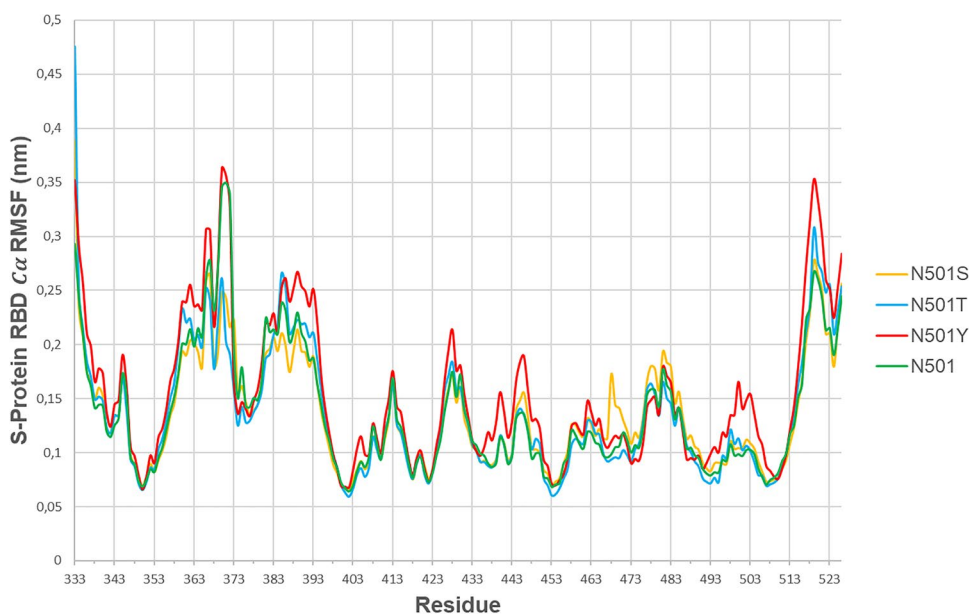


Figure 1. Backbone deviation analysis of N501[Y, T, S] hetero dimers (Gromacs *gmx rms* program).

Table 2. S-Protein RBD domain [499-505]: ProThrAsn(Tyr)GlyValGlyTyr RMSF analysis.

Position	499	500	501	502	503	504	505
Residue	Pro	Thr	Tyr/Asn	Gly	Val	Gly	Tyr
N501Y C_{α} RMSF (nm)	0.135	0.1657	0.1409	0.1488	0.1544	0.136	0.1159
N501T C_{α} RMSF (nm)	0.0978	0.0989	0.0973	0.1027	0.1029	0.0989	0.0858
Difference (nm)	0.0372	0.0668	0.0436	0.0461	0.0515	0.0371	0.0301

**Figure 2.** Solubility analysis of N501[Y, T, S] hetero dimers (Gromacs *gmx sasa* program).**Figure 3.** Mean C_{α} fluctuation per residue (S-protein RBD) (Gromacs *gmx rmsf* program).

than N501 and N501Y. On the other hand, the S-protein RBD domain [499-505] *ProThrTyrGlyValGlyTyr* in N501Y mutant exhibited unique behavior, as it tended to introduce higher C_{α} RMSF scores than any other variant of interest in this study (Figure 3). This indicates that

S-protein RBD motif [499-505] *ProThrTyrGlyValGlyTyr* deviation can be uniquely associated to the presence of Tyrosine (Y) at position 501, which results in local RBD pocket formation, nearby hACE2 Y41 (Figure 3). Table 2 shows the C_{α} RMSF scores for each amino acid

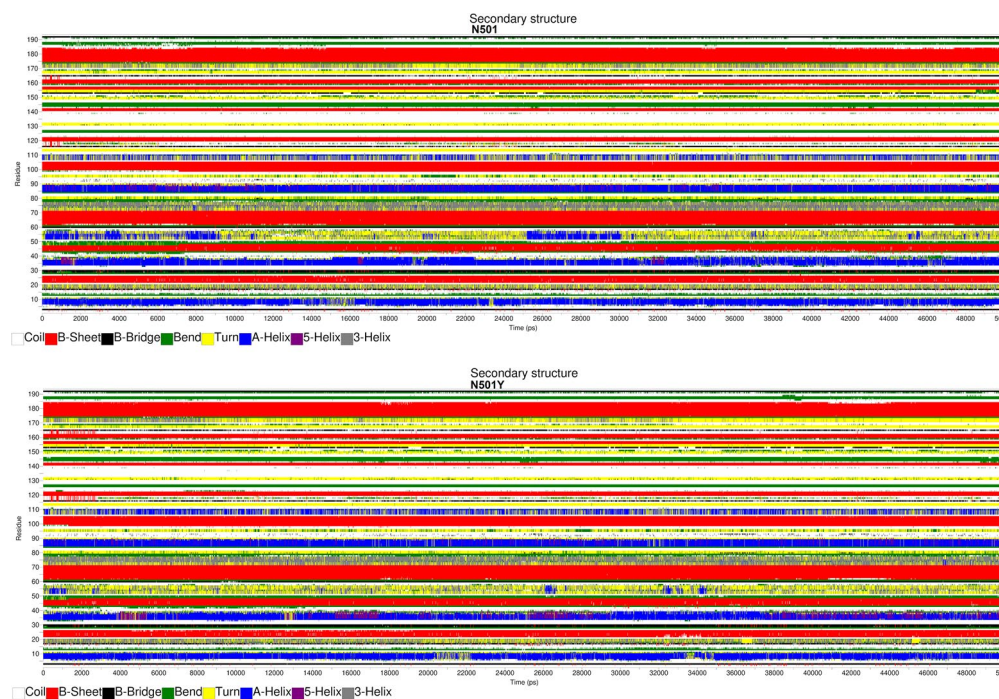


Figure 4. S-protein RBD secondary structure prediction (Gromacs *gmx do_dssp* module).

Note: Time is plotted on the horizontal axis (0–50 ns), while S-protein RBD residues in the range of 1–194 are shown on the vertical axis.

Table 3. Secondary structure assessment.

Range	S-protein RDB motif	N501 (wild type)	N501Y (mutant)
34–38 (Figure 4)	[366-370]: SerValLeuTyrAsn	Stable α -helix for $t > 38.5$ ns	Frequent α -helix to π -helix transitions for $t > 38.5$ ns
52–55 (Figure 4)	[384-387]: ProThrLysLeu	Turn to α -helix shifts during [0–9.25] ns and [25–30] ns	Mainly Turn
171–173 (Figure 4)	[503-505]: ValGlyTyr	3_0 -helix for $t > 33$ ns	Turn for $t > 33$ ns

in the domain of interest, for the N501(wild type)/N501Y mutation.

Figure 4 shows graphical outputs of *do_dssp* secondary structure analysis (31) for S-protein RBD in the case of N501 (wild type structure) and N501Y (mutant structure). The actual position p of every amino acid in the S-protein can be computed as: $p = i + 332$, such as i is the position of the corresponding amino acid on the vertical axis plotted on Figure 4. For example, the first RBD amino acid (**Thr**, $i = 1$) is at position 333 in the S-protein ($p = 333$), thus p is calculated as $p = 1 + 332 = 333$, the second RBD amino acid (**Asn**, $i = 2$) is at position 334 in the S-protein ($p = 334$) and is computed as $p = 2 + 332 = 334$ and so on.

Table 3 summarizes major RBD folding differences between N501 and N501Y dimers. S-protein RBD [366-370]: SerValLeuTyrAsn forms stable α -helix for $t > 38.5$ ns in N501, while frequent α -helix to π -helix transitions were observed in N501Y in $t > 38.5$ ns, Figure 4. RBD turn [384-387]: ProThrLysLeu was shifted to α -helix during [0-9.25] ns and [25-30] ns in N501,

while no turn to α -helix transitions were observed in N501Y, Figure 4. An interesting conformational change was observed in the domain of the interest S-protein RBD [499-505]: ProThrAsn(Tyr)GlyValGlyTyr. Namely S-protein RBD [503-505]: ValGlyTyr forms 3_0 -helix in N501 for $t > 33$ ns that seems to have been destabilized into a turn in N501Y in the same interval, Figure 4. In general, N501Y mutation has destabilized some of the most-stable secondary S-protein RBD conformations, such as the α -helix, Figure 4.

Figure 5 shows the number of intermolecular hydrogen bonds formed between S-protein RBD–hACE2 and S-protein RBD [499-505]–hACE2 for N501 and N501Y structures in $t = [0-50]$ ns. On average, during the simulation, the wild-type structure (N501) S-protein RBD formed a higher number of intermolecular polar contacts than N501Y. On the other hand, Tyrosine at position 501 increased the average number of h-bond interactions in domain [499-505] relative to the wild-type structure (Table 4 and Figure 5).



Figure 5. Hydrogen bonds dynamics in N501 and N501Y hetero dimers (Gromacs *gmx hbond* program).

Note: red, number of intermolecular hydrogen bonds in N501Y complex during [0–50] ns MD simulation; green, number of intermolecular hydrogen bonds in N501 complex during [0–50] ns MD simulation.

Table 4. Average number of hydrogen bonds formed in [0–50] ns, wild-type or mutant S-protein RBD [499-505]: *ProThrAsn(Tyr)GlyValGlyTyr*.

Domain	Number of intermolecular h-bonds, $t=[0-50]$ ns Average \pm S.D.*	
	S-protein–hACE2	S-Protein [499-505]–hACE2
Wild type (N501)	7.378 \pm 1.801	2.186 \pm 0.759
Mutant (N501Y)	5.735 \pm 1.323	2.375 \pm 0.71

*SD, standard deviation.

The intermolecular hydrogen bonds formed between S-protein RBD [499-505]: *ProThrAsn(Tyr)GlyValGlyTyr* and hACE2 receptor were subject of special analysis (Tables 5 and 6). Hydrogen bonds were identified as **DonorHydrogen...Acceptor** atomic structures in N501 and N501Y heterodimers. The impact of a hydrogen bond was measured by means of *occupancy* – life span of h-bond in % and r_{DA} (donor-acceptor distance, measured in nm). The prime aim was to identify major shifts in intermolecular polar contacts due to the

presence of Tyrosine at position 501 instead of Asparagine (Tables 5 and 6).

The obtained results (Tables 5 and 6) showed that the N501Y mutation shifts the intermolecular hydrogen bond $Gly502_{S-Protein} N-H \dots O=C Lys353_{hACE2}$ to permanent polar contact, stronger than the one observed in the wild type. In the N501 variant, a Gly502 to Lys353 main chain intermolecular h-bond was observed 79.5% of the time (83.1% occupancy during convergence period [42.3-50] ns) and it was

shifted to 98.4% occupancy (97.9% occupancy during [42.3-50] ns) in N501Y. A stronger h-bond $Gly502_{S-Protein} N-H \dots O = CLys353_{hACE2}$; $r_{NO} = 0.2909 \pm 0.0157 \text{ nm}$ in $t = [0-50]$ ns ($r_{NO} = 0.2915 \pm 0.0162 \text{ nm}$ in the convergence phase, for $t = [42.3-50]$ ns), compared to $r_{NO} = 0.3031 \pm 0.0195 \text{ nm}$ ($r_{NO} = 0.3061 \pm 0.0193 \text{ nm}$) in the wild type, was observed in the presence of the N501Y mutation.

The N501Y mutation also influenced the $Tyr505_{S-Protein} \dots Glu37_{hACE2}$ intermolecular hydrogen bond, resulting in much stronger and more persistent polar contact (Tables 5 and 6). The H-bond occupancy and strength significantly increased during the convergence phase, [42.3-50] ns: $(N501Y)r_{OO} = 0.2812 \pm 0.0208 \text{ nm}$ versus $(N501)r_{OO} = 0.3133 \pm 0.0242 \text{ nm}$, having increased contact occupancy for 10.6% relative to N501.

In the course of simulation, prior to position 501, Thr500 (S-protein) in N501 formed a highly persistent h-bond with Asp355 (hACE2), 86.8% occupancy in N501 versus 66.5% occupancy in N501Y (Tables 5 and 6). In terms of strength, the h-bond scored almost equally, $r_{OO} \approx 0.27 \text{ nm}$ in [0-50] ns ([42.3-50] ns). During the convergence state ([42.3-50] ns), the N501Y structure intensified the h-bond occupancy (88.8%), but still under 98.4% of occupancy in N501. However,

Thr500 (S-protein) also formed periodic contacts to Tyr41 (hACE2), by means of the same Thr500 side chain oxygen atom (acting as a donor), when not engaged in an h-bond to Asp355 (hACE2) (Table 6, Figure 6). Much of the time when no h-bond existed between Thr500 (S-protein) and Asp355 (hACE2), the N501Y structure favored hydrogen bond formation between the Thr500 (S-protein) side chain hydroxyl group and the Tyr41 (hACE2) phenolic oxygen (Figure 6), compensating for the less frequent polar contact to Asp355 in N501Y. The h-bond $THR500_{S-Protein} O-H \dots O(H)Tyr41_{hACE2}$ was highly intensified during the convergence state in N501Y: $(N501Y)r_{OO} = 0.2812 \pm 0.0208 \text{ nm}$ against $(N501)r_{OO} = 0.3133 \pm 0.0242 \text{ nm}$ in N501 (Tables 5 and 6).

Neither Asparagine, nor Tyrosine at position 501, formed highly persistent intermolecular hydrogen bonds. Only temporal polar contacts to Tyr41, Lys 353, Asp 355 in N501 and Asp38, Lys353 in N501Y complex were observed (Tables 5 and 6).

Binding energy shift (kcal/mol) of the most influenced hydrogen bond due to the N501Y mutation $Gly502_{S-Protein} N-H \dots O = CLys353_{hACE2}$ was computed applying Eq. 1 (Table 7). There was an increased binding affinity for 0.91 kcal/mol during the simulation and 1.06 kcal/mol during convergence state.

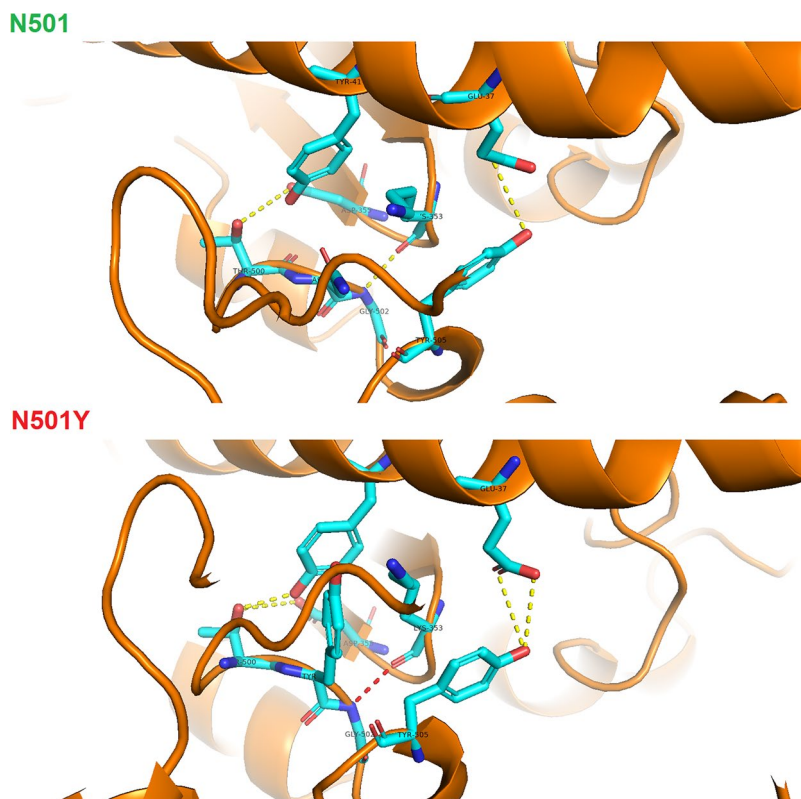


Figure 6. Visualization of hydrogen bonds between hACE2 and SARS-CoV-2 spike protein – N501 against N501Y heterodimer drawn in PyMol molecular graphics program.

Note: red, high occupancy of a hydrogen bond; yellow, average occupancy of a hydrogen bond.

Table 7. $Gly502_{S-Protein} N-H...O=C Lys353_{hACE2}$ energy (kcal/mol) in N501 and N501Y heterodimers.

$Gly502_{S-Protein} N-H...O=C Lys353_{hACE2}$	$r_{ON,avg}(\text{\AA})$	$r_{CH,a}(\text{\AA})$	$r_{OH,avg}(\text{\AA})$	$r_{CN,avg}(\text{\AA})$	Energy (kcal/mol)
[0-50] ns wild type (N501)	3.03053307	3.13261252	2.08334171	4.10371134	3.909532232
[0-50] ns mutant (N501Y)	2.9092403	3.06141174	1.94696323	4.02390412	4.816210332
[42.3-50] ns wild type (N501)	3.06087363	3.16656786	2.11826833	4.13800312	3.739474518
[42.3-50] ns mutant (N501Y)	2.91512583	3.06757616	1.95074172	4.03357616	4.803593311

Note: Eq. 1 was applied to compute hydrogen bond energy.

Discussion

As of March 3rd 2023, following GISAID database - EpiCoV module (<https://gisaid.org/>), the N501Y SARS-CoV-2 S-protein mutation was reported in 8 406 386 SARS-CoV-2 isolates worldwide, N501T in 6320 and N501S in 750. The primary aim of this study was to examine the impact of the three most frequent Asn501 polar uncharged amino acid mutations: N501Y, N501T and N501S in contact to the hACE2 receptor.

The findings from this study indicate that the presence of Tyrosine at position 501 in SARS-CoV-2 S-protein RBD implies major destabilization of the complex, which was proved in terms of recorded backbone oscillations during the course of simulation (0.49 nm peak amplitude, Figure 1). After 42.3 ns, all variants of interest reached convergence state, which confirmed the appropriate preparation of the structures (Figure 1).

Highly polar Serine at position 501 favors protein solubility over N501 and N501T complexes (Table 1), exhibiting approximately the same solvent accessibility surface area as N501Y: $SASA_{N501Y} = 364.93 \pm 3.3889 nm^2$, $SASA_{N501S} = 364.66 \pm 3.2842 nm^2$. From the detailed h-bonds analysis of S-protein RBD [499-500] (Table 6), the Tyr501 phenolic hydroxyl group serves as a donor complex to Asp38 (hACE2) in 3.5% of the time (0% during convergence state), while the Tyr501 phenolic oxygen atom acts as an acceptor in 2% of the time (2.2% during convergence state). This indicates that most of the time the Tyr501 phenolic -OH group formed polar contacts with surrounding solvent molecules, thus increasing the complex solubility.

Sveral S-protein RBD domains were highly destabilized in the presence of Tyr501 (Figure 3 and Table 2). Of all highly destabilized domains, special attention has been put on the S-protein RBD [499-505]: ProThrAsn(Tyr)GlyValGlyTyr domain, which includes a mutational spot. Following secondary structure assessment analysis (Figure 4 and Table 3), the N501Y mutation shifts the S-protein RBD [503-505]: ValGlyTyr 3_{10} -helix to turn, increasing the local flexibility after 33 ns. Two other domains were also destabilized in the

presence of Tyr501, shifting highly-stable α -helix to π -helix, α -helix to turn at some points (Table 3).

Hydrogen bond analysis of the RBD [499-505]: ProThrAsn(Tyr)GlyValGlyTyr region showed that some intermolecular polar contacts were intensified due to Tyr501 (Tables 5 and 6). Tyr501 shifts $Gly502_{S-Protein} N-H...O=C Lys353_{hACE2}$ into a permanent polar contact (Table 6), with an increased strength relative to the wild type: 3.909532232 kcal/mol in N501 versus 4.816210332 kcal/mol in N501Y during the course of simulation (3.739474518 kcal/mol versus 4.803593311 kcal/mol during convergence state) (Table 7). The computed results [kcal/mol] follow experimentally determined N-H...O=C intermolecular h-bond energy, which may range within [4 - 6] kcal/mol (39). While the wild-type Tyr505 phenolic -OH group participated in single h-bond formation (Table 5), the N501Y Tyr505 -OH group periodically formed polar contacts with two Glu37 side chain oxygen atoms, enhancing the interaction between Tyr505 and Glu37 (Tables 5 and 6, Figure 6). Although the hydrogen bond formed between Thr500 and Asp355 seems to be of lower occupancy in N501Y relative to the wild type, it was supplemented by frequent polar contacts to Tyr41 that were highly intensified during convergence state (Tables 5 and 6, Figure 6).

Other studies have also investigated the impact of this crucial N501Y mutation (11, 13,14). However, most of them only mention the occurrence of a specific hydrogen bond, without providing further details concerning the dynamics of a hydrogen bond, which is crucial for understanding the biomechanical implications of the N501Y substitution. Some studies (11, 14) explain that N501Y increased the SARS-CoV-2 infectivity due to Tyr501...Tyr41 π - π stacking interaction and Tyr501...Lys353 π -cation interaction. In terms of h-bond analysis, Tian et al. (11) have also confirmed that Thr500 forms two hydrogen bonds with Tyr41, Asp355 in the wild type, but no further details regarding the h-bonds dynamics were provided. Similarly, Thr500 to Asp355 polar connection was also confirmed by Rostami et al. (13).

The present study analyzed the hydrogen bonds dynamics in terms of occupancy and strength (estimated by r_{DA}) and revealed for the first time intensified Tyr505...Glu37 interaction and Gly502...Lys353 h-bond shift into permanent polar connection due to N501Y mutation. This study points to increased N501Y RBD [499-505] infectivity, nearby mutational spot 501, additionally boosting already confirmed Tyr501...Tyr41 π - π stacking interaction and Tyr501...Lys353 π -cation interaction.

Several Cryo-electron structural studies (40–42), evaluated the role of key spike protein mutations upon the hACE2 binding affinity. Experimental findings (40–42) are in line with in silico findings presented in this study, showing additional insights upon key acting residues.

Cryo-electron experiments (42) identified a pocket-binding structure in Y501- carrying variants, accommodating the Y501 aromatic ring in a sandwich between hACE2 Y41 and K353 residues. Following the findings from the present study, the spike protein RBD [499-505]: PTYGVGY motif nearby hACE2 Y41 is characterized by peak RMSF motif deviation only in the presence of Y501 – structural deviation, which actually leads to RBD pocket formation near the hACE2 Y41 binding interface.

The presence of a Y505_{S-Protein}...E37_{hACE2} hydrogen bond was identified by Saville et al. (41). The present study also confirm the existence of intermolecular polar contact between Y505 and E37 that has been intensified in terms of occupancy and strength, shifting the weak wild-type Y505_{S-Protein}...E37_{hACE2} hydrogen bond ($r_{DA} = 0.3133 \pm 0.0242 \text{ nm}$) into a stronger one ($r_{DA} = 0.2812 \pm 0.0208 \text{ nm}$) in the presence of Y501.

As the spike protein RBD [499-505]: PTYGVGY domain remains unchanged in PDB structures: 7SY2, 7MJN, 7SY4, 7SY6, the molecular dynamics findings from the present study are also applicable to the previous Y501-Cryo-EM bearing structures.

Many studies, (11, 14, 43) try to explain the impact of the N501Y mutation by considering the change of intermolecular contacts at position 501 only. The present study evaluates the change of contacts' affinities of 501-surrounding residues: P499, T500, G502, V503, G504, Y505, having substituted Asparagine to Tyrosine at position 501, which is a novel finding in the present research.

Despite of the well-documented findings for Y501-K353 π -cation interaction and Y501...Y41 π - π stacking interaction (11, 14) that result in increased binding affinity due to the formation of the Y501-binding pocket (43), the present study reveals that the binding affinity is additionally increased by

formation of a permanent and stronger G502-K353 hydrogen bond in the presence of Y501. Following the findings from the present study, Y501 also intensifies the Y505-E37 polar interaction. The present study has also found that two stable RBD α -helices are destabilized into a π -helix and turn respectively in the presence of Y501. Finally, an empirical proof for the presence of the Y501-binding pocket was presented, estimated by means of C_{α} 's RMSF RBD [499-505]: PTYGVGY peak formation, which is adjacent to the hACE2 Y41 binding interface.

Conclusions

This study reveals some important biomechanical aspects of SARS-CoV-2 S-protein RBD-hACE2 dynamics upon Asn501 mutation. While the N501S mutation favors hetero-dimer solubility, the N501Y mutation tends to shift some stable RBD secondary motifs into more flexible ones, such as π -helix and turn. The N501Y mutation shifts the Gly502_{S-Protein} N-H...O = CLys353_{hACE2} hydrogen bond into permanent intermolecular polar contact, also favoring enhanced Tyr505...Glu37 hydrogen bond interaction. The N501Y variant Thr500 forms periodically polar contacts with two hACE2 amino acids: Tyr41 and Asp355, which may be also seen as a form of increased electrostatic attraction towards the hACE2 receptor due to the N501Y mutation.

Author contributions

Done Stojanov: conceptualization, methodology, simulation, validation, investigation, writing—original draft preparation and revision, visualization.

Data availability

The data that support the findings reported in this paper are available in the Appendix files: *RMSD_analysis.xlsx*; *RMSF_analysis.xlsx*; *SASA_analysis.xlsx*; *H_bonds_analysis.xlsx*; *Gly502_Lys353_H-bond_Energy.xlsx*

Disclosure statement

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

Funding

The author(s) reported there is no funding associated with the work featured in this article.

References

- [1] Fleischmann WR Jr. Viral genetics. In Baron S, editor. Medical microbiology. 4th ed. Galveston (TX): University of Texas Medical Branch at Galveston; 1996.
- [2] Chen J, Wang R, Wang M, et al. Mutations strengthened SARS-CoV-2 infectivity. *J Mol Biol.* 2020;432(19):5212–5226.
- [3] Yang J, Petitjean SJ, Koehler M, et al. Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nat Commun.* 2020;11(1):1–10.
- [4] Kim Y, Kim EJ, Lee SW, et al. Review of the early reports of the epidemiological characteristics of the B. 1.1. 7 variant of SARS-CoV-2 and its spread worldwide. *Osong Public Health Res Perspect.* 2021;12(3):139–148.
- [5] Gómez CE, Perdiguero B, Esteban M. Emerging SARS-CoV-2 variants and impact in global vaccination programs against SARS-CoV-2/COVID-19. *Vaccines.* 2021;9(3):243.
- [6] Harvey WT, Carabelli AM, Jackson B, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol.* 2021;19(7):409–424.
- [7] Subramoney K, Mtileni N, Bharuthram A, et al. Identification of SARS-CoV-2 omicron variant using spike gene target failure and genotyping assays, Gauteng, South Africa, 2021. *J Med Virol.* 2022;94(8):3676–3684.
- [8] Nguyen HL, Thai NQ, Nguyen PH, et al. SARS-CoV-2 omicron variant binds to human cells more strongly than the wild type: evidence from molecular dynamics simulation. *J Phys Chem B.* 2022;126(25):4669–4678.
- [9] Ou J, Lan W, Wu X, et al. Tracking SARS-CoV-2 omicron diverse spike gene mutations identifies multiple inter-variant recombination events. *Sig Transduct Target Ther.* 2022;7(1):1–9.
- [10] Liu Y, Liu J, Plante KS, et al. The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature.* 2022;602(7896):294–299.
- [11] Tian F, Tong B, Sun L, et al. N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2. *Elife.* 2021;10:e69091.
- [12] Hou X, Zhang Z, Gao J, et al. SARS-CoV-2 spike protein N501Y mutation causes differential species transmissibility and antibody sensitivity: a molecular dynamics and alchemical free energy study. *Mol Syst Des Eng.* 2021;6(11):964–974.
- [13] Rostami N, Choupani E, Hernandez Y, et al. SARS-CoV-2 spike evolutionary behaviors; simulation of N501Y mutation outcomes in terms of immunogenicity and structural characteristic. *J Cell Biochem.* 2022;123(2):417–430.
- [14] Chakraborty S. E484K and N501Y SARS-CoV 2 spike mutants increase ACE2 recognition but reduce affinity for neutralizing antibody. *Int Immunopharmacol.* 2022;102:108424.
- [15] Kazybay B, Ahmad A, Mu C, et al. Omicron N501Y mutation among SARS-CoV-2 lineages: insilico analysis of potent binding to tyrosine kinase and hypothetical repurposed medicine. *Travel Med Infect Dis.* 2022;45:102242.
- [16] Lan J, Ge J, Yu J, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature.* 2020;581(7807):215–220.
- [17] Abraham MJ, Murtola T, Schulz R, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1-2:19–25.
- [18] Brooks BR, Brucoleri RE, Olafson BD, et al. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem.* 1983;4(2):187–217.
- [19] Darden T, Perera L, Li L, et al. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure.* 1999;7(3):R55–R60.
- [20] Ewald P. Evaluation of optical and electrostatic lattice potentials. *Ann Phys.* 1921;64:253–287.
- [21] Zhang Q, Xiang R, Huo S, et al. Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Sig Transduct Target Ther.* 2021;6(1):1–19.
- [22] Birtles D, Oh AE, Lee J. Exploring the pH dependence of the SARS-CoV-2 complete fusion domain and the role of its unique structural features. *Protein Sci.* 2022;31(9):e4390.
- [23] Curry HB. The method of steepest descent for non-linear minimization problems. *Quart Appl Math.* 1944;2(3):258–261.
- [24] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys.* 2007;126(1):014101.
- [25] Berendsen HJ, Postma JV, Van Gunsteren WF, et al. Molecular dynamics with coupling to an external bath. *J Chem Phys.* 1984;81(8):3684–3690.
- [26] Verlet L. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* 1967;159(1):98–103.
- [27] Aier I, Varadwaj P, Raj U. Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. *Sci Rep.* 2016;6:34984.
- [28] Benson NC, Daggett V. A comparison of multiscale methods for the analysis of molecular dynamics simulations. *J Phys Chem B.* 2012;116(29):8722–8731.
- [29] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982;157(1):105–132.
- [30] Eisenhaber F, Lijnzaad P, Argos P, et al. The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J Comput Chem.* 1995;16(3):273–284.
- [31] Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers Original Res Biomolecules.* 1983;22(12):2577–2637.
- [32] Meyers RA. *Encyclopedia of physical science and technology.* 3rd ed. Cambridge (MA): Academic Press, 2001.
- [33] Hubbard RE, Muhammad KH. Hydrogen bonds in proteins: role and strength. In *Encyclopedia of life sciences (ELS).* Chichester: John Wiley & Sons, Ltd; 2010.
- [34] Andrews D, Scholes G, Wiederrecht G. 2010. *Comprehensive nanoscience and technology.* Cambridge (MA): Academic Press.
- [35] Park SJ, Seo MK. *Interface science and composites.* 1st ed., Vol. 18. Cambridge (MA): Academic Press; 2011.

- [36] Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol.* 1984;44(2):97–179.
- [37] Vanommeslaeghe K, Hatcher E, Acharya C, et al. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem.* 2010;31(4):671–690.
- [38] Trevino SR, Scholtz JM, Pace CN. Amino acid contribution to protein solubility: Asp, Glu, and Ser contribute more favorably than the other hydrophilic amino acids in RNase Sa. *J Mol Biol.* 2007;366(2):449–460.
- [39] Deshmukh SA, Sankaranarayanan SK, Suthar K, et al. Role of solvation dynamics and local ordering of water in inducing conformational transitions in poly (N-isopropylacrylamide) oligomers through the LCST. *J Phys Chem B.* 2012;116(9):2651–2663.
- [40] Mannar D, Saville JW, Zhu X, et al. SARS-CoV-2 omicron variant: antibody evasion and cryo-EM structure of spike protein–ACE2 complex. *Science.* 2022;375(6582):760–764.
- [41] Saville JW, Mannar D, Zhu X, et al. Structural and biochemical rationale for enhanced spike protein fitness in Delta and kappa SARS-CoV-2 variants. *Nat Commun.* 2022;13(1):742.
- [42] Zhu X, Mannar D, Srivastava SS, et al. Cryo-electron microscopy structures of the N501Y SARS-CoV-2 spike protein in complex with ACE2 and 2 potent neutralizing antibodies. *PLoS Biol.* 2021;19(4):e3001237.
- [43] Çubuk H, Özbi LM. In silico analysis of SARS-CoV-2 spike protein N501Y and N501T mutation effects on human ACE2 binding. *J Mol Graph Model.* 2022;116:108260.