Tropical Journal of Natural Product Research

Available online at https://www.tjnpr.org

Original Research Article

Mutation Pattern in the Receptor Binding Motif of SARS-Cov-2 Variants and the Effect on Molecular Interactions in Docked Ligand Complexes

Israel E. Ebhohimen[1], Ojei H. Onyijen[2], Vaishali Arora[3], Vanisha Arora[4], Victoria T. Adeleke[5], Moses Okpeku[6]*

[1]*Department of Chemical Sciences, Samuel Adegboyega University, Ogwa, Edo State, Nigeria*
[2]*Department of Mathematical and Physical Sciences, Samuel Adegboyega University, Ogwa, Edo State, Nigeria*
[3]*Department of Botany, University of Delhi, India*
[4]*Department of Botany, Osmania University, Hyderabad, India*
[5]*Department of Chemical Engineering, Mangosuthu University of Technology, Umlazi, South Africa*
[6]*Discipline of Genetics, School of Life Sciences, University of KwaZulu-Natal, Westville Campus, Durban, South Africa*

ARTICLE INFO

ABSTRACT

The spike glycoprotein of SARS-Cov-2 is a therapeutic target for Covid-19 and mutations in the Receptor Binding Motif (RBM) may alter the binding properties of ligands proposed to inhibit viral entry. This study aimed to identify the existence of a mutation pattern in the RBMs of SARS-Cov-2 variants and study the effect on ligand binding interactions. RBM sequences were obtained using NCBI BLASTP and subjected to multiple and pairwise sequence alignment analysis. Hypothetical generations were drawn from the phylogenetic tree. The effect of mutation on ligand binding was studied by docking zafirlukast on selected RBMs. Molecular dynamics simulations were conducted to explain molecular interactions. The sequences at the same phylogenetic level showed higher similarity with the observed differences defined by the crystallized chain length. 6XDG_E, a leaf node sequence was 76% similar to 7NXA_E, a branch from the root, and had the highest mutation. Differences in sequence similarity across successive generations were based on mutations and crystallized chain length and the amino acid substitution is not predictable. Different bond types and binding affinities were observed as well as varying Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and Region of Gyration (RoG) values for the RBMs in different variants. The RMSD, RMSF, and RoG did not differ significantly in the bound and free states of RBM from specific variants suggesting that the observed differences are attributable to amino acid substitutions. This information is crucial for drug development intended to block SARS-Cov-2 entry.

Keywords: SARS-Cov-2 spike glycoprotein, receptor binding motif, Covid-19, Molecular dynamics simulation

Introduction

The coronavirus, SARS-CoV-2, discovered in Wuhan, China in December 2019 is the pathogen responsible for Covid-19 that was declared a pandemic by the World Health Organization on January 30, 2020.[1] SARS-COV-2 is a member of the coronavirus family, and the genome is a non-segmented positive single-stranded RNA, with a size of 27 - 32 kilobases. The viral genome encodes four main structural proteins; spike (S), nucleocapsid (N), membrane (M), and envelope (E), required to make complete viral particles.[1,2] Several therapeutic approaches have been proposed, such as modulation of immune defense, blocking viral entry, interfering with the endocytic pathway, targeting the cellular signaling pathway, and blocking polyprotein post-translational processing.[3]
To block viral entry, inhibiting spike glycoprotein binding has been established as a viable therapeutic window.

*Corresponding author. E mail: OkpekuM@ukzn.ac.za
Tel: +27641286372

The spike glycoprotein is the first contact with the host cell and is crucial for various processes, such as attachment, receptor binding, membrane fusion via conformational changes, internalization of the virus, and host tissue tropism.[4] Serological analyses of almost 650 individuals infected with SARS-CoV-2 indicated that ~90% of the plasma or serum neutralizing antibody activity targets the spike receptor binding domain.[5]
The spike protein comprises the S1 and S2 subunits; the S1 subunit contains a signal peptide (SS), followed by an N-terminal domain (NTD) and receptor-binding domain (RBD) comprising residues 333–438 and residues 507–527, and a core Receptor Binding Motif (RBM), residues 438–506. The virus recognizes the receptor through the receptor-binding motif (RBM), which binds to the outer surface of the angiotensin-converting enzyme 2 (ACE2).[6–8] The S2 subunit contains conserved fusion peptide (FP), heptad repeat 1 (HR1), a central helix (CH), a connector domain (CD), heptad repeat 2 (HR2), a transmembrane domain (TM), and a cytoplasmic tail (CT).[8]
Coronaviruses are RNA viruses thus are subject to mutations which represents the first step in the diversification of viruses in nature and may be useful for adaptation to changing environments.[9,10] RNA viruses have been classified as minimalists in terms of replication fidelity.[10,11]
Single or multiple mutations in RBM are of great concern as they may reduce the efficacy of drugs proposed to block viral entry.[12] If a missense mutation induces a synonymous amino acid substitution, the effect of the alteration will be less severe. However, if it is a non-synonymous amino acid substitution, it is likely to produce severe changes in protein stability, structure, and dynamics.[13] If these

mutations occur close to or in the active site, it may alter the binding properties.

This research aimed to determine if amino acid substitutions in the RBM of spike glycoproteins of SARS-Cov-2 variants follow a particular pattern and study the effect of the substitutions on ligand binding.[12,14]

Materials and Methods

To study the effect of binding therapeutic agents, RBMs from variants with ≥ 97% similarity were selected. Since no antiviral pills were approved for use at the time this study was conceptualized, zafirlukast reported to bind to the RBM was selected solely to explain molecular interactions between a therapeutic agent and the RBM of SARS-Cov-2 spike glycoprotein.[13]

Data retrieval
Structural and sequence data of published SARS-Cov-2 spike glycoprotein (Protein Data Bank ID: 7NXA) obtained by x-ray diffraction, with a resolution factor of 2.50 Å was downloaded from Protein Data Bank (PDB). 7NXA is the crystal structure of SARS-CoV-2 B.1.351 variant spike glycoprotein in complex with COVOX-222 and EY6A Fabs. It comprises 5 unique protein chains, a residue count of 1084 amino acids, and a total structure weight, of 120.08 kDa. 7NXA_E [auth E] (S1) comprises the RBM with 205 residues.[15] This randomly selected spike glycoprotein RBM was used as a reference sequence to select closely related variants to reduce the effect of multiple amino acid substitutions on protein structure. The 7NXA was cleaned using Discovery Studio Visualizer 20 to obtain 7NXA_E. The variants were classified into generations based on clades formed in the evolutionary sequence. Zafirlukast was docked on a variant in every four serial clades and the sample sequence 7NXA_E.

Orthologous analysis
The FASTA sequence of 7NXA_E was used to obtain similar sequences using BLASTP on NCBI database (description = spike protein S1 SARS-CoV-2; percent identity ≥ 97%; E value = 2e-151 – 9e-135). The multiple sequence alignment was done using the Clustal Omega web server (https://www.ebi.ac.uk/Tools/msa/clustalo/). The percentage similarity was calculated and the phylogenetic tree was drawn using Ugene Pro 39.0. The spike glycoprotein RBMs were categorized into hypothetical generations based on clades formed by internal nodes represented with the numbers; 1, 2, 3, 4 and a sequence from the branch in the root node (Supplementary Figure 2) in the phylogenetic tree. The percentage similarity between the 6XDG_E, a leaf node sequence, and the older generations, as well as the similarity between successive generations, were determined by pairwise sequence analysis on Ugene Pro 39.0.[16]

Molecular docking
Zafirlukast, a chemical agent with high binding affinity for the RBM of SARS-CoV2 spike glycoprotein[4] was docked on auto-grids (XYZ = -51.784182 -128.297091 1.370636 [7CM4]; 60.726545 -36.543818 36.588182[7DMU]; 8.719091 -10.637455 23.514545[7M7W]; 29.498091 -55.342727 -21.959364[7NX9]; 29.053455 -55.279182 -22.050455[7NXA]) around ARG357 in the RBM that was reported as an interacting amino acid residue using Discovery Studio Visualizer 20. Docking was done using AutoDock Vina 1.1.2 and the best pose for each complex was documented.

Molecular dynamics simulation
The proteins and ligands in the docked complexes were prepared for molecular dynamics simulation using UCSF Chimera 1.14. Molecular dynamics simulation was done with the AMBER 14 package. The input topologies were generated using the LEAP module of AMBER 14. This was done by introducing ions (Cl⁻) into the solvation box of water molecules (8 Å). The energy minimization to obtain the lowest energy for high energy configurations in the protein was done. This step was initially performed with 10,000 steps (500 steepest descents with 9500 conjugate gradient) and followed by full minimization of 2000 steps. The system was gradually heated (from 0 to 300 K) for 2 ns in a canonical ensemble (NVT) with a Langevin thermostat. The collision frequency applied to the system was 1.0 ps^{-1}, with the density of the water system regulated with 4 ns of NPT simulation. The molecular dynamics production was run 200 ns of NPT (constant number N, pressure P, and temperature T), where equilibration of the system was reached at 300 K for another 2 ns at a pressure of 1 bar. After molecular dynamic simulation, the PTRAJ and CPPTRAJ modules in AMBER 14 were used to analyse parameters: Root Mean Square Deviation (RMSD), Root Mean Square Fluctuations (RMSF), and Region of Gyration (RoG)[17].

Results and Discussion

Multiple sequence analysis
The analysis of similar biological sequences termed Multiple Sequence Alignment (MSA) is an important step for homology modeling and is used to infer evolutionary relationships between sequences. From NCBI BLASTP, fifteen homologous sequences of SARS-Cov-2 RBDs were downloaded and analyzed. The results of the MSA revealed a point mutation in the sequences (Supplementary Figure 1). The docked sequences es are presented in Figure 1.

Phylogeny
The evolutionary relationships between the sequences and the genetic distances are presented in Supplementary Figure 2. The percentage similarity of the sequences in successive generations was obtained by pairwise sequence analyses (Supplementary Table 1). Although chain length was a factor in the observed percentage similarity of the sequences, there was no mutation in the latest generations within the aligned regions. However, a higher degree of amino acid substitution was seen when 6XDG_E, a sequence in the leaf node was compared to a sequence in a branch from the root (Supplementary Table 2). The variants in the longest lineage revealed by the phylogenetic tree were categorized into generations based on clades formed by internal nodes represented with the numbers; 1, 2, 3, and 4 (Supplementary Figure 2).

Figure 1: Multiple sequence alignment of docked sequences

Table 1: Average RMSD, RMSF and RoG Values for RBMs in the Bound and Free State and Binding Affinity of Docked Complexes during MDS

Variant	Root Mean Square Deviation (Å)		Root Mean Square Fluctuation (Å)		Region of Gyration (Å)		Affinity (kcal/mol)
	Bound (Mean ± SEM)	Free (Mean ± SEM)	Bound (Mean ± SEM)	Free (Mean ± SEM)	Bound (Mean ± SEM)	Free (Mean ± SEM)	
7CM4	1.6 ± 7e-4	1.32 ± 7e-4	13.29 ± 0.41	12.62 ± 0.32	31.18 ± 6.78e-5	31.19 ± 5.44e-5	-34.58±0.13
7DMU	1.78 ± 5e-4	1.71 ± 7e-4	13.45 ± 0.33	13.33 ± 0.34	33.36 ± 5.84e-5	33.28 ± 5.67e-5	-20.35±0.10
7M7W	1.80 ± 5e-4	1.70 ± 8e-4	15.28 ± 0.46	13.74 ± 0.43	35.61 ± 5.65e-5	35.63 ± 5.61e-5	-26.45±0.10
7NX9	1.74 ± 5e-4	1.21 ± 4e-4	11.22 ± 0.32	10.27 ± 0.28	31.58 ± 6.07e-5	31.60 ± 5.76e-5	-19.50±0.10
7NXA	1.54 ± 4e-4	1.63 ± 4e-4	13.44 ± 0.36	15.89 ± 0.48	31.49 ± 7.51e-5	31.30 ± 5.26e-5	-30.40±0.07

Molecular docking

The docked complexes showing interacting amino acid residues in the RBMs from selected variants are presented in Figure 2. The docking simulation revealed that the interaction of zafirlukast with the RBMs varied in binding affinity, bond type, and interacting amino acid residues. The binding affinities of the best poses for each complex are presented in Table 1 and the bond type for each amino acid residue is presented in Figure 2.

The error-prone replication of viruses in host cells has been reported as a survival strategy and may be associated with disease manifestations. It is of concern if these mutations occur in the binding region or residues that contribute to the binding interaction.[18,19] Point mutations alter the binding properties of proteins. Missense mutations can also alter protein-protein interactions,[20] protein stability,[21] and the characteristics of the active site.[22]

The spike glycoprotein due to its significance in the SARS-CoV-2 binding to the human ACE2[4] has been reported as a therapeutic window for the treatment of Covid-19.[5,12] However, mutations have been reported to occur in the spike glycoprotein including the human ACE2 RBD. This could alter infectivity, the severity of disease, host immunity[14] and the binding of therapeutic ligands.[4,12,18]

The crystal structure of SARS-CoV-2 (variant B.1.351) spike glycoprotein comprising five unique protein chains with a residue count of 1084 amino acids was used as the base sequence for this study. The RBM comprising amino acid residues 324 - 526 (7NXA_E [auth E])[15] was extracted and the FASTA sequences of RBMs were analysed by multiple sequence analysis (Supplementary Figure 1). The phylogenetic tree showed that the variants used for the study are genetically related (Supplementary Figure 2). The RBM variants were then categorized into hypothetical generations (Supplementary Figure 2) and varying percentage dissimilarities were revealed by pairwise sequence analysis (Supplementary Table 1). From the results, the most common mutations were at residues 427, 488, 494, and 511. In each position, amino acid substitutions were observed to be restricted (only a pair of amino acids substituted at each point). These amino acid substitutions and crystalized chain length were responsible for the observed percentage differences in the sequence analysis. The observed point mutations did not happen sequentially across the generations of variants studied.

If the observed mutations occurred in a particular pattern across the lineage, then it will be possible to predict future mutations as well as predict the effect on the binding property. This will aid the development of chemical agents that can inhibit the ACE2 binding of predicted specific variants or multiple variants.[23] Pairwise sequence analyses of a leaf node sequence with preceding generations (Supplementary Table 2) revealed more mutations compared to the pairwise sequence analyses of successive generations. Sequences at the same level in the phylogenetic tree showed higher percentage similarity and fewer amino acid substitutions in the aligned regions. The leaf node in the phylogeny studied showed no amino acid substitutions in the same region. A similar pattern was seen in the earliest generation. This may be useful to an extent in the prediction of mutation patterns. However, the absence of a direct mutation pattern that indicates non-predictability may complicate the development of therapeutic agents targeting viral entry by binding the RBD.[20–22]

Molecular docking was used to study molecular interactions between zafirlukast and the RBM variants. The observed binding affinities (Table 1) were higher than the affinity (-10.3 kcal/mol) reported for the whole spike glycoprotein up conformation reported by Ramirez-Salinas and co-workers.[4] The range of binding affinities observed was significantly higher than the values reported by Nag and co-workers who studied the effect of mutations in the RBM on the binding affinity for therapeutic agents.[12] In that study, a homology model of the RBM was built based on the reported mutations in spike glycoprotein variants reported from different regions of the world.

Several intermolecular forces are involved in molecular interactions between ligands and receptors. These forces have different thermodynamic profiles which may directly affect specificity and affinity.[24] To improve the correctness of the outcome, the XYZ coordinates for the binding interaction in the docking study were set around ARG357 for all RBMs. The results presented in Figure 2 and Table 2 show the interacting residues and their energy decomposition, indicating that point mutations in the RBM can affect the binding of chemical agents to several SARS-CoV-2 variants.

Table 2: Total energy decomposition (Generalized Born solvent) of the interacting spike glycoprotein RBM-zafirlukast complex during molecular dynamics simulation

Variant	Interacting amino acid residue	Decomposition Energy (kcal/mol)
7CM4	TYR396	-1.01 ± 0.54
	PRO426	-1.45 ± 0.62
	PHE429	-1.51 ± 0.54
	THR430	-1.45 ± 0.71
	PHE464	-1.62 ± 0.57
	PHE515	-1.06 ± 0.78
	LEU517	-2.15 ± 1.89
7DMU	ARG355	-1.54 ± 1.38
	ARG357	-1.17 ± 1.33
7M7W	TYR396	-2.90 ± 1.40
	PHE464	-1.57 ± 0.61
	GLU516	-1.69 ± 1.25
	LEU517	-1.33 ± 1.05
7NX9	LEU517	-1.12 ± 1.37
7NXA	TYR396	-1.92 ± 0.58
	PHE429	-1.31 ± 0.60
	THR430	-1.85 ± 0.88
	GLU516	-1.29 ± 0.77
	LEU517	-1.07 ± 0.62

Molecular dynamics simulations were performed to evaluate complex stability and interactions between the docked complexes.[17] The atom-positional root-mean-square deviation (RMSD), a measure for the positional divergence of one or multiple atoms, is one of the most commonly used plot types in the field of biophysical simulations. The RMSD of all complexes were comparable indicating stable interactions of zafirlukast with the RBM variants (Figure 2). The positional root-mean-square fluctuation (RMSF) represents the degree of variation of a given atom over time. The RMSF values were plotted per residue for the RBMs in the bound and unbound states (Figure 2). Atomic fluctuations varied significantly in both bound and free RBMs between residues 450 -500 (Figure 2, Supplementary Figure 4).

The RoG of the ligand-protein complexes was also studied to ascertain dynamic adaptability and compactness in the aqueous environment. The RoG for all complexes were within very small ranges ~0.2 – 0.3 Å (Figure 2, Supplementary Figure 5) which may account for the relative stability revealed by the RMSD.[12] The average values obtained for RMSD, RMSF, and RoG for the RBMs in the bound and free states are presented in Table 1.

Amino acid substitutions in the binding site alter binding characteristics.[20–22] The results obtained from this study revealed that the substitutions altered bond type and binding residues in the complex. The effect of mutation on the structure, function, and dynamics of the receptor-binding domain of human SARS-CoV-2 with host cell receptors has been reported. These mutations alter the flexibility of the RBM and alter significant binding characteristics.[25] A similar study on breast cancer cells showed that for proteins where mutations occur, drug binding characteristics are altered.[8] The binding affinity and specificity of a drug molecule for its target depend on the type of attractive forces as well as the energy of reacting residues. In this study, the total energy decomposition of binding residues in the RBMs and the bond types varied (Table 2; Figure 2). The RMSD, RMSF, and RoG values for the variants in the bound and free states (Supplementary Figure 3) were not significantly different in each case. This suggests that the binding of a ligand to the active site did not significantly alter molecular interactions in the RBM. The observed variations in the parameters across the variants (Figure 3) are attributable to amino acid substitutions in the RBM.[12,13]

Despite the very close genetic relationship (Supplementary Figure 2) between the RBMs used in this study, differences were observed in the molecular interactions in complexes formed between the RBMs and zafirlukast. The therapeutic efficacy of a single compound against several variants when the inhibition of SARS-CoV-2 host cell entry is the target may not be equal. Furthermore, the plethora of compounds that target blocking viral entry that have been reported by several authors in older generations of the virus may not be active against current and future variants of concern. To design drugs that can effectively block viral entry, the effect of mutations in the spike glycoprotein RBD on the protein structure, binding affinity, specificity, and molecular interactions between binding residues must be taken into consideration.
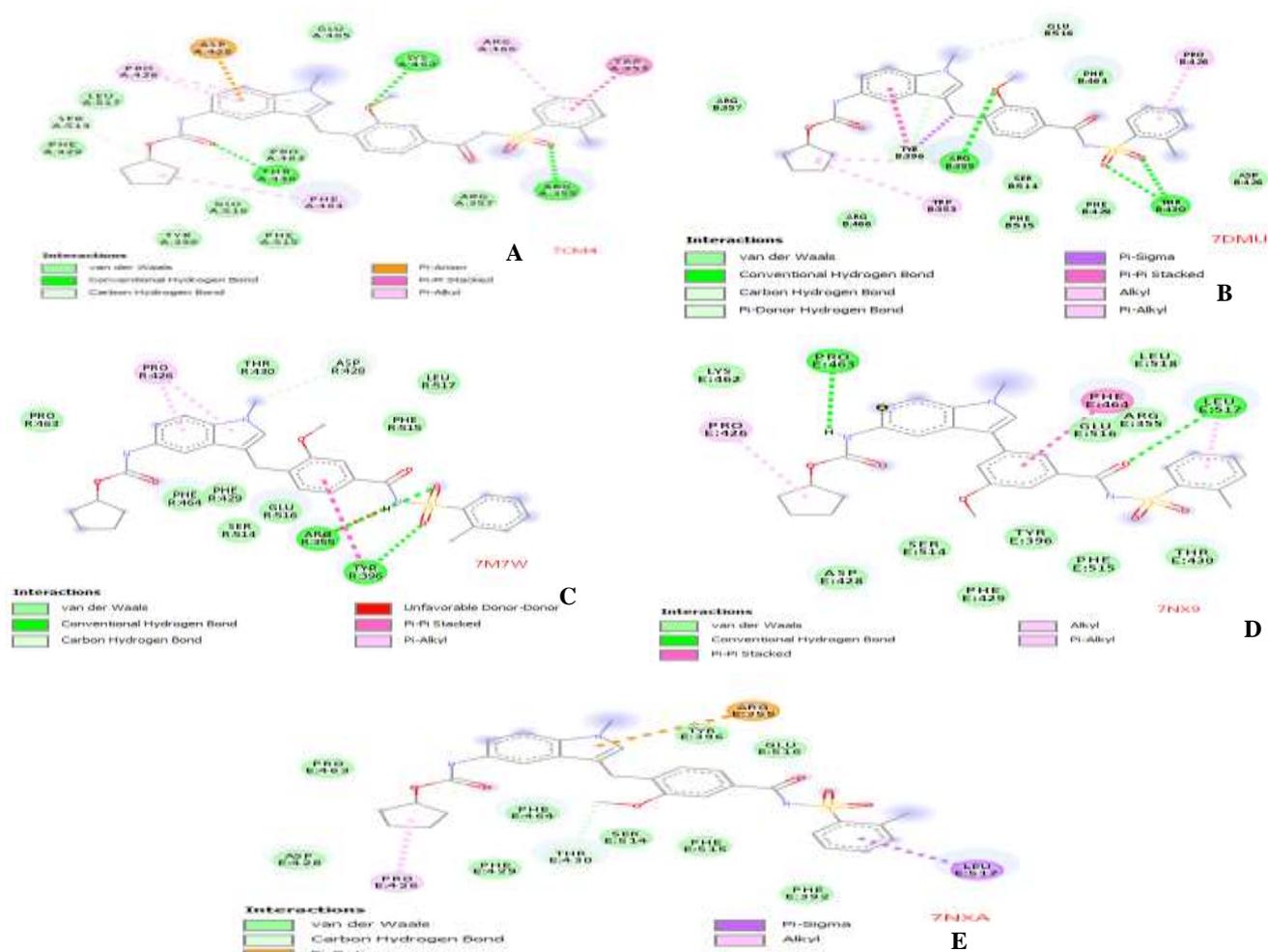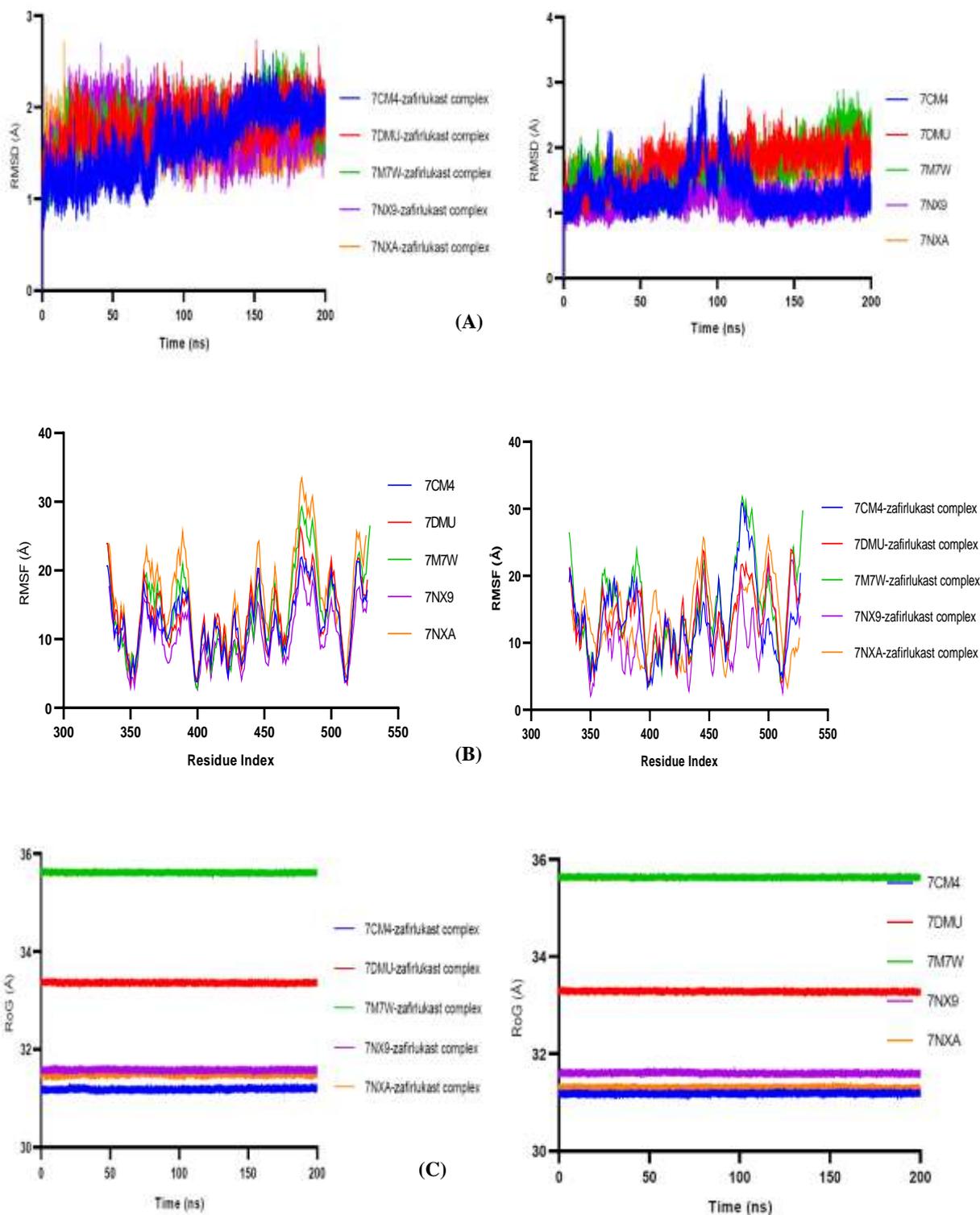

Figure 2: Molecular interactions between zafirlukast and spike glycoprotein RBMs. (A) 7CM4- Zafirlukast complex (B) 7DMU- Zafirlukast complex (C) 7M7W- Zafirlukast complex (D) 7NX9- Zafirlukast complex (E) 7NXA- Zafirlukast complex

(A)

(B)

(C)

Conclusion

The result suggests that an equal therapeutic effect may not be achieved by a single therapeutic agent on several SARS-Cov-2 variants since the differences in molecular interactions are attributable to the amino acid substitutions and not the bound compound. Therapeutic activity may only be achieved repeatedly if a therapeutic agent is applied to the same variant

Conflict of Interest

The authors declare no conflict of interest.

Authors' Declaration

The authors hereby declare that the work presented in this article is original and that any liability for claims relating to the content of this article will be borne by them.

References

1. Onyijen OH, Hamadani A, Awojide S, Ebhohimen IE. Prediction of deaths from Covid-19 in Nigeria using various machine learning algorithms. SAU Sci-Tech J. 2021; 6(1):109-117. https://www.journals.sau.edu.ng/index.php/sjbas/article/view/491/352

2. Palacios Cruz M, Santos E, Velázquez Cervantes MA, León Juárez M. COVID-19, a worldwide public health emergency. Rev Clin Esp. 2021; 221(1):55-61.

3. Nitulescu, G. M., Paunescu, H., Moschos, S. A., Petrakis, D., Nitulescu, G., Ion, G. N. D., Spandidos, D. A., Nikolouzakis, T. K., Drakoulis, N., Tsatsakis, A. Comprehensive analysis of drugs to treat SARS-CoV-2 infection: Mechanistic insights into current COVID-19 therapies (Review). Int J Mol Med. 2020; 46(2):467-488.

4. Ramírez-Salinas GL, Martínez-Archundia M, Correa-Basurto J, García-Machorro J. Repositioning of ligands that target the spike glycoprotein as potential drugs for sars-cov-2 in an in silico study. Molecules. 2020;25(23).

5. Piccoli L, Park YJ, Tortorici MA, Czudnochowski N, Walls AC, Beltramello M, Silacci-Fregni C, Pinto D, Rosen LE, Bowen JE, Acton OJ, Jaconi S, Guarino B, Minola A, Zatta F, Sprugasci N, Bassi J, Peter A, De Marco A, Nix JC, Mele F, Jovic S, Rodriguez BF, Gupta SV, Jin F, Piumatti G, Lo Presti G, Pellanda AF, Biggiogero M, Tarkowski M, Pizzuto MS, Cameroni E, Havenar-Daughton C, Smithey M, Hong D, Lepori V, Albanese E, Ceschi A, Bernasconi E, Elzi L, Ferrari P, Garzoni C, Riva A, Snell G, Sallusto F, Fink K, Virgin HW, Lanzavecchia A, Corti D, Veesler D. Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. Cell. 2020; 183(4):1024-1042.e21.

6. Omotuyi IO, Nash O, Ajiboye OB, Iwegbulam CG, Oyinloye EB, Oyedeji A, Kashim ZA, Okaiyeto K. Atomistic simulation reveals structural mechanisms underlying D614G spike glycoprotein-enhanced fitness in SARS-COV-2. J Comput Chem. 2020;41(24):2158-2161.

7. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell. 2020;181(2):281-292.e6.

8. Wan S, Kumar D, Ilyin V, Ussama Al H, Gulab S, Alexander K, Peter VC. The effect of protein mutations on drug binding suggests ensuing personalised drug selection. Sci Rep. 2021;11:1-10.

9. Domingo E and Perales C. Viral quasispecies. PLoS Genet. 2019; 15(10):1-20.

10. Smith EC, Sexton NR, Denison MR. Thinking Outside the Triangle : Replication Fidelity of the Largest RNA Viruses. Annu Rev Virol. 2014; 1:111-132.

11. Jácome R, Vidal YL, León SP De, Darwin C. Are RNA Viruses Candidate Agents for the Next Global Pandemic ? A Review. 2017; 58(3):343-358.

12. Nag A, Paul S, Banerjee R, Kundu R. In silico study of some selective phytochemicals against a hypothetical SARS-CoV-2 spike RBD using molecular docking tools. Comput Biol Med. 2021;137(August):104818.

13. Zhang Z, Miteva MA, Wang L, Alexov E. Analyzing effects of naturally occurring missense mutations. Comput Math Methods Med. 2012; 2012:1-15.

14. Harvey WT, Carabelli AM, Jackson B, Ravindra KG, Emma CT, Ewan HM, Catherine L, Richard R, Andrew R, Sharon PJ, David RL. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol. 2021; 19(7):409-424.

15. Dejnirattisai W, Zhou D, Supasa P, Liu C, Mentzer AJ, Ginn, HM, Zhao Y, Duyvesteyn HME, Tuekprakhon A, Nutalai R, Wang B, López-Camacho C, Slon-Campos J, Walter TS, Skelly D, Costa Clemens SM, Naveca FG, Nascimento V, Nascimento F, Fernandes da Costa C, Resende PC, Pauvolid-Correa A, Siqueira MM, Dold C, Levin R, Dong T, Pollard AJ, Knight JC, Crook D, Lambe T, Clutterbuck E, Bibi S, Flaxman A, Bittaye M, Belij-Rammerstorfer S, Gilbert SC, Carroll MW, Klenerman P, Barnes E, Dunachie S, Paterson NG, Williams MA, Hall DA, Hulswit JG, Bowden TA, Fry EE, Mongkolsapaya J, Ren J, Stuart DI, Screaton GR. Antibody evasion by the P.1 strain of SARS-CoV-2. Cell. 2021; 184(11):2939-2954.e9.

16. Okonechnikov K, Golosova O, Fursov M, Varlamov A, Vaskin Y, Efremov I, German Grehov OG, Kandrov D, Rasputin K, Syabro M, Tleukenov T. Unipro UGENE: A unified bioinformatics toolkit. Bioinformatics. 2012; 28(8):1166-1167.

17. Madlala T, Adeleke VT, Fatoba AJ, Okpeku M, Adeniyi AA, Adeleke MA. Designing multiepitope-based vaccine against Eimeria from immune mapped protein 1 (IMP-1) antigen using immunoinformatic approach. Sci Rep. 2021; 11(1):1-18.

18. Domingo E, Garc C, Lobo-vega R. Proofreading-Repair Activities in RNA Virus Genetics. Viruses. 2021; 13(1882):1-15.

19. Dixit A, Torkamani A, Schork NJ, Verkhivker G. Computational modeling of structurally conserved cancer mutations in the RET and MET kinases: The impact on protein structure, dynamics, and stability. Biophys J. 2009; 96(3):858-874.

20. Teng S, Madej T, Panchenko A, Alexov E. Modeling effects of human single nucleotide polymorphisms on protein-protein interactions. Biophys J. 2009; 96(6):2178-2188.

21. Zhang Z, Teng S, Wang L, Schwartz CE, Alexov E. Computational analysis of missense mutations causing Snyder-Robinson syndrome. Hum Mutat. 2010; 31(9):1043-1049.

22. Zhang Z, Norris J, Schwartz C, Alexov E. In silico and in vitro investigations of the mutability of disease-causing missense mutation sites in spermine synthase. PLoS One. 2011; 6(5):1-10.

23. Salama MA, Hassanien AE, Mostafa A. The prediction of virus mutation using neural networks and rough set techniques. EURASIP J Bioinforma Syst Biol. Published online 2016.

24. Kawasaki Y and Freire E. Finding a better path to drug selectivity. Drug Discov Today. 2011; 16(21-22):985-990.

25. Dehury B, Raina V, Misra N, Suar M. Effect of mutation on structure , function and dynamics of receptor binding domain of human SARS-CoV-2 with host cell receptor ACE2 : a molecular dynamics simulations study. J Biomol Struct Dyn. Published online 2020: 1-15. https://doi.org/10.1080/07391102.2020.1802348