33 (2022) 201040

Contents lists available at ScienceDirect

Human Gene

journal homepage: www.journals.elsevier.com/human-gene

The impacts of 13 novel mutations of SARS-CoV-2 on protein dynamics: In silico analysis from Turkey

Sezin Unlu [a,*], Aylin Uskudar-Guclu [a], Isli Cela [b]

[a] Baskent University, Medical Faculty, Department of Medical Microbiology, Ankara, Turkey
[b] Bilkent University, Department of Molecular Biology and Genetics, Ankara, Turkey

ARTICLE INFO

Keywords:
SARS-CoV-2
Spike Glycoprotein
Mutations
In silico
Protein dynamics
Turkey

ABSTRACT

SARS-CoV-2 inherits a high rate of mutations making it better suited to the host since its fundamental role in evolution is to provide diversity into the genome. This research aims to identify variations in Turkish isolates and predict their impacts on proteins. To identify novel variations and predict their impacts on protein dynamics, in silico methodology was used. The 411 sequences from Turkey were analysed. Secondary structure prediction by Garnier-Osguthorpe-Robson (GOR) was used. To find the effects of identified Spike mutations on protein dynamics, the SARS-CoV-2 structures (PDB:6VYB, 6M0J) were uploaded and predicted by Cutoff Scanning Matrix (mCSM), DynaMut and MutaBind2. To understand the effects of these mutations on Spike protein molecular dynamics (MD) simulation was employed. Turkish sequences were aligned with sequences worldwide by MUSCLE, and phylogenetic analysis was performed via MegaX. The 13 novel mutations were identified, and six of them belong to spike glycoprotein. Ten of these variations revealed alteration in the secondary structure of the protein. Differences of free energy between the reference sequence and six mutants were found below zero for each of six isolates, demonstrating these variations have stabilizing effects on protein structure. Differences in vibrational entropy calculation revealed that three variants have rigidification, while the other three have a flexibility effect. MD simulation revealed that point mutations in spike glycoprotein and RBD:ACE-2 complex cause changes in protein dynamics compared to the wild-type, suggesting possible alterations in binding affinity. The phylogenetic analysis showed Turkish sequences distributed throughout the tree, revealing multiple entrances to Turkey.

1. Introduction

SARS-CoV-2 is a single-stranded RNA virus having a positive polarity. Its size differs between 26 to 32 kb (Ouassou et al., 2020) and its genome comprises 6-11 open reading frames (Kumar et al., 2020). RNA viruses undergo 10^{-4} substitutions per year. The high mutation rate of RNA viruses is considered beneficial for viruses since it makes them better suited to the host due to their fundamental role in evolution by providing diversity into the genome. SARS-CoV-2 inherits mutations at a high rate, and most of them evolve quickly. Tracking these mutations is pivotal for developing vaccines and intervention strategies since they may affect the efficacy of vaccines (Wu et al., 2020).

Amino acid alterations are caused by non-synonymous mutations, which can change the protein structure of the virus. Alteration in protein structure can affect viral particle association during the replication and assembly process (Li et al., 2021). Besides, these amino acid substitutions may affect the binding affinity of the virus to the host receptor, angiotensin-converting enzyme-2 (ACE-2) (Yang et al., 2020). The spike protein of SARS-CoV-2 have two subunits; S1 and S2. While S1 is responsible for recognizing and binding ACE-2, S2 is responsible for membrane fusion (Unlu et al., 2021). Due to spike protein's crucial roles, assessing the wide mutational profile of spike protein would be vital for designing vaccines targeting spike protein (Begum et al., 2021).

Since the protein structure of a virus plays a pivotal role in its function, alteration in protein structure can affect the virus's characteristics such as its function, virulence, infectivity and transmissibility (Nguyen et al., 2020). Protein dynamics studies give significant insights regarding protein structure, protein stability and flexibility and protein-protein interaction (Mahtarin et al., 2020). Previous molecular dynamics (MD) simulations demonstrated that tested drugs could contact residues on RBD, thereby disrupting its interaction with ACE-2 and thus providing drug alternatives against SARS-CoV-2 (Deganutti et al., 2020).

* Corresponding author at: Bağlıca Kampüsü, Dumlupınar Blv. 20. Km, 06810 Etimesgut, Ankara, Turkey.
E-mail address: sezinnunlu@gmail.com (S. Unlu).

https://doi.org/10.1016/j.humgen.2022.201040
Received 1 November 2021; Received in revised form 19 February 2022; Accepted 27 April 2022
Available online 10 May 2022

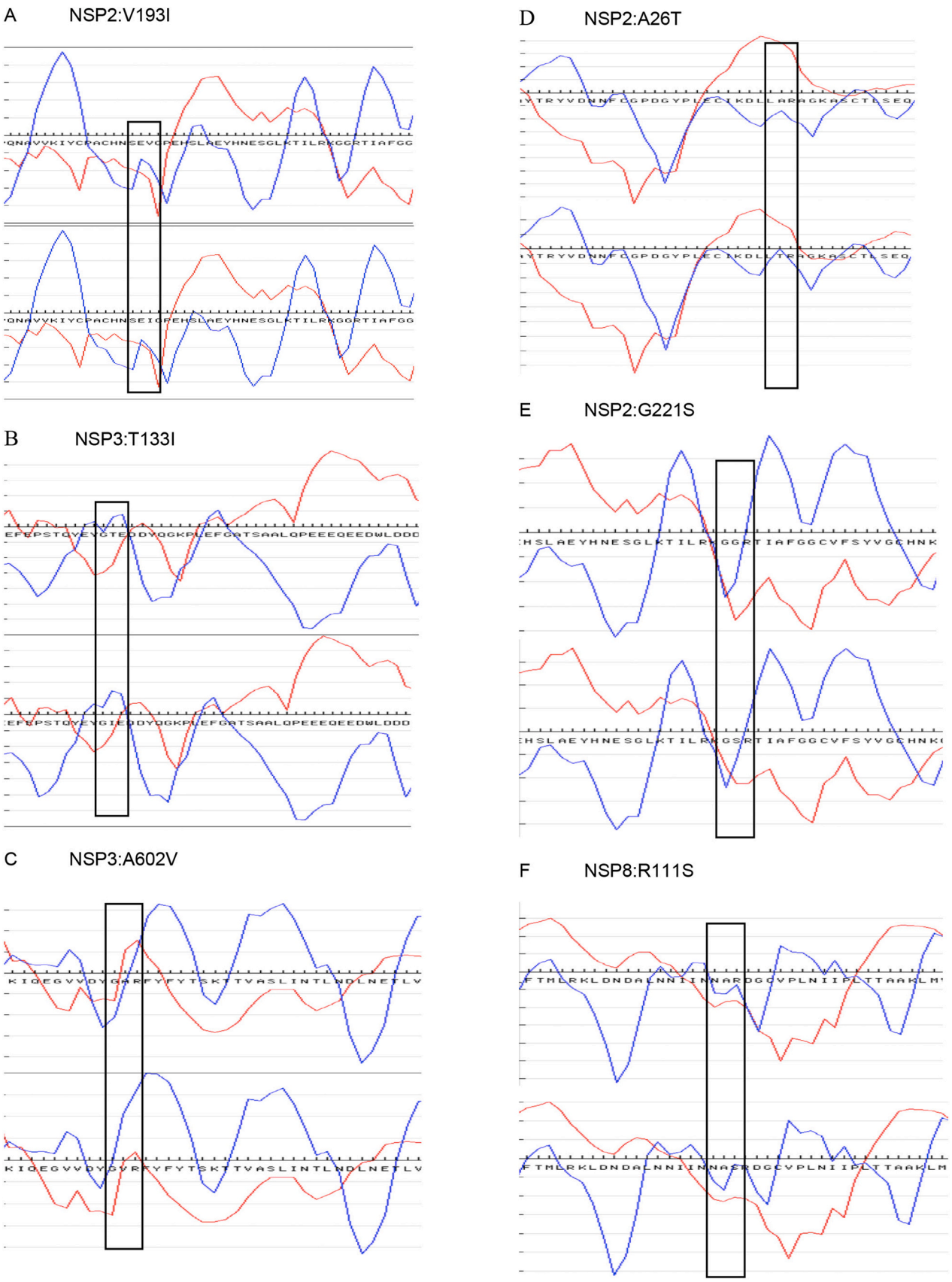

Fig. 1. Secondary structure analysis of Wuhan reference genome vs. (a,b,c) EPI_ISL_428713, (d) EPI_ISL_429863, (e,f) EPI_ISL_509415, (g) EPI_ISL_735349, (h) EPI_ISL_735339,(i) EPI_ISL_730569, (j) EPI_ISL_730573, (k) EPI_ISL_730576, (l) EPI_ISL_735312, (m) EPI_ISL_812895.

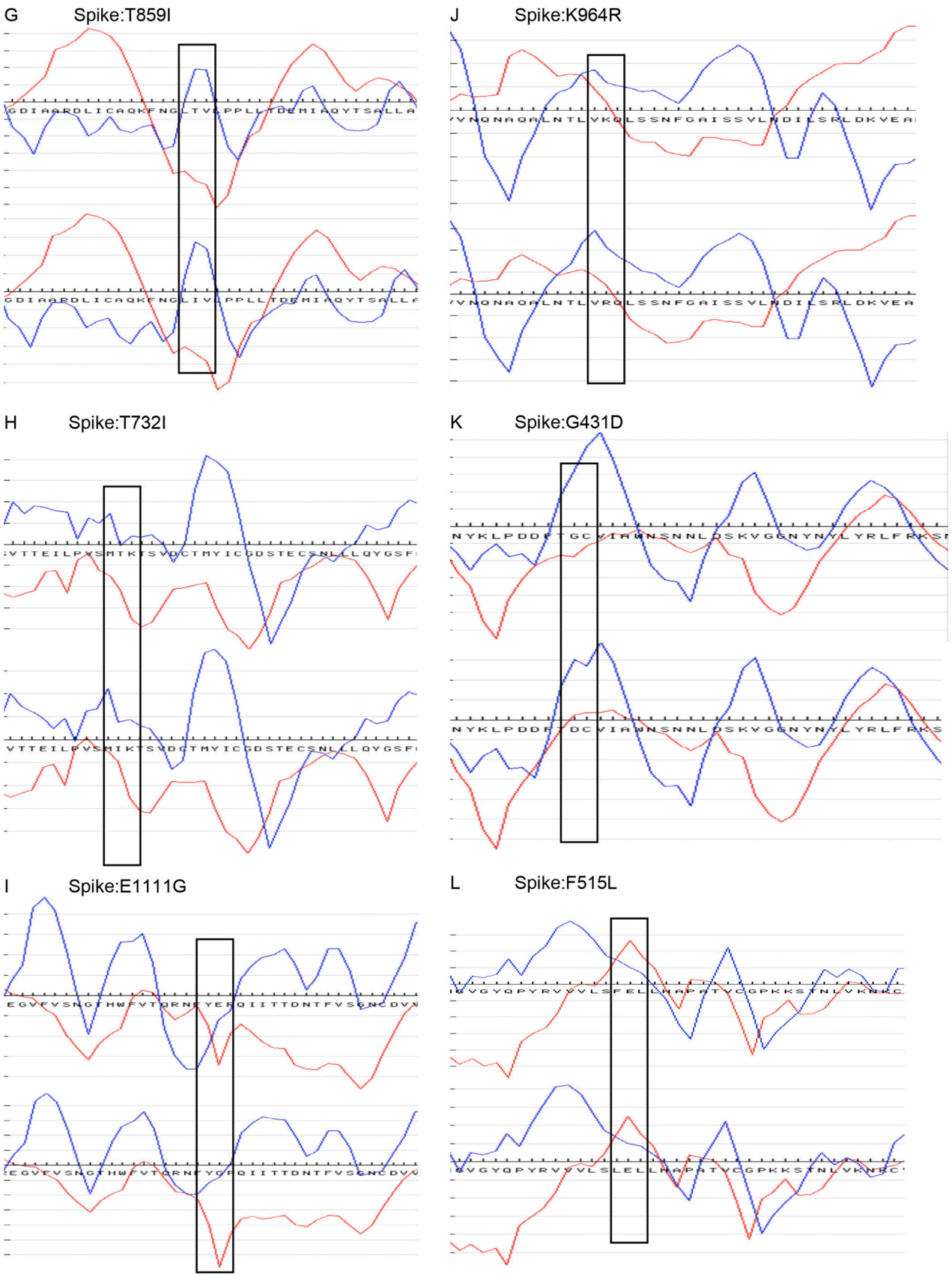

Fig. 1. (*continued*).

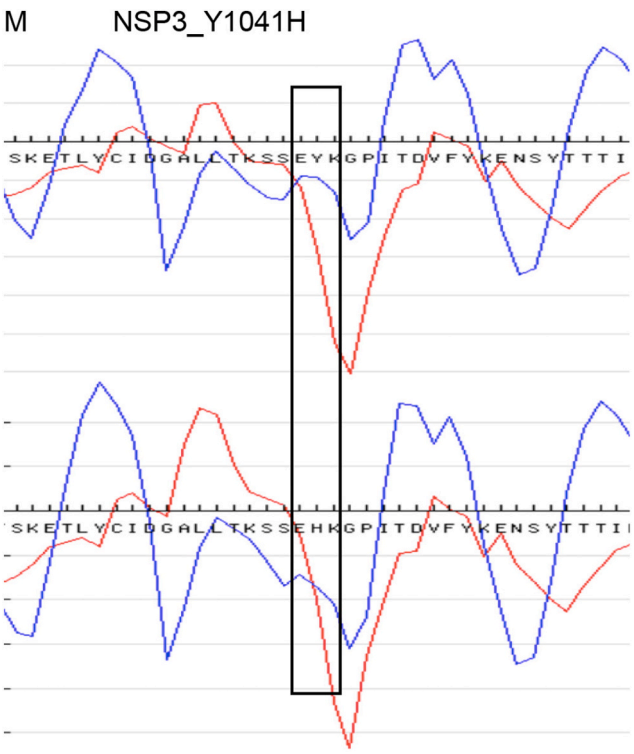

Fig. 1. (*continued*).

Recently, the N501Y variant emerged independently in lineage B.1.1.7 and is located in RBD, which enhances viral fitness for viral replication in human airway cells, resulting in enhanced transmission (Liu et al., 2021). A recent MD study investigated how this variant can cause alteration in conformation for residues located at RBD:ACE-2 interface and revealed that contact areas of RBD and ACE-2 were almost constant. The experimental screening revealed that the value of differences of free energy (ΔΔG) is -0.81 kcal·mol^{-1}, suggesting that this variation can provide more stable binding with hACE-2 by increasing binding affinity (Luan et al., 2021).

This study aims to identify novel non-synonymous variations in Turkish isolates and their impacts on protein structure and dynamics of SARS-CoV-2. Furthermore, to better understand the transmission pattern of SARS-CoV-2 in Turkey, a phylogenetic analysis was performed.

2. Material & method

This study was approved by Baskent University Institutional Review Board and Ethics Committee (Project no: KA21/81).

2.1. Dataset construction

From the beginning of the pandemic to the 5th of February, 2021, 411 sequences from Turkish isolates were uploaded to GISAID (https://www.gisaid.org/). In order to compare these 411 sequences with the Wuhan reference sequence (EPI_ISL_402124), they were aligned by the MUSCLE algorithm on MEGAX software (Kumar et al., 2018) to find point mutations.

2.2. Secondary structure prediction

411 isolates fasta sequences were analysed. Nucleotide sequences were translated into amino acid sequences by the ExPaSy tool (https://www.expasy.org/). Amino acid sequences of each of the Turkish isolates were compared with the Wuhan reference sequence to identify impacts of these variations on secondary protein structure which were predicted by the GOR method (http://cib.cf.ocha.ac.jp/bitool/GOR/).

2.3. Computational analysis

To assess the effects of identified Spike mutations on protein stability and flexibility, and conformation, SARS-CoV-2 open configuration structure (PDB:6VYB) was obtained from RCSB Protein Data Bank (PDB) (https://www.rcsb.org/; Walls et al., 2020). The 6VYB structure was uploaded to the Cutoff Scanning Matrix (mCSM) server, which calculates the effects of mutations according to atomic distance patterns (Pires et al., 2014). DynaMut web server for calculating vibrational entropy differences ($\Delta\Delta S^{Vib}$), was used (Rodrigues et al., 2018). MutaBind2 webserver (Zhang et al., 2020) was used to evaluate the effects of mutations on spike in complex with human ACE-2 receptor with PDB ID:6M0J. The methodology and results were validated with the known highly transmissible Delta variant from a patient from Turkey (EPI_ISL_2232671). Identified mutations were generated using mutagenesis tool in PyMOL version 2.5.1 (https://pymol.org/2/) by using PDB: 6VYB and 6M0J as a template. Visualization of wild type and the mutants were performed on VMD (Humphrey et al., 1996)

2.4. Molecular dynamics (MD) simulations

To further investigate how these variants located at RBD:ACE-2 may cause alteration in conformation and binding stability, 50 ns MD simulations were performed by using GROMACS 5.20.1 software package and CHARMM27 force field was employed (Abraham et al., 2015). To reconstruct missing atoms at the side chain SwissPDB viewer 4.1.0 was used (Guex et al., 2009). The systems were solvated by using SPC/E water model and centered in a cubic box of 1.0 nm^3 distance in each site. The systems were neutralized with NaCl; after that, each of the systems was subjected to energy minimization using steepest descent algorithm until the maximum force of 1000 kJ mol^{-1} nm^{-1} was achieved. The minimized systems were well-equilibrated for 100 ps at 310 K and 1 bar pressure in NVT and NPT. The electrostatic term was described through

Table 1
Co-existing mutations of Turkish Isolates.

Co-existing Mutations	N (%)
NSP12_P323L, Spike_D614G, N_G204R, N_R203K	107 (26.0)
NSP12_P323L, Spike_D614G	231 (56.2)
NSP9_L42F, NSP12_P323L, Spike_D614G, N_G204R, N_R203K	64 (15.6)
NSP2_V198I, NSP2_R27C, NSP3_P1228L, NSP4_M33I, NSP6_L37F	12 (2.9)
NSP12_P323L, NSP13_S485L, NS3_Q57H	65 (15.8)
NSP12_P323L, NSP13_S485L, Spike_D614G, NS3_Q57H	60 (14.6)
(NSP3_T749A, NSP3_T183I, NSP3_A890D, NSP3_I1412T, NSP6_F108del, NSP6_G107del, NSP6_S106del, NSP8_Q24R, NSP12_P323L, NSP13_K460R, Spike_A570D, Spike_S982A, Spike_P681H, Spike_D614G, Spike_T716I, Spike_Y144del, Spike_N501Y	8 (1.9)
NSP3_T183I, NSP3_A890D, NSP3_I1412T, NSP6_F108del, NSP6_G107del, NSP6_S106del, NSP12_P323L, NSP13_K460R, Spike_A570D, Spike_S982A, Spike_P681H, Spike_D614G, Spike_T716I, Spike_Y144del, Spike_N501Y, Spike_D1118H, Spike_V70del	10 (2.4)

Table 2
Calculations of ΔΔG and $\Delta\Delta S^{Vib}$ between references genome spike protein and mutants isolated from Turkish isolates.

Accession ID	Location	Mutation	Predicted Stability Change (ΔΔG):	$\Delta\Delta S^{Vib}$
EPI_ISL_735349	Spike	T859I	-0.230 kcal/mol	0.191 kcal. $mol^{-1}.K^{-1}$
EPI_ISL_735339	Spike	T732I	-0.118 kcal/mol	-0.492 kcal. $mol^{-1}.K^{-1}$
EPI_ISL_730569	Spike	E1111G	-1.047 kcal/mol	0.869 kcal. $mol^{-1}.K^{-1}$
EPI_ISL_730573	Spike	K964R	-0.553 kcal/mol	-1.154 kcal. mol-1.K-1
EPI_ISL_730576	Spike	G431D	-1.008 kcal/mol	-0.975 kcal. mol-1.K-1
EPI_ISL_735312	Spike	F515L	-0.399 kcal/mol	0.152 kcal. mol-1.K-1
EPI_ISL_2232671	Spike	L452R	-1.176 kcal/mol	0.250 kcal. mol-1.K-1

the particle mesh Ewald method (Darden et al., 1993) using a 10.0 Å cut-off, The SHAKE algorithm (van Gunsteren and Berendsen, 1977) was used to constrain bond lengths at their equilibrium values, and a time step was set to 2.0 fs. The MD trajectory analysis was performed to evaluate root mean square deviation (RMSD), root mean square fluctuations (RMSF), radius of gyrations (Rg) and solvent accessible surface area (SASA) based on the resulting trajectories by GROMACS tools.

2.5. Phylogenetic analysis

Retrieved SARS-CoV-2 whole-genome sequences from Turkey were aligned by the MUSCLE algorithm, and the reference sequence from Wuhan was used as an outgroup for the phylogenetic tree. For the statistical method, Maximum Likelihood was used via MEGAX (Kumar et al., 2018). Bootstrap 1000 replication was used as a test of the phylogeny. Poisson was used as a substitution model.

3. Results

3.1. Dataset construction

In total, 411 sequences from Turkey from the beginning of the pandemic to the 5th of February 2021 were retrieved from GISAID. These sequences were aligned with Wuhan reference sequence and 13 novel point mutations were identified. Three of these mutations were found in non-structural protein (NSP) 2, one in NSP8, and three in NSP3. Six of these variations were located in the spike protein of SARS-CoV-2 (Fig. 1). Mutation analysis revealed that 91.4% of Turkish isolates harbour D614G mutation in spike protein, followed by P323L mutation in NSP12 at a frequency of 89.8%. R203K and G204R mutation in nucleocapsid protein were detected in 57.2% and 56.7% of uploaded sequences, respectively. Besides, 10.5% of sequences harboured N501Y mutation in spike protein. The frequency of co-existing mutations was demonstrated in table 1.

3.2. Secondary structure prediction

In order to determine impacts of these identified mutations on protein secondary structure, GOR method was applied. Novel variations NSP3_T133I, Spike_T859I and Spike_T732I, causes an amino acid change from threonine (T) to isoleucine (I) at position 133, 859 and 732, respectively (Fig. 1b, 1g, 1h). This change lead to the loss of β-sheet formation at position 133 and favored coil formation in the mutant portion. While, at position 732 and neighboring two amino acids, coil formation was replaced with β-sheet formation. At position 859, no change was detected in secondary structure, yet turn formation altered to coil formation detected in the neighboring amino acid. A novel spike mutation E1111G causes the amino acid change from glutamic acid (E) to glycine (G) at position 1111. Change from glutamic acid to glycine caused loss of α-helix structures and favored β-sheet formation (fig. 1i). Another spike variation, G431D, causes the change from glycine to aspartic acid (D), leading to a change in secondary structure by favoring the α-helixes and loss of β-sheet formation at the mutation site (fig. 1k). In K964R and F515L mutations cause the amino acid change from lysine (K) to arginine (R) and from phenylalanine (F) to leucine (L), respectively; there was no significant alteration detected (fig. 1j, 1l). Three amino acid substitutions were identified in NSP2; A26T, G221S and V193I. While A26T causes loss of α-helixes in neighboring amino acid, there was no significant alteration in the secondary structure in the presence of A26T, G221S and V193I mutations (fig. 1d, 1e, 1a). A novel variation in NSP3, A602V, causes the amino acid change from alanine to valine, leading to a change from α-helix to β-sheet and affects the secondary structure of NSP3 of the virus (fig. 1c). No significant alteration was detected in the presence change from tyrosine (Y) to histidine (H) at position 1041 in NSP3 (NSP3_Y1041H) (fig. 1m) The last novel mutation in NSP8 is R111S; changes from arginine to serine at position 111. Alterations of these amino acids did not reveal a change at the mutation site. However, alteration in secondary structure detected in neighboring amino acids (fig. 1f).

3.3. Computational analysis

Six novel spike protein variations were identified in Turkish populations, and they were further analysed to reveal their impacts on protein dynamics. ΔΔG between the reference sequence and six Spike mutants were found below zero for each of six isolates, demonstrating these variations stabilizing effects on protein structure. Similarly, ΔΔG for the Delta variant spike mutation (L452R) was found to be -1.176 kcal/mol, supporting its stabilizing effect on protein structure. Further, differences in vibrational entropy ($\Delta\Delta S^{Vib}$) were calculated to detect whether these variations have a rigidification or flexibility effect. Three variants were found to have rigidification, while the other three have a flexibility effect on protein structure (table 2). $\Delta\Delta S^{Vib}$ for the L452R mutation showed an increment in the molecule flexibility.

Among these six spike glycoprotein mutations, two of them were located in the domain which forms a complex with human ACE-2, G431D and F515L. Binding affinity ($\Delta\Delta G^{binding}$) > 0 signifies a decrement in binding affinity, while $\Delta\Delta G^{binding} < 0$ corresponds to an increment in binding affinity and thereby greater susceptibility to SARS-CoV-2 infection. $\Delta\Delta G^{binding}$ for G431D mutation was found to be 1.41, while for F515L, it was 0.04. Both of these mutations decrease the binding affinity of spike glycoprotein and human ACE-2. These methods

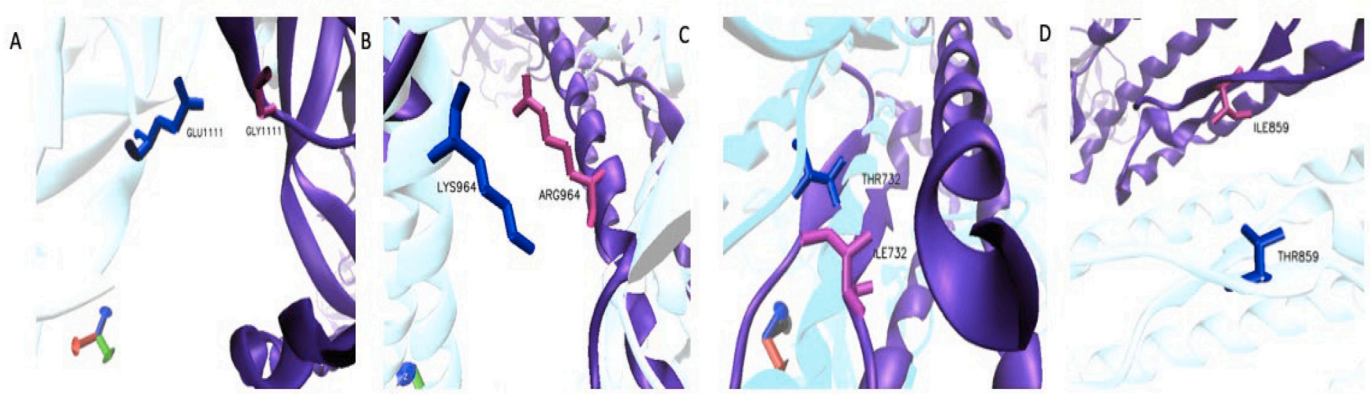

Fig. 2. 3D structure and trajectory from MD simulation of spike crystal structure 6vyb and mutants. 3D Structural conformation of (a) E1111G, (b) K964R, (c) T732I and (d) T859I (cyan) and the amino acid of G1111, R946, I732 and I859 are shown in magenta while the wild-type residues are indicated in blue.

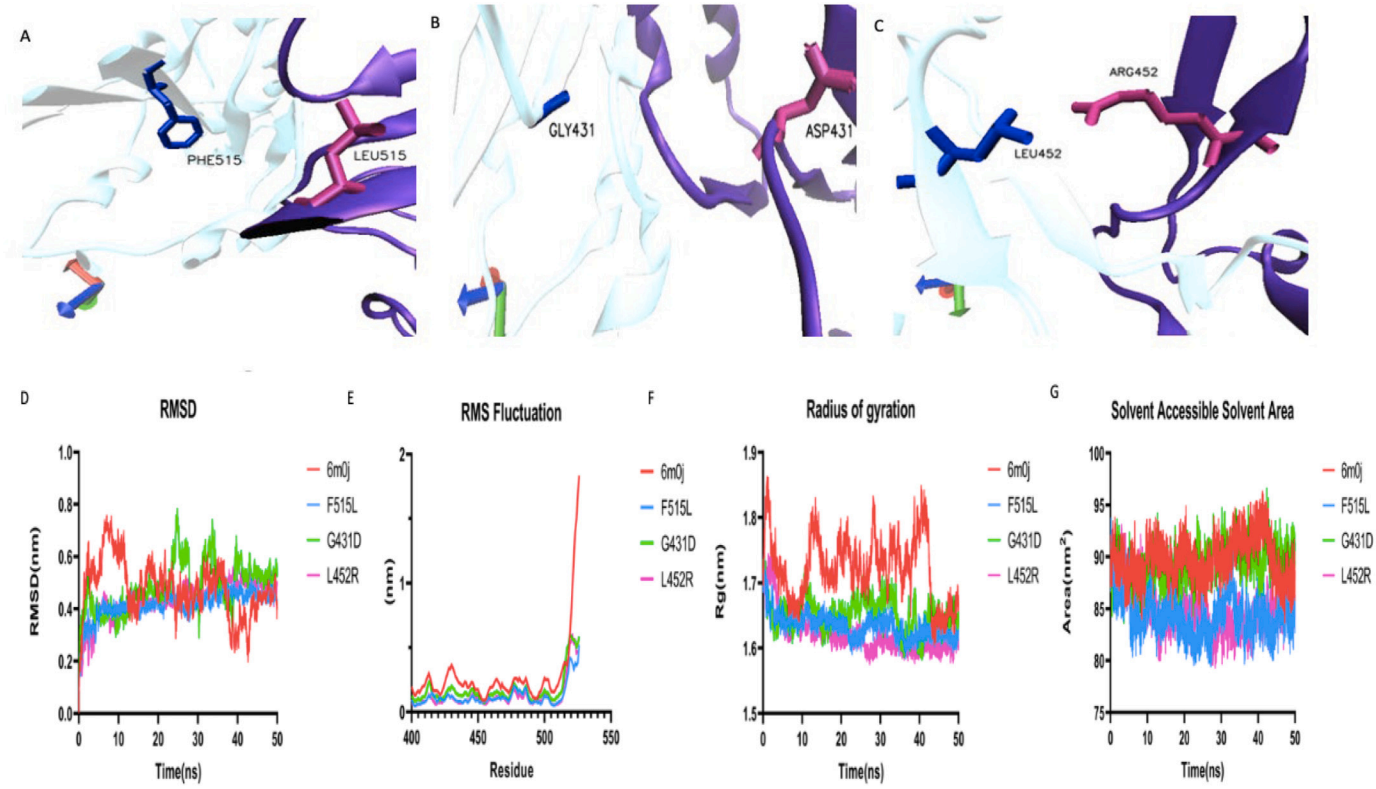

Fig. 3. 3D structure and trajectory from MD simulation of spike crystal structure 6m0j and mutants. 3D Structural conformation of (a) F515L, (b) G431D and (c) L452R (cyan) and the amino acid of L515, D431 and 452R are shown in magenta while the WT residues are indicated in blue. The structural analysis was carried out by (d) RMSD (e) RMSF and (f) the radius of gyration and (g) SASA.

were validated with highly transmissible Delta variant spike mutation L452R, located in the domain in complex with human ACE-2, and $\Delta\Delta G^{binding}$ for L452R found to -0.79, demonstrating greater susceptibility to SARS-CoV-2 infection.

3.4. Molecular dynamics (MD) simulations

The 3D structures of mutant spike proteins were obtained by point mutation of PDB: 6VYB using PyMol software and mutants vs. wild-type were visualized by using VDM (fig. 2). The PBD:6M0J and variants located in RBD:ACE-2 complex were submitted to MD simulation for 50 ns (fig 3). The first and last frames of MD simulation are available in the supplementary file. Average RMSD values of backbone atoms are 4.78 Å, 4.82 Å, 4.17 Å and 4.35 Å respectively for wild-type (6M0J), G431D, F515 and L452R mutants. RMSD of all systems underwent an increase during the first 15 ns and then converged to reach equilibrium after 45 ns of MD simulation which demonstrates a good convergence for each system (fig 3d). In RMSF analysis five main fluctuation peaks were observed in the wild-type. Even though the fluctuations were shared between the wild-type and the mutants, in the wild-type a higher peak between the 500 and 525th residue was observed as compared to mutants (fig 3e). Despite this change in the fluctuation between the wild-type and the mutants, the flexibility pattern in other residues is similar until the 500th residue. R_g average for wild-type was 17.23 Å while for the mutants it varied between 16.22 Å to 16.53 Å suggesting that the mutants have a similar degree of compactness even though slightly different from that of the wild-type (fig 3f). As it can be observed in the SASA graph, the solvent accessibility area of the G431D variant is

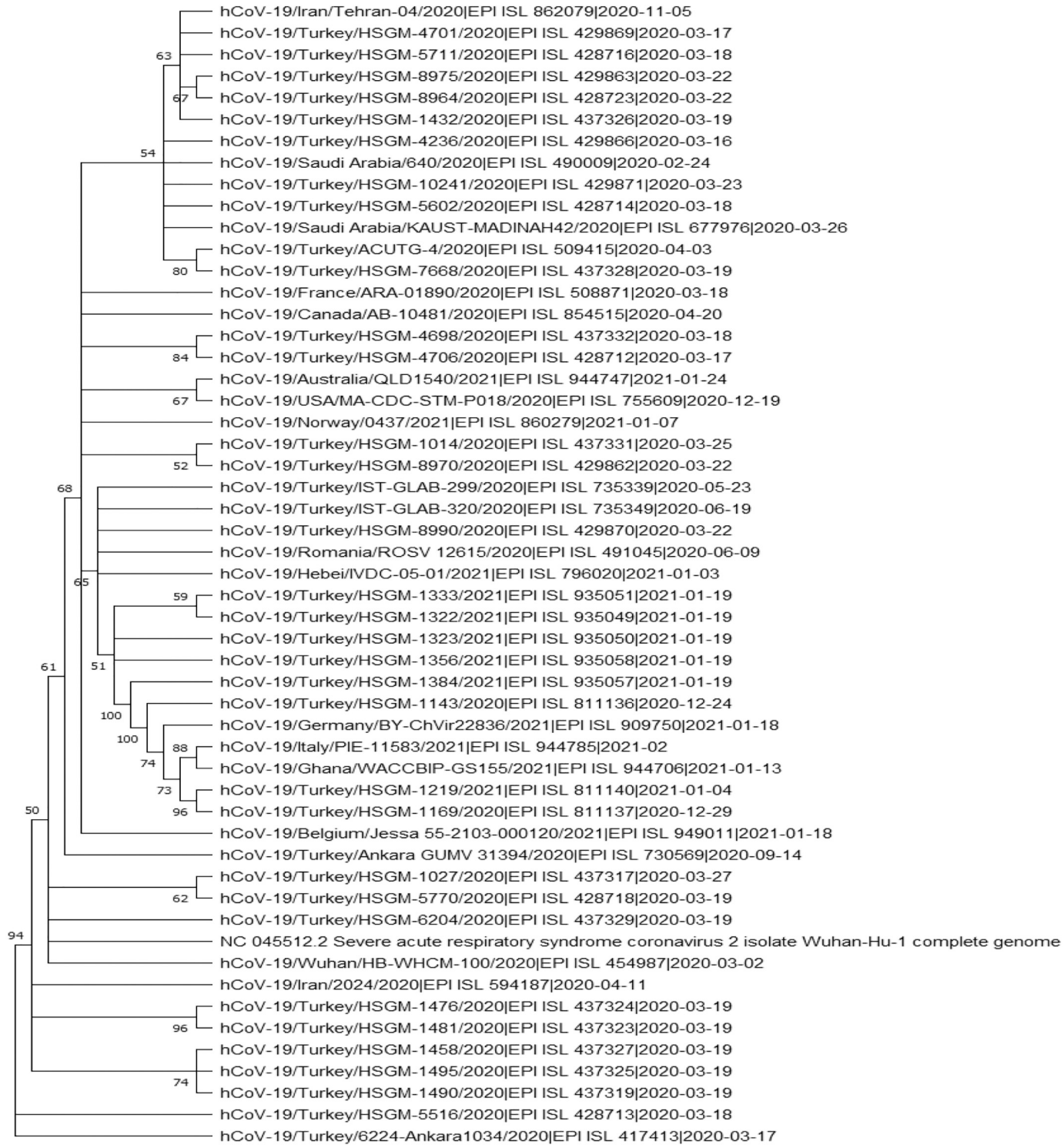

Fig. 4. Phylogenetic analysis of Turkish Isolates.

similar to that of the WT. On the other hand, the accessibility area of two other variants (F515L, L452R) is lower than the WT, indicating an increase in the hydrophobic interactions among non-polar residues. From these results, it could be inferred that F515L and L452R mutations have a stabilizing effect on the spike protein as they induce higher interactions in the hydrophobic core (fig 3g).

3.5. Phylogenetic analysis

Turkish patients with an ID number EPI_ISL_429869, EPI_ISL_428716, EPI_ISL_428723, EPI_ISL_429866, EPI_ISL_429871, EPI_ISL_428714, EPI_ISL_437328 had travel history from Saudi Arabia, patient with an ID number EPI_ISL_437326 had travel history from Iran and patients with an ID number EPI_ISL_509415, EPI_ISL_429863 had no travel history, are clustered together with the patient from Iran and Saudi Arabia. Two Turkish isolates (EPI_ISL_811140, EPI_ISL_811137)

with N501Y variation belonging to the B.1.1.7 lineage which mainly identified in the UK (Cella et al., 2021) were found to be clustered together with sequences from Germany, Italy and Ghana, which are then further grouped with the patient from Belgium and Turkey (fig 4).

4. Discussion

It is known that mutations have significant roles in evolution since they can introduce diversity into the genome. These mutations may provide an advantage to SARS-CoV-2 by altering the protein structure and interfering with its relations to other proteins (Pires et al., 2014). Since SARS-CoV-2's function is determined by its protein structure, predicting the effects of mutation on protein structure may provide pivotal information about SARS-CoV-2 virulence, infectivity and transmission (Nguyen et al., 2020).

Through the mutation analysis, 13 novel mutations were identified, and many of them caused alterations in protein structure. Among them, three non-synonymous mutations were detected to cause an amino acid change from polar amino acid threonine to a non-polar and hydrophobic amino acid isoleucine (fig. 1b, 1g, 1h). These alterations make this position hydrophobic which may be the reason for alterations in secondary structure. A change from glutamic acid to glycine was identified in a spike protein mutation (fig. 1i). Since glutamic acid is hydrophilic, polar amino acid and glycine is neutral, non-polar amino acid, maybe this change was the reason for the loss of β-sheet structures and favoring of α-helixes formation. Another spike variation, G431D, might have caused a secondary structure change, due to glycine being a neutral, non-polar amino acid, and aspartic acid being hydrophilic, polar amino acid (fig. 1k). Also, the loss of α-helixes was identified in neighboring amino acid in the presence of NSP2 mutation A26T; can be due to the alteration of the protein's hydrophobicity (fig. 1d). Change from hydrophilic amino acid to polar one was identified in NSP8 variation R111S, causing an alteration in the secondary structure of NSP8 in neighboring amino acids (fig. 1f).

As expected, the alteration between similar amino acids in nature was less susceptible to change protein secondary structure. For instance, in the presence of K964R, F515L, V193I, G221S and Y1041H mutations, where amino acids were similar in nature no significant alteration was detected on the secondary structure of the protein (fig. 1j, 1l, 1a, 1e, 1m). However, although both alanine and valine are hydrophobic and non-polar amino acids, alterations to valine led to the loss of β-sheet formation and affected the secondary structure of NSP3 in the presence of A602V mutation (fig. 1c).

Further, ΔΔG were calculated between spike protein of the wild-type and mutants for evaluating the impact of these mutations. Six mutant forms of spike protein analysed demonstrated that each of these mutations has stabilizing effects. Maximum negative ΔΔG was detected for mutation E1111G (table 2). Protein stability is determined by several factors, and differences in vibrational entropy are the major contributor to protein stability. $\Delta\Delta S^{Vib}$ below zero means that the mutation has rigidification effects on protein stability. On the other hand, $\Delta\Delta S^{Vib}$ above zero signifies flexibility effects on protein stability (Azad, 2021). The maximum positive $\Delta\Delta S^{Vib}$ was obtained for E1111G, while the maximum negative was obtained for K964R mutation.

The computational results were further supported by MD simulations to detect RBD:ACE-2 complex molecular changes associated with stability loss due to the amino acid substitutions. The RMSD of all mutants except G431D had a smaller deviation in RMSD and the lower RMSD average as compared to the wild-type, suggesting an increment in the stability, confirming the computational analysis results obtained from the current study highlighting an increase in protein stability due to these mutations. On the other hand, G431D showed an increment in RMSD value as compared to the wild-type suggesting a lower stability. A previous study has shown that alterations in flexibility are the pivotal contributors to changes in entropy upon binding as they also affect binding interface (Dehury et al., 2020). The mutant G431D exhibited a decrement in fluctuation suggesting a higher rigidity as compared to the wild-type which is also supported by the $\Delta\Delta S^{Vib}$ value.

Actively tracking and characterizing of SARS-CoV-2 lineages and sub-lineages provide significant knowledge for a more precise diagnosis and genetic diversity of the virus. This genomic surveillance helps to better understand the transmission pattern of the virus (Cella et al., 2021). The phylogenetic analysis demonstrated that sequences from Turkey were distributed throughout the tree, revealing multiple entrances to the country. Furthermore, the close relationship between Saudi Arabia and Iran with Turkey was also observed in the patients who travelled to Saudi Arabia or Iran. However, two Turkish isolates belonging to B.1.1.7 lineage were not close to isolates from Saudi Arabia and Iran. In contrast, these two were found to share similarities to isolates from Europe and Africa.

5. Conclusion

In conclusion, to the best of our knowledge, 13 novel mutations were identified in Turkish isolates and reported for the first time. This study revealed these mutations' potential impacts on protein dynamics. Predictions regarding these variations' effects are required to investigate for the further understanding of their exact roles in SARS-CoV-2 transmission, virulence, and infectivity. Since viruses are utilized mutations to adjust to the host, mutation studies provide critical insight for better therapeutic targeting.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

This study was approved by Ministry of Health of Turkey with an application ID: 2021-02-06T23_03_34

CRediT authorship contribution statement

Sezin Unlu: Conceptualization, Methodology, Software, Writing – original draft, Validation. **Aylin Uskudar-Guclu:** Data curation, Supervision, Validation, Writing – review & editing. **Isli Cela:** Visualization, Writing – original draft, Investigation, Software.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to acknowledge all the Turkish and worldwide hospitals and laboratories that sequenced and submitted the viral genomes on GISAID.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humgen.2022.201040.

References

Abraham, M.J., Murtola, T., Schulz, R., Pall, S., Smith, J.C., Hess, B., 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 1-2, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

Azad, G.K., 2021. Identification and molecular characterization of mutations in nucleocapsid phosphoprotein of SARS-CoV-2. PeerJ. 9 https://doi.org/10.7717/peerj.10666.

Bank RCSBPD. Homepage. RCSB PDB. https://www.rcsb.org/. Accessed May 18, 2022.

Begum, F., Mukherjee, D., Thagriki, D., Das, S., Tripathi, P.P., Banerjee, A.K., Ray, U., 2021. Analyses of spike protein from first deposited sequences of SARS-CoV2 from West Bengal, India. F1000Research 9, 1–13. https://doi.org/10.12688/f1000research.23805.1.

Cella, E., Benedetti, F., Fabris, S., Borsetti, A., Pezzuto, A., Ciotti, M., Pascarella, S., Ceccarelli, G., Zella, D., Ciccozzi, M., Giovanetti, M., 2021. SARS-CoV-2 Lineages and Sub-Lineages Circulating Worldwide: A Dynamic Overview. Chemotherapy 66 (1-2), 3–7. https://doi.org/10.1159/000515340.

Darden, T., York, D., Pedersen, L., 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J. Chem. Phys. 98 (12), 10089–10092. https://doi.org/10.1063/1.464397.

Deganutti, G., Prischi, F., Reynolds, C.A., 2020. Supervised molecular dynamics for exploring the druggability of the SARS-CoV-2 spike protein. J. Comput. Aided. Mol. Des. 35, 195–207. https://doi.org/10.1007/s10822-020-00356-4.

Dehury, B., Raina, V., Misra, N., Suar, M., 2020. Effect of mutation on structure, function and dynamics of receptor binding domain of human sars-cov-2 with host cell receptor ace2: A molecular dynamics simulations study. J. Biomol. Struct. Dyn. 1–15 https://doi.org/10.1080/07391102.2020.1802348.

Guex, N., Peitsch, M.C., Schwede, T., 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis. 30 (S1) https://doi.org/10.1002/elps.200900140. S162–S73.

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: Visual molecular dynamics. J. Mol. Graph. 14 (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

Initiative. GISAID. https://www.gisaid.org/.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. Mega x: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35 (6), 1547–1549. https://doi.org/10.1093/molbev/msy096.

Kumar, S., Nyodu, R., Maurya, V.K., Saxena, S.K., 2020. Morphology, Genome Organization, Replication, and Pathogenesis of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), 2, pp. 23–31. https://doi.org/10.1007/978-981-15-4814-7_3.

Li, J., Guo, M., Tian, X., 2021. Virus-Host Interactome and Proteomic Survey Reveal Potential Virulence Factors Influencing SARS-CoV-2 Pathogenesis. Med. 2, 99–112. e7. https://doi.org/10.1016/j.medj.2020.07.002.

Liu, Y., Liu, J., Plante, K.S., Plante, J.A., Xie, X., Zhang, X., 2021. The N501Y spike substitution enhances SARS-CoV-2 transmission. bioRxiv. https://doi.org/10.1101/2021.03.08.434499, 9:2021.03.08.434499.

Luan, B., Wang, H., Huynh, T., 2021. Enhanced binding of The N501Y-mutated SARS-COV-2 spike protein to the Human ACE2 Receptor: Insights from molecular dynamics simulations. FEBS. Lett. 595, 1454–1461. https://doi.org/10.1002/1873-3468.14076.

Mahtarin, R., Islam, S., Islam, M.J., 2020. Structure and dynamics of membrane protein in SARS-CoV-2. J. Biomol. Struct. Dyn. https://doi.org/10.1080/07391102.2020.1861983.

Nguyen, T.T., Pathirana, P.N., Nguyen, T., 2020. Genomic mutations and changes in protein secondary structure and solvent accessibility of SARS-CoV-2 (COVID-19 virus). bioRxiv 1–16. https://doi.org/10.1101/2020.07.10.171769.

Ouassou, H., Kharchoufa, L., Bouhrim, M., 2020. The Pathogenesis of Coronavirus Disease 2019 (COVID-19): Evaluation and Prevention. J. Immunol. Res. 2020 https://doi.org/10.1155/2020/1357983.

Pires, D.E.V., Ascher, D.B., Blundell, T.L., 2014. MCSM: Predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics. 30, 335–342. https://doi.org/10.1093/bioinformatics/btt691.

Prediction of the secondary structure by gor, 2022. Retrieved June 15, 2021, from. http://cib.cf.ocha.ac.jp/bitool/GOR/.

Rodrigues, C.H.M., Pires, D.E.V., Ascher, D.B., 2018. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic. Acids. Res. 46, W350–W355. https://doi.org/10.1093/nar/gky300.

SIB Swiss Institute of Bioinformatics. Expasy. https://www.expasy.org/ (Accessed March 14, 2021).

The PyMOL Molecular Graphics System. Version 1.5.0.4 Schrödinger, LLC. https://pymol.org/2/.

Unlu, S., Uskudar-Guclu, A., Basustaoglu, A., 2021. Predominant Mutations of SARS-CoV-2: Geographical Distribution and Their Potential. Mediterr. J. Infect. Microb. Antimicrob. 10, 15. https://doi.org/10.4274/mjima.galenos.2021.2021.15.

van Gunsteren, W.F., Berendsen, H.J.C., 1977. Algorithms for macromolecular dynamics and constraint dynamics. Mol. Phys. 34, 1311–1327.

Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., Veesler, D., 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell. 181 (2), 281–292.e6. https://doi.org/10.1016/j.cell.2020.02.058.

Wu, S., Tian, C., Liu, P., Guo, D., Zheng, W., Huang, X., Zhang, Y., Liu, L., 2020. Effects of SARS-CoV-2 mutations on protein structures and intraviral protein-protein interactions. J. Med. Virol. 93 (4), 2132–2140. https://doi.org/10.1002/jmv.26597.

Yang, J., Petitjean, S.J.L., Koehler, M., 2020. SARS-CoV-2 binding to the ACE2 receptor. Nat. Commun. 11, 4541. https://doi.org/10.1038/s41467-020-18319-6.

Zhang, N., Chen, Y., Lu, H., Zhao, F., Alvarez, R.V., Goncearenco, A., Panchenko, A.R., Li, M., 2020. MutaBind2: Predicting the Impacts of Single and Multiple Mutations on Protein-Protein Interactions. IScience. 23 (3), 100939 https://doi.org/10.1016/j.isci.2020.100939.