

Article

Molecular and Structural Analysis of Specific Mutations from Saudi Isolates of SARS-CoV-2 RNA-Dependent RNA Polymerase and their Implications on Protein Structure and Drug–Protein Binding

Mubarak A. Alamri ^{1,*}, Muhammad Tahir ul Qamar ², Alhumaidi B. Alabbas ¹, Safar M. Alqahtani ¹, Manal A. Alossaimi ¹, Sikandar Azam ³, Muhammad Harris Hashmi ⁴ and Muhammad Shahid Riaz Rajoka ^{5,*}

- ¹ Department of Pharmaceutical Chemistry, College of Pharmacy, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
- ² Integrative Omics and Molecular Modeling Laboratory, Department of Bioinformatics and Biotechnology, Government College University Faisalabad, Faisalabad 38000, Pakistan
- ³ Department of Virology, Graduate School of Medicine, Tohoku University, Sendai 980-8577, Japan
- ⁴ University of Alabama at Birmingham, Birmingham, AL 35294, USA
- ⁵ Department of Microbial Pathogenesis, University of Maryland, Baltimore, MD 21201, USA
- * Correspondence: m.alamri@psau.edu.sa (M.A.A.); shahidrajoka@yahoo.com (M.S.R.R.)



Citation: Alamri, M.A.; Tahir ul Qamar, M.; Alabbas, A.B.; Alqahtani, S.M.; Alossaimi, M.A.; Azam, S.; Hashmi, M.H.; Rajoka, M.S.R. Molecular and Structural Analysis of Specific Mutations from Saudi Isolates of SARS-CoV-2 RNA-Dependent RNA Polymerase and their Implications on Protein Structure and Drug–Protein Binding. *Molecules* **2022**, *27*, 6475. <https://doi.org/10.3390/molecules27196475>

Academic Editors: Xue Zhi Zhao and Terrence R. Burke

Received: 14 August 2022

Accepted: 24 September 2022

Published: 1 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The COVID-19 pandemic caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has stressed the global health system to a significant level, which has not only resulted in high morbidity and mortality but also poses a threat for future pandemics. This situation warrants efforts to develop novel therapeutics to manage SARS-CoV-2 in specific and other emerging viruses in general. This study focuses on SARS-CoV2 RNA-dependent RNA polymerase (RdRp) mutations collected from Saudi Arabia and their impact on protein structure and function. The Saudi SARS-CoV-2 RdRp sequences were compared with the reference Wuhan, China RdRp using a variety of computational and biophysics-based approaches. The results revealed that three mutations—A97V, P323I and Y606C—may affect protein stability, and hence the relationship of protein structure to function. The apo wild RdRp is more dynamically stable with compact secondary structure elements compared to the mutants. Further, the wild type showed stable conformational dynamics and interaction network to remdesivir. The net binding energy of wild-type RdRp with remdesivir is -50.76 kcal/mol, which is more stable than the mutants. The findings of the current study might deliver useful information regarding therapeutic development against the mutant RdRp, which may further furnish our understanding of SARS-CoV-2 biology.

Keywords: SARS-CoV-2; COVID-19; mutations; RdRp; molecular dynamics simulation

1. Introduction

COVID-19 is a major pandemic of the 21st century and is caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) [1,2]. The virus poses a real threat to the global health care system and has resulted in millions of deaths. As of 14:49 CEST on 14 September 2022, WHO had received reports of 607,083,820 confirmed cases of COVID-19, including 6,496,721 deaths. COVID-19 symptoms vary, but mild cases frequently experience fatigue, cough, and fever. Moderate cases may experience mild pneumonia or breathing difficulties, while severe cases may experience organ failure, severe pneumonia and death [3]. Although the COVID-19 disease is managed in many countries around the world, hundreds of cases are still reported every day [4]. Evidence suggests that SARS-CoV-2 mortality can vary significantly depending on geographic location. Differences in viral infection rates can be attributed to a variety of factors, including different national

policies governing people's movement restrictions, isolation and quarantine, as well as differences in genetic population herd immunity [5]. Therefore, substantial efforts are needed to better understand the virus's biology to obtain useful insights for the discovery and development of novel therapeutics.

Different variants of the virus were reported in late 2020, which enabled the virus to spread quickly in the population and made it more virulent compared to the one reported in Wuhan, China [6,7]. The World Health Organization classified these variants as variants of concern and are as follows. The alpha variant (lineage B.1.1.7), which emerged in the United Kingdom in the year 2020 [8]. Notable mutations in the alpha lineage are N501Y and P681H [9]. Initial analysis based on matched case-control studies in the United Kingdom revealed that the alpha variant was not significantly associated with an elevated risk of hospitalization or mortality among infected individuals. However, further investigation revealed that the variant is linked to an increase in severity, and as a result, a 61% increase in mortality rate [10]. The beta variant (lineage B.1.351) was reported from South Africa in the month of May 2020 [11]. The gamma variant was discovered in Brazil and was revealed to be more virulent and associated with transmissibility [12]. According to reports, the Gamma variant may be more severe, with a greater number of younger people presenting with advanced disease and succumbing to the virus. The delta variant (B.1.617.2), a dominant variant with eight S protein changes [13], which emerged in India, was found responsible for killing thousands of people, as the variant was seen with more transmissibility and changed antigenicity [14]. The omicron variant emerged in Botswana in 2021. Several other variants have been witnessed by WHO and are predicted to continue this cycle of genetic changes [15]. Therefore, constant efforts are required to monitor these genetic changes so they can be better controlled.

Vaccines provide protection without the risk of infection and subsequent severe symptoms, and they continue to be the most effective strategy for reducing disease burden and acquiring immune protection against SARS-CoV-2. Updating or developing new COVID-19 vaccines may thus be required to control viral variants. Current mRNA, protein and viral vaccines could be updated by replacing older S protein variants with emerging variants [10]. It is possible that both old and new forms of the S protein could be included in a single vaccine, referred to as a multivalent vaccine [16]. Although vaccinated people are at risk of reinfection, their transmission rates are lower than those of unvaccinated people. Vaccinated people also have lower disease severity, are more likely to recover and require less hospitalization. Some countries have administered booster shots to fully vaccinated people, and they have been shown to reduce severe COVID-19 and SARS-CoV-2 infections [17].

SARS-CoV-2 is an RNA virus, the genome of which is about 29.8 Kb in length [18,19]. The virus genome comprises 14 open reading frames (ORFs) and codes for 29 proteins [20]. The first two ORFs (1a and 1b) comprise two-thirds of the genome and encode polyproteins that are proteolytically cleaved into 16 nonstructural proteins (NSP1–NSP16), including the SARS-CoV2 RNA-dependent RNA polymerase (RdRp; nsp12). The ORF1a encodes for polyprotein 1a (pp1a), which is cleaved into 11 nonstructural proteins (nsp1–nsp11), whereas ORF1b produces nonstructural proteins nsp12–16. The last third of the genome encodes other proteins, including Envelope (E), Membrane (M), Nucleocapsid (N) and Spike (S) protein [21]. The SARS-CoV2 RNA-dependent RNA polymerase (RdRp) from the structural perspective contains multiple domains and catalyzes RdRp synthesis [22]. Due to its vital functionality in the virus's biology and survival plus the absence of its homology sequence in the host, it is an ideal target for the design and development of antiviral drugs [23]. As such, this enzyme has been successfully targeted by the remdesivir drug, which interferes with the enzyme's functionality and is capable of inhibiting the virus [24]. In response to the drug action, the virus mutates the enzyme sequence, which results in decreased binding, making the virus escape from the drug action.

Genomic surveillance is emerging as a vital necessity to achieve effective mitigation and containment. Since SARS-CoV-2 variants have already been identified, it is critical to

obtain reliable evidence about whether they are more virulent, contagious or resistant to COVID-19 vaccines before they spread around the world. Genomic surveillance increases the availability of whole-genome data, advances phylogenetic methods and makes use of next-generation sequencing applications. These methods provide novel ways to detect variants that differ phenotypically or antigenically. Genomic surveillance allows for earlier detection and implementation of effective strategies to mitigate and contain SARS-CoV-2 variants and other novel virus outbreaks [25,26].

In this study, we evaluated several mutations of the RdRp enzyme extracted from Saudi Arabia SARS-CoV-2 isolates and aligned them with the Wuhan RdRp enzyme to highlight key mutations that effect the enzyme structure, function and dynamics and overall virus pathogenicity, transmissibility and disease mortality. The findings of this study were found to be consistent with those of Pachetti et al. [5] and Yashvardhini et al. [27], who described the RdRp of SARS-CoV-2 acquiring drug-resistance properties due to the occurrence of a high frequency of mutations in the RdRp of infected populations worldwide and in India, respectively. The findings will not only give key findings to the society of the virus pathogenic strains circulating in Saudi Arabia but also will guide the discovery of novel drugs.

2. Materials and Methods

2.1. Sequence Retrieval from NCBI Virus Database

The NCBI Virus database (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/find-data/virus>) of NCBI stores sequence and proteomic data of viruses, which can be easily accessed and can be used in biological research. The study was commenced by retrieving all RNA-dependent RNA polymerase isolated from Saudi Arabia SARS-CoV-2. The Wuhan SARS-CoV-2 RdRp was considered as reference for comparative analysis.

2.2. Multiple Sequence Alignment by Clustal Omega Program

Multiple sequence alignment was performed for identification of mutations in the SARS-CoV-2 RNA-dependent RNA polymerase protein. For comparative analysis, the RNA-dependent RNA polymerase sequence from the first sequenced Chinese genome (YP_009724397) was used as a reference. Clustal Omega was used for alignment of the sequences [28]. In order to measure the impact of mutation on the enzyme structure, dynamics and stability, the free-energy difference as well as vibrational entropy energy were estimated. The mutational impact study was conducted using DynaMut program [29]. During this analysis, the crystal structure of polymerase with pdb ID of 6VYO was employed for RBD (receptor-binding domain) modeling. RBD is a virus immunogenic fragment that binds to a specific endogenous receptor sequence to gain entry into host cells, whereas RCSB protein ID: 6WJI was considered for domain dimerization analysis of N-protein. The DynaMut software guides visual understanding of fluctuations in protein structure due to mutations.

2.3. Secondary Structure Predictions

The secondary structure analysis was performed using an online CFSSP software [30]. CFSSP employ Chou–Fasman algorithm for predicting secondary structure elements of given amino sequence, and the output is in the form of alpha helix, beta sheet and turns.

2.4. Structural Homology Modeling

The structure modeling of the mutants was performed using SwissModel [31]. Selection of the template for the modeled structure was carried out based on sequence identity score and query coverage. The Qualitative Model Energy Analysis score (QMEANDisCo Global) for models were 0.88, 0.89 and 0.89 for A97V, P323I and Y606C structures.

2.5. Molecular Docking Studies

Molecular docking was performed using Autodock Vina software [32]. The X-ray structure of RNA-dependent-RNA polymerase (RdRp) with 2.50 Å resolution (BDP ID: 7BV2) [24], as well as the three models of RdRp structures with mutations were used. The protein structures were optimized by removing unwanted cocrystallized water molecules and further saved in PDB format utilizing BIOVIA Discovery Studio Visualizer 2019. The structure of remdesivir was obtained from PubChem (CID: 121304016) in PDB format. The polar hydrogen within the protein and ligand were added by Autodock Tools [33] and then all structures were saved in Autodock PDBQT format. The Autogrid tool was used for the measuring the docking grid box with 1.00 Å spacing and box dimensions of 24 Å × 22 Å × 28 Å (x, y and z) and center of 113.78 Å × 118.474 Å × 132.008 Å (x, y and z). The binding modes and interactions of remdesivir with RdRp binding site were analyzed and visualized using PyMOL [34] and UCSF Chimera 1.14 [35].

2.6. Molecular Dynamics (MD) Simulation

MD simulations were conducted for a production run of 20 ns for wild enzyme and mutants using AMBER20 [36]. Preprocessing of the systems involved the application of AMBER antechamber program [37]. Topology and parameter files were built using the ff14SB force field [38] and recorded through the leap module. Both systems were neutralized by adding sodium ions and placed in 12 Å TIP3P water box. Energy of the systems was minimized by running 1500 steps of the steepest descent method and 1000 rounds of the conjugate gradient method. For nonbounded interactions, the threshold distance was set to 8 Å. The systems were then heated for 50 ps at constant volume and temperature (300 K). Systems equilibration was achieved for 100 ps where periodic boundary conditions along with constant pressure and Langevin thermostat were applied [39]. Lastly, both systems were subjected to a production run of 20 ns in an explicit solvent model using the isothermal-isobaric ensemble where temperature and pressure were set to 300 K and 1 atm, respectively. To model the long-range electrostatic effects, periodic boundary conditions and particle-mesh Ewald method were used along with a coupling algorithm to couple temperature with an external bath. SHAKE algorithm [40] was employed to apply hydrogen bond constraints whereas constant temperature was achieved through the Langevin coupling integration algorithm. A time step of 2 ps was used to Newton's equations for complex dynamics and trajectory files were created after 1 ps. The PTRAJ module [41] of AMBER was used to analyze MD trajectories.

2.7. Binding Free-Energy Calculations Using MMGBSA

AMBER20 MMPBSA.py module was used to calculate the binding free energy of the simulated complexes [42]. This tool estimated different components of the net binding free energy significant in interactions between the ligand and the biomolecule target. MM/GBSA is now a commonly employed technique in drug discovery practices to evaluate the affinity of ligand molecules to a specific biological target and underline key chemical intermolecular interaction energies. Estimation of the MM/GBSA binding free energy can be concluded as:

$$\Delta G_{\text{bind}} = \Delta G_{\text{complex}} - (\Delta G_{\text{protein}} - \Delta G_{\text{ligand}})$$

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta E_{\text{polar}} + \Delta E_{\text{SASA}} - T\Delta S$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{elect}} + \Delta E_{\text{vdw}}$$

$$\Delta E_{\text{SASA}} = \gamma \text{SASA}$$

In the above equations, ΔG_{bind} represents the net free binding energy, $-T\Delta S$ corresponds to entropic energy and ΔE_{MM} indicates molecular mechanics potential energy and can be split into van der Waals (ΔE_{vdw}) energy and electrostatic (ΔE_{elect}) energy. ΔE_{polar}

and ΔE_{SASA} are used for polar contributions and nonpolar solvation energy, respectively. The latter is calculated using a solvent-accessible surface area (SASA) nonpolar model.

3. Results and Discussion

All of the valid submitted Saudi Arabia SARS-CoV-2 RdRp enzyme sequences found in the viral database of NCBI were retrieved (21 May 2022) for comparative sequence analysis to highlight mutations that may alter the enzyme structure and function and may contribute to the virus's overall survival in terms of transmissibility, infectious potential and escaping from the currently available therapeutics. The accession number of the available Saudi Arabia SARS-CoV-2 RdRp enzymes is listed in Table 1. For alignment purposes, Clustal Omega was employed to perform multiple sequence alignment of the sequence whereas RdRp from Wuhan SARS-CoV-2 was used as a reference (accession number: YP_009724389). The MSA analysis identified three mutants of the RdRp in Saudi Arabia RdRp sequences. These mutants are: A97V, P323I and Y606C. A97V and Y606C mutations were observed in a single sample, while P323I was seen in 56 samples, as shown in Table 2.

Table 1. Details of SARS-CoV-2 ORF1ab polyprotein sequences used in the analysis.

S. No.	Accession Number
1	YP_009724389
2	QMS50985
3	QMS50997
4	QMS51009
5	QMS51021
6	QMS51033
7	QMS51045
8	QMS51057
9	QMS51069
10	QMS51081
11	QMS51093
12	QMS51105
13	QMS51117
14	QMS51129
15	QMS51141
16	QMS51153
17	QMS51165
18	QMS51177
19	QMS51189
20	QMS51201
21	QMS51213
22	QMS51225
23	QMS51237
24	QMS51249
25	QMS51261
26	QMS51273
27	QMS51285
28	QMS51297
29	QMS51309
30	QMS51321
31	QLH56060
32	QLH56072
33	QLH56084
34	QLH56096
35	QLH56108
36	QLH56120
37	QLH56132
38	QLH56144

Table 1. *Cont.*

S. No.	Accession Number
39	QLH56156
40	QLH56168
41	QLH56180
42	QLH56192
43	QLH56204
44	QLH56216
45	QLH56228
46	QLH56240
47	QLH56252
48	QKU37019
49	QKU37031
50	QKU37043
51	QKU37055
52	QKU37067
53	QKU37079
54	QKU37091
55	QKU37103
56	QKU37115
57	QKU37127
58	QKU37139
59	QKU37151

Table 2. Details of identified mutations.

Mutation ID	Wild-Type Residue	Position of Mutation	Mutated Residue	Frequency
A97V	A	97	V	1
P323I	P	323	I	56
Y606C	Y	606	C	1

3.1. Identification of Mutations in Structural Proteins Present in Saudi Isolates

Enzymes as proteins in nature performed their function in dynamics, which is significant for functionality and protein–protein interactions as well as for protein–ligand interactions. The impact of the identified mutations on the RdRp enzyme structure stability and functionality was determined by the mean of the free-energy differences ($\Delta\Delta G$) between the Wuhan RdRp and mutants. The study of this free-energy difference can highlight the impact of mutation on the secondary structure’s stability and tertiary structure’s dynamics. In the process, vibrational entropy was estimated between the wild-type RdRp and mutants. The entropy energy anticipated the stability changes in kcal/mol, as when the $\Delta\Delta G$ is ≥ 0 it demonstrates a stabilizing effect of the mutation on the protein stability, while when $\Delta\Delta G$ is ≤ 0 it is an indication of the mutation’s destabilizing effect on the enzyme structure and overall functionality. The vibrational entropy energy (termed as $\Delta\Delta S_{VibENCoM}$) between the mutants and wild type was assessed by the EN-CoM tool. The $\Delta\Delta S_{VibENCoM}$ of ≥ 0 conferred flexibility to the RdRp enzyme, whereas $\Delta\Delta S_{VibENCoM} \leq 0$ illustrates rigidity to the enzyme due to mutation. The free-energy difference among the mutants and between the mutants and wild type is tabulated in Table 3. The A97V and P323I were found to have a stabilizing impact on the RdRp enzyme structure and functionality, while and Y606C revealed a destabilizing impact.

Table 3. The values of change in $\Delta\Delta S$ ENCoM and $\Delta\Delta G$ (kcal/mol) due to the mutations in RdRp.

S. No	Wuhan Isolate	Saudi Isolate	AA Position	$\Delta\Delta S$ ENCoM	$\Delta\Delta G$ DynaMut	$\Delta\Delta G$ mCSM	$\Delta\Delta G$ SDM	$\Delta\Delta G$ DUET	Effect
1	A	V	97	4.117	1.397	−0.271	−1.270	−0.242	Stabilizing
2	P	I	323	0.406	1.017	−0.251	1.500	0.454	Stabilizing
3	Y	C	606	−0.984	−0.675	−1.675	−1.200	−1.721	Destabilizing

3.2. Δ Vibrational Entropy Energy between Wild-Type and Mutant

The vibrational entropy energy was determined further to find out whether the identified mutant(s) can either increase or decrease flexibility of the RdRp. The $\Delta\Delta S_{Vib}$ ENCoM value of A97A RdRp mutant in comparison to the wild RdRp is $1.228 \text{ kcal mol}^{-1} \text{ K}^{-1}$, while that of P323I is $-0.508 \text{ kcal mol}^{-1} \text{ K}^{-1}$. Both these mutants were estimated to decrease RdRp flexibility. On the other hand, the Y606C $\Delta\Delta S_{Vib}$ ENCoM value is $0.844 \text{ kcal mol}^{-1} \text{ K}^{-1}$, which is an indication of increased RdRp flexibility. The vibrational entropy energy difference between the wild-type RdRp and mutated RdRp is presented in Figure 1.

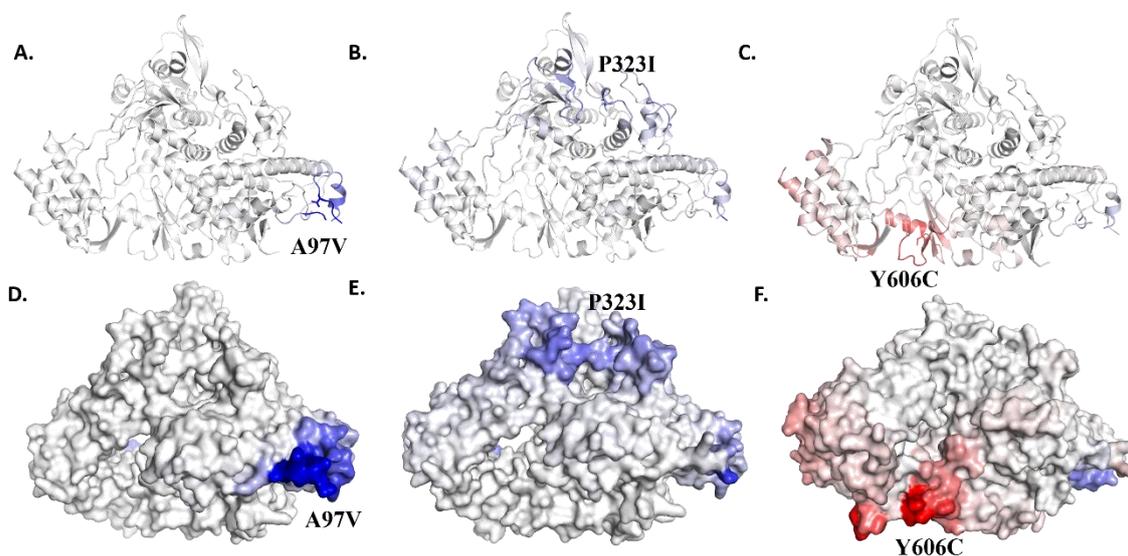


Figure 1. Δ Vibrational entropy energy between wild-type and mutant RdRp. Ribbon representation showing the change in vibrational entropy energy between wild-type and mutant RdRp due to (A) A97V and (B) P323I (C) Y606C. Surface representation showing the change in vibrational entropy energy between wild-type and mutant PLpro due to (D) A97V and (E) P323I (F) Y606C. Amino acids are colored according to the vibrational entropy change as a consequence of mutation of PLpro protein. Blue represents a rigidity of the structure and red represents a gain in flexibility.

3.3. Mutations Cause Alteration in Secondary Structure of Proteins

In order to evaluate the mutations' effect on RdRp secondary structure elements, secondary structure analysis was carried out. Comparative analysis of secondary structure elements between the wild RdRp and mutants can be depicted in Figure 2. All the identified mutations in general do not affect the secondary structure elements of the position of the wild type. However, the A97V mutations were found to alter the five-helix string (present in the wild) and shorten it to the three-helix string upon mutation. P232I seems to have less impact on the overall structure element, whereas Y606C was revealed to produce a three-helix string right after the single mutation position in comparison to the wild type. The secondary structure analysis indicates that the mutation might affect the structure elements' composition and patterning, which in turn might effect the RdRp structure and function, but this needs to be validated and studied further.

WT	Query	VVKRHTFSNYQHEETIYNLLKDCPAVAKHDFKFRIDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCD
	Helix	HHHHH HHHHHHHHHH HHHHHH HH H
	Sheet	EE EE
	Turns	T T T TT TT
	Struc	EEHEEEEEEEHHHEEEEEHHHEHEHHHHHHHEEEEEHHTHEEEEEHHHEEEEEHHHEEEEEHHHHHTTCEH
A97V	Query	VVKRHTFSNYQHEETIYNLLKDCPAVVKHDFKFRIDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCD
	Helix	HHHHH HHHHHHHHHH HHHHHH HH H
	Sheet	EE EE
	Turns	T T T TT TT
	Struc	EEHEEEEEEEHHHEEEEEHHHEHEEEEEHHHEEEEEHHHEEEEEHHHEEEEEHHHHHTTCEH
WT	Query	KLFDTRYFKYWDQTYHPNCVNLDDRCILHCANFNWLFSTVFPPTISFGPLVRKIFVDGVPPFVWSTGYHFRE
	Helix	HHHHHHHHHHH HHHHHHHHHHHHHHHH HHHHHHHHHHHHHH HHH
	Sheet	EE
	Turns	T T T T T
	Struc	HHEEEEEEEEEEEEETEEEEEEEEEEHHHEEEEEEEEEEEETEEEEEEEEEEEEEEEEETEEHHH
P323I	Query	KLFDTRYFKYWDQTYHPNCVNLDDRCILHCANFNWLFSTVFIPTISFGPLVRKIFVDGVPPFVWSTGYHFRE
	Helix	HHHHHHHHHHH HHHHHHHHHHHHHHHH HHHHHHHHHHHHHH HHH
	Sheet	EE
	Turns	T T T T T
	Struc	HHEEEEEEEEEEEEETEEEEEEEEEEHHHEEEEEEEEEEEETEEEEEEEEEEEEEEEEETEEHHH
WT	Query	SICSTMTRQGFHQKLLKSIAATRGATVWIGTSKIFYGGWNNMLKTVYSDEVNPHLMGWDPKCDRAMPNML
	Helix	HHHH HHHHHHHHHHHHHHHHHH HHHHHHH HHHHHHH HHHHHH
	Sheet	EE EEEEE EEEEE
	Turns	T T T T T
	Struc	EEEEEEEEEEHHHHHEHHHHHHHEEEEEEEEEEEEEEEHHHEEEEEEECCCTHHEHEEECCCCCTHHHHEE
Y606C	Query	SICSTMTRQGFHQKLLKSIAATRGATVWIGTSKIFYGGWNNMLKTVCSDEVNPHLMGWDPKCDRAMPNML
	Helix	HHHH HHHHHHHHHHHHHHHHHH HHHHHHH HHHHHHH HHHHHH
	Sheet	EE EEEEE EEEEE
	Turns	T T T T T
	Struc	EEEEEEEEEEHHHHHEHHHHHHHEEEEEEEEEEEEEEEHHHEEEEEHHHHHTHHEHEEECCCCCTHHHHEE

Figure 2. Prediction of secondary structure of RdRp protein. The difference in secondary structure between Wuhan (WT) and Saudi isolates are highlighted with position of dashed box in respective panels.

3.4. Intramolecular Interactions Are Altered Due to Mutations in Proteins

A mutation within the amino acid sequence may bring about a change in the protein's structure and function [43]. Therefore, the alteration of interatomic intramolecular interactions caused by identified mutations within the RdRp enzyme was assessed using DynaMut webserver. The substitution of the hydrophobic residue alanine (A) with hydrophobic valine (V) changes the interaction pattern, leading to the formation of two hydrogen bonds between mutant V97 and N136 and L40 (Figure 3A,B). These changes can provide more stability and less flexibility to the protein, which supports previous experiments. Likewise, mutation of P with hydrophobic I residue at position 323 induces the formation of hydrogen bonds with F317, which may result in the protein stability increasing (Figure 3C,D). However, a change from small hydrophobic V to C residue at position 606 seems to disrupt several intramolecular interactions, causing a destabilization effect (Figure 3E,F). These results were inconsistent with the above obtained results.

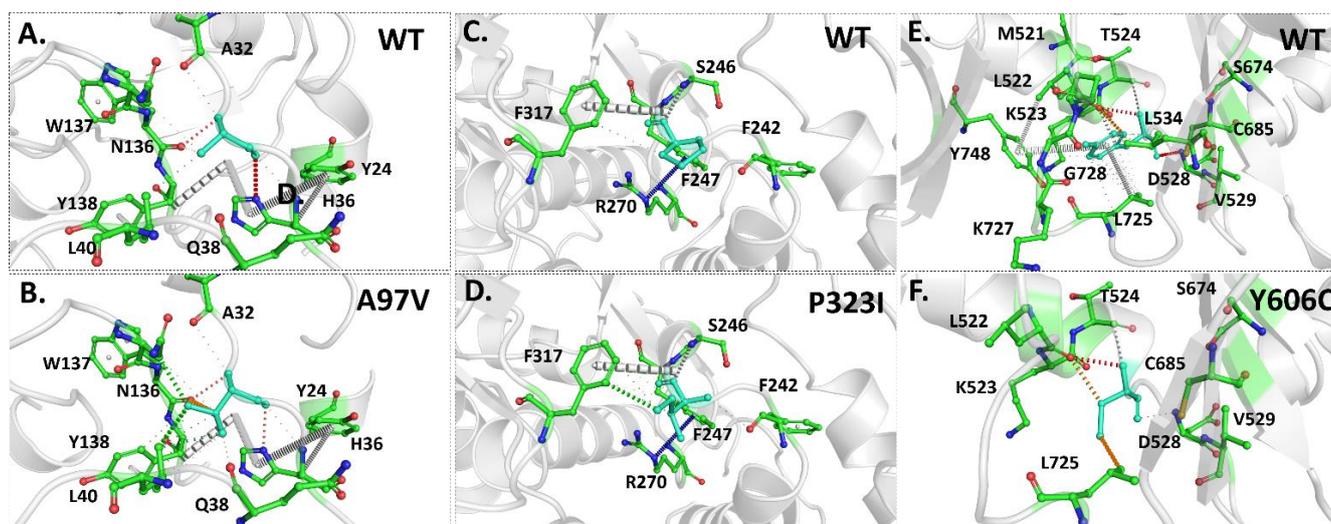


Figure 3. Effect of amino acid substitution on interatomic interactions of RdRp. (A,B) show the interatomic interactions mediated by wild type and A97V. (C,D) show the interatomic interactions mediated by wild type and P323I. (E,F) show the interatomic interactions mediated by wild type and Y606C. Wild-type and mutant residues are colored in cyan and are also represented as sticks alongside the surrounding residues (green), which are involved in any type of interactions.

3.5. Effect of Mutations on Protein Structural Conformation and Dynamic Stability

To understand the dynamic stability of the apo wild-type and mutant RdRp enzymes, molecular dynamics simulations were carried out for length of 20 ns. The potential energy of the systems was determined first, which revealed that the mutant is in contrast to the wild type. The mutant RdRp was revealed to be more relaxed and unstrained compared to the wild-type RdRp (Figure 4A). The carbon alpha root-mean-square deviation ($C\alpha$ RMSD) was measured first for all systems (Figure 4B). In general, the wild as well as the mutants were seen in relative stable equilibrium. The wild RdRp in particular was observed in low dynamics across the length of simulation time, except towards the end where the enzyme achieved a slightly higher dynamic compared to the mutant enzymes. The A97V among the mutants was the most stable in the simulated time, and a good overall stability was noticed. These results were validated by calculating the radius of gyration (RoG) for the systems (Figure 4C). The RoG findings are mostly in line with the $C\alpha$ RMSD and demonstrate that the wild RdRp has a much lower RoG pattern. This points to greater compactness of the enzyme and more stable behavior. On contrast, the mutants have higher RoG and are reported to have a highly relaxed structure. Further, to access the stability, the folding and accessibility of water to the protein were calculated by solvent-accessible surface area (SASA) analysis (Figure 4D). Lower SASA values demonstrate tight packing of the protein folding and less flexibility. This further defines less access of solvent molecules and other interacting partners to bind the enzyme. It can be interpreted that upon mutations, the RdRp showed more solvent accessibility and impacted the enzyme's intermolecular interaction dynamics and folding, functionality and overall stability dynamics.

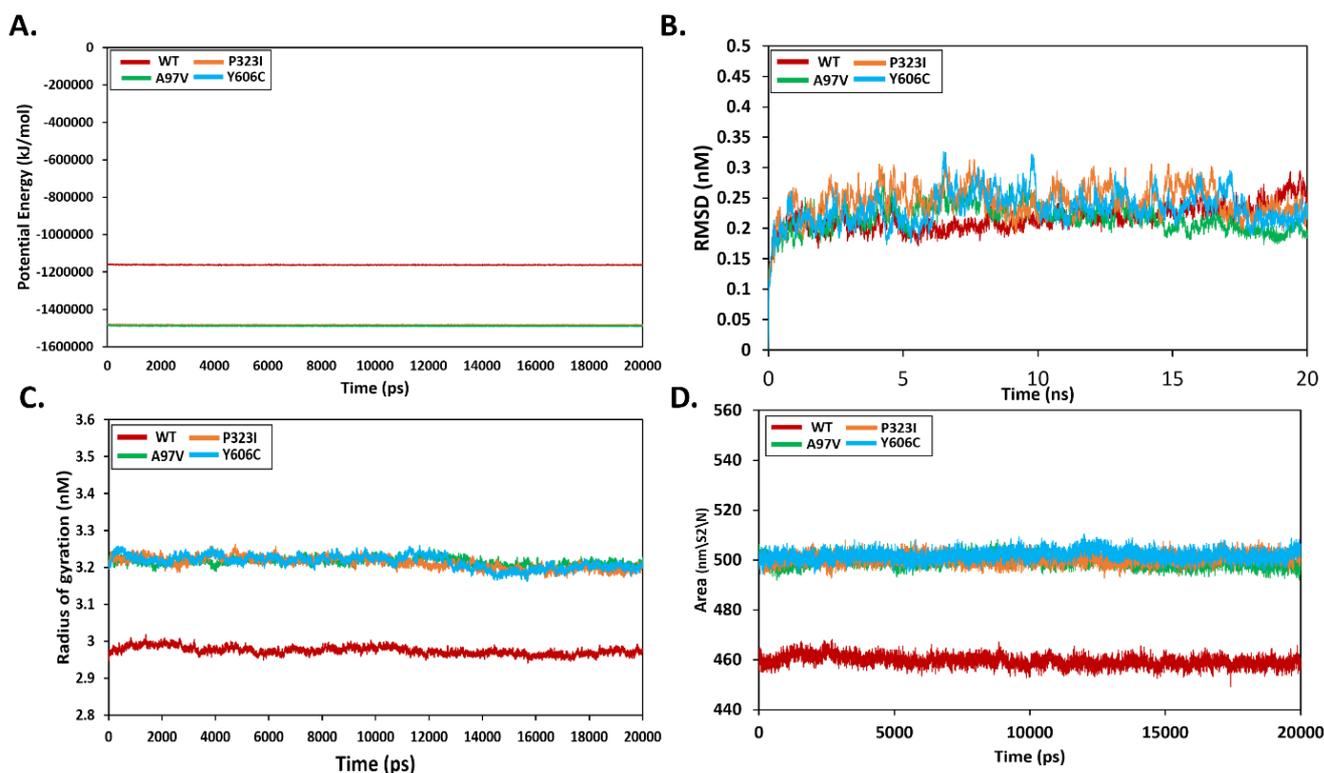


Figure 4. Molecular dynamics simulation analysis of wild-type and mutant RdRp enzymes. (A)—potential energy, (B)—RMSD, (C)—gyration, (D)—SASA.

3.6. Effect of Mutations on Binding and Stability of Remdesivir

The effect of mutations on remdesivir with RdRp was elucidated by docking the remdesivir with both the wild type and mutants. It has been reported that the $C\alpha$ RMSD of the WT-RdRp complex with remdesivir is more stable compared to the rest of the mutants. The mutation in RdRp results in more flexibility being conferred to the enzyme's flexible loops, which in turn make the enzyme more dynamic. The wild type, on the other hand, makes the enzyme structure more compact and allows for stable dynamics. The WT-RdRp and mutants' RMSD are given in Figure 5A. The residue-wise flexibility of the WT-RdRp and mutants is depicted in Figure 5B, where it can be clearly observed that the WT-RdRp has more stable residue fluctuations compared to the rest of the mutants.

Next, the molecular dynamics simulation trajectories were used for the calculation of binding free energies for the wild type and mutants with remdesivir using the MMGBSA method. The net binding free energies of the WT-RdRp complex with remdesivir, A97V-RdRp complex with remdesivir, P232I-RdRp complex with remdesivir and Y606C-RdRp complex with remdesivir are -50.76 kcal/mol, -44.96 kcal/mol, -47.2 kcal/mol and -41.34 kcal/mol, respectively. The wild-type system, as indicated by molecular dynamics analysis, was found the most stable in terms of intermolecular interaction and docked conformation, followed by P232I, A97V and Y606C. In all the systems, van der Waals energy was found dominant, and also the electrostatic energy plays a significant role in overall stability. Complete details of binding energies are tabulated in Table 4.

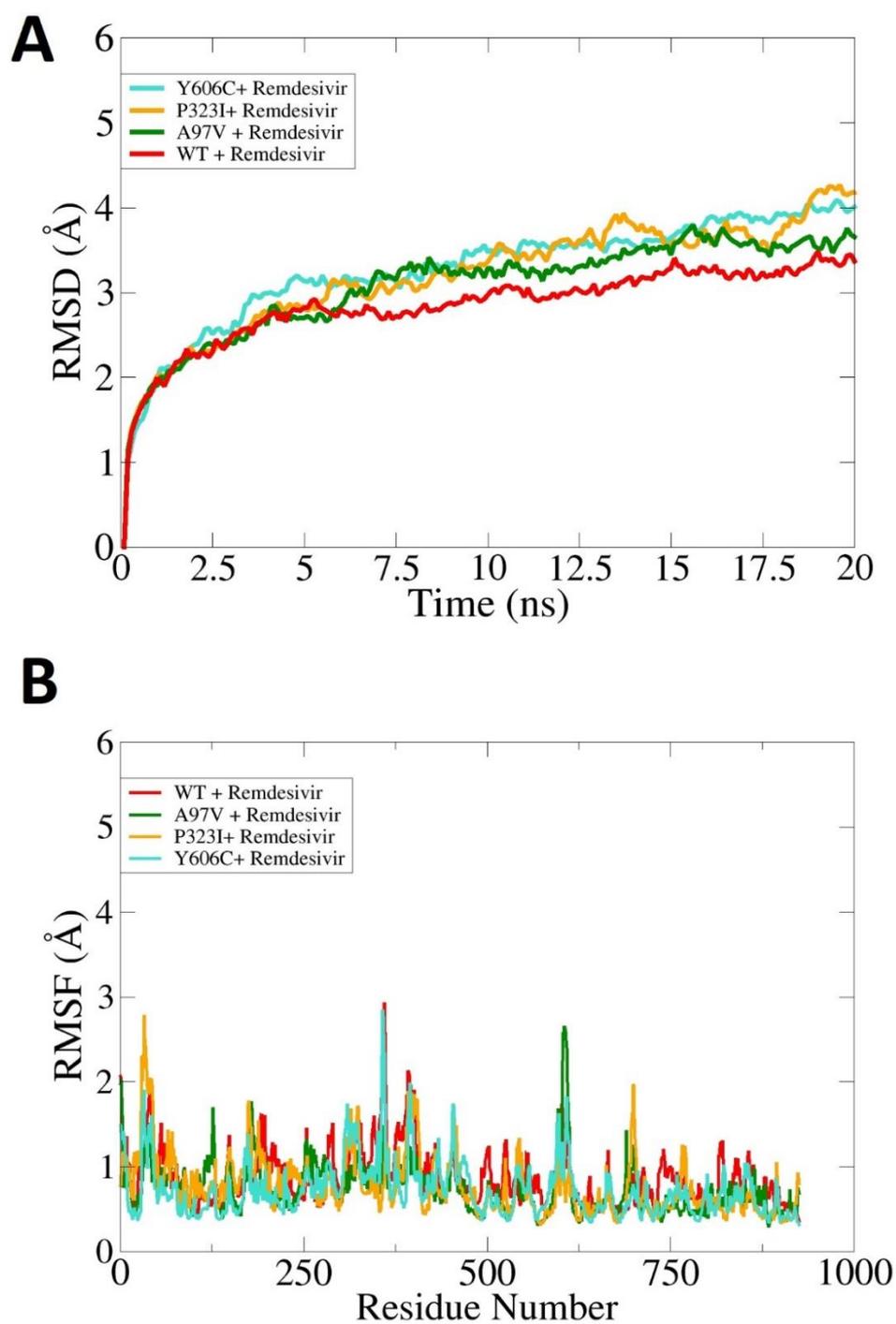


Figure 5. Molecular dynamics simulation analysis of WT-RdRp and mutants. (A)—RMSD and (B)—RMSF.

Table 4. Molecular dynamics simulation trajectory-based binding free energies in kcal/mol for wild type and mutants with remdesivir.

Energy Parameter	WT-RdRp Complex with Remdesivir	A97V-RdRp Complex with Remdesivir	P232I -RdRp Complex with Remdesivir	Y606C-RdRp Complex with Remdesivir
VDWAALS	−40.68	−39.33	−41.49	−37.51
EEL	−25.28	−22.00	−23.22	−19.67
Delta G gas	−65.96	−61.33	−64.71	−57.18
Delta G solv	15.20	16.37	17.51	15.84
Delta Total	−50.76	−44.96	−47.2	−41.34

4. Conclusions

Due to the potential impact of some of these mutations on the specificity of diagnostic tests and vaccines and therapeutic development aimed at these sites, it is crucial to closely monitor the emergence of mutations throughout the duration of this ongoing pandemic. High numbers of mutations in the SARS-CoV-2 genome allow for efficient survival of the virus, allowing the virus to fit itself to the changing host immune responses, which in turn make it difficult to effectively manage the virus. Therefore, it is vital to understand the virus mutations capable of evading the virus and escaping from being therapeutically managed. In this study, three mutations—A97V, P323I and Y606C—were identified in different RdRp enzyme sequences extracted from Saudi Arabia isolates when aligned with reference Wuhan SARS-CoV-2 RdRp. The mutations were identified to affect the enzyme structure dynamics and stability and to alter interactions with the remdesivir ligand. Hence, studies are needed to explore the mutations' impact on the RdRp structure–function activity and use that knowledge in further studies to design better therapeutics.

Author Contributions: Conceptualization, M.A.A. (Mubarak A. Alamri) and M.S.R.R.; data curation, M.A.A. (Mubarak A. Alamri), M.T.u.Q., A.B.A., S.M.A., M.A.A. (Manal A. Alossaimi), S.A. (Sikandar Azam) and M.H.H.; formal analysis, M.A.A. (Mubarak A. Alamri), M.T.u.Q., A.B.A., S.M.A., M.A.A. (Manal A. Alossaimi), S.A. (Sikandar Azam) and M.H.H.; funding acquisition, M.A.A. (Mubarak A. Alamri); investigation, M.A.A. (Mubarak A. Alamri); methodology, M.A.A. (Mubarak A. Alamri), M.T.u.Q. and M.S.R.R.; project administration, M.A.A. (Mubarak A. Alamri); resources, M.A.A. (Mubarak A. Alamri), M.T.u.Q. and M.S.R.R.; software, M.A.A. (Mubarak A. Alamri), M.T.u.Q. and M.S.R.R.; supervision, M.A.A. (Mubarak A. Alamri) and M.S.R.R.; validation, A.B.A., S.M.A., M.A.A. (Manal A. Alossaimi), M.H.H. and M.S.R.R.; visualization, M.A.A. (Mubarak A. Alamri), M.T.u.Q. and S.A. (Sikandar Azam); writing—original draft, M.A.A. (Mubarak A. Alamri) and M.T.u.Q.; writing—review and editing, A.B.A., S.M.A., M.A.A. (Manal A. Alossaimi), S.A. (Sikandar Azam), M.H.H. and M.S.R.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was financially supported by the Prince Sattam bin Abdulaziz University (PSAU) under the umbrella of “Deanship of Scientific Research” with project ID 2021/03/19145.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available within the article.

Acknowledgments: This research work was financially supported by the Prince Sattam bin Abdulaziz University (PSAU) under the umbrella of “Deanship of Scientific Research” with project ID 2021/03/19145, therefore, we are extremely grateful to PSAU for their kind and generous support.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Amanat, F.; Krammer, F. SARS-CoV-2 vaccines: Status report. *Immunity* **2020**, *52*, 583–589. [[CrossRef](#)]
2. Tahir ul Qamar, M.; Alqahtani, S.M.; Alamri, M.A.; Chen, L.-L.L.; ul Qamar, M.T.; Alqahtani, S.M.; Alamri, M.A.; Chen, L.-L.L. Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. *J. Pharm. Anal.* **2020**, *10*, 313–319. [[CrossRef](#)]
3. Mehmood, I.; Ijaz, M.; Ahmad, S.; Ahmed, T.; Bari, A.; Abro, A.; Allemailem, K.S.; Almatroudi, A.; Tahir ul Qamar, M. SARS-CoV-2: An update on genomics, risk assessment, potential therapeutics and vaccine development. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1626. [[CrossRef](#)] [[PubMed](#)]
4. Zhu, Z.; Lian, X.; Su, X.; Wu, W.; Marraro, G.A.; Zeng, Y. From SARS and MERS to COVID-19: A brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. *Respir. Res.* **2020**, *21*, 224. [[CrossRef](#)]
5. Pachetti, M.; Marini, B.; Benedetti, F.; Giudici, F.; Mauro, E.; Storici, P.; Masciovecchio, C.; Angeletti, S.; Ciccozzi, M.; Gallo, R.C.; et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J. Transl. Med.* **2020**, *18*, 179. [[CrossRef](#)]
6. Davies, N.G.; Abbott, S.; Barnard, R.C.; Jarvis, C.I.; Kucharski, A.J.; Munday, J.; Pearson, C.A.B.; Russell, T.W.; Tully, D.C.; Washburne, A.D.; et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **2021**, *372*, 3055. [[CrossRef](#)]
7. Chand, M.; Hopkins, S.; Dabrera, G.; Achison, C.; Barclay, W.; Ferguson, N. *Investigation of Novel SARS-COV-2 Variant: Variant of Concern 202012/01 (PDF)*; Public Health England PHE: London, UK, 2020; Volume 21.
8. Grint, D.J.; Wing, K.; Houlihan, C.; Gibbs, H.P.; Evans, S.J.W.; Williamson, E.; McDonald, H.I.; Bhaskaran, K.; Evans, D.; Walker, A.J.; et al. Severity of SARS-CoV-2 alpha variant (B.1.1.7) in England. *Clin. Infect. Dis.* **2021**, *75*, ciab754.
9. Islam, A.; Sayeed, M.A.; Kalam, M.A.; Ferdous, J.; Rahman, M.K.; Abedin, J.; Islam, S.; Shano, S.; Saha, O.; Shirin, T.; et al. Molecular epidemiology of SARS-CoV-2 in diverse environmental samples globally. *Microorganisms* **2021**, *9*, 1696. [[CrossRef](#)]
10. Mistry, P.; Barmania, F.; Mellet, J.; Peta, K.; Strydom, A.; Viljoen, I.M.; James, W.; Gordon, S.; Pepper, M.S. SARS-CoV-2 variants, vaccines, and host immunity. *Front. Immunol.* **2021**, *12*, 809244. [[CrossRef](#)]
11. Viana, R.; Moyo, S.; Amoako, D.G.; Tegally, H.; Scheepers, C.; Althaus, C.L.; Anyaneji, U.J.; Bester, P.A.; Boni, M.F.; Chand, M.; et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* **2022**, *603*, 679–686. [[CrossRef](#)]
12. Khateeb, J.; Li, Y.; Zhang, H. Emerging SARS-CoV-2 variants of concern and potential intervention approaches. *Crit. Care* **2021**, *25*, 244. [[CrossRef](#)] [[PubMed](#)]
13. Brown, C.M.; Vostok, J.; Johnson, H.; Burns, M.; Gharpure, R.; Sami, S.; Sabo, R.T.; Hall, N.; Foreman, A.; Schubert, P.L. Outbreak of SARS-CoV-2 infections, including COVID-19 vaccine breakthrough infections, associated with large public gatherings—Barnstable County, Massachusetts, July 2021. *Morb. Mortal. Wkly. Rep.* **2021**, *70*, 1059. [[CrossRef](#)] [[PubMed](#)]
14. Singh, P.; Sharma, K.; Singh, P.; Bhargava, A.; Negi, S.S.; Sharma, P.; Bhise, M.; Tripathi, M.K.; Jindal, A.; Nagarkar, N.M. Genomic characterization unravelling the causative role of SARS-CoV-2 Delta variant of lineage B.1.617.2 in 2nd wave of COVID-19 pandemic in Chhattisgarh, India. *Microb. Pathog.* **2022**, *164*, 105404. [[CrossRef](#)] [[PubMed](#)]
15. Meo, S.A.; Meo, A.S.; Al-Jassir, F.F.; Klonoff, D.C. Omicron SARS-CoV-2 new variant: Global prevalence and biological and clinical characteristics. *Eur. Rev. Med. Pharmacol. Sci.* **2021**, *25*, 8012–8018. [[PubMed](#)]
16. Eyre, D.W.; Taylor, D.; Purver, M.; Chapman, D.; Fowler, T.; Pouwels, K.; Walker, A.S.; Peto, T.E.A. The impact of SARS-CoV-2 vaccination on Alpha and Delta variant transmission. *medRxiv* **2021**. [[CrossRef](#)]
17. Callaway, E.; Ledford, H. How to redesign COVID vaccines so they protect against variants. *Nature* **2021**, *590*, 15–16. [[CrossRef](#)]
18. Elshabrawy, H.A. SARS-CoV-2: An Update on Potential Antivirals in Light of SARS-CoV Antiviral Drug Discoveries. *Vaccines* **2020**, *8*, 335. [[CrossRef](#)] [[PubMed](#)]
19. Alamri, M.A.; Tahir Ul Qamar, M.; Mirza, M.U.; Bhadane, R.; Alqahtani, S.M.; Muneer, I.; Froeyen, M.; Salo-Ahen, O.M.H.H. Pharmacoinformatics and molecular dynamics simulation studies reveal potential covalent and FDA-approved inhibitors of SARS-CoV-2 main protease 3CLpro. *J. Biomol. Struct. Dyn.* **2020**, *39*, 4936–4948. [[CrossRef](#)] [[PubMed](#)]
20. Yang, Y.; Peng, F.; Wang, R.; Guan, K.; Jiang, T.; Xu, G.; Sun, J.; Chang, C. The deadly coronaviruses: The 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China. *J. Autoimmun.* **2020**, *109*, 102434. [[CrossRef](#)]
21. Walls, A.C.; Park, Y.-J.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Velesler, D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **2020**, *181*, 281–292. [[CrossRef](#)] [[PubMed](#)]
22. Elfiky, A.A. SARS-CoV-2 RNA dependent RNA polymerase (RdRp) targeting: An in silico perspective. *J. Biomol. Struct. Dyn.* **2020**, *39*, 3204–3212. [[CrossRef](#)] [[PubMed](#)]
23. Mirza, M.U.; Vanmeert, M.; Froeyen, M.; Ali, A.; Rafique, S.; Idrees, M. In silico structural elucidation of RNA-dependent RNA polymerase towards the identification of potential Crimean-Congo Hemorrhagic Fever Virus inhibitors. *Sci. Rep.* **2019**, *9*, 6809. [[CrossRef](#)]
24. Yin, W.; Mao, C.; Luan, X.; Shen, D.-D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020**, *368*, 1499–1504. [[CrossRef](#)] [[PubMed](#)]
25. Marotta, C.; Stefanelli, P.; Tramuto, F.; Angeloni, U.; Maida, C.M.; Cernigliaro, A.; Barone, T.; Vitale, F.; Rezza, G.; Mazzucco, W. The dual/global value of SARS-COV-2 genome surveillance on migrants arriving to Europe via the mediterranean routes. *Ann. Glob. Health* **2021**, *87*, 71. [[CrossRef](#)]

26. Robishaw, J.D.; Alter, S.M.; Solano, J.J.; Shih, R.D.; DeMets, D.L.; Maki, D.G.; Hennekens, C.H. Genomic surveillance to combat COVID-19: Challenges and opportunities. *Lancet Microbe* **2021**, *2*, e481–e484. [[CrossRef](#)]
27. Yashvardhini, N.; Jha, D.K.; Bhattacharya, S. Identification and characterization of mutations in the SARS-CoV-2 RNA-dependent RNA polymerase as a promising antiviral therapeutic target. *Arch. Microbiol.* **2021**, *203*, 5463–5473. [[CrossRef](#)]
28. Bawono, P.; Dijkstra, M.; Pirovano, W.; Feenstra, A.; Abeln, S.; Heringa, J. Multiple sequence alignment. *Methods Mol. Biol.* **2017**, *1525*, 167–189. [[CrossRef](#)]
29. Rodrigues, C.H.M.; Pires, D.E.V.; Ascher, D.B. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* **2018**, *46*, W350–W355. [[CrossRef](#)]
30. Kumar, T.A. CFSSP: Chou and Fasman secondary structure prediction server. *Wide Spectr.* **2013**, *1*, 15–19.
31. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [[CrossRef](#)]
32. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)] [[PubMed](#)]
33. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [[CrossRef](#)] [[PubMed](#)]
34. Pasi, M.; Tiberti, M.; Arrigoni, A.; Papaleo, E. XPyder: A PyMOL plugin to analyze coupled residues and their networks in protein structures. *J. Chem. Inf. Model.* **2012**, *52*, 1865–1874. [[CrossRef](#)]
35. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)] [[PubMed](#)]
36. Case, D.A.; Belfon, K.; Ben-Shalom, I.; Brozell, S.R.; Cerutti, D.; Cheatham, T.; Cruzeiro, V.W.D.; Darden, T.; Duke, R.E.; Giambasu, G.; et al. *Amber 2020*; University of California: San Francisco, CA, USA, 2020.
37. Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Antechamber: An accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc.* **2001**, *222*, U403.
38. Case, D.A.; Babin, V.; Berryman, J.T.; Betz, R.M.; Cai, Q.; Cerutti, D.S.; Cheatham III, T.E.; Darden, T.A.; Duke, R.E.; Gohlke, H.; et al. The FF14SB force field. *Amber* **2014**, *14*, 29–31.
39. Izaguirre, J.A.; Catarello, D.P.; Wozniak, J.M.; Skeel, R.D. Langevin stabilization of molecular dynamics. *J. Chem. Phys.* **2001**, *114*, 2090–2098. [[CrossRef](#)]
40. Kräutler, V.; Van Gunsteren, W.F.; Hünenberger, P.H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **2001**, *22*, 501–508. [[CrossRef](#)]
41. Roe, D.R.; Cheatham III, T.E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [[CrossRef](#)]
42. Miller, B.R.; McGee, T.D.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [[CrossRef](#)]
43. Studer, R.A.; Dessailly, B.H.; Orengo, C.A. Residue mutations and their impact on protein structure and function: Detecting beneficial and pathogenic changes. *Biochem. J.* **2013**, *449*, 581–594. [[CrossRef](#)]