

The N15Y mutation in the E protein of severe acute respiratory syndrome coronavirus 2 affects the binding of doxycycline and rutin

Ali Adel Dawood^{1*} ORCID: <https://orcid.org/0000-0001-8988-5957>

Bassam Ismael Jasim² ORCID: <https://orcid.org/0000-0002-1144-7444>

Omar Al-jalily¹ ORCID: <https://orcid.org/0000-0002-4653-0141>

¹ Department of Anatomy, College of Medicine, University of Mosul. Mosul. Iraq.

² Department of Medicine, College of Medicine, University of Nineveh, Mosul, Iraq.

email: aad@uomosul.edu.iq

The severe acute respiratory syndrome coronavirus 2 genetic variation must be closely monitored. Viral transmission can inevitably result from mutations in the viral genome and functional proteins that aid in the virus's adaptation to the host. This study aimed to look for mutations in the E protein and see how they affected the ligands' molecular docking. Mathematical saturation mutagenesis and other informatics techniques were used. Fourteen severe acute respiratory syndrome coronavirus 2 isolates from Iraq were selected. Doxycycline and rutin were chosen as ligands. In four strains of severe acute respiratory syndrome coronavirus 2, the N15Y mutation was detected in the envelope protein. Depending on the calculation of the amount of energy of the atoms, this mutation is critical in modifying the shape of the protein as well as increasing protein stability. In the single-chain mutation, one pocket was determined, while all pentamer chains had two pockets. The N15Y mutation altered the degree of doxycycline binding by affecting the residue of attachment of the ligands. It also altered the position of the rutin's attachment to the E protein, which has a clear impact on the virion particle.

Keywords: COVID-19; coronavirus; mutation; doxycycline; rutin.

Introduction

Coronaviridae family has resurfaced as a result of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic.⁽¹⁾ It has been shown that a minor alteration in the SARS-CoV-2 genome will cause a significant change in the structure of drug target proteins, rendering some currently available drugs ineffective.⁽²⁾ As a result, in this current pandemic, it is critical to monitor SARS-CoV-2 genetic diversity. Mutations in the viral genome and functional proteins that help in the virus's adaptation to a new host would inevitably result in such a case of virus dissemination.⁽³⁾ Many attempts have been made to combat the SARS-CoV-2 infection and its global expansion. Natural selection works on variants, and these variations also result in novelty.⁽⁴⁾ The ability of the virus to evolve and become more compatible with the host may be determined by the age and gender of those infected, as well as the length of time it takes to recover from the disease.⁽⁵⁾ The nature and frequency of virus mutations remain under study.⁽⁶⁾ The spike (S), membrane (M), and envelope (E) proteins are essential SARS-CoV-2 proteins

that aid in the infection and propagation of the virus within the host cell.⁽⁷⁾

The E protein is a minor component of the virus membrane that plays a role in virus replication and infection. Coronaviruses without the E protein are promising vaccine candidates because this protein is active in essential aspects of the viral life cycle.⁽⁸⁾ A hydrophilic N-terminal domain (NTD 7–12 aa), a broad hydrophobic transmembrane domain (25 aa), and a hydrophilic C-terminal domain (CTD) make up the secondary structure of E protein.⁽⁹⁾

The current study focuses on the E protein because this protein regulates the maturation and retention of the spike protein. We used the available database of SARS-CoV-2 proteins to focus on studying of various mutations in proteins from different isolates in Iraq. Knowing the mutations in SARS-CoV-2's proteins may help to solve the mystery of higher COVID-19 transmission rates that led to the pandemic, as well as provide a helping hand in targeting the virus directly.

* Ph.D. Microbiology, Dept. of Anatomy, College of Medicine, University of Mosul, Mosul, Iraq.

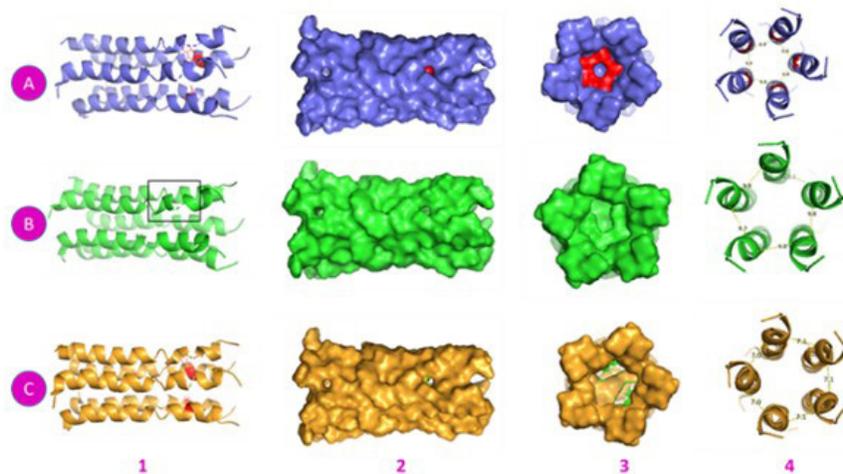


Fig. 2. Prediction of the pentamer E protein (7K3G) and constructing mutation. A: Wild type 1- Cartoon with N15 residue in all chains, red color. 2-Surface side view. 3- Cross-section. 4- Distances between chains of atoms N15 residue. B- One chain mutation of Y15. 1- Cartoon with Y15. 2- Surface side view. 3- Cross-section shows a pocket in Y15 residue. 4- Distances between chains of atoms Y15 residue. C: All chains mutated with Y15. 1- Cartoon with Y15 in all chains, red color. 2- Surface side view shows a large hole. 3- Cross-section presents 2 pockets. 4- Distance between chains of atoms Y15 residue.

In this mutation, the amino acid asparagine is substituted by tyrosine.

When building the 3D structure, PDB (2MM4) was chosen for the E protein. The best recognition scores between template and proteins under analysis were 93%. On the template chosen for the E protein, we constructed a 3D structure of the mutant type. The proportion of modifications in the protein structure and the morphological changes associated with each mutation are depicted in the cartoon and ribbon surface models. The study's results were used to calculate the distance between wild and mutant atoms. The current study revealed that the distance between atoms differs. The distances between the atoms around residue 15 in the wild and mutant strains are different (6.6 Å, 7.6 Å, 3.0 Å, 4.0 Å, and 5.6 Å), (5.9 Å, 5.7 Å, 4.5 Å, 6.2 Å, and 6.1 Å), respectively. The main residue and the whole protein have a less pronounced difference (35.9 Å, 30.3 Å, and 37.5 Å), (35.9 Å, 30.3 Å, and 36.5 Å), respectively.

A pentamer structure of E protein (PDB ID: 7K3G) consists of five chains. The mutation was first loaded onto one chain, then onto all links. By homology modeling and constructing the topology structure, we established that there is a difference in how a mutation affects the overall shape of the protein, especially if the mutation occurs in all chains. The new protein has a completely different topology from the old one, which

has an effect on the protein and, of course, the virion particle (Fig. 2).

Before beginning the molecular fusion process, the E protein (7K3G) was miniaturized, yielding the following: the initial energy is (-47628.5kJ/mol) with a score of (-5.29) and the final energy is (-71100.2 kJ/mol) with a score of (-1.77). According to the docking results between E protein and ligands, the molecular docking of nine models of each type of ligand has a high value when compared to the wild and mutant types. When the root-mean-square deviation (RMSD) is zero, the optimal degree of coalescence energy has been chosen. Model number 1 was selected for each kind because it had the maximum degree of fusion, as indicated in Table 1. The binding affinity of the wild type was greater than the mutant type for rutin: -8.6 kcal/mol, -8.3 kcal/mol respectively. This study found that the mutant type's binding energy for the doxycycline ligand is higher than the wild type's: -8.3kcal/mol and -6.8kcal/mol, respectively.

According to the 2D structure of (7K3G-Doxycycline), the main interacting residues in the E protein are GLU8, THR11, and ASN15 for the wild type. The resultant mutant type (7K3GM-Doxycycline), in which the amino acid asparagine was replaced with tyrosine, shows a shift in the degree of contact. It possesses an unstable bond link, but one with greater energy than in the wild type, and the fusion now encompasses the residues

Table 1. Binding affinity scores of molecular docking models (protein-ligand interaction) and the degree of RMSD for both rutin and doxycycline ligands.

Wild type-Ligand	Binding Affinity	RMSD/ub	RMSD/lb	Mutant type-Ligand	Binding Affinity	RMSD/ub	RMSD/lb
7k3g-Rutin Model 1*	-8.6	0	0	7k3gM-Rutin Model 1	-8.3	0	0
7k3g-Rutin Model 2	-8.5	3.925	1.723	7k3gM-Rutin Model 2	-8.1	4.799	2.071
7k3g-Rutin Model 3	-8.5	2.416	1.682	7k3gM-Rutin Model 3	-7.9	7.403	2.095
7k3g-Rutin Model 4	-8.5	2.425	1.69	7k3gM-Rutin Model 4	-7.8	4.242	2.255
7k3g-Rutin Model 5	-8.2	15.214	13.009	7k3gM-Rutin Model 5	-7.6	4.327	1.818
7k3g-Rutin Model 6	-8.2	15.927	13.441	7k3gM-Rutin Model 6	-7.6	6.98	2.419
7k3g-Rutin Model 7	-8.2	15.504	13.207	7k3gM-Rutin Model 7	-7.5	5.569	2.639
7k3g-Rutin Model 8	-8.2	15.622	13.196	7k3gM-Rutin Model 8	-7.5	7.021	1.87
7k3g-Rutin Model 9	-8.2	15.88	13.453	7k3gM-Rutin Model 9	-7.5	7.191	2.465
7k3g-Doxycycline Model 1**	-6.8	0	0	7k3gM-Doxycycline Model 1	-8.3	0	0
7k3g-Doxycycline Model 2	-6.8	6.532	3.718	7k3gM-Doxycycline Model 2	-7.8	1.787	1.543
7k3g-Doxycycline Model 3	-6.8	4.037	2.462	7k3gM-Doxycycline Model 3	-7.4	5.785	3.029
7k3g-Doxycycline Model 4	-6.8	4.035	2.448	7k3gM-Doxycycline Model 4	-7.3	7.104	2.251
7k3g-Doxycycline Model 5	-6.7	6.404	3.516	7k3gM-Doxycycline Model 5	-7	6.961	3.102
7k3g-Doxycycline Model 6	-6.7	4.885	3.116	7k3gM-Doxycycline Model 6	-7	6.375	2.308
7k3g-Doxycycline Model 7	-6.7	6.05	2.827	7k3gM-Doxycycline Model 7	-6.9	3.586	2.208
7k3g-Doxycycline Model 8	-6.7	5.454	3.742	7k3gM-Doxycycline Model 8	-6.8	2.024	1.509
7k3g-Doxycycline Model 9	-6.3	38.61	36.046	7k3gM-Doxycycline Model 9	-6.7	6.259	2.867

Model 1*: the highest score for 7K3G-Rutin for wild and mutant type. Model 1**: the highest score for 7K3G-Doxycycline for wild and mutant type.

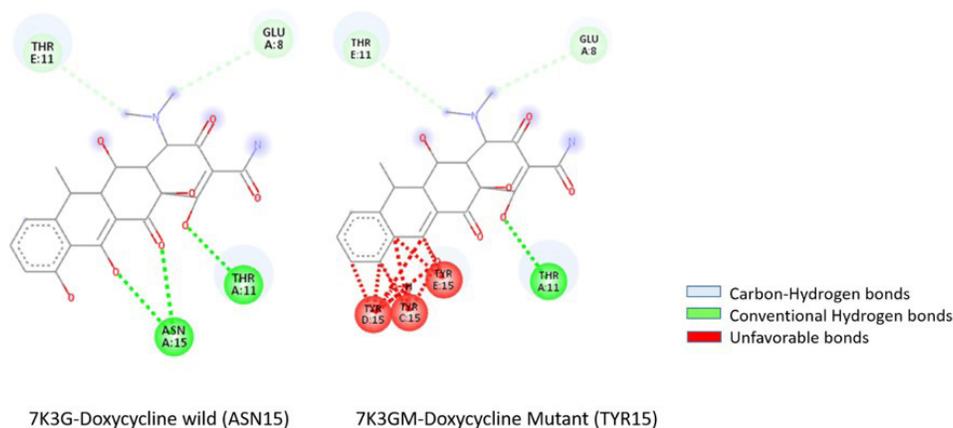


Fig. 3. 2D structure of molecular docking interaction residues showing the types of hydrogen, conventional and unfavorable bonds of the wild strain 7K3G-Doxycycline and mutant strain 7K3GM-Doxycycline.

GLU8, THR11, and TYR15, as seen in Figure 3. VAL25, LEU28, and ALA32 are the major interacting residues in the wild type (7K3G-Rutin). On the other hand, the interaction between residues of mutant strain (7K3GM-Rutin) resulted in a significant alteration in the docking site. The interacting residues (GLU8, THR11, and TYR15) share the same interaction with a mutant type of the doxycycline ligand (Fig. 4). This finding suggests that the new mutation in the E protein has altered the degree of binding and impacted the ligand-binding site.

We detected a conformational change in the shape of the mutant protein, in relation to the wild type, after constructing the 3D structure. In addition, as shown in Figure 5, the degree of conferring hydrogen bonding was greater for the wild type of doxycycline docking. In the case of the ligand rutin, it was registered that a mutation in position N15Y changed the structure of the protein dramatically. In addition to altering the docking location, the mutation residue became part of the docking site, as seen in Figure 6.

Discussion

According to popular belief, the latest COVID-19 pandemic was caused by the cross-species transmission of alpha-coronavirus, commonly found in bats and possibly pangolins, to humans. E protein has been linked to viral entry, replication, and particle assembly in human cells.⁽¹¹⁾ Understanding the development of this modern coronavirus and assuring the performance of new diagnostic tests, vaccinations, and therapies against COVID-19 requires tracking SARS-CoV-2 genetic variation and evolving mutations in this continuing pandemic. The amino acid heterogeneity of the SARS-CoV-2 E and M structural proteins is investigated in this descriptive analysis.⁽¹²⁾

This study detected a mutation in the E protein that significantly impacted the protein's conformational transition. The total percentage of mutation occurring in our current research does not align with previous studies. For example, a study revealed that 99.99% of the 103,419 E sequences studied were conserved.⁽¹³⁾ Furthermore, E protein had fewer amino acid modifications, despite having almost equal retention 99.98%.⁽¹⁴⁾ The mutation rate was less than 1% in all geographic regions except Africa, where 92% of the epiweeks with African sequences available belonged to the most recent epiweeks. According to epiweek's review, there was no consistent growth over time, either nationally or regionally.^(15,16)

A short (7–12 aa) hydrophilic N-terminal domain (NTD), a broad hydrophobic transmembrane domain (25 aa) with a high proportion of valine and leucine, and a hydrophilic C-terminal domain (CTD) make up the coronaviruses E secondary structure.⁽¹⁷⁾ Despite the mutation location changed from helices to coils, these characteristics were found in the secondary structure of the E protein. Consequently, the frequency of residues was altered in the wild type compared to the mutagenic variant.

The structure of the E protein changed dramatically as a result of the mutation. Our results confirm that the rate of mutant variants is 6% from all sequences isolated in Iraq. This demonstrates that the measurements vary between the atoms at the mutation residue and the protein tip site. Furthermore, the amount of energy measured for the atoms at different levels was low for covalent bonds and strong for solvated bonds. The protein becomes more stable as a result of this mutation.

In terms of the E protein, the pentamer model (7K3G) was selected to confirm the impact of the single mutation on the protein's overall form. For a single chain and all chains, the 3D structure of the protein in the mutation state was built. The existence of a new pocket appeared within the protein in the case of a single chain mutation. Still, the analysis discovered two internal pockets as a result of the consequence of the combined five chains mutation. We assume that this effect would impact the virion's overall shape and effects on the host cell.

According to our findings, the docking of the doxycycline revealed the presence of additional connections in the mutant position and a rise in the degree of fusion. This may call attention to the possibility that doxycycline has a direct effect on the virus by enhancing its binding to the E protein. The effect of the mutation was proven in the case of rutin by altering the ligand-protein binding site, which is also helpful in practice.⁽¹⁸⁾

It's worth noting that the mutant type has a lower degree of hydrogen bonding than the wild type for both ligands. We believe the study will continue looking for other ligands that affect the protein, perhaps opening up a new sector in COVID-19 therapy.

Unfortunately, it is impossible to know if the positions of these changes are exposed to the membrane's internal or external hand. In either case, the replacement and deletion are major alterations that may affect conformational properties and potentially protein-protein interactions. More systemic research is needed. However, these

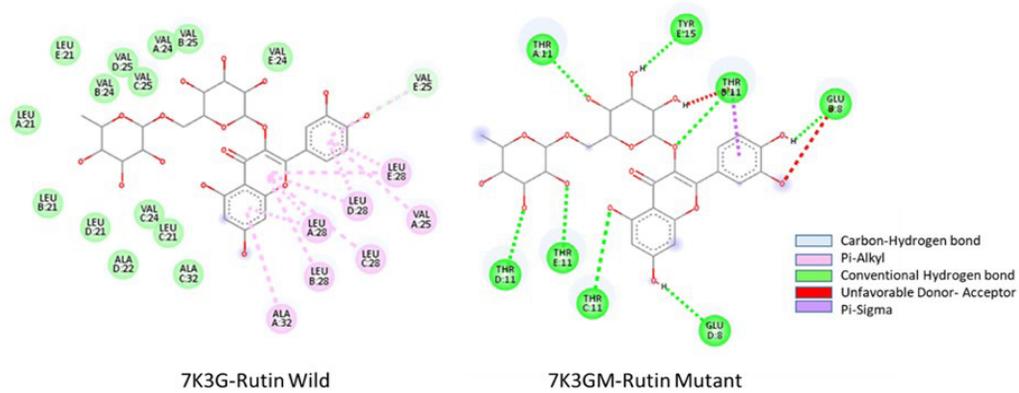


Fig. 4. 2D structure of molecular docking interaction residues showing the types of hydrogen, Pi-Alkyi, unfavorable donor and acceptor, and Pi-Sigma bonds of the wild strain 7K3G-Rutin and mutant strain 7K3GM-Rutin.

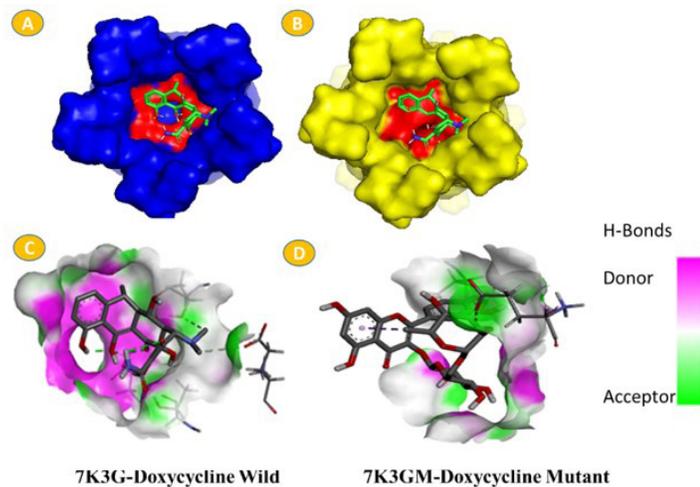


Fig. 5. Molecular docking of the E protein and doxycycline ligand. A: Wild E protein surface shows ASN15 in red color reacting with the doxycycline in green ribbon stick. B: Mutant E protein surface shows TYR15 in red color reacting with doxycycline in green ribbon stick. C: Wild docking residue shows the degree of difference between the donor and the acceptor of the hydrogen bonds. D: Mutant docking residue shows the degree of difference between the donor and the acceptor of the hydrogen bonds.

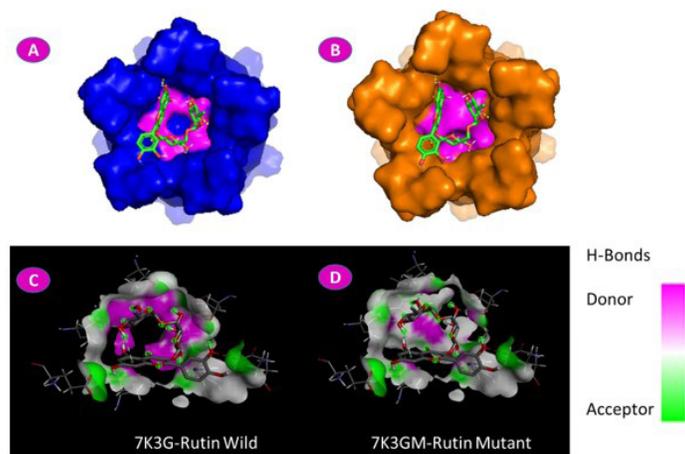


Fig. 6. Molecular docking of the E protein and rutin ligand. A: Wild E protein surface shows ASN15 in magenta color reacting with the rutin in green ribbon stick. B: Mutant E protein surface shows TYR15 in magenta color reacting with rutin in green ribbon stick. C: Wild docking residue shows the degree of difference between the donor and the acceptor of the hydrogen bonds. D: Mutant docking residue.

modifications may affect the oligomerization mechanism that leads to the formation of a transmembrane ion channel.^(19,20)

Understanding amino acid modifications' global and regional effects would require further research into related mutations, virus-host protein interactions, and protein structure. Comparative researches may, therefore, shed light on the molecular processes underlying the emergence of an outbreak of epizootic origin, as well as propose molecular targets for therapeutics or reverse vaccinology experiments.

Conclusions

This study investigated a significant mutation in the E protein that directly impacted the protein's shape and, as a result, on the virion particle as a whole. The mutation altered the degree of binding for doxycycline by directly affecting the ligands' attachment residues. The mutation altered the location of the rutin's attachment to the E protein, which has an impact on the virion particle.

Conflict of interest

The authors declare that there is no conflict of interest.

Acknowledgment

This research was funded by the vice-chancellor for research of Jahrom University of Medical Sciences and Health Services. The authors are appreciative to all the patients and their families for their kind cooperation in this research.

Author's contributions

Ali Adel Dawood: designing the idea, choosing programs, writing the manuscript, using software and writing the results.

Bassam Ismael Jasim: article language review, analyze and clarifying the purpose of the study.

Omar Al-jalily: rearrangement and finishing style of the full text

All authors reviewed and approved the final version of this manuscript.

References

1. Yoshimoto FK. The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2 or n-COV19), the Cause

of COVID-19. *Protein J.* 2020; 39(3):198-216. doi: <https://10.1007/s10930-020-09901-4>.

2. Yang Y, Xiong Z, Zhang S, Yan Y, Nguyen J, Ng B, et al. Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors. *Biochem J.* 2005; 392 (Pt 1):135-43.
3. Venkatagopalan P, Daskalova SM, Lopez LA, Dolezal KA, Hogue BG. Coronavirus envelope (E) protein remains at the site of assembly. *Virology.* 2015; 478:75-85. doi: <https://10.1016/j.virol.2015.02.005>.
4. Ruch TR, Machamer CE. The hydrophobic domain of infectious bronchitis virus E protein alters the host secretory pathway and is important for the release of infectious virus. *J Virol.* 2011; 85(2):675-85. doi: <https://10.1128/JVI.01570-10>.
5. Zheng Y, Zhuang MW, Han L, Zhang J, Nan ML, Zhan P, et al. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) membrane (M) protein inhibits type I and III interferon production by targeting RIG-I/MDA-5 signaling. *Signal Transduct Target Ther.* 2020; 5(1):299. doi: <https://10.1038/s41392-020-00438-7>.
6. Thomas S. The Structure of the Membrane Protein of SARS-CoV-2 Resembles the Sugar Transporter SemiSWEET. *Pathog Immun.* 2020; 5(1):342-63. doi: <https://10.20411/pai.v5i1.377>.
7. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. *Nat Med.* 2020; 26:450-2. doi: <https://10.1038/s41591-020-0820-9>.
8. Bhowmik D, Nandi R, Jagadeesan R, Kumar N, Prakash A, Kumar D. Identification of potential inhibitors against SARS-CoV-2 by targeting proteins responsible for envelope formation and virion assembly using docking based virtual screening, and pharmacokinetics approaches. *Infect Genet Evol.* 2020; 84:104451. doi: <https://10.1016/j.meegid.2020.104451>.
9. Walensky RP, Walke HT, Fauci AS. SARS-CoV-2 Variants of Concern in the United States Challenges and Opportunities. *JAMA.* 2021; 325(11):1037-8. doi: <https://10.1001/jama.2021.2294>.
10. Pandey U, Yee R, Shen L, Judkins AR, Bootwalla M, Ryutov A. High prevalence of SARS-CoV-2 genetic variation and D614G mutation in pediatric patients with COVID-19. *Open Forum Inf Dis.* 2021; 8(6): ofaa551. doi: <https://10.1093/ofid/ofaa551>.
11. Schein CH. Repurposing approved drugs on the pathway to novel therapies. *Med Res Rev.* 2020; 40(2):586-605. doi: <https://10.1002/med.21627>.
12. Shen L, Dien Bard J, Biegel JA, Judkins AR, Gai X. Comprehensive Genome Analysis of 6,000 USA SARS-CoV-2 Isolates Reveals Haplotype Signatures and Localized Transmission Patterns by State and by Country. *Front Microbiol.* 2020; 11:573430. doi: <https://10.3389/fmicb.2020.573430>.
13. Azeez SA, Alhashim ZG, Al Otaibi WM, Alsuwat HS, Ibrahim AM, Almandil NB, et al. State-of-the-art tools to identify druggable protein-ligand of SARS-CoV-2. *Arch Med Sci.* 2020; 16(3):497-507. doi: <https://10.5114/aoms.2020.94046>.
14. Dawood A, Altobje M. Inhibition of N-linked Glycosylation by Tunicamycin May Contribute to The Treatment of SARS-CoV-2. *Microbiol Path.* 2020; 149:104586. doi: <https://10.1016/j.micpath.2020.104586>.

15. Li Q, Wu J, Nie J, Li X, Huang W, Wang Y, et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell*. 2020; 182(5): 1248-94. doi: <https://10.1016/j.cell.2020.07.012>.
16. Dawood A, Altobje M, Alrassam Z. Molecular Docking of SARS-CoV-2 Nucleocapsid Protein with Angiotensin-Converting Enzyme II. *Mikrobiol Z*. 2021; 83(2):82-92. doi: <https://10.15407/mikrobiolj83.02.082>.
17. Elbe S, Buckland-Merrett G. Data, disease, and diplomacy: GISAID's innovative contribution to global health. *Glob Chall*. 2017; 1(1):33-46. doi: <https://10.1002/gch2.1018>.
18. Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Comm*. 2020; 11:1620. doi: <https://10.1038/s41467-020-15562-9>.
19. Dawood A. Increasing the frequency of omicron variant mutations boosts the immune response and may reduce the virus virulence. *Microb Pathog*. 2022; 164: 105400. doi: <https://10.1016/j.micpath.2022.105400>.
20. Shen L, Maglinte DT, Ostrow D, Pandey U, Bootwalla M, Ryutov A, et al. Children's Hospital Los Angeles COVID-19 Analysis Research Database (CARD) - A Resource for Rapid SARS-CoV-2 Genome Identification Using Interactive Online Phylogenetic Tools. *bioRxiv* 2020.05.11.089763. doi: <https://10.1101/2020.05.11.089763>.

La mutación N15Y en la proteína E del coronavirus 2 del síndrome respiratorio agudo severo afecta la unión de doxiciclina y rutina

Resumen

La variación genética del coronavirus 2 del síndrome respiratorio agudo severo debe ser monitoreada de cerca. La transmisión viral puede resultar inevitablemente de mutaciones en el genoma viral y proteínas funcionales que ayudan en la adaptación del virus al hospedero. Este estudio tuvo como objetivo buscar mutaciones en la proteína E y ver cómo afectaban el acoplamiento molecular de los ligandos. Se utilizó la mutagénesis de saturación matemática y otras técnicas informáticas. Se seleccionaron 14 aislamientos del coronavirus 2 del síndrome respiratorio agudo severo en Irak. Se eligieron doxiciclina y rutina como ligandos. En cuatro cepas del coronavirus 2 del síndrome respiratorio agudo severo, se detectó la mutación N15Y en la proteína de la envoltura. Dependiendo del cálculo de la cantidad de energía de los átomos, esta mutación es fundamental para modificar la forma de la proteína y aumentar la estabilidad de la proteína. En la mutación de cadena única, se determinó un bolsillo, mientras que todas las cadenas de pentámero tenían dos bolsillos. La mutación N15Y alteró el grado de unión de doxiciclina al afectar el residuo de unión de los ligandos. También alteró la posición de unión de la rutina a la proteína E, lo que tiene un claro impacto en la partícula del virión.

Keywords: COVID-19; coronavirus; mutación; doxiciclina; rutina.

Recibido: 14 de abril del 2022

Aceptado: 15 de agosto de 2022