Glycosylation as a key for enhancing drug recognition into spike glycoprotein of SARS-CoV-2

Georcki Ropón-Palacios[1], Jhon Pérez-Silva[1], Ricardo Rojas-Humpire[1], Gustavo E. Olivos-Ramírez[1], Manuel Chenet-Zuta[2], Victor Cornejo-Villanueva[3], Sheyla Carmen-Sifuentes[1], Kewin Otazu[1], Yaritza L. Ramirez-Díaz[1], Ihosvany Camps[1,4]

[1] Laboratório de Modelagem Computacional, Instituto de Ciências Exatas, Universidade Federal de Alfenas. Brasil.
[2] Escuela de Posgrado, Universidad San Ignacio de Loyola, Perú.
[3] Escuela de Genética y Biotecnología, Universidad Nacional Mayor de San Marcos. Perú.
[4]High Performance & Quantum Computing Labs. Waterloo, Canada

Corresponding author:
Ihosvany Camps : icamps@unifal-mg.edu.br ,
Georcki Ropón-Palacios: groponp@gmail.com

Abstract

The emergence of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) and its spread since 2019 represents the major public health problem worldwide nowadays, generating a high number of infections and deaths. That's why, in addition to vaccination campaigns, the design of a drug to help in the treatment of severe cases of COVID-19 is being investigated. In relation to SARS-CoV-2, one of its most studied proteins is the spike protein (S protein), which mediates host-cell entry and is heavily glycosylated. Regarding the latter, several investigations have been carried out, since it plays an important role in the evasion of the host's immune system and contributes to protein folding and the thermostability of the viral particle. For that reason, our objective was to evaluate the impact of glycosylations on the drug recognition on two domains of the S protein, the receptor-binding domain (RBD) and the N-terminal domain (NTD) through molecular dynamics simulations and computational biophysics analysis. Our results show that glycosylations in the S protein induce structural stability and changes in rigidity/flexibility related to the number of glycosylations in the structure. These structural changes are important for its biological activity as well as the correct interaction of ligands in the RBD and NTD regions. Additionally, we evidenced a roto-translation phenomenon in the interaction of the ligand with RBD in the absence of glycosylation, which disappears due to the influence of glycosylation and the convergence of metastable states in RBM. Similarly, glycosylations in NTD promote an induced-fit phenomenon, which is not observed in the absence of glycosylations; this process is decisive for the activity of the ligand at the cryptic site. Altogether, these results provide an explanation of glycosylation relevance in biophysical properties and drug recognition to S protein of SARS-CoV-2 which must be considered in the rational drug development and virtual screening targeting S protein.

Keywords: *COVID-19, Molecular Dynamics, N- and O-Glycosylation, Cryptic pocket, Induced-Fit binding, Roto-translation phenomenon, Free Energy Landscape*

1. Introduction

Coronaviruses (CoVs) are a diverse group of enveloped, positive sense, single-stranded RNA viruses causing mild to severe respiratory infections in humans (Pal et al., 2020). In December 2019, a novel coronavirus, designated as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), outbroke and rapidly spread causing an epidemic of unusual pneumonia in Wuhan, China (Wu & McGoogan, 2020). This novel coronavirus disease was declared as COVID-19 by the World Health Organization (WHO) at the beginning of 2020, since then has spread fast all over the world and has become the major public health problem. Such has been the impact of this pandemic that, until this submission, more than 204 million confirmed cases and more than 4 million deaths have been reported (JHU, 2021). For that reason, governments and the scientific community developed strategies to contain the impact of COVID-19, based on epidemiological surveillance, vaccination, and infection mechanisms research (Ibrahim, 2020). This has led to an accelerated development of vaccines and the initiation of a worldwide vaccination campaign, which to date has seen more than 4.56 billion people vaccinated (Randall et al., 2021). However, many countries continue to face serious problems in obtaining vaccines and continue to record numerous cases. The problem is exacerbated by the lack of a drug that can be used in cases of severe symptoms. This situation highlights the need to discover a drug to help treat severe cases of COVID-19.

Currently, one of the most studied biological processes in SARS-CoV-2 is the recognition of the host cell-mediated by a membrane glycoprotein named Spike protein (S protein). The S protein is a trimeric class I fusion protein. Its extracellular domain contains the subunits S1 (residues 14-685) and S2 (residues 86-1273) (Li, 2016), associated with receptor binding and membrane fusion, respectively (Song et al., 2018). The S1 subunit contains the Receptor Binding Domain (RBD, residues 333-527) that interacts with the human Angiotensin-converting enzyme 2 (ACE2) (Monteil et al., 2020). This region also contains the Recognition Binding Motif (RBM, residues 438-506) that directly contacts ACE2 (Yang et al., 2020). In such interaction, the spike protein undergoes conformational changes that allow an asymmetric rearrangement that drives one of its three RBDs to its "up" conformation. Only RBDs in this conformation can bind ACE2 (Roy et al., 2020). Since this is a vital process for the initiation of infection, a lot of research has pointed to the S protein as one of the most important therapeutic targets (Krumm et al., 2021).

The earliest studies of cryo-electron microscopy (cryo-EM) structure of the S protein in SARS-COV-2 revealed that it is heavily glycosylated (Yao et al., 2020). Each promoter contains 22 canonical N-linked glycosylation and 3 O-glycosylation sequons (Zhao et al., 2020). Glycosylation plays a pivotal role in the viral life cycle, as it aids in evading the host immune system (Kasuga et al., 2021). Furthermore, glycosylations contribute to protein folding and the thermostability of the viral particle (Kawase et al., 2019). They also can participate in viral entry by specific interactions with the host surface (Raman et al., 2016). Therefore, understanding the molecular mechanisms of these structures during drug recognition in SARS-CoV-2 may contribute to therapeutic alternatives.

In addition, exploring new druggable sites in the spike protein could potentially lead to inhibition of its function. The discovery of cryptic sites precisely focuses on cryptic site forms a pocket in a holo-structure that can be identified transiently due to conformational changes (Sztain et al., 2020). These sites are generally not detected during the atomic resolution process (Beglov et al., 2018); however, new algorithms have been developed to help predict potential cryptic sites, for instance, CryptoSite (Cimermancic et al., 2016). Evaluating new druggable sites is a recent alternative being studied in SARS-CoV-2 with the aim of expanding its druggable proteome (Cavasotto et al., 2021).

Since there is a need to develop and find effective drugs to treat COVID-19. Here, we evaluated the impact of glycosylation in the biophysical properties and drug recognition in RBD and N-

terminal domain (NTD) of SARS-CoV-2 compared to non-glycosylated structures, through molecular dynamics simulations and computational biophysics analysis.

2. Material and Methods

2.1. Data set. The data used in this paper was taken from work of Otazu et al. (2020), where it shows the importance of the two ligands selected in this work, through molecular docking data. Molecular complexes in PDB format were shared with us.

2.2. MD simulation. To explore the biological role of glycosylations on small molecules, like drugs, we used all-atom molecular dynamics (MD) simulations. The system was built on the CHARMM-GUI server (Jo et al., 2008) using a Spike glycosylated protein available on the COVID-19 Proteins archive (Woo et al., 2020). The solution builder module was used to generate the system topology on a cubic box with a padding of 1.5 nm. The TIP3P water was used to solvate the box, following ionization with sodium (Na+) and chlorine (Cl-) ions to neutralize the system at 154 mM. The CHARMM36m force field was selected for the calculation of the interactions.

MD simulations were performed in GROMACS v2019.3 (Abraham et al., 2015) in four steps. First, energy minimization using the steepest descent algorithm with 5000 steps or until reaching an energy <10 kJ/mol/nm to eliminate bad contacts. Second, an NVT equilibrium phase at 310 K for 2 ns to equilibrate the system temperature. Third, an NPT equilibrium phase at 1 bar for 4 ns to equilibrate the system pressure. he Berendsen thermostat (Berendsen, 1991) and the Parrinello-Rahman barostat (Parrinello & Rahman, 1981) were used in the equilibrium phases. Fourth, a production simulation for 50 ns with integration steps of 2 fs, under constant pressure and temperature using leap-frog integration algorithm (van Gunsteren & Berendsen, 1988). To generate the trajectories, the LINCS algorithm was used to constrain the interactions during equilibrium, while the Particle-Mesh Ewald algorithm was used to constrain the long-range ionic interactions.

2.8. Data analysis and biomolecular graphics. The trajectories were analyzed using geometric and structural properties to determine the influence of glycosylations on molecular recognition. These were done using a pool of *in-house* scripts based on tcl embedded in VMD (Humphrey et al., 1996) and the python library MDAnalysis (Michaud-Agrawal et al., 2011). All images were rendered with VMD.

3. Results

This work explores the impact of glycosylation and ligand interactions on RBD and NTD of SARS-CoV-2. In figure 1, six systems built in the presence or absence of glycosylations or ligands are shown. Two systems contain the interaction of ligand TCMDC-124223 to RBD **(figure 1a)**, two others the interaction of ligand TCMDC-133766 to NTD **(figure 1b)**. Finally, two other systems without ligands were built as controls **(figure 1c).**

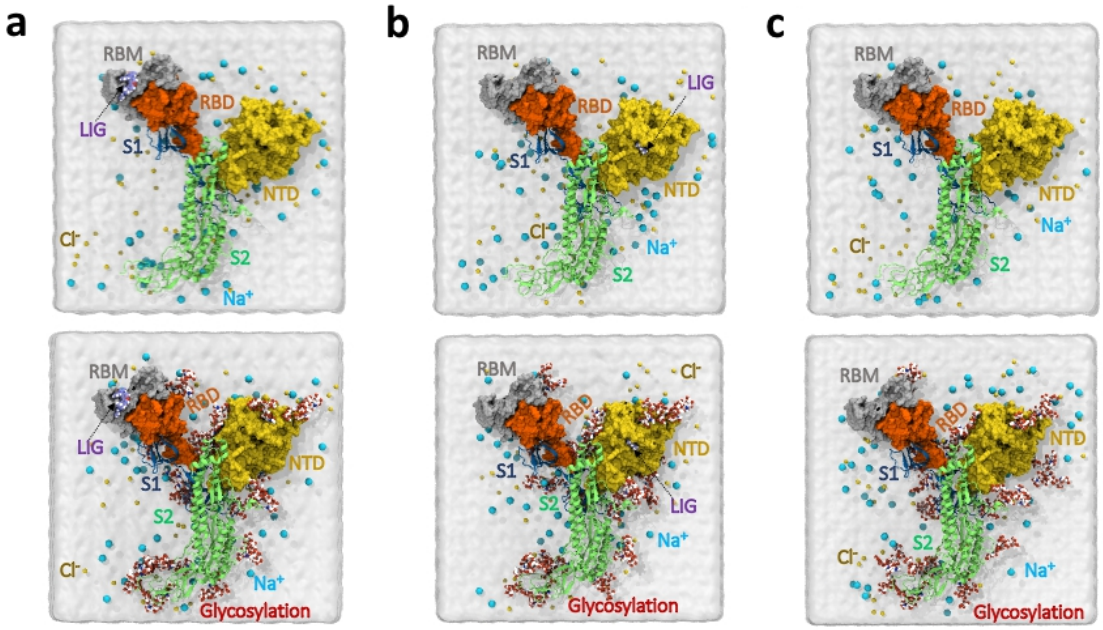

Figure 1. System setup for molecular dynamics simulation. (a) Representation of S protein with interaction of TSMDC-124223 in RBD in presence or absence of glycosylations (b) Representation of S protein with interaction of TCMDC-133766 in NTD in presence or absence of glycosylations (c) Representation of S protein without interaction of ligands in RBD or NTD in presence or absence of glycosylations. All systems contain Na+ and Cl-, the water is represented as the surface in white color.

3.2. Glycosylation induces important protein stability. To evaluate structural changes in the Spike protein, an RMSD analysis was performed through *in-house* scripts based on MDanalysis. These results correspond to the analysis of RBD and NTD in presence or absence of ligands or glycosylations. The glycosylated RBD without ligand (Apo) reaches significant stability at 20ns compared to non-glycosylated Apo structure. This stability increased in the glycosylated RBD with ligand (Holo) in contrast to non-glycosylated Holo structure that no reaches stability at the final of simulation, all RBD Apo and Holo structures presented structural changes until 4Å of RMSD **(figure 2a)**. The impact of glycosylations is very important for the biological properties of RBD (Reis et al., 2021). In this sense, our results reveal that glycosylations introduce important structural changes in the receptor-binding motif (RBM) compared to non-glycosylated Apo structure, this is determinant for the recognition of ligands to RBD **(figure 2b, d)**. On the other hand, the analysis of NTD reveals that glycosylations maintain stable structures Apo or Holo along the time simulation, despite ligand interaction induces light structural disturbances. This phenomenon explains the constant changes and structural instability of non-glycosylated Apo or Holo structures **(figure 2a)**. Therefore, local rigid changes by glycosylations control are needed to the fit-induced of ligand in the cryptic pocket found in NTD, this is correlated by highly glycosylation observed on NTD compared to other components of Spike protein **(figure 2c, d)**

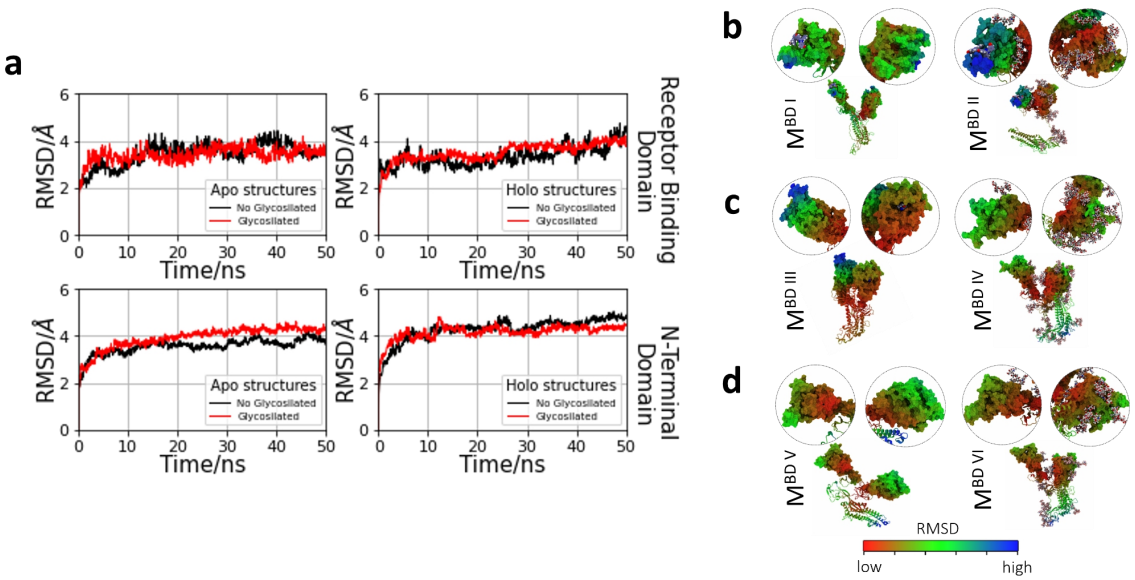

Figure 2. Conformational stability induced by glycosylations. (a) Comparison of structural evolution of RBD and NTD with (Holo structure) or without (Apo structure) ligand interaction in presence or absence of glycosylations. (b) Local structural changes in RBD with ligand interaction; left structure without glycosylation and right with glycosylation. (c) Local structural changes in NTD with ligand interaction; left structure without glycosylation and right with glycosylation. (d) Local structural changes in RBD and NTD without ligand interaction; left structure without glycosylation and right with glycosylation. Hue changes shown in the color bar indicate minimal (red), intermediate (green), or high (blue) structural changes.

3.3. Rigid/flexibility changes in RBD and NTD by glycosylation.

To evaluate changes in the rigid/flexibility of protein structure, RMSF analysis of the backbone was performed using *in-house* scripts based on MDanalysis. The RBD and NTD regions of the spike protein were analyzed in the presence and absence of glycosylations or ligand interactions. The glycosylated and non-glycosylated RBD structures without ligand (Apo) presented similar changes in flexibility with maximum RMSF values of 30 Å in the RBM region. In contrast to the RBD structures with ligand (Holo), the observed changes are attributed to the introduction of the ligand and not so much to the glycosylations, so that the interaction of ligand with the non-glycosylated structure significantly increased the structural rigidity **(figure 3a)**. This phenomenon is important given that the flexibility of the RBD is determinant for the interaction with receptors such as ACE2 (Chen et al., 2020), thus the importance of local changes in rigidity/flexibility induced by glycosylations in RBD is evident **(figure 3b, d)**. On the other hand, analysis of the NTD region shows that the Apo structures are stiffer in the presence of glycosylations than the non-glycosylated structure; these changes are more evident in the region from I100 to Y200. In contrast, ligand interaction induces local flexibility to the glycosylated structure similar to the flexibility values of the non-glycosylated structure. This change in the flexibility of the NTD is very important for the fit-induced by ligand recognition on the cryptic site **(figure 3c, d)**.

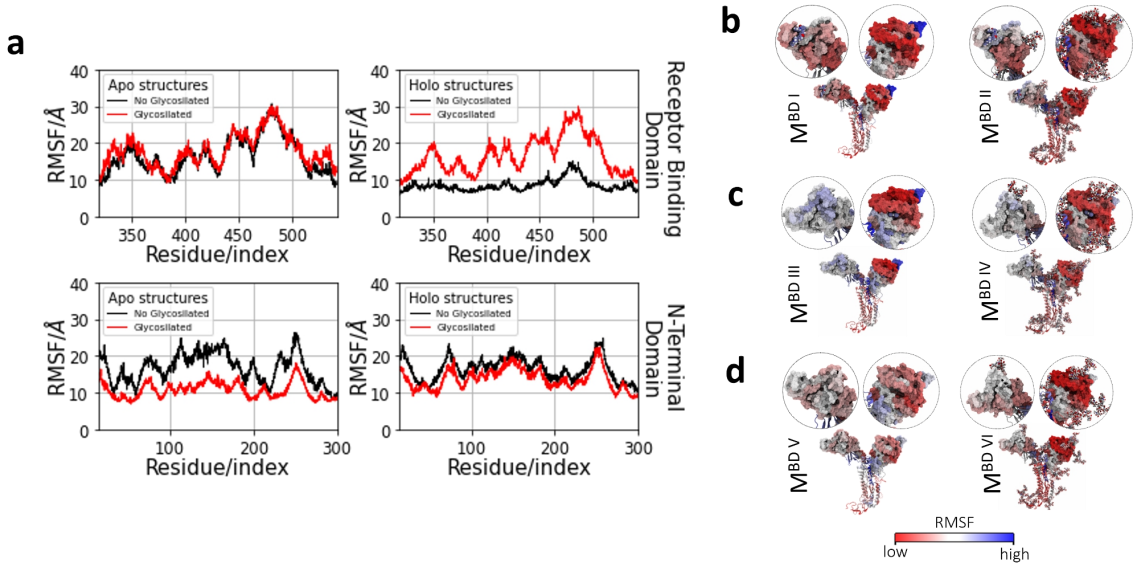

Figure 3. Local changes in rigid/flexibility induced by glycosylations. (a) Comparison of rigid/flexibility on backbone of RBD and NTD with (Holo structure) or without (Apo structure) ligand interaction in presence or absence of glycosylations. (b) Rigid/flexibility changes in RBD with ligand interaction; left structure without glycosylation and right with glycosylation. (c) Rigid/flexibility changes in NTD with ligand interaction; left structure without glycosylation and right with glycosylation. (d) Rigid/flexibility changes in RBD and NTD without ligand interaction; left structure without glycosylation and right with glycosylation. Hue changes shown in the color bar indicate minimal (red), intermediate (white), or high (blue) flexibility.

3.4. Glycosylation changes roto-translation to the fix binding mode. The molecular dynamics of the glycosylated and non-glycosylated RBD structures revealed that the ligand undergoes through different conformational states promoted by roto-translational phenomena in the RBM region in the absence of glycosylation, the displacement was observed from the first ns, followed by a perpendicular rotation at the opposite end of the origin, ending in a slight return of the structure almost at the end of the simulation **(figure 4a, d)**. In contrast, glycosylation induces a sustained ligand interaction in loop/beta structure formed by residues T470 to F490, which remains stable until 40ns where a major conformational change is generated before the end of the simulation **(figure 4b,d)**. To determine the characteristics of roto-translation phenomenon, the distance of the RBM residue G446 to the central benzofuranic structure of the ligand was calculated, while for the calculation of the rotation angle the residue Q493 was used as the origin to generate angles θ with respect to the central structure of the ligand. Finally, the generation of metastable states was evaluated through free energy landscape (FEL) matrices, all analysis was performed by *in-house* scripts based on MDanalysis **(figure c, g, h)**. The displacement to 10 Å of the G446 residue of the ligand in the non-glycosylated structure at 20ns which was maintained throughout the simulation was evidenced, this is related to the important rotational changes generated in the first ns of the simulation and the generation of metastable states from 18Å to 10Å at θ angles from 80° to 150° **(figure e-g)**. In contrast, in the glycosylated structure the ligand remains approximately 22Å from the G446 residue for a longer time, this is related to the generation of metastable states at 22Å and θ angle between 100° to 140°, then a shift is observed before 30 ns and at the end of the simulation, this shift occurs in parallel to the rotational changes in time **(figure e, f, h)**. The roto-translation phenomenon can be reviewed in the supplementary videos.

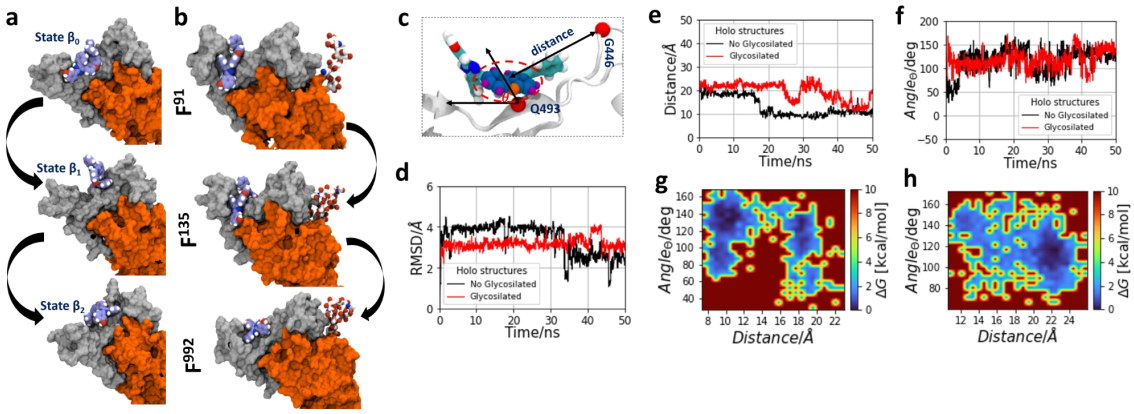

Figure 4. Interaction of ligand TCMDC-124223 on RBM. Evolution of ligand conformational states along simulation time; gray structure represents RBM and complete gray and orange structure represent RBD (a) Ligand representation β_0 at frame 91, β_1 at frame 135, and β_2 at frame 992 in absence of glycosylations (b) Ligand representation β_0 at frame 91, β_1 at frame 135 and β_2 at frame 992 in presence of glycosylations. (c) Ligand distance to G466 and θ angle formed in the roto-translation. (d) Conformational changes of ligand in presence and absence of glycosylations. (e) Translation changes of ligand in presence and absence of glycosylations. (f) Rotational changes of ligand in presence and absence of glycosylations. (g) Free energy landscape (FEL) matrix of ligand in absence of glycosylations. (h) Hue changes shown in the color bar indicate minimal (red), intermediate (green), or high (blue) probability of metastable states in FEL matrix.

3.5. Glycosylation promotes recognition by induced fit.

The molecular dynamics of the glycosylated and non-glycosylated NTD structures revealed that ligand remains inside the cryptic pocket, throughout the simulation. In this sense, we observed that the glycosylated structure promotes the induced fit, encapsulating the ligand at the end of the simulation. In contrast, the non-glycosylated structure does not present this phenomenon since it maintains the structure of the cryptic pocket without significant changes **(figure 5a, b)**. To determine the characteristics of the ligand in the induced fitting process we evaluated the RMSD of the ligand, the solvent accessible surface areas (SASA), and the generation of metastable states through FEL matrices at a temperature of 310K between the number of contacts and the SASA, all analysis was performed by *in-house* scripts based on MDanalysis. The ligand presented an early reduction of the RMSD until reaching equilibrium at 20 ns, this represents the progressive encapsulation process of the cryptic site in the glycosylated structure, which is related to the variations of the SASA before 20 ns and subsequent fall until reaching equilibrium **(figure 5c, d)**. This process of encapsulation by the induced fit significantly restricts the number of ligand contacts from 20 ns to the end of the simulation **(figure 5e)**. On the contrary, the non-glycosylated structure maintains the ligand without significant variations until 40ns where it reaches another stable conformation. Related to this, the induced fit phenomenon is not evidenced since the SASA presents slight variations increasing the solvent area that reaches high values until the end of the simulation. This process generates an increase in the degrees of freedom, evidenced by the increase in the number of contacts. Additionally, the glycosylated structure generates fewer metastable states than the non-glycosylated structure, due to the confinement by the induced fit, unlike the non-glycosylated structure that presents dispersed metastable states due to the higher SASA and freedom of contacts **(figure g, f)**. The induced fit phenomenon can be reviewed in the supplementary videos.

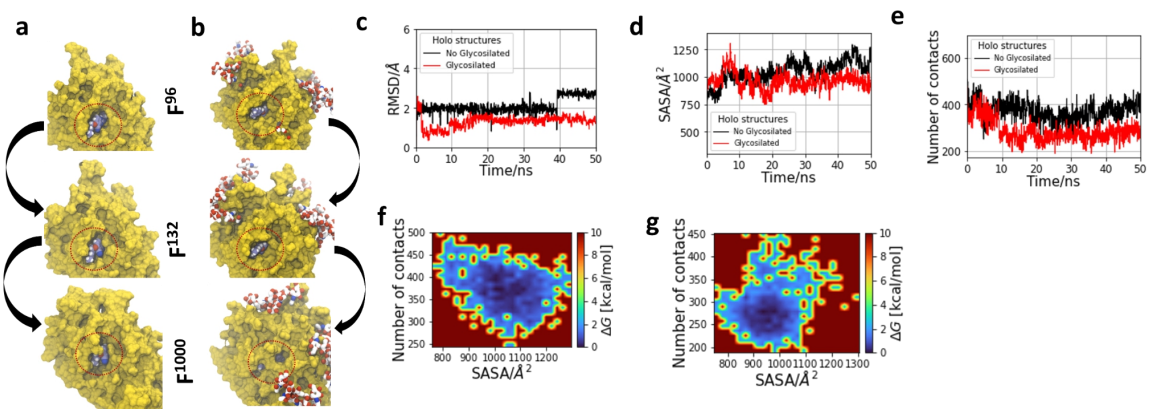

Figure 5. Interaction of ligand TCMDC-133766 on the cryptic pocket of NTD. (a) Evolution of ligand interaction at frames 96, 132, and 1000 in absence of glycosylations. (b) Evolution of ligand interaction at frames 96, 132, and 1000 in presence of glycosylations. (c) Conformational changes of ligand in presence and absence of glycosylations. (d) Solvent accessible surface areas changes of ligand in presence and absence of glycosylations. (e) Changes in the number of protein contacts of ligand in presence and absence of glycosylations. (f) Free energy landscape (FEL) matrix of ligand in absence of glycosylations. (g) FEL matrix of ligand in presence of glycosylations. Hue changes shown in the color bar indicate minimal (red), intermediate (green), or high (blue) probability of metastable states in FEL matrix.

4. Discussion

The COVID-19 pandemic represents one of the biggest problems for public health today (Wang et al., 2021). For this reason, researchers around the world focus their studies on drug development targeting key structural components of SARS-COV-2 such as RBD and NTD (Hu et al., 2021). In this sense, computer-aided drug design is a widely used approach for this purpose (Gurung et al., 2021). However, simulating biological systems for the investigation of potential drugs requires the correct configuration of the biophysical properties granted by macromolecules such as glycans or other post-translational modifications.

Our results show that glycosylations are important for structural stability and local changes in rigid/flexibility of proteins such as RBD and NTD of SARS-CoV-2. This can be explained by the steric restrictions that O- and N-glycans generate in different protein systems (Solá & Griebenow, 2009; Watanabe et al., 2004). For instance, the study of Lee et al. showed that N-glycosylations does not induce significant structural changes in mammalian proteins, this was attributed to the reduction of structural dynamics and therefore increased stability (Lee et al., 2015). Similarly, glycosylations of viral envelopes serve a wide range of functions, including regulation of viral tropism, host immunity, and protein stability (Bagdonaite & Wandall, 2018). On the other hand, changes in rigid/flexibility are apparently influenced by the structural complexity of the glycosylated proteins, since in some systems it is significant while in others it is not transcendent (Sarkar and Wintrode, 2011; Weiß et al., 2021). In our study, the local flexibility and stability induced by glycosylations were decisive for the correct recognition of ligands in structural targets of SARS-CoV-2.

In addition, the roto-translation phenomenon in the RBM was a characteristic pattern of ligand interaction in the absence of RBD glycosylations. This was related to the local structural flexibility of the pocket that plays a fundamental role in allowing or restricting this phenomenon by modifying entropy of the ligand (Chang et al., 2007; Chang and Gilson, 2004). In this sense, interactions of ligands with flexible binding sites can limit their roto-translation freedom, allowing stable conformations long enough for successful activation of the ligand-receptor complex. Moreover, stable interactions from a physicochemical view depend on other factors such as the

potential of hydrogens, temperature of the system, the number and types of bonds formed (Majewski et al., 2019). Therefore, glycosylations induce metastable states in the loop/beta structure formed by residues T470 to F490 to restrict the roto-translation phenomenon in a physiological environment.

Furthermore, another widely documented form of ligand interaction was evidenced on the NTD in presence of the glycosylations, the induced-fit phenomenon. The comparison between holo structures demonstrates that with glycosylations, the NTD undergoes certain stable conformational changes when forming the complex structure with its binding partner. This important type of interaction was reported to generate a tight binding between a molecular host and guests (Kasai, Aoyagi and Fujita, 2000; Wang et al., 2016; Hong et al., 2017), to induce signal transduction (Maiti et al., 2014; Langton et al., 2017) and to confer allosteric regulation through conformational changes after the binding process (Sevcsik et al., 2011; Suzuki et al., 2016). However, for successful induced-fit ligand binding, it is necessary a conformational flexibility that enables molecular hosts to respond to the shape and electrostatic surface of guest molecules (Zhan et al., 2018). In this sense, glycosylations in NTD introduce the necessary conformational rigid/flexibility to create energetically favorable states within a fit-induced pocket that encapsulates the ligand upon binding.

Finally, our study shows that glycosylations are important components for the biophysical properties of SARS-CoV-2 protein S. This has implications for drug recognition in the RBD and NTD domain, therefore future virtual screening and docking studies need to consider N- and O-glycosylations to obtain more accurate results on the interaction of ligands on druggable targets to inhibit SARS-CoV-2 infection.

5. Conclusion

We conclude that non-protein structural components such as glycosylations play an important role in the biophysical structural properties in S protein of SARS-CoV-2. These biophysical changes enhance drug recognition in RBD and NTD, key structures for the development of new drugs.

Our main findings were the induction of structural stability and local changes in rigidity/flexibility related to the number of glycosylations in RBD and NTD. These structural changes are important for its biological activity and drug recognition, evidenced by roto-translation phenomenon in the interaction of the ligand with RBD in the absence of glycosylation and the induced fit phenomenon of ligand in NTD in presence of glycosylations. Therefore, glycosylations must be placed into account for rational drug development and virtual screening targeting S protein.

Supplementary material

Movies of molecular dynamics trajectories made with VMD can be found in https://doi.org/10.5281/zenodo.5295791 .

Acknowledge

We would like to acknowledge financial support from the Brazilian agencies CNPq, CAPES and FAPEMIG. Part of the results presented here were developed with the help of CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenho em São Paulo) grant UNICAMP/FINEP-MCT and CENAPAD-UFC (Centro Nacional de Processamento de Alto Desempenho, at Universidade Federal do Ceará) and Compute Canada Database (CCDB).

6. References

Abraham, M.J. *et al.* (2015) 'GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers', *SoftwareX*, 1-2, pp. 19–25. doi:10.1016/j.softx.2015.06.001.

Armijos-Jaramillo, V. *et al.* (2020) 'SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals undiscovered amino acids necessary for complex stability', *Evolutionary Applications*, 13(9), pp. 2168-2178. doi:10.1101/2020.03.21.001933.

Bagdonaite, I. and Wandall, H.H. (2018) 'Global aspects of viral glycosylation', *Glycobiology*, 28(7), pp. 443–467. doi:10.1093/glycob/cwy021.

Banck, M. *et al.* (2011) 'Open Babel: An open chemical toolbox', *J Cheminform*, 3(1), pp. 33. doi:10.1186/1758-2946-3-33.

Beglov, D. *et al*. (2018) 'Exploring the structural origins of cryptic sites on proteins', *Proc Natl Acad Sci USA*, 115(15), pp. 3416–3425. doi:10.1073/pnas.1711490115.

Berendsen, H.J.C. (1991) 'Transport Properties Computed by Linear Response through Weak Coupling to a Bath', in M. Meyer and V. Pontikis (eds), *Computer Simulation in Materials Science*, Springer, Dordrecht, doi:10.1007/978-94-011-3546-7_7

Bertram, S. *et al*. (2011) 'Cleavage and Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by Human Airway Trypsin-Like Protease', *Journal of Virology*, 85(24), pp. 13363–13372. doi:10.1128/JVI.05300-11.

Cavasotto, C.N., Lamas, M.S., and Maggini, J. (2021) 'Functional and druggability analysis of the SARS-CoV-2 proteome', *Eur J Pharmacol*, 890, pp. 1-21. doi:10.1016/j.ejphar.2020.173705.

Chan, J.F. *et al*. (2015) 'Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease', *Clin Microbiol Rev*, 28(2), pp. 465–522. doi:10.1128/CMR.00102-14.

Chang, C.E. and Gilson, M.K. (2004) 'Free energy, entropy, and induced fit in host-guest recognition: calculations with the second-generation mining minima algorithm', *J Am Chem Soc*, 126(40), pp. 13156–13164. doi: 10.1021/ja047115d.

Chen, T.M. *et al.* (2020) 'A mathematical model for simulating the phase-based transmissibility of a novel coronavirus', *Infectious Diseases of Poverty*, 9, pp. 1-8. doi:10.1186/s40249-020-00640-3.

Chen, W. *et al.* (2020) 'The N-glycosylation sites and Glycan-binding ability of S-protein in SARS-CoV-2 Coronavirus', *bioRxiv,* pp. 1-20. doi:10.1101/2020.12.01.406025.

Chen, Y. *et al.* (2020) 'Structure analysis of the receptor binding of 2019-nCoV', *Biochemical and Biophysical Research Communications*, 525(1), pp. 135–140. doi:10.1016/j.bbrc.2020.02.071.

Cimermancic, P. *et al*. (2016) 'CryptoSite: Expanding the Druggable Proteome by Characterization and Prediction of Cryptic Binding Sites', *Journal of Molecular Biology*, 428(4), pp. 709–719. doi:10.1016/j.jmb.2016.01.029.

Coppola, M. and Mondola, R. (2020) 'Phytotherapeutics and SARS-CoV-2 infection: Potential role of bioflavonoids', *Medical Hypotheses*, 140, pp. 109766. doi:10.1016/j.mehy.2020.109766.

Corrêa, C., Laaksonen, A. and Barroso da Silva, F.L. (2020) 'On the interactions of the receptor-binding domain of SARS-CoV-1 and SARS-CoV-2 spike proteins with monoclonal antibodies and the receptor ACE2', Virus Res, 285, pp. 1-13. doi:10.1016/j.virusres.2020.198021.

Coutard, B. *et al.* (2020) 'The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade', *Antiviral Res*, 176, pp. 1-5. doi:10.1016/j.antiviral.2020.104742.

DeLano, W.L. (2002) *The PyMOL molecular graphics system on world wide web*. Available at: http://www.pymol.org (Accessed: 2 June 2021).

Dickson, A. (2018) 'Mapping the Ligand Binding Landscape', *Biophys J*, 115(9), pp. 1707–1719. doi:doi.org/10.1016/j.bpj.2018.09.021.

Dong, N. *et al*. (2020) 'Genomic and protein structure modelling analysis depicts the origin and infectivity of 2019-nCoV, a new coronavirus which caused a pneumonia outbreak in Wuhan, China', *F1000Research*, 9, pp. 1-9. doi:10.12688/f1000research.22357.1.

Enriquez, M.I. (1996) 'Diseño racional de nuevos agentes antiviricos. Desarrollo de un modelo de farmacoforo por aplicacion de tecnicas de modelizacion molecular'. Master's Thesis. Universidad de Navarra, Spain. Available at: https://dialnet.unirioja.es/servlet/tesis?codigo=279289 (Accessed: 2 June 2021).

Fehr, A.R. and Perlman, S. (2015) 'Coronaviruses: an overview of their replication and pathogenesis', *Methods Mol Biol*, 1282, pp. 1–23. doi:10.1007/978-1-4939-2438-7_1.

Gao, Y. *et al*. (2020) 'Structure of the RNA-dependent RNA polymerase from COVID-19 virus', *Science*, 368(6492), pp. 779–782. doi:10.1126/science.abb7498.

Gierer, S. *et al*. (2013) 'The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies', *J Virol*, 87(10), pp. 5502–55011. doi:10.1128/JVI.00128-13.

Glowacka, I. *et al*. (2011) 'Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response', *J Virol*, 85(9), pp. 4122–4134. doi:10.1128/JVI.02232-10.

Gurung, A.B. *et al.* (2021) 'An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19', *BioMed Research International*, 2021, pp. 1–18. doi:10.1155/2021/8853056.

Hariharan, P. and Guan, L. (2014) 'Insights into the inhibitory mechanisms of the regulatory protein IIA(Glc) on melibiose permease activity', *J Biol Chem*, 289(47), pp. 33012–33019. doi: 10.1074/jbc.M114.609255.

Ho, T.Y. *et al*. (2007) 'Emodin blocks the SARS coronavirus spike protein and angiotensin-converting enzyme 2 interaction', *Antiviral Res*, 74(2), pp. 92–101. doi:10.1016/j.antiviral.2006.04.014.

Hong, C.M. *et al*. (2017) 'Conformational Selection as the Mechanism of Guest Binding in a Flexible Supramolecular Host', *J Am Chem Soc,* 139(23), pp. 8013–8021. doi: 10.1021/jacs.7b03812.

Hu, X. *et al*. (2021) 'The study of antiviral drugs targeting SARS-CoV-2 nucleocapsid and spike proteins through large-scale compound repurposing', *Heliyon*, 7(3), pp. 1-9. doi:10.1016/j.heliyon.2021.e06387.

Humphrey, W., Dalke, A. and Schulten, K. (1996) 'VMD: visual molecular dynamics', *J Mol Graph*, 14(1), pp. 33–38, 27–28. doi:10.1016/0263-7855(96)00018-5.

Hussain, M. *et al.* (2020) 'Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein', *J Med Virol*, 92(9), pp. 1580–1586. doi:10.1002/jmv.25832.

Ibrahim, N.K. (2020) 'Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications', *Journal of Infection and Public Health*, 13(11), pp. 1630-1638. doi:10.1016/j.jiph.2020.07.019.

Jo, S. *et al*. (2008) 'CHARMM-GUI: a web-based graphical user interface for CHARMM', J Comput Chem, 29(11), pp. 1859–1865. doi:10.1002/jcc.20945.

Johns Hopkins University - JHU (2021). *COVID-19 Dashboard.* Available at: https://coronavirus.jhu.edu/map.html. (Accessed: 13 August 2021).

Kasai, K., Aoyagi, M. and Fujita, M. (2000) 'Flexible Coordination Networks with Fluorinated Backbones. Remarkable Ability for Induced-Fit Enclathration of Organic Molecules', *J Am Chem Soc*, 122(9), pp. 2140–2141. doi: 10.1021/ja992553j.

Kasuga, Y. *et al*. (2021) 'Innate immune sensing of coronavirus and viral evasion strategies', *Exp Mol Med,* 53, pp. 723–736. doi:10.1038/s12276-021-00602-1.

Kawase, M. *et al.* (2019) 'Biochemical Analysis of Coronavirus Spike Glycoprotein Conformational Intermediates during Membrane Fusion', *J Virol*, 93(19), pp. 1-19. doi:10.1128/JVI.00785-19.

Krumm, Z.A. *et al*. (2021) 'Precision therapeutic targets for COVID-19', *Virol J*, 18, pp. 1-22. doi:10.1186/s12985-021-01526-y.

Lan, J. *et al*. (2020) 'Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor', *Nature*, 581(7807), pp. 215–220. doi:10.1038/s41586-020-2180-5.

Langton, M.J. *et al.* (2017) 'Triggered Release from Lipid Bilayer Vesicles by an Artificial Transmembrane Signal Transduction System', *J Am Chem Soc*, 139(44), pp. 15768–15773. doi: 10.1021/jacs.7b07747.

Lee, H.S., Qi, Y. and Im, W. (2015) 'Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study', *Scientific Reports*, 5(1), pp. 1-7. doi:10.1038/srep08926.

Letko, M., Marzi, A. and Munster, V. (2020) 'Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses', *Nat Microbiol*, 5, pp. 562-569. doi:10.1038/s41564-020-0688-y.

Li, F. *et al*. (2005) 'Structure of SARS coronavirus spike receptor-binding domain complexed with receptor', *Science*, 309(5742), pp. 1864–1868. doi:10.1126/science.1116480.

Li, F. (2016) 'Structure, Function, and Evolution of Coronavirus Spike Proteins', *Annu Rev Virol*, 3(1), pp. 237–261. doi:10.1146/annurev-virology-110615-042301.

Lipinski, C.A. *et al.* (2001) 'Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings', *Adv Drug Deliv Rev*, 46(1–3), pp. 3–26. doi:10.1016/s0169-409x(00)00129-0.

Liu, P. *et al*. (2020) 'Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)?', *PLOS Pathogens*, 16(5), pp. 1-13. doi:10.1371/journal.ppat.1008421.

López, V. *et al*. (2020) 'Recommendations on management of the SARS-CoV-2 coronavirus pandemic (Covid-19) in kidney transplant patients', *Nefrologia*, 40(3), pp. 265–271. doi:10.1016/j.nefro.2020.03.002.

Maiti, S. *et al*. (2014) 'Multivalent Interactions Regulate Signal Transduction in a Self-Assembled Hg2+ Sensor', *J Am Chem Soc*, 136(32), pp. 11288–11291. doi: 10.1021/ja506325e.

Matsuyama, S. *et al.* (2010) 'Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2', *J Virol*, 84(24), pp. 12658–12664. doi:10.1128/JVI.01542-10.

Michaud-Agrawal, N., *et al.* (2011) 'MDAnalysis: a toolkit for the analysis of molecular dynamics simulations', *J Comput Chem*, 32(10), pp. 2319–2327. doi:10.1002/jcc.21787.

Millet, J.K. and Whittaker, G.R. (2015) 'Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis', *Virus Res*, 202, pp. 120–134. doi:10.1016/j.virusres.2014.11.021.

Monteil, V. *et al*. (2020) 'Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2', *Cell*, 181(4), pp. 905-913. doi:10.1016/j.cell.2020.04.004.

Morris, G.M. *et al*. (2009) 'AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility', *J Comput Chem*, 30(16), pp. 2785–2791. doi:10.1002/jcc.21256.

Otazu, K. *et al.* (2020) 'Targeting Receptor Binding Domain and Cryptic Pocket of Spike glycoprotein from SARS-CoV-2 by biomolecular modeling', *Biomolecules*, pp. 1-24. arXiv:2006.06452.

Pal, M. *et al.* (2020) 'Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update', *Cureus*, 12(3), pp. 1-13. doi:10.7759/cureus.7423.

Palacios, M. *et al.* (2020) 'COVID-19, a worldwide public health emergency', *Rev Clin*, 221(1), pp. 55-61. doi:10.1016/j.rceng.2020.03.001.

Parrinello, M. and Rahman, A. (1981) 'Polymorphic transitions in single crystals: A new molecular dynamics method', *Journal of Applied Physics*, 52(12), pp. 7182–7190. doi:10.1063/1.328693.

Raman, R. *et al.* (2016) 'Glycan-protein interactions in viral pathogenesis', *Current opinion in structural biology*, 40, pp. 153–162. doi:10.1016/j.sbi.2016.10.003.

Randall, T. *et al.* (2021). *More Than 4.56 Billion Shots Given: Covid-19 Tracker. In the U.S., 353 million doses have been administered*. Available at: https://www.bloomberg.com/graphics/covid-vaccine-tracker-global-distribution/ (Accessed: 2 June 2021).

Reis, C.A., Tauber, R. and Blanchard, V. (2021) 'Glycosylation is a key in SARS-CoV-2 infection', *J Mol Med (Berl)*, 99(8), pp. 1023–1031. doi:10.1007/s00109-021-02092-0.

Robertson, J.G. (2007) 'Enzymes as a special class of therapeutic target: clinical drugs and modes of action', *Curr Opin Struct Biol*, 17(6), pp. 674–679. doi:10.1016/j.sbi.2007.08.008.

Ropón-Palacios, G. *et al*. (2020) 'Potential novel inhibitors against emerging zoonotic pathogen Nipah virus: a virtual screening and molecular dynamics approach', *J Biomol Struct Dyn*, 38(11), pp. 3225–34. doi:10.1080/07391102.2019.1655480.

Roy, S., Jaiswar, A. and Sarkar, R. (2020). 'Dynamic Asymmetry Exposes 2019-nCoV Prefusion Spike'. *J Phys Chem Lett*, 11(17), pp. 7021–7027. doi:10.1021/acs.jpclett.0c01431.

Salentin, S. *et al*. (2015) 'PLIP: fully automated protein-ligand interaction profiler', *Nucleic Acids Res*, 43(W1), pp. W443-447. doi:10.1093/nar/gkv315.

Santos-Martins, D. *et al*. (2019) 'Accelerating AutoDock4 with GPUs and Gradient-Based Local Search', *J Chem Theory Comput,* 17(2), pp. 1060-1073. doi:10.1021/acs.jctc.0c01006.

Sarkar, A. and Wintrode, P.L. (2011) 'Effects of glycosylation on the stability and flexibility of a metastable protein: The human serpin α1-antitrypsin', *International Journal of Mass Spectrometry*, 302(1–3), pp. 69–75. doi:10.1016/j.ijms.2010.08.003.

Sevcsik, E. *et al*. (2011) 'Allostery in a Disordered Protein: Oxidative Modifications to α-Synuclein Act Distally To Regulate Membrane Binding', *J Am Chem Soc*, 133(18), pp. 7152–7158. doi: 10.1021/ja2009554.

Shulla, A. *et al.* (2011) 'A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry', *J Virol*, 85(2), pp. 873–882. doi:10.1128/JVI.02062-10.

Solá, R.J. and Griebenow, K. (2009) 'Effects of glycosylation on the stability of protein pharmaceuticals', *J Pharm Sci*, 98(4), pp. 1223–1245. doi:10.1002/jps.21504.

Song, W. *et al*. (2018) 'Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2', *PLoS Pathog*, 14(8), pp. 1-19. doi:10.1371/journal.ppat.1007236.

Suryamohan, K. *et al*. (2021) 'Human ACE2 receptor polymorphisms and altered susceptibility to SARS-CoV-2', *Commun Biol*, 4, pp. 1-11. doi:doi.org/10.1038/s42003-021-02030-3.

Suzuki, Y. *et al*. (2016) 'Allosteric Regulation of Unidirectional Spring-like Motion of Double-Stranded Helicates', *J Am Chem Soc*, 138(14), pp. 4852–4859. doi: 10.1021/jacs.6b00787.

Sztain, T., Amaro, R. and McCammon, J.A. (2020) 'Elucidation of cryptic and allosteric pockets within the SARS-CoV-2 protease', *J Chem Inf Model*, 61(7), pp. 3495–3501. doi:10.1021/acs.jcim.1c00140.

Tai, W. *et al*. (2020) 'Identification of SARS-CoV RBD-targeting monoclonal antibodies with cross-reactive or neutralizing activity against SARS-CoV-2', *Antiviral Res*, 179, pp. 1-6. doi:10.1016/j.antiviral.2020.104820.

Trott, O. and Olson, A.J. (2010) 'AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading', *J Comput Chem*, 31(2), pp. 455–461. doi:10.1002/jcc.21334.

Vajda, S. *et al*. (2018) 'Cryptic binding sites on proteins: definition, detection, and druggability', *Curr Opin Chem Biol*, 44, pp. 1–8. doi:10.1016/j.cbpa.2018.05.003.

Van Gunsteren, W.F., and Berendsen, H.J.C. (1988) 'A Leap-frog Algorithm for Stochastic Dynamics', *Molecular Simulation*, 1(3), pp. 173–185. doi:10.1080/08927028808080941.

Van Voorhis, W.C. *et al*. (2016) 'Open Source Drug Discovery with the Malaria Box Compound Collection for Neglected Diseases and Beyond', *PLoS Pathog*, 12(7), pp. 1-23. doi:10.1371/journal.ppat.1005763.

Wang, C. *et al.* (2021) 'COVID-19 in early 2021: current status and looking forward', *Signal Transduct Target Ther*, 6(1), pp. 1-14. doi:10.1038/s41392-021-00527-1.

Wang, Q. *et al*. (2020) 'Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2', *Cell*, 181(4), pp. 894-904. doi:10.1016/j.cell.2020.03.045.

Wang, S. *et al*. (2016) 'Capsule–Capsule Conversion by Guest Encapsulation', *Angew Chem Int Ed Engl*, 128(6), pp. 2103–2106. doi: 10.1002/ange.201509278.

Wang, W., Wang, Y.X. and Yang, H.B. (2016) 'Supramolecular transformations within discrete coordination-driven supramolecular architectures', *Chem Soc Rev*, 45(9), pp. 2656–2693. doi: 10.1039/C5CS00301F.

Watanabe, I. *et al.* (2004) 'Glycosylation Affects the Protein Stability and Cell Surface Expression of Kv1.4 but Not Kv1.1 Potassium Channels', *J Biol Chem*, 279(10), pp. 8879–8885. doi:10.1074/jbc.m309802200.

Waterhouse, A. *et al*. (2018) 'SWISS-MODEL: homology modelling of protein structures and complexes', *Nucleic Acids Res*, 46(W1), pp. W296–303. doi:10.1093/nar/gky427.

Weiß, R.G. *et al.* (2021) 'N-Glycosylation Enhances Conformational Flexibility of Protein Disulfide Isomerase Revealed by Microsecond Molecular Dynamics and Markov State Modeling', *J Phys Chem B*. doi:10.1021/acs.jpcb.1c04279.

Woo, H. *et al.* (2020) 'Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane', *J Phys Chem B*, 124(33), pp. 7128–7137. doi:10.1021/acs.jpcb.0c04553.

Wrapp, D. *et al*. (2020) 'Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation', *Science*, 367(6483), pp. 1260–1263. doi:10.1126/science.abb2507.

Wu, Z. and McGoogan, J.M. (2020) 'Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention', *JAMA*, 323(13), pp. 1239–1242. doi:10.1001/jama.2020.2648.

Xia, S. *et al*. (2020) 'Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein', *Cell Mol Immunol*, 17(7), pp. 765–767. doi:10.1038/s41423-020-0374-2.

Yan, R. *et al*. (2020) 'Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2', *Science*, 367(6485), pp. 1444–1448. doi:10.1126/science.abb2762.

Yang, M. *et al*. (2020) 'Characteristics of registered studies for Coronavirus disease 2019 (COVID-19): A systematic review', *Integr Med Res*, 9(3), pp. 1-9. doi:10.1016/j.imr.2020.100426.

Yao, H. *et al.* (2020) 'Molecular Architecture of the SARS-CoV-2 Virus', *Cell*, 183(3), pp. 730–738. doi:10.1016/j.cell.2020.09.018.

Yi, C. *et al*. (2020) 'Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies', *Cell Mol Immunol*, 17(6), pp. 621–630. doi:10.1038/s41423-020-0458-z.

Yu, W. *et al*. (2012) 'Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations', *J Comput Chem*, 33(31), pp. 2451–2468. doi:10.1002/jcc.23067.

Zhan, Y.Y. *et al*. (2018) 'Induced-fit expansion and contraction of a self-assembled nanocube finely responding to neutral and anionic guests', *Nature Communications*, 9(1), pp. 1-6. doi: 10.1038/s41467-018-06874-y.

Zhao, P. *et al*. (2020) 'Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor', *Cell host & microbe*, 28(4), pp. 586–601. doi:10.1016/j.chom.2020.08.004.