



OPEN

In silico mutational analysis of ACE2 to check the susceptibility of lung cancer patients towards COVID-19

Zumama Khalid^{1✉}, Abeeđa Tu-Allah Khan¹, Radwan Alnajjar^{2,3}, Eman Santali⁴ & Abdul Rauf Shakoori¹

Being the second major cause of death worldwide, lung cancer poses a significant threat to the health of patients. This worsened during the era of pandemic since lung cancer is found to be more prone to SARS-CoV-2 infection. Many recent studies imply a high frequency of COVID-19 infection associated severe outcome. However, molecular studies are still lacking in this respect. Hence the current study is designed to investigate the binding affinities of ACE2 lung cancer mutants with the viral spike protein to find the susceptibility of respective mutants carrying patients in catching the virus. Quite interestingly, our study found lesser binding affinities of all the selected mutants thus implying that these cancer patients might be less affected by the virus than others. These results are opposed to the recent studies' propositions and open new avenues for more in-depth studies.

Lung cancer is the second cause of death in males and females after cardiovascular diseases worldwide¹. Among cancers, lung cancer is the first leading cause of death among men and the second leading cause of death among women². Deaths due to lung cancer in women of developed countries are higher as compared to developing countries³. The worldwide annual estimate of people diagnosed with lung cancer is 1.8–2.0 million, and 1.6 million people fail to survive⁴. Lung cancer is considered one of the deadliest diseases due to its delayed diagnosis, disease setback, and lack of curative medication⁵.

Lung cancer has been one of the widely malignant cancers in the entire world because of its diagnosis at advanced stages. Multiple steps are involved in the mechanism of metastasis of such wide-spreading cancer. These steps require crucial mechanisms such as angiogenesis, the transition of cells from endothelial form to mesenchymal form (EMT) by the interruption in cell to cell and cell to matrix attachment, apoptosis, and movement of cells to a secondary site. Genetic involvement in these mechanisms and pathways ensures lung cancer to be invasive and highly metastatic⁶. Lung cancer is a multifactorial disease; hence together, environmental and genetic elements play a vital role in its susceptibility⁷. Various genome-wide association studies have pointed out the multiple single nucleotide polymorphisms (SNP) in association with lung carcinoma, especially on 6p21.33, 15q25.1, and 5p15.33^{8,9}.

In November 2019, a novel virus emerged from Wuhan (China). This virus was labeled as Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses¹⁰. The viral infection starts from the binding of viral spike protein to the receptor site on the surface of the host. For spike protein (SARS-CoV-2 glycoprotein) of COVID-19, the identified receptor is ACE2 (Angiotensin Converting Enzyme 2) (Fig. 1A)¹¹. Spike protein has two subunits S1 and S2; the S1 subunit bears the receptor-binding-domain (RBD). RBD binds directly to the peptidase domain of the ACE2 receptor. After this binding site on the S2 subunit exposes, which gets cleaved by the host protease, and viral infection is initiated (Fig. 1B)¹². ACE2 enzyme catalyzes the conversion of angiotensin to other different forms^{13,14}.

Cancer patients are seen to be more prone to COVID-19 than normal people due to their immunocompromised health status. Another reason found of being lung cancer patients more prone to COVID-19 infection is the healthcare system where contact with COVID-19 infection was more during cancer treatment¹⁵. There is a high frequency of severe symptoms due to COVID-19 in patients with having lung carcinoma, hematologic

¹School of Biological Sciences, University of the Punjab, Quaid-i-Azam Campus, Lahore 54590, Pakistan. ²Department of Chemistry, Faculty of Science, University of Benghazi, Benghazi, Libya. ³Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa. ⁴College of Pharmacy, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia. ✉email: zumama.khalid@gmail.com

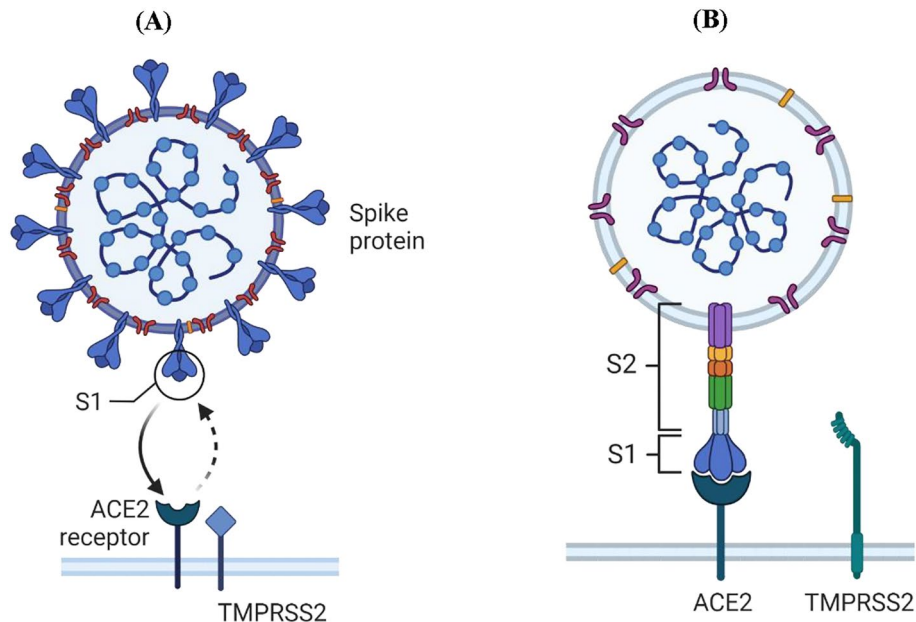


Figure 1. The structure of SARS-CoV-2 and the ACE2 receptor (Created with BioRender.com).

cancer, and metastatic cancer¹¹. A high fatality rate was observed in lung cancer patients as compared to other cancer patients having COVID-19^{16–19}. Although recent studies have stated that COVID-19 patients who have cancer have a high possibility of severe outcomes than patients without cancer, but there is no study pointing towards the possible molecular mechanism behind it. However, Yu et al. reported that lung cancer patients have more severity of the viral disease than other patients as the major entry route for the virus is through the lungs, and that too with reasons unknown^{16–19}.

Therefore, this work has been designed to investigate the binding affinities of the viral spike protein toward mutated ACE2 proteins to decipher the susceptibility and probable molecular mechanism of the COVID-19 infection concerning ACE2.

Materials and methods

ACE2 mutants and SARS-CoV2 spike protein. The 3D structures of ACE2 and SARS-CoV2 spike protein were retrieved from RCSB-PDB. The RCSB-PDB code for the structures of SARS-CoV2 spike protein and ACE2 is 6M17²⁰. Missense mutations and frameshift mutations in the ACE2 gene in lung tissues were curated from the COSMIC cancer database. All the mutations in the study were used in accordance with the database regulations and guidelines (<https://cancer.sanger.ac.uk/cosmic>). These mutations were induced in the 3D structure of ACE2 through mutagenesis in PyMOL²¹. Structures were visualized via UCSF Chimera Version 1.14²².

Protein–protein docking. Protein–protein docking was performed on the HADDOCK server (<https://wenmr.science.uu.nl/haddock2.4/>), a web server for the docking of biomolecular structures, and the obtained binding affinities of the docked poses of the spike glycoprotein of SARS-CoV-2 with all the cancer mutants of ACE2 were analyzed²³. Binding residues of ACE2 (24,27,28,30,31,34,35,37,38,41,42,79,82,83,330,353,354,355,357,393) were docked with binding residues of RBD (Receptor Binding Domain) of SARS-CoV2 spike protein (486,487,489,456,455,475,417,493,453,505,501,449,496,502,500,496,446) along with wild type ACE2 and spike protein²⁴. The docking postures of ACE2 cancer mutants and spike glycoprotein were visualized using PyMOL and UCSF Chimera Version 1.14^{21,22}.

Analysis of docking. Based on HADDOCK scores, the top 5 docked complexes of ACE2 cancer mutants and spike protein with higher docking scores were assessed for stability and binding affinity (kcal mol^{-1}). This analysis was performed on Protein Binding Energy Prediction (PRODIGY) server^{25,26}. PRODIGY predicts stability and binding affinity of docking complex based on structural properties of both interacting molecules in complex. Binding affinity was demonstrated by ΔG (kcal mol^{-1}) and stability by the dissociation constant K_d (M). Analysis was done at different temperature ranges. Docked complexes with higher binding affinity were subjected to molecular dynamics simulation.

Molecular dynamic simulation and MM-GBSA. The molecular dynamic simulations were carried out using the Desmond simulation package of Schrödinger LLC²⁷. Molecular dynamic simulations and MM-GBSA calculations were described in detail in the supplementary information, and it was conducted according to previous work^{28–30}.

Serial no.	Position	Mutation (CDC)	Mutation (AA)	Mutation type	References
1	34	c.39T>C	p.H34N	Substitution—missense	31
2	47	c.140 C>G	p.S47C	Substitution—missense	32
3	60	c.178 C>G	p.Q60E	Substitution—missense	31
4	84	c.250 C>G	p.P84A	Substitution—missense	33
5	99	c.295 G>T	p.A99S	Substitution—missense	31
6	147	c.440 G>T	p.G147V	Substitution—missense	31
7	190	c.568 A>T	p.M190L	Substitution—missense	32
8	211	c.631 G>T	p.G211W	Substitution—missense	31
9	219	c.656 G>C	p.R219P	Substitution—missense	31
10	253	c.757 C>A	p.P253T	Substitution—missense	32
11	256	c.768 C>G	p.I256M	Substitution—missense	31
12	320	c.958 C>T	p.L320F	Substitution—missense	31
13	357	c.1071 G>T	p.R357S	Substitution—missense	34
14	394	c.1180 A>G	p.N394D	Substitution—missense	35
15	395	c.1184 G>T	p.G395V	Substitution—missense	36
16	398	c.1192 G>A	p.E398K	Substitution—missense	37
17	44	c.1330 C>T	p.L444F	Substitution—missense	38
18	477	c.1429 T>C	p.W477R	Substitution—missense	36
19	491	c.1471 G>T	p.V491L	Substitution—missense	36
20	670	c.2008 G>T	p.V670L	Substitution—missense	31
21	693	c.2077 G>A	p.D693N	Substitution—missense	39
22	710	c.2082 C>A	p.R710C	Substitution—missense	40
23	720	c.2128 C>T	p.N720D	Substitution—missense	32
24	768	c.2303 G>T	p.R768L	Substitution—missense	34
25	798	c.2392 A>C	p.T798P	Substitution—missense	39

Table 1. Mutations for the ACE2 gene curated from the COSMIC database.

Ethical declaration. Ethical approval was not required since no animal or human subjects were used directly in the current study. Online public computational data was used in accordance with the guidelines of the COSMIC cancer database (<https://cancer.sanger.ac.uk/cosmic>).

Results and discussion

ACE2 mutants. Mutations found in ACE2 were of different types, including deletion, substitution, and insertion. The COSMIC database provided a set of mutations in the ACE2 gene in lung tissues, as depicted in Table 1; out of them, only missense and frameshift mutations were curated because they lead to change the resultant protein.

Predicted structures of SARS-CoV-2 spike protein. The 3D model of SARS-CoV-2 spike glycoprotein and ACE2 were derived from RCSB PDB (6M17), Fig. 2. The spike protein contains the receptor-binding domain (RBD) in either the open or closed conformations. RBD of SARS-CoV-2 spike glycoprotein was further used for docking purposes (Fig. 3).

Protein to protein docking. Docking results from the HADDOCK were derived. Since minimum HADDOCK scores depict the more binding affinity of two interacting proteins, so the best poses of the docking were obtained (Figs. 4 and 5). The top five complexes out of the 5 complexes were selected based on their high HADDOCK scores than others (Table 2). Inter and intra hydrogen bonds were visualized between SARS-CoV-2 spike glycoprotein and ACE2 mutant in docking complexes.

Protein to protein docking analysis. Due to simple scoring methods, protein docking is not reliable enough for the prediction of binding affinity between two interacting proteins in complexes⁴¹. The binding affinity of two proteins in a complex relies on dissociation constant (Kd), temperature and pH also⁴². These factors are not included in docking scoring methods of docking servers. Therefore, the binding affinity of frontier complexes (top 5) of interacting partners with the lowest docking scores was checked through the PRODIGY server (Tables 3, 4, 5, 6), binding affinity of all 5 complexes at core body temperature (37 °C) is represented in Table 3. High stability and strong binding affinity between two interacting proteins in a complex are indicated by a smaller Kd value⁴³. Spike protein-ACE mutants showed high stability and binding affinity as the temperature rises, and at the optimum temperature, stability and binding affinity became constant. Whereas the lowest protein stability and binding affinity were seen at 25 °C, and highest on the optimum temperature (Table 3, Tables

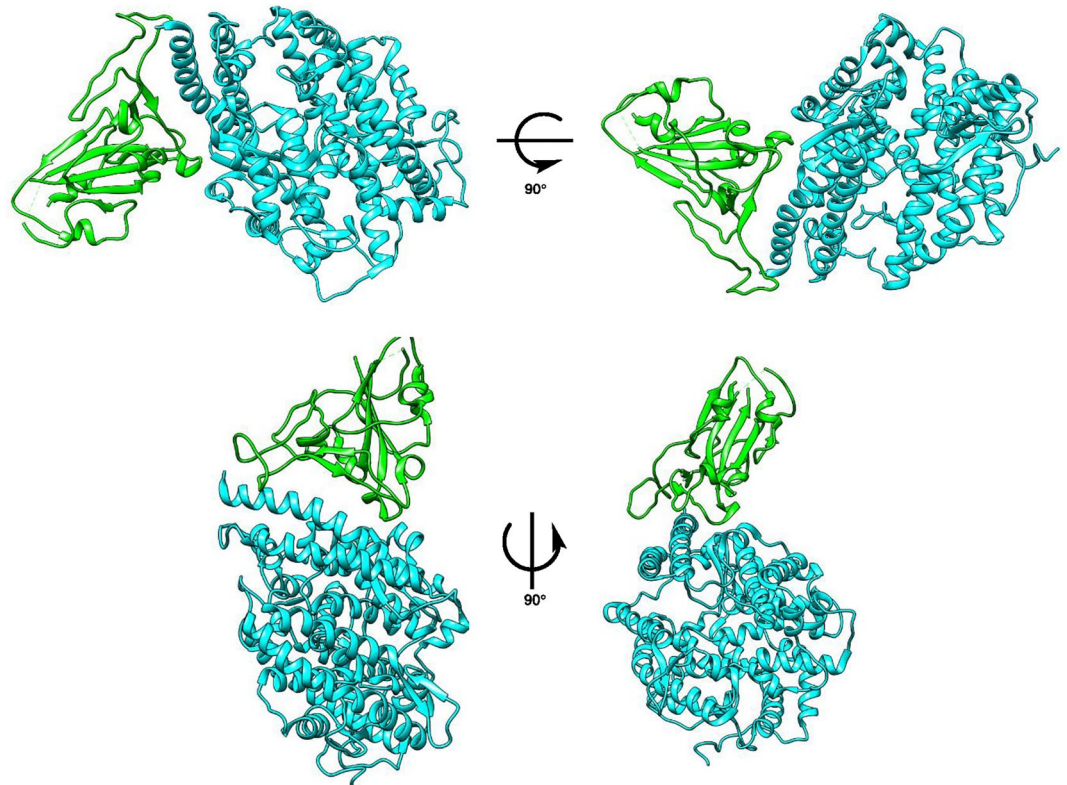


Figure 2. ACE2 and SARS-CoV-2 spike glycoprotein (6M17) (Cyan: ACE2; Green: RBD).

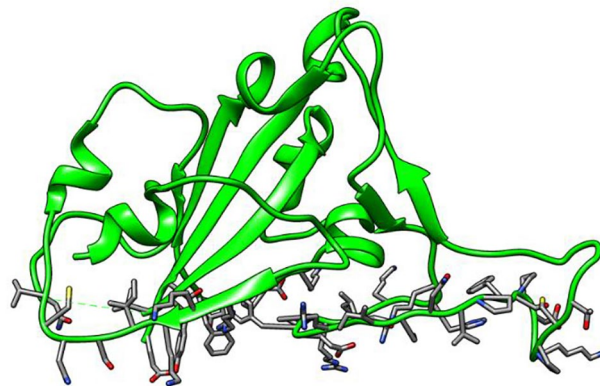


Figure 3. Receptor binding domain of SARS-CoV-2 spike protein.

SI 3–8). Other type of interactions which may affect the binding affinity (K_d), and stability (ΔG) of docked complexes are also represented in Table 4.

A change in the dissociation constant of different mutants was noted. While a decrease of 1 log fold was observed for V491L, an increase of 2 log folds was present in L320F and Q60E. However, the K_d for mutants P84A and D693N were similar to those of controls (Table 3). Moreover, a slight increase in the K_d was also observed with increasing temperatures for mutants D693N, V491L, L320F, and Q60E (Tables SI 3–8). The reason for this might be increased energy imparted with the increasing temperatures, which might be making the complexes less stable, hence an increase in K_d . Another important point to infer from these is the change in binding affinities with physiological temperatures of the human body under different states, such as normal hypothermic and hyperthermic conditions. The results also imply the less binding affinities of these ACE2 mutants under hyperthermic conditions, such as fever, and vice versa. A similar study has been performed by Basit and colleagues, where the binding affinity of ACE2 was shown to be constant at different temperatures; however, the dissociation constant was shown to be increased as the temperature increased⁴⁴.

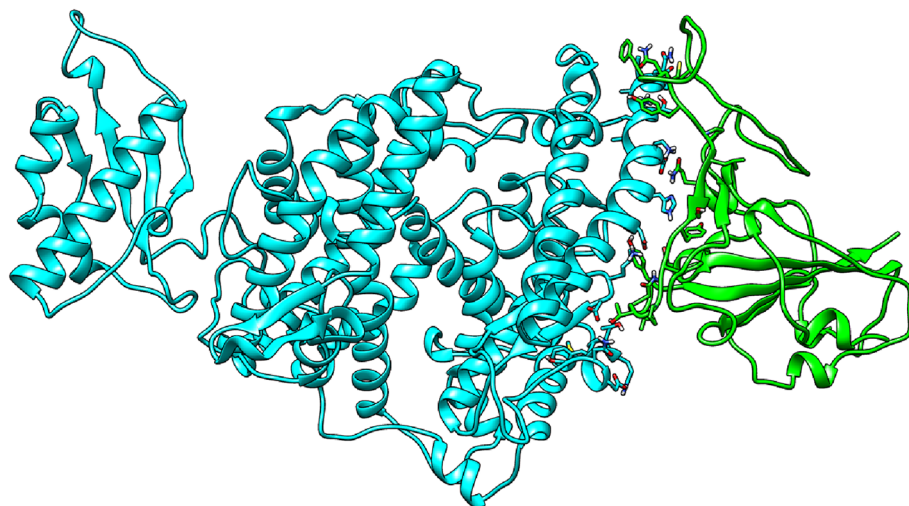


Figure 4. Spike–ACE2 protein interactions.

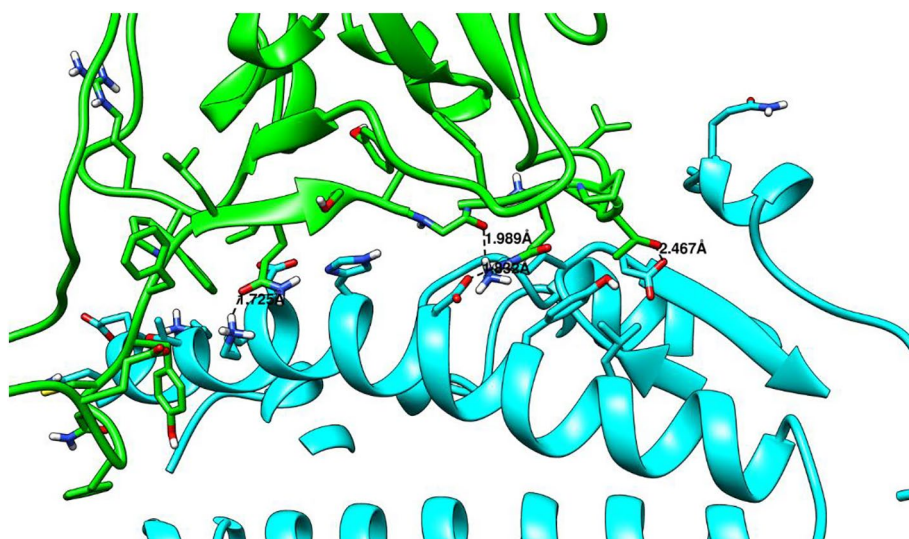


Figure 5. The interactions of Spike–ACE2 with highlighted hydrogen bonds.

Complex	HADDOCK score	Cluster size	RMSD from the overall lowest-energy structure	Z-Score	Van der Waals energy
CNT	-141.4 ± 6.6	38	0.9 ± 0.5	-1.0	-67.3 ± 5.7
P84A	-143.7 ± 9.5	135	1.3 ± 0.7	-1.7	-64.9 ± 8.1
D693N	-139.2 ± 3.2	133	0.8 ± 0.7	-1.1	-58.5 ± 4.5
V491L	-138.0 ± 2.6	90	1.7 ± 1.4	-1.3	-61.7 ± 9.3
Q60E	-136.8 ± 6.7	66	0.7 ± 0.6	-1.1	-55.9 ± 5.7
L320F	-136.3 ± 6.6	133	1.2 ± 1.3	-1.6	-57.6 ± 8.1

Table 2. Docking results of control complex (ACE2 and SARS-CoV2 spike protein) and top 5 docked molecules (different ACE2 mutants and normal SARS-CoV-2 spike protein). Docking scores are represented by HADDOCK scores (the more negative docking scores, the better the binding affinity of two proteins).

Molecular dynamic simulations. Docking procedures consider rigid and lack the free movement of the protein; thus, molecular dynamics simulation was conducted on the best six Spike-ACE2 complexes in an attempt to better understanding and validation of the docking results. The MD (Molecular Dynamics) simulations were running for a period of 500 ns, and the root means square deviation RMSDs for the complex, the

Temperature (°C)	Protein–protein complex	ΔG (kcal mol ⁻¹)	Kd (M)
37	CNT	-11.5	7.8×10^{-9}
37	P84A	-11.8	5.0×10^{-9}
37	D693N	-12.0	3.6×10^{-9}
37	V491L	-12.9	8.3×10^{-10}
37	L320F	-10.9	2.1×10^{-8}
37	Q60E	-10.9	1.9×10^{-8}

Table 3. Binding affinity of docking complex and its dissociation constant (Kd) at 37 °C. The binding affinity and stability of docked proteins are calculated in the form of ΔG (kcal mol⁻¹) and Kd (M), respectively. Smaller Kd value is showing high stability and strong binding affinity between two proteins.

Protein–protein complex	Number of interfacial contacts (ICs) per property						Non interacting surface (NIS) per property	
	ICs charged-charged	ICs charged-polar	ICs charged-apolar	ICs polar-polar	ICs polar-apolar	ICs apolar-apolar	NIS charged (%)	NIS apolar (%)
CNT	2	10	22	6	20	16	26.7	24.83
P84A	2	12	31	7	20	19	25.39	38.08
D693N	3	12	26	4	20	12	25.46	37.98
V491L	5	11	22	2	23	22	25.78	37.46
L320F	3	11	25	8	19	17	25.78	37.61
Q60E	3	11	26	7	18	16	26.07	37.46

Table 4. Type of interactions affecting the binding affinity (Kd), and stability (ΔG) of docked complexes at 37 °C.

Complex	CNT	D693N	L320F	P84A	Q60E	V491L
Average H-bonds	6.1	10.6	18.9	12.8	15.6	4.7
H-bond at 500 ns	5	11	14	12	15	4

Table 5. The average of H-bond interactions between the SARS-CoV-2 spike and ACE 2 in the last 50 ns of simulation time.

Complex	ΔG binding	Coulomb	Covalent	H-bond	Lipo	Packing	Solv_GB	VdW
CNT	-87.62	-58.59	4.64	-5.87	-25.38	-2.32	101.79	-102.82
D693N	-69.30	-65.86	5.37	-6.76	-9.70	-1.35	99.72	-92.29
Q60E	-79.59	-69.59	11.16	-7.70	-20.11	-2.97	108.04	-98.95
V491L	-72.51	-7.00	3.72	-4.90	-20.73	-2.61	53.28	-94.95
L320F	-86.61	-46.38	0.19	-6.54	-20.96	-2.84	80.00	-90.16
P84A	-75.96	-28.72	6.28	-7.48	-20.32	-4.02	64.48	-86.32

Table 6. Prime MM-GBSA energies for Ligands binding at the active site of COVID-19 main protease. *Coulomb* Coulomb energy, *Covalent* Covalent binding energy, *VdW* Van der Waals energy, *Lipo* Lipophilic energy, *Solv_GB* Generalized Born electrostatic solvation energy, *H-bond* Hydrogen-bonding energy.

SARS-CoV-2 Spike, and the ACE2 were reported and analyzed. The total energy along with the potential energy of each system was monitored during the simulation and the average was reported in Table S1.

Proteins and complex Root Mean Square Deviation (RMSD) analysis. To monitor the impact of simulation on the stability of the SARS-CoV-2 Spike–ACE2 complexes, the root mean square deviations values were reported as a function of time for all C_α atoms of the proteins with respect to their initial positions. The RMSD results for SARS-CoV-2 Spike–ACE2 complexes were plotted as a function of simulation time and presented in Fig. 6, while the RMSD for the SARS-CoV-2 spike protein was presented in Fig. 7; finally, the RMSD of the ACE2 was plotted in Fig. 8. As it can be seen from Fig. 6, most complexes showed a stable most complexes showed stability with RMSD around 3–4 Å except for CNT-complex and L320F-complex. The CNT-complex, D693N-complex, and the L320F-complex fluctuated till around 100 ns, 180 ns, and 450 ns, respectively. Other

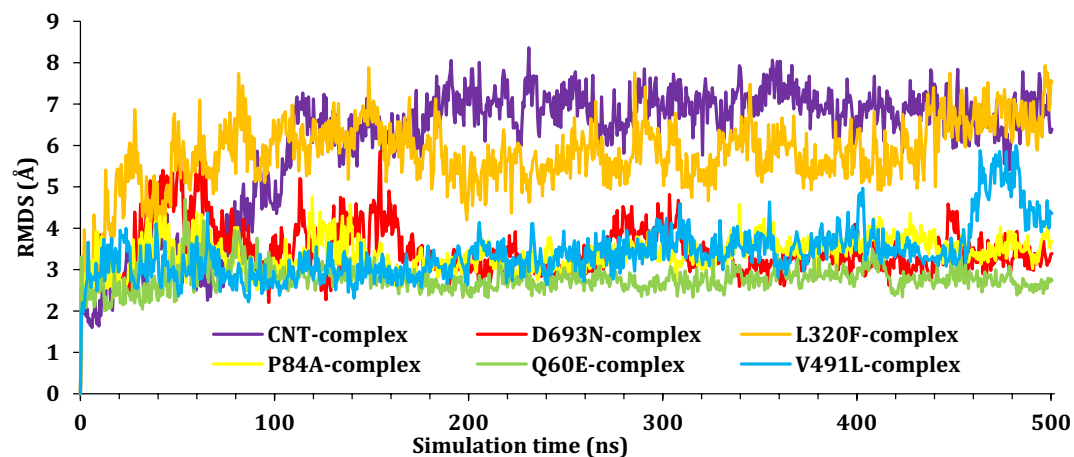


Figure 6. The RMSD for Ca atoms (Å) with respect to the initial structure as a function of simulation time (ns) for the six complexes.

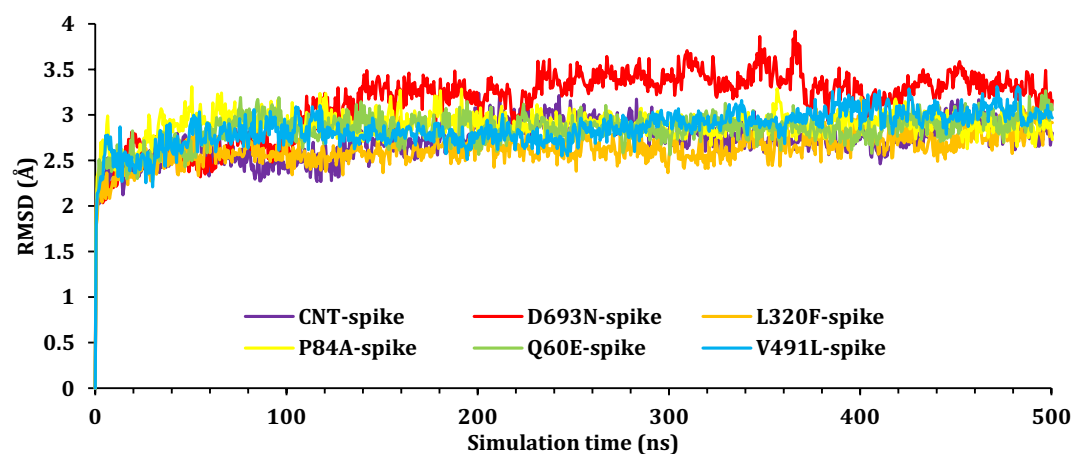


Figure 7. The RMSD for Ca atoms (Å) with respect to the initial structure as a function of simulation time (ns) for the SARS-CoV-2 spike protein.

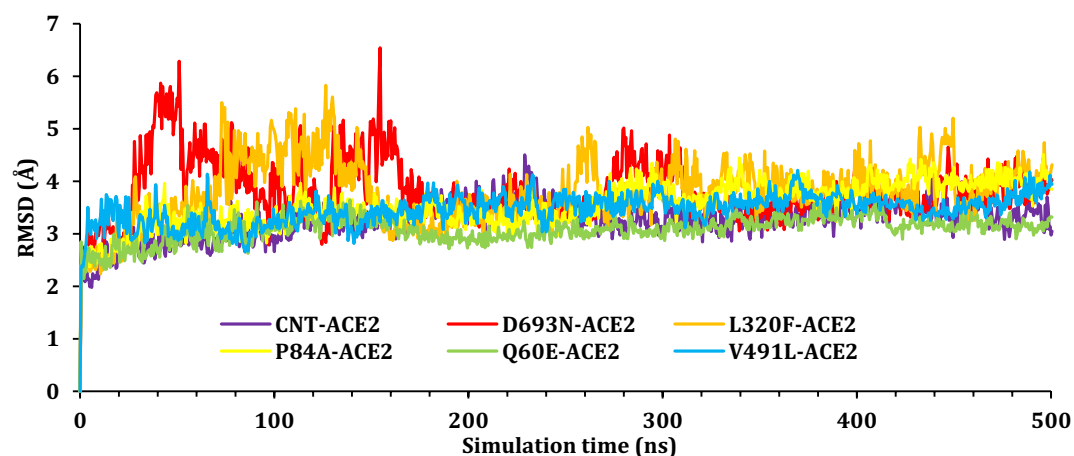


Figure 8. The RMSD for Ca atoms (Å) with respect to the initial structure as a function of simulation time (ns) for the ACE2 protein.

complexes showed stability at an early stage of the simulation and held that stability during the simulations; at around 470–490 ns. A jump in the RMSD for V491L was observed and due to the movement of the ACE domain.

Further, the RMSD of each chain was studied individually with respect to its original position within the complex. Figure 7 shows the fluctuation of the spike domain of the protein–protein complex; as it can be seen most spike domains are at around 3 Å, which is acceptable fluctuations for proteins with such size. The only notable change occurs within the D693N Spike, and this fluctuation is due to the movement of the nonstructural part (loops) of the protein Figure S11 (“Supplementary information”).

Figure 8 presents the RMSD of the ACE2 protein moiety; again, most proteins reached a plateau at around 200 ns, with an RMSD of ~ 3.5 Å. ACE2–D693N showed a high fluctuation at about 20–180 ns due to the instability of the tail (N-terminal) of the ACE2 during simulation, Fig. S12 (“Supplementary information”). Figure 4 shows the position of the SARS-CoV-2 Spike with regard to ACE 2 at the beginning of the simulation (0 ns) and the end of the simulations (500 ns). In general, all ACE 2 hold position to the SARS-CoV-2 Spike; the hydrogen bonding breaking, and formation was monitored during the simulation and will be discussed later.

Figure 9 shows a surface presentation of the first and last frame of the trajectories for a better presentation of the movement of the two proteins with respect to simulation time, and the residuals interactions will be discussed later through the course of the manuscript.

Root mean square fluctuation (RMSF). The RMSF helps characterize local changes along the protein chain. The RMSF plot peaks indicate the part of the protein that fluctuates the most during the simulation. Typically, the N- and the C- terminal fluctuates the most. Also, secondary structures like α -helices and β -strand consider more rigid than the unstructured part of the protein and fluctuate less than loops regions. Figure 10 presents the RMSF of all complexes, residue index from 0 to 182 corresponding to the SARS-CoV-2 Spike, while the residue index from 183 to 890 corresponding to ACE 2. The fluctuations between 15 and 65 corresponding to residues ARG357 to PHE400 of the SARS-CoV-2 Spike, and the peak at 182 presents the C- terminal of the SARS-CoV-2 Spike. The peak at 298 refers to the fluctuation of ASN134, PRO135, and ASP136 of the ACE 2 protein, and the peak at 790 is due to the fluctuations of the following residues ALA627, LEU628, and GLY629. The RMSF of L320F is provided in the supplementary information for the residual numbering example. Protein secondary structure elements (SSE) like α -helices and β -strands are monitored throughout the simulation for all complexes and reported in SI 9–20.

Hydrogen bond analysis. Hydrogen bonds consider the most important interactions in protein chemistry; the protein structure depends on this interaction, including 2D, 3D, and quaternary structures. The hydrogen bond between the SARS-CoV-2 Spike and the ACE 2 was monitored during the simulation time and plotted as a function of time (Fig. 8). As can be seen from Fig. 6, the H-bond between the two proteins fluctuated between 0 and 25 H-bonds, the average hydrogen bond interaction for the last 50 ns was calculated and reported in Table 5. L320F showed the most H-bond interactions with an average of 18.9 H-bonds, followed by Q60E and P84A with H-bonds interactions of 15.6 and 12.8, respectively. The number of H-bond interactions in the last frame was also reported in Table 5, and some of the residuals involved in these H-bonds were reported in Table SI 2. H-bond interactions between the SARS-CoV-2 Spike and ACE 2 during simulations, as seen in Fig. 11. Using Ligplot, all the complexes were examined for hydrogen bonding interactions between 0 and 500 ns. (Fig. 11 and SI 3–8).

MM-GBSA calculations. Schrodinger software has a python script called thermal_mmgsba.py, which was used to calculate the MM-GBSA from the trajectories and extract the average binding energies, including the average MM-GBSA binding energy, average Coulomb energy, average Covalent binding energy, average Van der Waals energy, average Lipophilic energy, average Generalized Born electrostatic solvation energy, and average Hydrogen-bonding energy. All the obtained energies are shown in Table 6.

As it can be seen from Table 6, CNT and L320F showed the highest binding energy with -87.62 and -86.61 kcal mol⁻¹, respectively, while D693N showed the lowest binding energy -69.30 kcal mol⁻¹. Coulomb energy is associated with electrostatic forces of the system and reflects the ionic interactions of the system. CNT, D693N, and Q60E showed good ionic interactions, with Q60E having the highest value, while V491L seem to lose ionic interactions. Most complexes showed good H-bond energies, as well as good Van der Waals energies.

Conclusion

In this research, the relationship between lung cancer and COVID-19 was addressed at a molecular level through computational study. The binding affinity of the viral spike protein of the SARS-CoV-2 glycoprotein towards a mutate ACE2 was investigated, using both docking and molecular dynamic simulation approaches. Among 25 selected mutations, it was found that five mutations have higher binding energies than others from a docking perspective. These five complexes' stability was studied and investigated further considering molecular dynamics simulations. Finally, the binding free energy calculations using the MM-GBSA approach were implemented and showed that these mutations have a binding energy of the following order CNT > L320F > Q60E > P84A > V491L > D693N. These findings suggest that some cancer patients will be less affected than others, even though most reported mutations were not within the active site of interactions; the binding energy was affected by these mutations. These results somehow oppose those clinical studies which state that lung cancer patients are more prone to COVID-19 infection based on the healthcare scenario they have observed. That might be due to the contact of healthcare system with COVID-19 to lung cancer patients during treatment. However, more in-depth studies are needed to be performed to reach at a valid endpoint.

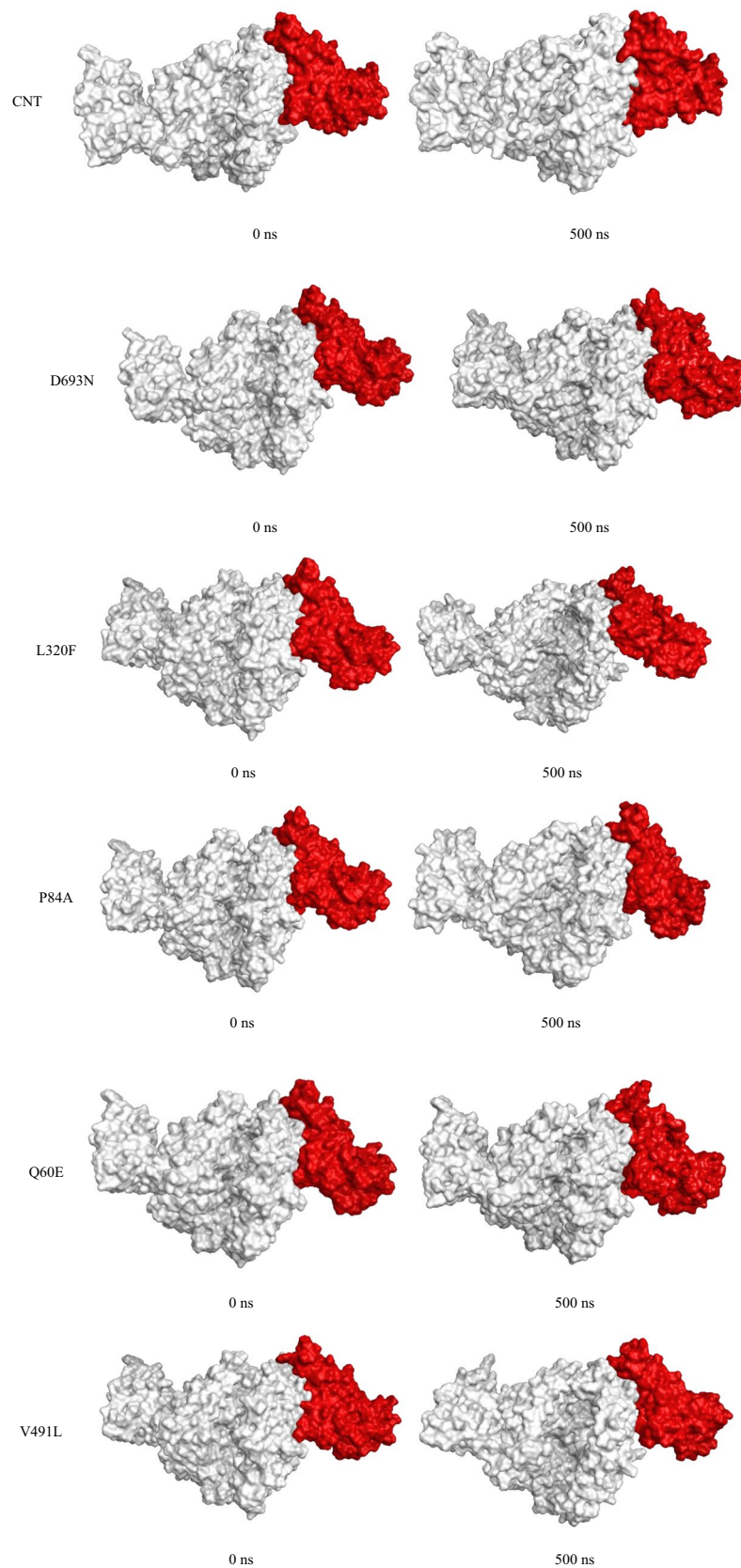


Figure 9. Snapshot of each complex (ACE 2, white; SARS-CoV-2 spike, red) at 0 ns and 500 ns.

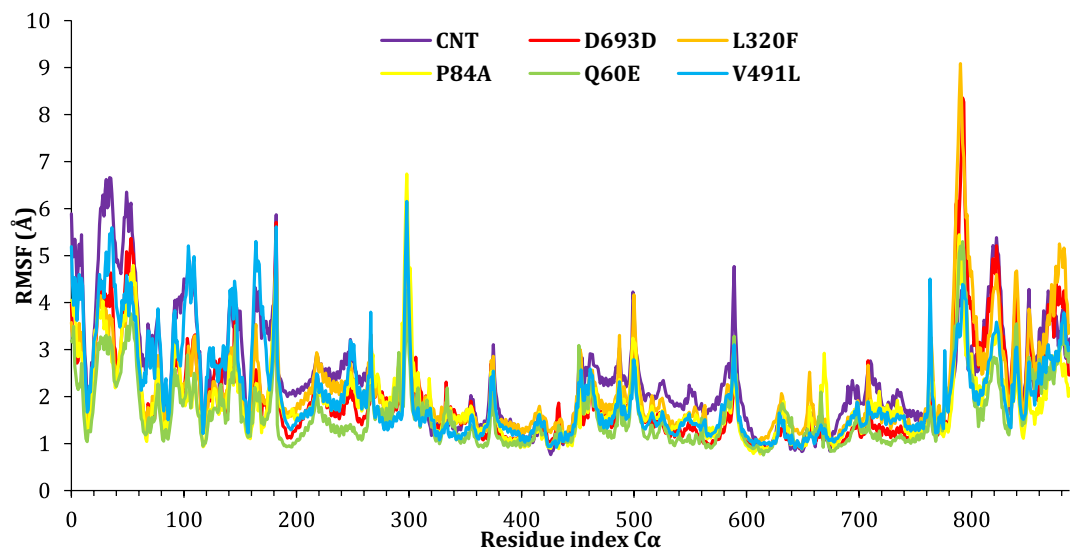


Figure 10. The RMSF of the complexes during simulation trajectories.

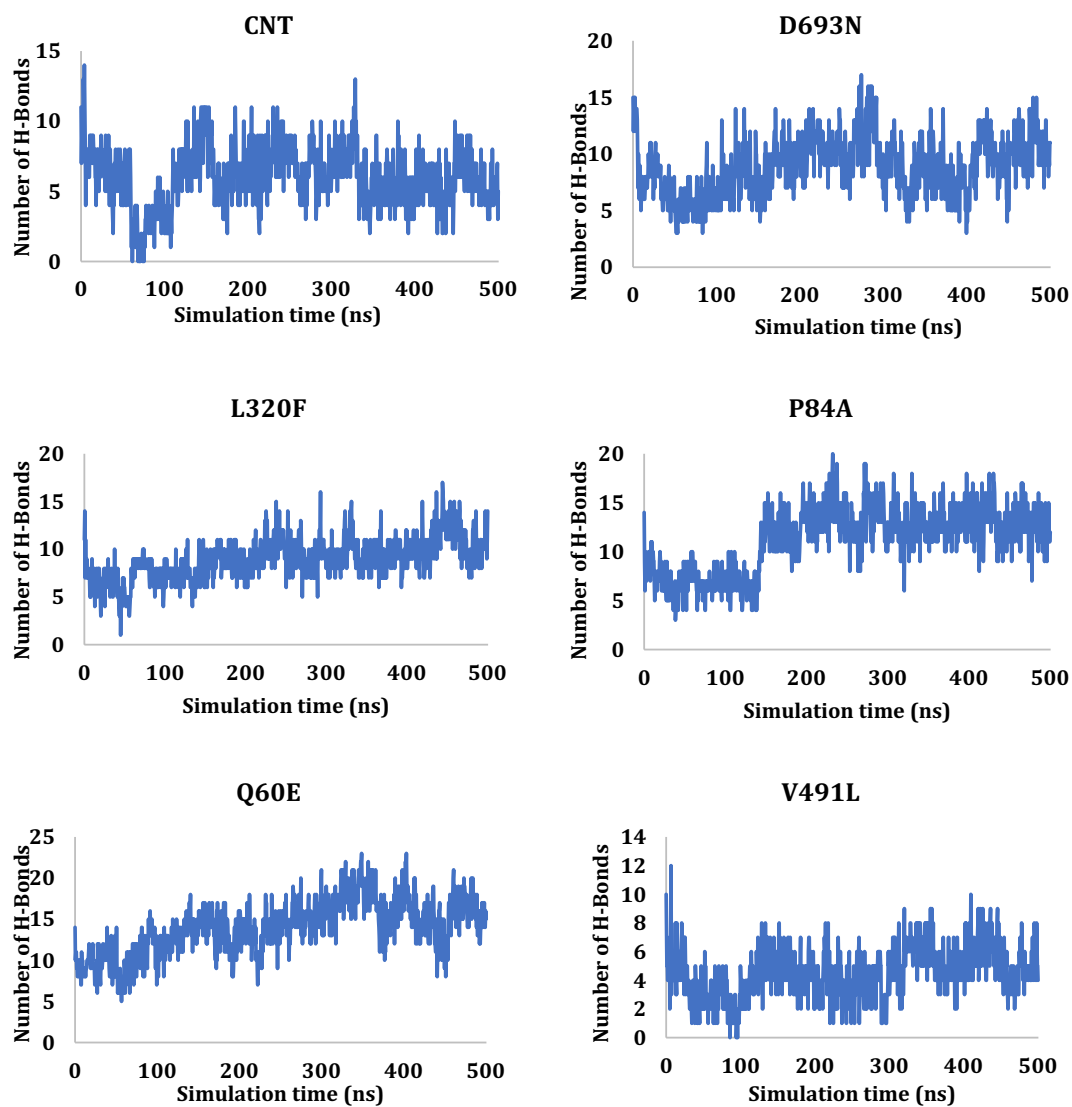


Figure 11. H-bonds interactions between the SARS-CoV-2 Spike and ACE 2 during simulations.

Data availability

COSMIC Cancer Database (<https://cancer.sanger.ac.uk/cosmic>) was used to curate ACE2 lung cancer mutations. These mutations were publicly available on the website of database to be used for research work. These mutations were further analyzed to check their susceptibility for COVID-19. Throughout the manuscript used mutations were cited properly (Table 1). According to the rules and regulations of the database, data can be used for research purposes and correct citation is required, and so it was the done the same.

Received: 16 December 2021; Accepted: 25 April 2022

Published online: 12 May 2022

References

1. Mc Namara, K., Alzubaidi, H. & Jackson, J. K. Cardiovascular disease as a leading cause of death: How are pharmacists getting involved?. *Integr. Pharm. Res. Practice* **8**, 1 (2019).
2. Ahmad, A & Gadgeel, S (ed). Lung cancer statistics. in *Lung Cancer and Personalized Medicine*, 1–19. (Springer, 2016).
3. Barta, J. A., Powell, C. A., & Wisnivesky, J. P. Global epidemiology of lung cancer. *Ann. Glob. Health.* **85**(1) (2019).
4. Hirsch, F. R. *et al.* Lung cancer: Current therapies and new targeted treatments. *Lancet* **389**(10066), 299–311 (2017).
5. Zhong, S. *et al.* Physical activity and risk of lung cancer: A meta-analysis. *Clin. J. Sport Med.* **26**(3), 173–181 (2016).
6. Perlikos, F., Harrington, K. J. & Syrigos, K. N. Key molecular mechanisms in lung cancer invasion and metastasis: A comprehensive review. *Crit. Rev. Oncol. Hematol.* **87**(1), 1–11 (2013).
7. Amos, C., Xu, W., & Spitz, M. Is there a genetic basis for lung cancer susceptibility? in *Chemoprevention of Cancer*, 3–12. (Springer, 1999).
8. Hung, R. J. *et al.* A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* **452**(7187), 633–637 (2008).
9. Wang, Y. *et al.* Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat. Genet.* **40**(12), 1407–1409 (2008).
10. Gorbalenya, A. E., Baker, S. C., Baric, R., Groot, R. J. de, Drosten, C., Gulyaeva, A. A., Haagmans, B. L., Lauber, C., Leontovich, A. M., & Neuman, B. W. (2020). *Severe Acute Respiratory Syndrome-Related Coronavirus: The Species and Its Viruses—A Statement of the Coronavirus Study Group.*
11. Dai, M. *et al.* Patients with cancer appear more vulnerable to SARS-CoV-2: A multicenter study during the COVID-19 outbreak. *Cancer Discov.* **10**(6), 783–791 (2020).
12. Millet, J. K. & Whittaker, G. R. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Res.* **202**, 120–134. <https://doi.org/10.1016/j.virusres.2014.11.021> (2015).
13. Donoghue, M. *et al.* A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1–9. *Circ. Res.* **87**(5), e1–e9 (2000).
14. Ferrario, C. M., Trask, A. J. & Jessup, J. A. Advances in biochemical and functional roles of angiotensin-converting enzyme 2 and angiotensin-(1–7) in regulation of cardiovascular function. *Am. J. Physiol.-Heart Circ. Physiol.* **289**(6), H2281–H2290 (2005).
15. Passaro, A. *et al.* Severity of COVID-19 in patients with lung cancer: Evidence and challenges. *J. Immunother. Cancer* **9**(3), e002266. <https://doi.org/10.1136/jitc-2020-002266> (2021).
16. Jacobo, R. *et al.* COVID-19 and lung cancer: A greater fatality rate?. *Lung Cancer* **146**, 19–22 (2020).
17. Liang, W. *et al.* Cancer patients in SARS-CoV-2 infection: A nationwide analysis in China. *Lancet Oncol.* **21**(3), 335–337 (2020).
18. Sidaway, P. COVID-19 and cancer: What we know so far. *Nat. Rev. Clin. Oncol.* **17**(6), 336 (2020).
19. Zhang, L. *et al.* Clinical characteristics of COVID-19-infected cancer patients: A retrospective case study in three hospitals within Wuhan, China. *Ann. Oncol.* **31**, 894–901 (2020).
20. Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**(6485), 1444–1448 (2020).
21. Schrodinger, L. J. V. (2010). The PyMOL molecular graphics system. *1*(5).
22. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**(13), 1605–1612 (2004).
23. Van Zundert, G. C. P. *et al.* The HADDOCK2.2 web server: User-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* **428**(4), 720–725 (2016).
24. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**(7807), 215–220 (2020).
25. Vangone, A. & A.M.J.E. Bonvin. Contacts-based prediction of binding affinity in protein–protein complexes. *elife.* **4**, e07454 (2015).
26. Xue, L. C., Rodrigues, J. P., Kastriitis, P. L., Bonvin, A. M. & Vangone, A. PRODIGY: A web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics* **32**(23), 3676–3678 (2016).
27. Bowers, K. J., Chow, D. E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., *et al.* Scalable algorithms for molecular dynamics simulations on commodity clusters. in *SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, 43–43. (IEEE, 2006).
28. Elkwaifi, G., Mohamed, N., Elabbar, F., & Alnajjar, R. Flavonoid content of the Libyan OnosmaCyrenaicum: Isolation, identification, electronic chemical reactivity, drug likeness, docking, and MD study. *J. Biomol. Struct. Dyn. AHEAD-OF-PRINT*: 1–16 (2021).
29. Alnajjar, R., Mohamed, N. & Kawafi, N. Bicyclo [1.1.1] Pentane as phenyl substituent in atorvastatin drug to improve physicochemical properties: drug-likeness, DFT, pharmacokinetics, docking, and molecular dynamic simulation. *J. Mol. Struct.* **1230**, 129628 (2021).
30. Alnajjar, R., Mostafa, A., Kandeil, A. & Al-Karmalawy, A. A. Molecular docking, molecular dynamics, and in vitro studies reveal the potential of angiotensin II receptor blockers to inhibit the COVID-19 main protease. *Heliyon* **6**(12), e05641 (2020).
31. Toonkel, R. L., Borczuk, A. C. & Powell, C. A. Tgf-beta signaling pathway in lung adenocarcinoma invasion. *J. Thorac. Oncol.* **5**(2), 153–157 (2010).
32. McMillan, E. A. *et al.* Chemistry-first approach for nomination of personalized treatment in lung cancer. *Cell* **173**(4), 864–878. e829 (2018).
33. Liu, P. *et al.* Identification of somatic mutations in non-small cell lung carcinomas using whole-exome sequencing. *Carcinogenesis* **33**(7), 1270–1276 (2012).
34. Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**(6), 1107–1120 (2012).
35. Seo, J.-S. *et al.* The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome Res.* **22**(11), 2109–2119 (2012).
36. He, L. & Hannon, G. J. MicroRNAs: Small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5**(7), 522–531 (2004).
37. Rudin, C. M. *et al.* Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat. Genet.* **44**(10), 1111–1116 (2012).
38. Wagner, A. H. *et al.* Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. *Nat. Commun.* **9**(1), 1–11 (2018).

39. Campbell, P. J., Getz, G., Korbel, J. O., Stuart, J. M., Jennings, J. L., Stein, L. D., Perry, M. D., Nahal-Bose, H. K., Ouellette, B. F. F., Li, C. H., Rheinbay, E., Nielsen, G. P., Sgroi, D. C., Wu, C.-L., Faquin, W. C., Deshpande, V., Boutros, P. C., Lazar, A. J., Hoadley, K. A., Louis, D. N., Dursi, L. J., Yung, C. K., Bailey, M. H., Saksena, G., Raine, K. M., Buchhalter, I., Kleinheinz, K., Schlesner, M., Zhang, J., Wang, W., Wheeler, D. A., Ding, L., Simpson, J. T., O'Connor, B. D., Yakneen, S., Ellrott, K., Miyoshi, N., Butler, A. P., Royo, R., Shorsler, S. I., Vazquez, M., Rausch, T., Tiao, G., Waszak, S. M., Rodriguez-Martin, B., Shringarpure, S., Wu, D.-Y., Demidov, G. M., Delaneau, O., Hayashi, S., Imoto, S., Habermann, N., Segre, A. V., Garrison, E., Cafferkey, A., Alvarez, E. G., Heredia-Genestar, J. M., Muias, E., Drechsel, O., Bruzos, A. L., Temes, J., Zamora, J., Baez-Ortega, A., Kim, H.-L., Mashl, R. J., Ye, K., DiBiase, A., Huang, K.-l., Letunic, I., McLellan, M. D., Newhouse, S. J., Shmaya, T., Kumar, S., Wedge, D. C., Wright, M. H., Yellapantula, V. D., Gerstein, M., Khurana, E., Marques-Bonet, T., Navarro, A., Bustamante, C. D., Siebert, R., Nakagawa, H., Easton, D. F., Ossowski, S., Tubio, J. M. C., De La Vega, F. M., Estivill, X., Yuen, D., Mihaiescu, G. L., Omberg, L., Ferretti, V., Sabarinathan, R., Pich, O., Gonzalez-Perez, A., Taylor-Weiner, A., Fittall, M. W., Demeulemeester, J., Tarabichi, M., Roberts, N. D., Van Loo, P., Cortés-Ciriano, I., Urban, L., Park, P., Zhu, B., Pitkänen, E., Li, Y., Saini, N., Klimczak, L. J., Weischenfeldt, J., Sidiropoulos, N., Alexandrov, L. B., Rabionet, R., Escaramis, G., Bosio, M., Holik, A. Z., Susak, H., Prasad, A., Erkek, S., Calabrese, C., Raeder, B., Harrington, E., Mayes, S., Turner, D., Juul, S., Roberts, S. A., Song, L., Koster, R., Mirabello, L., Hua, X., Tanskanen, T. J., Tojo, M., Chen, J., Aaltonen, L. A., Rättsch, G., Schwarz, R. F., Butte, A. J., Brazma, A., Chanock, S. J., Chatterjee, N., Stegle, O., Harismendy, O., Bova, G. S., Gordenin, D. A., Haan, D., Sieverling, L., Feuerbach, L., Chalmers, D., Joly, Y., Knoppers, B., Molnár-Gábor, F., Phillips, M., Thorogood, A., Townend, D., Goldman, M., Fonseca, N. A., Xiang, Q., Craft, B., Piñeiro-Yáñez, E., Muñoz, A., Petryszak, R., Füllgrabe, A., Al-Shahrour, F., Keays, M., Haussler, D., Weinstejn, J., Huber, W., Valencia, A., Papatheodorou, I., Zhu, J., Fan, Y., Torrents, D., Bieg, M., Chen, K., Chong, Z., Cibulskis, K., Eils, R., Fulton, R. S., Gelpi, J. L., Gonzalez, S., Gut, I. G., Hach, F., Heinold, M., Hu, T., Huang, V., Hutter, B., Jäger, N., Jung, J., Kumar, Y., Lalasingh, C., Leshchiner, I., Livitz, D., Ma, E. Z., Maruvka, Y. E., Milovanovic, A., Nielsen, M. M., Paramasivam, N., Pedersen, J. S., Puiggròs, M., Sahinalp, S. C., Sarrafí, I., Stewart, C., Stobbe, M. D., Wala, J. A., Wang, J., Wendl, M., Werner, J., Wu, Z., Xue, H., Yamaguchi, T. N., Yellapantula, V., Davis-Dusenbery, B. N., Grossman, R. L., Kim, Y., Heinold, M. C., Hinton, J., Jones, D. R., Menzies, A., Stebbings, L., Hess, J. M., Rosenberg, M., Dunford, A. J., Gupta, M., Imielinski, M., Meyerson, M., Beroukhi, R., Reimand, J., Dhingra, P., Favero, F., Dentro, S., Wintersinger, J., Rudneva, V., Park, J. W., Hong, E. P., Heo, S. G., Kahles, A., Lehmann, K.-V., Soulette, C. M., Shiraishi, Y., Liu, F., He, Y., Demircioğlu, D., Davidson, N. R., Greger, L., Li, S., Liu, D., Stark, S. G., Zhang, F., Amin, S. B., Bailey, P., Chateigner, A., Frenkel-Morgenstern, M., Hou, Y., Huska, M. R., Kilpinen, H., Lamaze, F. C., Li, C., Li, X., Li, X., Liu, X., Marin, M. G., Markowski, J., Nandi, T., Ojesina, A. I., Pan-Hammarström, Q., Park, P. J., Pedamallu, C. S., Su, H., Tan, P., Teh, B. T., Wang, J., Xiong, H., Ye, C., Yung, C., Zhang, X., Zheng, L., Zhu, S., Awadalla, P., Creighton, C. J., Wu, K., Yang, H., Göke, J., Zhang, Z., Brooks, A. N., Fittall, M. W., Martincorena, I., Rubio-Perez, C., Juul, M., Schumacher, S., Shapira, O., Tamborero, D., Mularoni, L., Hornshøj, H., Deu-Pons, J., Muiños, F., Bertl, J., Guo, Q., Gonzalez-Perez, A., & Xiang, Q. Pan-cancer analysis of whole genomes. *Nature*. **578**(7793), 82–93 (2020).
40. Li, C. *et al.* Whole exome sequencing identifies frequent somatic mutations in cell–cell adhesion genes in Chinese patients with lung squamous cell carcinoma. *Sci. Rep.* **5**, 14237 (2015).
41. Ramirez, D. & Caballero, J. Is it reliable to use common molecular docking methods for comparing the binding affinities of enantiomer pairs for their protein target?. *Int. J. Mol. Sci.* **17**(4), 525 (2016).
42. Kastriitis, P. L. & Bonvin, A. M. Are scoring functions in protein–protein docking ready to predict interactomes? Clues from a novel binding affinity benchmark. *J. Proteome Res.* **9**(5), 2216–2225 (2010).
43. Johnson, R. J., McCoy, J. G., Bingman, C. A., Phillips, G. N. Jr. & Raines, R. T. Inhibition of human pancreatic ribonuclease by the human ribonuclease inhibitor protein. *J. Mol. Biol.* **368**(2), 434–449 (2007).
44. Basit, A., Ali, T. & Rehman, S. U. Truncated human angiotensin converting enzyme 2; a potential inhibitor of SARS-CoV-2 spike glycoprotein and potent COVID-19 therapeutic agent. *J. Biomol. Struct. Dyn.* **39**(10), 3605–3614. <https://doi.org/10.1080/0739102.2020.1768150> (2021).

Acknowledgements

The authors would like to thank Taif University Researchers Supporting Project number (TURSP-2020/330), Taif University, Taif, Saudi Arabia for providing computer facilities.

Author contributions

Z.K. and A.T.K. designed the studied and drafted the manuscript. R.A. and E.S. performed molecular dynamic simulation. A.R.S. reviewed and supervised the study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-11805-5>.

Correspondence and requests for materials should be addressed to Z.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022