Reports of Biochemistry & Molecular Biology
Vol.13, No.1, Apr 2024

Original article

www.RBMB.net

Investigation of the Mutations in the SARS-CoV-2 Envelope Protein and Its Interaction with the PALS1 by Molecular Docking

Maryam Ajel[1,2], Seyed Mohammad Jazayeri[3], Emad Behboudi[4], Mansour Poorebrahim[5], Mahin Ahangar Oskouee[1,2], Hossein Bannazadeh Baghi[1,2], Alka Hasani[2], Mojtaba Varshochi[1], Ali Akbar Shekarchi[6], Mohammad Sabbaghian[2], Vahdat Poortahmasebi*[1,2]

Abstract

Background: The envelope (E) protein of globally circulating severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is highly conserved. This study aimed to find the mutation rate of the E genes in COVID-19 patients, and also to evaluate the conformational characteristics of viral E protein.
Methods: In this study, 120 patients with SARS-CoV-2 positive test results were selected according to real-time PCR assay. Specific primers for conventional PCR have been used to amplify E gene; furthermore, to identify the E gene mutations, direct sequencing of the E genes was also done. Bioinformatics techniques were used to investigate the possible effects of antigenic changes and 3D characteristics of amino acid substitutions. Also, the immunogenicity of wild-type and mutant E was analyzed utilizing a ClusPro docking server and the IEDB online platform.
Results: A total of 120 COVID-19 patients were included (57.5% were male and 42.5% female), with an overall mean age of 55.70±10.61 years old. Of 10 nucleotide changes, 8 (80%) were silent. Also, 2 (20%) missense mutations (amino acid altering) were found in the E gene (L73F and S68F).
Conclusion: These mutations insert some new helix structures in the E mutants. Also, the results of molecular docking studies indicated that both S68F and L73F mutations could notably enhance the stability and binding affinity of protein E's C-terminal motif to the Protein Associated with LIN7 1, MAGUK P55 Family Member (PALS1) which may probably increase local viral spread, and infiltration of immune cells into lung alveolar spaces.

Keywords: COVID-19, Envelope, Immunogenicity, Mutation, SARS-CoV-2.

Introduction

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) resulted in universal research in this regard in the same year, which affected the classification of discovered viruses (1). SARS-CoV-2 is the last known member of the coronavirus family to cause a severe pandemic (2, 3). This virus shares similarities with earlier strains in the terms of its homology and pathogenic mechanism, and it

1: Infectious and Tropical Diseases Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.
2: Department of Bacteriology and Virology, School of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran.
3: Research Center of Pediatric Infectious Diseases, Institute of Immunology and Infectious Diseases, Iran University of Medical Sciences, Tehran, Iran.
4: Department of Basic Medical Sciences, Khoy University of Medical Sciences, Khoy, Iran.
5: Targeted Tumor Vaccines Group, Clinical Cooperation Unit Applied Tumor Immunity, German Cancer Research Center (DKFZ), Heidelberg, Germany.
6: Department of Pathology, School of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran.
***Corresponding author: Vahdat Poortahmasebi; Tel: +98 4133364661; E-mail: poortahmasebiv@tbzmed.ac.ir.**
Received: 22 Apr, 2024; Accepted: 25 Aug, 2024

[DOI: 10.61186/rbmb.13.1.124]

is more temperature-resistant than earlier species (4-6).

Its envelope (E) protein is a 9-12 kDa protein with 76-109 amino acids that is found in extremely modest quantities in the virion (7, 8). The viral E protein plays a critical role in virion production, assembly, budding, and envelope creation. During particle maturation, the E protein may aid in the entry of the virion into the lumen of the endoplasmic reticulum and the Golgi membranes (9). Protein E comprises a 30-residue hydrophobic helical transmembrane domain bordered by an 8-residue N-terminal domain and a 37-residue C-terminal domain (CTD). The E protein's C-terminal cytoplasmic region is essential for interactions with a variety of partners, including its PDZ (PSD-95/Dlg/ZO-1)-binding motif (PBM) sequence for interacting with PDZ-containing proteins (10-12). Protein Associated With LIN7 1, MAGUK P55 Family Member (PALS1) regulates cell polarity through binding to the PBM of CRB3 (Crumbs Cell Polarity Complex Component 3) and loss of PALS1 may leads to disruption of cell polarity including loss of integrity of tight junctions (TJ) (13). The investigations show that PALS1 can play a critical role in SARS-CoV-2 virulence (14). The replication and transcription of the virus genome create a large load of the E protein in SARS-CoV-2-infected lung epithelial cells, which can be localized in the endoplasmic-reticulum-Golgi intermediate compartment (ERGIC) for viral assembly and budding. It is suggested that the interactions between E and PALS1 can attract PALS1 to the location of viral assembly, potentially disrupting the polarity complex and vascular structure. As a result, the inter-epithelial connections become loose and leak. The weak junctions may probably increase local viral spread, the fluid flow, and several types of immune cells into lung alveolar spaces (15).

The envelope is a highly conserved protein SARS-CoV-2 (mutation rate <2%) (16, 17). The existence of various mutations in the E protein may lead to conformational change, virus stabilization, and evolutions, resulting in vaccine and therapeutic efficiency (18). Furthermore, the virus may be more severe in the future, and consequently, novel antiviral agents and vaccines against SARS-CoV-2 variants might be more practical and helpful (19). Moreover, mutations in the primer and probe binding sites of the E gene may affect the Real-time-PCR (RT-PCR) based SARS-CoV-2 detection resulting in false negative reports (20, 21). This study aimed to investigate SARS-CoV-2 envelope mutations in the Iranian COVID-19 patients and also to predict the 3D conformational changes of the envelope mutants.

Materials and Methods

COVID-19 Patients and Diagnosis

In this cross-sectional study, a total of 120 patients who had SARS-CoV-2 RT-PCR positive result was recruited and was conducted in the Central Laboratory of East Azerbaijan Province, Department of Virology, Tabriz (38°7′ N, 46°20′ E), the capital city of East Azerbaijan province, northwest Iran.

All patients signed an informed consent issued by the Local Ethics Committee. The work was examined and authorized by the Ethical Committee of Tabriz University of Medical Sciences (TBZMED) [ID number: IR.TBZMED.REC.1400.152].

Nasopharyngeal/oropharyngeal specimens were obtained from inpatient and outpatient diagnoses referred to Tabriz medical centers. All enrolled patients in this study were diagnosed with COVID-19 according to the criteria issued by the National Center for Disease Control (CDC) of Iran's Ministry of Health and Medical Education (MOHME) (22). Patients were screened for SARS-CoV-2 infection based on molecular diagnosis and clinical presentation.

One-step RT-PCR was performed for screening of patients for SARS-CoV-2 through RdRp (labelled with FAM) and N (labelled with HEX) targeting genes according to manufacturer instruction (Pishtaz Teb Diagnostics kit, Iran). In addition, the second RT-PCR was performed to confirm the presence of the virus genome by E, labeled

[DOI: 10.61186/rbmb.13.1.124]

with FAM, and S, labeled with ROX, amplification (Covitech kit, Iran). For Pishtaz Teb and Covitech kits, a primer-probe set targeting human RNase P, labelled with ROX and HEX, respectively on a separate channel was used as an internal control. Thermal cycling conditions are included in previous studies (23, 24). Based on the kit instructions, samples with cycle of threshold (CT)-value < 35 were considered positive. In the present study, for better sequencing of samples, all clinical specimens with CT-values ≤ 30 were selected for E gene sequencing. The range of CT values in all COVID-9 patients was between 18-25. Direct sequencing was performed to display the E gene mutational patterns of 30 nasopharyngeal/oropharyngeal swab specimens of patients with Wuhan-1 variants (Classic variants), 30 samples of patients with Alpha variants (B.1.1.7), 30 samples of patients with Delta variants (B.1.617.2), and 30 samples of COVID-19 patients with Omicron variants (B.1.1.529). Data and characteristics of patients are shown in Table 1. Samples of people suspected of having COVID-19 disease were collected using nasopharyngeal and oropharyngeal swabs. In this study, Roche's High Pure Viral Nucleic Acid Kit was used to extract RNA. Then cDNA synthesis was performed using SinaClon First Strand cDNA Synthesis Kit (Sinaclon, Iran).

Table 1. The demographic and clinical properties of the enrolled COVID-19 subjects for envelope gene sequencing.

Variables	Total patients	Classic variant (Wuhan-1)	Alpha variant	Delta variant (B.1.617.2)	Omicron variant (B.1.1.529)
Sex					
Male	69 (57.5)	18 (60)	20 (66.7)	15 (50)	16 (53.3)
Female	51 (42.5)	12 (40)	10 (33.3)	15 (50)	14 (46.7)
Age, year (Mean±SD)	55.70±10.61	55.12±10.71	48.60±8.59	54.20±11.33	45.11±8.12
Disease status					
Inpatient	56 (46.7)	14 (46.7)	13 (43.3)	19 (63.3)	10 (33.3)
Outpatient	64 (53.3)	16 (53.3)	17 (56.7)	11 (36.7)	20 (66.7)
Travel to foreign countries	0	0	0	0	0

Identification of SARS-CoV-2 variants

As the samples were gradually introduced at different time intervals and with the emergence of dominant variants during the SARS-CoV-2 pandemic, the variant detection kit (Delta vSNiP Kit, Delsa Diagnosis Aria Company, Iran) was employed to assess the SARS-CoV-2 variants using the TaqMan RT-PCR method, following the manufacturer's instructions.

Conventional PCR and sequencing

For all samples, RNA was converted to cDNA by the cDNA Reverse Transcription SinaClon First Strand cDNA Synthesis Kit (Sinaclon, Iran). Then, using specific primers (Forward: 5'-CAAATTCACACAATCGACGGTTC-3' and Reverse: 5'-CCATAACAGCCAGAGGAAAATTAAC-3'), the SARS-CoV-2 E gene's full-length sequence (228 bp) was amplified for all the samples. The thermal-cycling profile included Taq activation at 95 °C for 4′, followed by 35 cycles of PCR amplification at the temperatures listed below: denaturation at 94 °C for 30 S, annealing at 57 °C for 30 S, and extension at 72 °C for 60 S, with a final extension at 72 °C for 5'. The ABI-3130 Genetic Analyzer machine (DNA Sequencer, Applied Biosystems 3130, Foster City, CA,

[DOI: 10.61186/rbmb.13.1.124]

USA) was used to sequence the PCR products bilaterally. Electropherograms were analyzed with the Chromas software (version 2.6.6) and checked manually to confirm the base assignment. Subsequently, all envelope sequence alignments were conducted via BioEdit software (version 7.2.6.1). Finally, to evaluate amino acid /nucleotide substitution in the envelope gene, the SARS-CoV-2 isolate Wuhan-Hu-1 (Accession number NC_045512.2) obtained from the GenBank were considered the reference, and the sequences were compared with it (Supplementary Data 1).

Structural Analysis of SARS-CoV-2 Envelope Gene Mutant

Homology modeling

To explore the effect of mutations on the protein structure, we predicted the three-dimensional (3D) structure of proteins using a homology modeling algorithm. To this end, the GalaxyTBM server was used (25) (https://galaxy.seoklab.org/cgi-bin/help.cgi?key=METHOD&type=TBM). Predicted structures were further refined by the Galaxy Refine server, which repacks side chains and performs subsequent overall structure relaxation by molecular dynamics (MD) simulation (26) (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE). Since achieving an ideal quality parameter is challenging, we also utilized two distinct tools for evaluating the quality of the protein structure. The ProQ Protein Quality Predictor and Chimera were both employed to validate the best model produced by each tool. The protein side chains optimization was achieved, followed by energy minimization and an assessment of stability. The results of various verification procedures consistently indicated the outstanding quality of the proposed models. Evaluation of the correct models was based on an LG score exceeding 1.5 and a MaxSub greater than 0.1. Meanwhile, MaxSub generates a single normalized score, reflecting the quality of the model by determining the greatest subset of Cα atoms in a model with an experimental structure (10). All observed mutations were assessed for their impact on the E protein structure using the DynaMut tool (16). This server incorporates both mutation effect and normal modes, providing a comprehensive analysis of variants influencing protein stability and flexibility. To evaluate the antigenicity of the SARS-CoV-2 envelope mutants, all mutant sequences were subjected to the antibody immune epitope database (IEDB) (27) (https://www.iedb.org/).

Molecular docking

The most recent characterized structure of wild-type protein E in complex with the PALS1 (PDB code: 7M4R) was used to investigate the effect of identified amino acid substitutions on the stability of interaction and binding affinity of protein E to its ligand, PALS1. The molecular docking studies were done utilizing the ClusPro server (28). PRODIGY webserver was used to calculate the binding affinity of docked complexes (29).

Statistical Analysis

SPSS version 24 was used for statistical analysis (SPSS Inc, Chicago, Illinois, USA). The findings of the experiments were provided as mean±standard deviation (SD), and absolute number (and percentage).

Results

Demographic and Mutational Characterization

Demographic information and dispersion of individuals in each variant have been mentioned in Table 1 in detail. Of 120 patients, 69 (57.5%) and 51 (42.5%) were male and female, respectively. The mean age of all patients was 55.70±10.61 years old. After sequencing alignment and compared with the reference sequence, mutations occurred in 10 samples. Out of 10 samples, 8 samples had silent (synonymous) mutations and the 2 others had missense (non-synonymous) mutations.

According to the results, at the nucleotide sequences, 6 nucleotide substitutions (C12U, C96U, A189G, C203U, C217U and C225A)

[DOI: 10.61186/rbmb.13.1.124]

have been observed. The total mutations and number of each mutant in each variant are shown in Table 2. C96U, C12U, C225A, and A189G have caused silent mutations. The C12U mutation was found in 4 samples, one of which was found in the classic variant, 2 in the alpha variant, and 1 in each of the Delta and Omicron variants. The C96U mutation, in which the uracil nucleotide replaced the cytosine at position 96, was observed in a total of 2 samples in each of the delta and omicron variants. The A189G mutation is found in the classic variant in 1 sample. The C225A mutation was observed in only one sample and the alpha variant. C203U and C217U mutations cause missense (non-synonymous) mutations and are more important because they can cause changes in the amino acid position in the SARS- CoV- 2 E protein sequence. The C203U caused the S68F mutation and the amino acid serine (S) was substituted with phenylalanine (F) at position 68. Missense mutations are divided into two categories: nonconservative mutation and conservative mutation. The S68F mutation is considered a nonconservative mutation. Because the amino acid phenylalanine, which has non-polar properties, was substituted for the amino acid serine, which has polar properties, it can be concluded that the physicochemical properties of the E protein have also changed. The C217U caused the L73F mutation and the amino acid leucine (L) was substituted with phenylalanine at position 73. The L73F mutation is categorized as conservative mutation because phenylalanine was substituted for the amino acid leucine, both of which are nonpolar.

Table 2. Results from SARS-CoV-2 envelope nucleotide sequencing.

Mutations	Type of mutation	Total number of mutations	Classic (n= 30)	Alpha (n= 30)	Delta (n= 30)	Omicron (n= 30)
C12U	Silent	4	1	2	0	1
C96U	Silent	2	0	0	1	1
A189G	Silent	1	1	0	0	0
C203U	Missense (S68F)	1	0	0	1	0
C217U	Missense (L73F)	1	0	1	0	0
C225A	Silent	1	0	1	0	0

Evolutionary Changes within SARS-CoV-2 Envelope Gene

The E protein encoding fragment contains 228 bases that encode 75 amino acids. After comparing with the reference sequence, it was found that there were 10 nucleotide substitutions in all 120 sequences (silent and missense mutations). Of these, only two mutations cause amino acid changes, and the rest cause silent mutations. Missense mutations in E protein at amino acid positions were 68 and 73, respectively.

In the 68th position, phenylalanine was substituted for serine (S68F) and in the 73rd position, phenylalanine was substituted for the amino acid leucine (L73F). The L73F mutant was identified in April 2021 and the S68F mutant in August 2021, and the samples were alpha and delta variants, respectively. The amino acid position 73 in protein E is located in a motif called DLLV and is essential for interaction with the PALS1 cellular protein.

The frequency of nucleotide mutation (the probability of nucleotide substitution per

[DOI: 10.61186/rbmb.13.1.124]

individual site) was 0.083. The amino acid mutation frequency (the probability of amino acid mutation per individual site) was 0.016. (Table 3). The ratio of missense mutations to silent mutations (*dN/dS*) in COVID-19 patients was 0.25. The *dN/dS* ratio is a very important value in the field of natural selection and evolution, which is obtained by dividing the total missense mutations by the total of silent mutations. If this ratio is greater than 1, it indicates a positive selection pressure, in which the sequences have not maintained their stability and have been selected due to appropriate changes. Also, if this ratio is less than 1, it indicates that the selection is negative (negative selection) and due to the high stability of such mutations, they are selected over time.

Table 3. Details of evolutionary changes within the SARS-CoV-2 E gene studied.

Characteristic variables	Number (%)
Number of SARS-CoV2 samples	120
Nucleotide mutations	10
Nucleotide mutation frequency	0.083
Amino acid mutations	2
Amino acid mutation frequency	0.016
Silent nucleotide mutations	8 (80)
Missense nucleotide mutations	2 (20)
Missense/silent ratio	0.25

Antigenic properties of envelope mutant proteins

The amino acid replacement at sites 68 (S68F) and 73 (L73F) of the envelope protein did not alter the antigenicity characters. The results of the IEDB server showed that amino acid replacement leading to S68F and L73F mutations could not reduce the antigenicity of the E protein compared to the reference strain (Fig. 1 and Table 4).

Table 4. Effects of amino acid substitution (S68F and L73F) on envelope antigenicity.

Wild type		Mutants	
Position	**Score**	**Position**	**Score**
S68	**0.578**	**F68**	**0.587**
L73	**0.405**	**F73**	**0.412**

Identified mutations induce new helix structures

Alterations of E protein stability could be effective in the viral assembly, conformation that eventually affects the pathogenesis of SARS-CoV-2. The resultant structures revealed that some new helix structures are added to the protein structure upon amino acid substitutions. This change might result in enhanced stability. Using DynaMut server, it was revealed that S68F mutation was stabilizing (ΔΔG: 0.360 kcal/mol), but L73F was destabilizing mutation (ΔΔG: −0.415 kcal/mol). Interestingly, although mutations occurred in the C-terminal of the protein sequence, the new helix structures were also found to be induced in the protein's N-terminal. This indicates that the whole intramolecular interaction network is also changed through amino acid substitutions. The highest helical content was found in the L73F mutant. Nonetheless, the original helix structures seemed to be conserved in the mutant proteins (Fig. 2A-C).

The protein E mutants showed higher binding affinity to PALS1

Docking analysis aims to identify potentially beneficial or detrimental effects of E protein mutants with consideration to the wild-type protein. This study proposed a model of interaction between the C-terminal DLLV motif (amino acid: 72-75) of protein E and a hydrophobic pocket from the PDZ and SH3 domains of PALS1 (10). Totally, residues from PDZ including Phe318, Leu321, Leu267, Pro266, and Val284, and residues from SH3 including Leu369 and Leu403 are involved in forming the hydrophobic-binding pocket. The findings of molecular docking studies indicated that both S68F and L73F

[DOI: 10.61186/rbmb.13.1.124]

substitutions could notably enhance the binding affinity of protein E's C-terminal motif to the PALS1. The ΔG of interaction for S68F-PALS1 and L73F-PALS1 complexes were calculated as -9.2 and -8.4 kcal/mol, respectively, which is considerably higher than that obtained for the wild-type peptide-PALS1 complex (-6.6 kcal/mol). In molecular docking, more negative score corresponds to a more stable and stronger binding affinity. Although the interaction mode of the peptide with the PALS1 was not significantly changed in the S68F mutant when compared to the wild-type peptide, the binding pocket of PALS1 was occupied by the L73F mutant in a different form (Fig. 1D-F). Together, the S68F and L73F mutations might induce a change in the protein E structure that favors its physicochemical properties and binding affinity to partner proteins. Therefore, our results patronage the hypothesis that the pathogenesis of SARS-CoV-2 could appertain on the enhanced interaction between the SARS-CoV-2 E protein (especially in L73F mutation) and PALS1 inducing an emphatic alteration of the TJ.

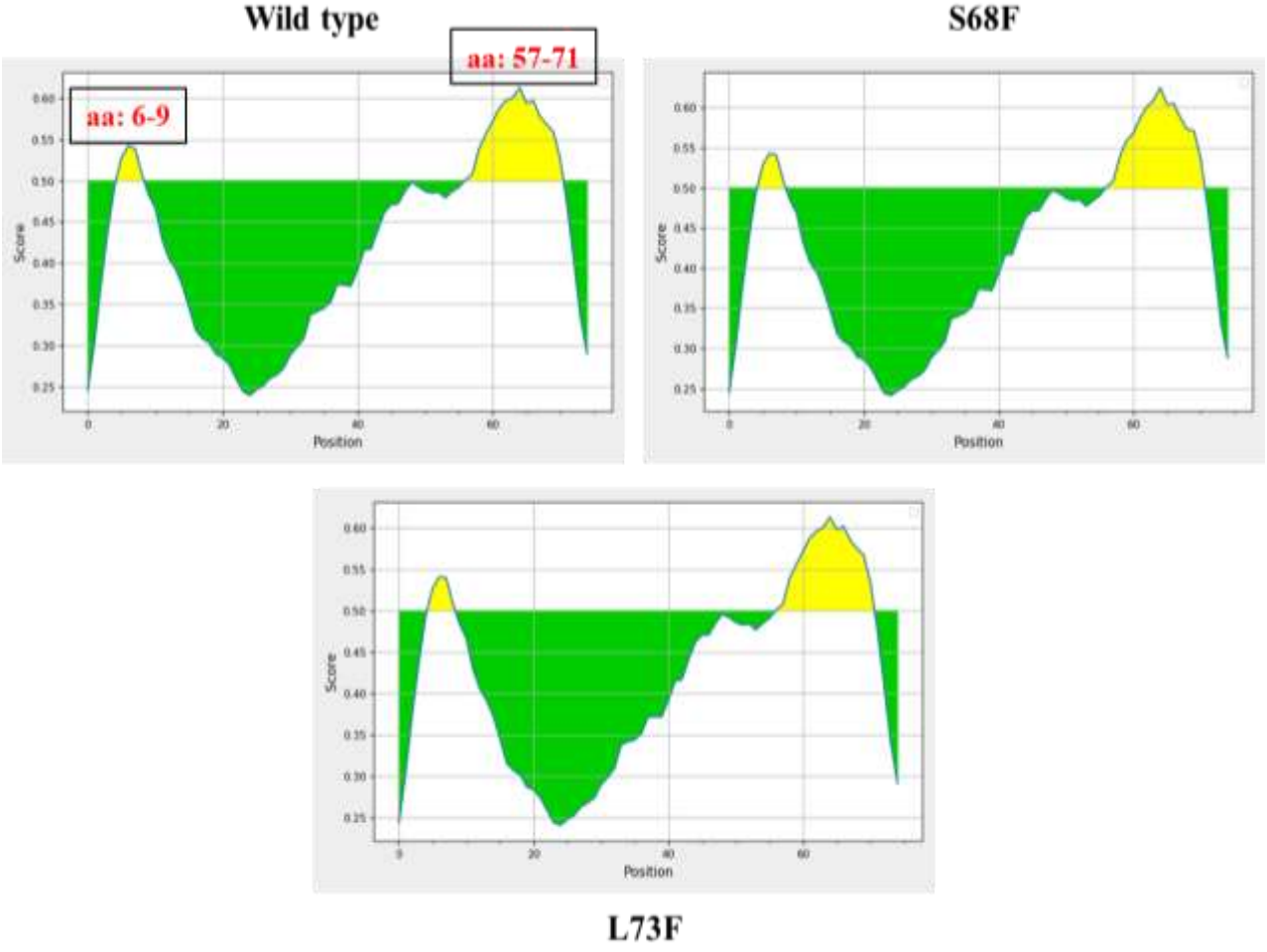

Fig. 1. Effects of amino acid substitution (S68F and L73F) on the antigenicity of SARS-CoV-2 envelope protein. Antigenic and non-antigenic areas are marked in yellow and green, respectively. Mutated regions are also marked with vertical (red) rectangles. The results from the IEDB server show that the antigenic properties of Envelope mutants have not changed compared to the wild type.

[DOI: 10.61186/rbmb.13.1.124]

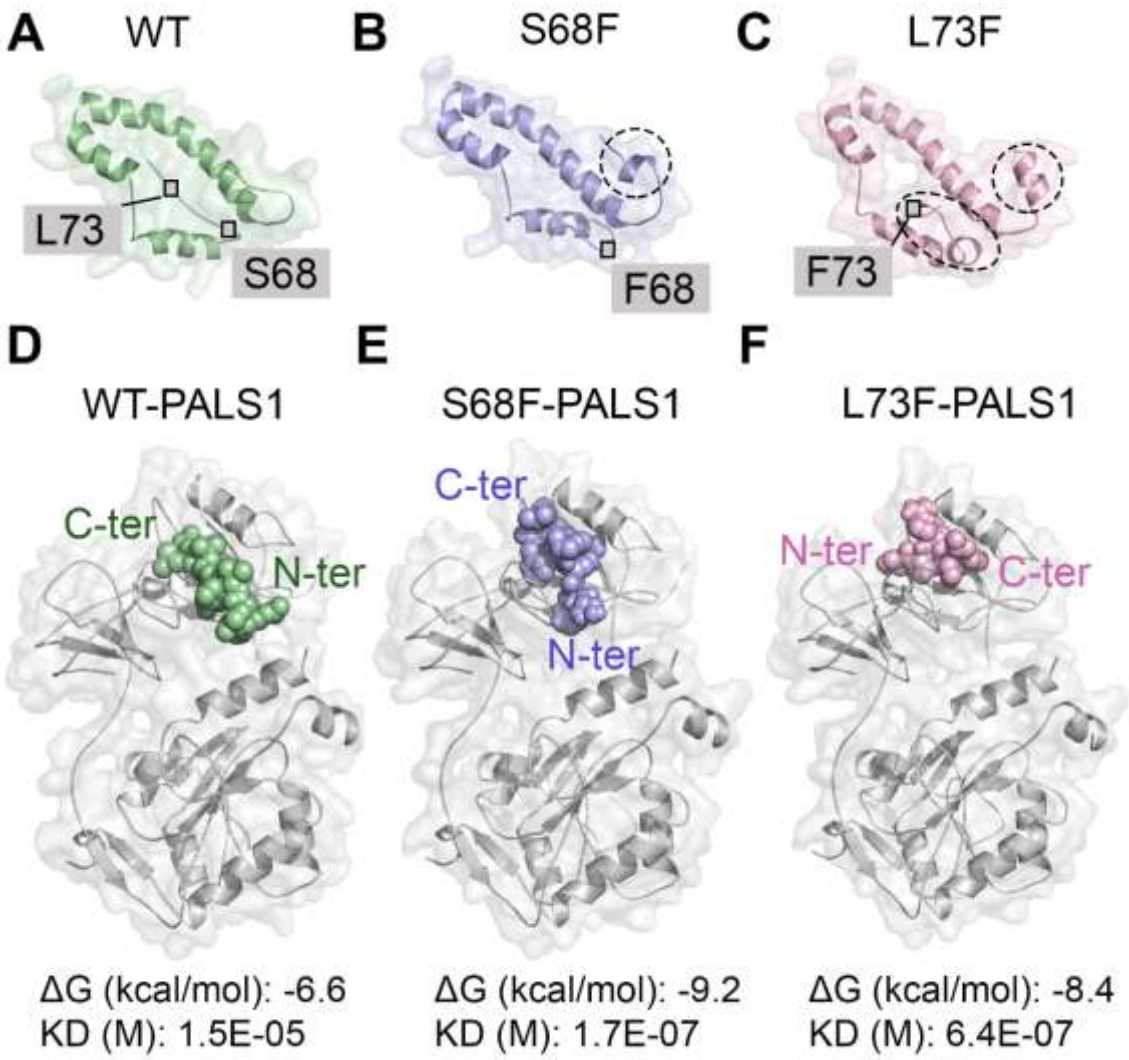

Fig. 2. Structural analysis of E proteins. 3D structures of (**A**) the wild-type, (**B**) S68F, and (**C**) L73F mutants were modeled using template-based homology modeling. Dotted circles show the inserted helix structures in the mutants. The position of substituted amino acids is shown in squares. WT: wild-type. The binding pocket of PALS1 was docked to the C-terminal DLLV motif of protein E obtained from the (**A**) wild-type, (**B**) S68F, and (**C**) L73F proteins. ΔG refers to the binding energy and KD indicates the dissociation constant. WT: wild-type. The interaction information of the complex of WT was obtained from the PDB database (ID: 7M4R).

Discussion

In the present study, we aimed to investigate SARS-CoV-2 envelope mutations in COVID-19 patients. In total, 2 of 120 patients (1.66%) had one mutated amino acid. 110 (91.66%) patients did not have any nucleotide substitutions at all. In total, 10 nucleotide substitutions occurred, of which 8 (80%) were silent (no amino acid change), and 2 (20%) were missense (amino acid altering) (Table 2). As it is known in the present study, the ratio (*dN/dS*) in the coding region E is less than one, which indicates that the E protein sequence has maintained its stability over time, and so-called such mutations have been negatively selected. E proteins are highly conserved with low mutation rates and therefore serve as important screening markers for SARS-CoV-2 infection (30). Comprehensive mutational analysis from the Global Initiative on Sharing All Influenza Data (GISAID) database reports less than 2% E protein mutant strains (17).

A large proportion (98.8%) of the E protein of SARS-CoV-2 strains is highly conserved and evolved slowly, despite the rapid

[DOI: 10.61186/rbmb.13.1.124]

transmission, evolution, and global spread(17, 31). Recently, it has been reported that the S68F, L73F, R69I, and P71L mutations have the highest frequency in the E protein of the SARS-CoV-2 strains and are located in the CTD of the E protein, indicating the continuous evolution of this region (17, 18, 32). On the other hand, the C-terminal DLLV motif of protein E plays a key role in creating interaction with the PALS1 protein (17). The mutations analyzed in this study, S68F, and L73F, are located in this region. It was found that there is an increased affinity for the interaction between the E protein and PALS1, which can lead to increased pathogenicity and virus spread. Thus, this study gives a structural insight into the mutation function related to the interaction efficiency of SARS-CoV-2 E protein with PALS1 protein, and by creating other mutations in the future, these protein-protein interaction (PPI) should be investigated by further studies. An MD simulations study indicated that 3 mutants (S68F, P71S, and L73F) exhibited a stabilizing effect on the envelope protein structure (18). However, our results indicated that the antigenicity scores for the S68 and F68 residues were calculated as 0.578 and 0.587, respectively (Table 4). Also, antigenicity scores for the L73 and F73 residues were calculated as 0.405 and 0.412, respectively. These results revealed a non-significant alteration of antigenicity in the mutant types. The mutational analysis of the previous study indicated that only 1.2% SARS-CoV-2 strains possessed missense mutations. Furthermore, unlike the other structural proteins (M, N and S), the E protein evolved slowly (17). Therefore, highly conserved E protein can be a potential target for SARS-CoV-2 vaccine and treatment approaches.

In the present study, the effect of mutations in the protein function of SARS-CoV-2 E protein over some time was investigated using high-accuracy *in silico* biological tools. A significant issue is the formation of new helix structures resulting from those mutations that can increase the stability of the protein. The emergence and effects of mutations may play a critical role in virus protein stabilization and evolution. Also, according to docking results (Fig. 2), these mutations (S68F and L73F) significantly increased the affinity of protein E interaction with PALS1. The modeling method used in this research was template-based modeling (TBM), which uses experimental structures of homologous proteins as a model. However, in front of the higher accuracy of bioinformatics tools, there are distinctions between in silico protein modeling and experimental capacity. In fact, further experimental structure studies are required to analyze and evaluate protein interactions accurately. The SARS-CoV-2 E protein is highly conserved and has ion channel activity (33). The expression of the E gene is essential for the assembly and release of the virus (34). Recent studies have shown that the rate of mutation in SARS-CoV-2 is 10^{-3} per site per year. These mutations can be challenging to produce vaccines by altering virus antigen regions (35, 36). An essential issue with E protein mutations is that the changes increase its binding specificity to its target proteins, which can affect the inflammatory and pathogenic processes during infection. As a result of multiple mutations in the SARS-CoV-2 genome, different strains have been created that have affected the pathological behavior of the virus and the immune system's response to virus activity (17).

The envelope is a crucial multifunctional protein. It mediates immune responses. The N terminus domain has a critical role in inflammasome activation and the C terminus contains a PDZ binding motif that targets cellular PDZ domains that regulates host immune responses. Therefore, PBM has a role in the pathogenicity of SARS-CoV-2 (10, 37). By targeting cellular junctions and polarity components, viruses can hijack cell machinery. The mutations in the E protein localized near the PBM could have important effects on the structure and ion channel activity of the E protein as well as on the host machinery targeted by the variants during the infection (37). PDZ domains are specific sequence motifs found in various species and

[DOI: 10.61186/rbmb.13.1.124]

are very versatile in their PPIs. PDZ-mediated interactions could affect regulating several critical cellular pathways such as physical scaffolding, signal transduction, and cellular trafficking (38). For this reason, the dysfunction of the PDZ domain often causes a pathological condition in the cell. Accordingly, cellular physiological interactions will be impaired if the interactions mediated by PDZ domains are disrupted (39). Extensive studies have been performed to identify interactions mediated by the PDZ domain. However, since PDZ-containing proteins interact with a wide range of proteins within protein networks, they need further investigation for PDZ-mediated interaction. Therefore, bioinformatics tools can help us to understand the regulatory mechanisms of PDZ PPI and structural alteration (40-42).

PDZ-domain proteins are now known to be important targets for viruses. Some viruses, such as adenovirus, SARS coronavirus, and influenza, have a PBM in their nonstructural and envelope proteins (43). Envelope protein E in SARS-CoV and SARS-CoV-2 has C-terminal PBM and ion channel (IC) activity that can interact with the PALS1 PDZ domain. During this process, PALS1 remains in the Golgi and causes functional disruption of the cellular TJ (44). Accordingly, intracellular signaling transduction and the extracellular messengers between cells are disrupted and lead to damage to lung epithelial tissue (33). With the help of the PDZ-binding motif, the SARS coronavirus E protein can interact with more than 400 host cell proteins and has a high potential for impaired metabolism and cell function (45). Given the importance of protein-regulated processes, new complementary therapeutic strategies can be developed alongside vaccination.

One of the major limitations of the present study was the lack of *in vitro* investigation for the phenotypic pattern of each variant. The proposed model for modified envelope structure was based on the *in silico* analysis and no experimental study was carried out to find out the exact modification of structures by complementary methods. Additional cohort studies should be utilized in the future to evaluate further the effects of these and other mutations on envelope structure and viral replication.

Taken together, regardless of variants, a low prevalence of naturally occurring SARS-CoV-2 envelope mutants were observed in Iranian COVID-19 patients. COVID-19-positive Iranian patients showed a low focused distribution of amino acid replacement with various physicochemical features and hypothetical conformations. The present study will significantly help to increase the knowledge of evolutionary approaches and the pathogenicity of SARS-CoV-2, which might be helpful for better combat of COVID-19 and the progress of vaccines in the future.

Ethical approval and consent to participate

All patients signed an informed consent issued by the Local Ethics Committee. The work was examined and authorized by the Ethical Committee of Tabriz University of Medical Sciences (TBZMED) [ID number: IR.TBZMED.REC.1400.152].

Conflict of interest

None of the authors declare any conflict of interest.

Acknowledgments

The authors sincerely appreciate the staff of the Central Laboratory of East Azerbaijan Province and Research Center of Pediatric Infectious Diseases (RCCV). The present project was supported by the Infectious and Tropical Diseases Research Center, Tabriz University of Medical Sciences, Tabriz, Iran (grant no: 66115). The study also was supported by a grant from the Research Center for Clinical Virology, Tehran University of Medical Sciences (grant no. 48684).

[DOI: 10.61186/rbmb.13.1.124]

References

1. Rostami-Far Z, Rahmani K, Mansouri K, Khadem Erfan MB, Shaveisi-Zadeh F, Nikkhoo B. Genetic Regulation of Interleukin-6 and Interleukin-10 in COVID-19 Infection. Rep Biochem Mol Biol. 2023 Jul;12(2):284-293.
2. Emadi MS, Soltani S, Noori B, Zandi M, Shateri Z, Tabibzadeh A, et al. Highly Conserve Sequences in Envelope, Nucleoprotein and RNA-Dependent RNA Polymerase of SARS-CoV-2 in Nasopharyngeal Samples of the COVID-19 Patients; a Diagnostic Target for Further Studies. J Cell Mol Anesth. 2022;7(2):78-83.
3. Zandi M, Soltani S, Tabibzadeh A, Nasimzadeh S, Behboudi E, Zakeri A, et al. Partial sequence conservation of SARS-CoV-2 NSP-2, NSP-12, and Spike in stool samples from Abadan, Iran. Biotechnol Appl Biochem. 2023;70(1):201-209.
4. Behboudi E, Hamidi V, Gholizadeh F, Grala EM, Ghelmani Y, Nakhaie M, et al. Association between ABO blood groups and rhesus antigen and susceptibility to COVID-19 in the Yazd hospital. New Microbes New Infect. 2021;44:100934.
5. Chan JF, To KK, Tse H, Jin DY, Yuen KY. Interspecies transmission and emergence of novel viruses: lessons from bats and birds. Trends Microbiol. 2013;21(10):544-55.
6. Mostafa-Hedeab G. ACE2 as Drug Target of COVID-19 Virus Treatment, Simplified Updated Review. Rep Biochem Mol Biol. 2020;9(1):97-105.
7. Zandi M, Behboudi E, Soltani S. Role of Glycoprotein Hemagglutinin-Esterase in COVID-19 Pathophysiology? Stem Cell Rev Rep. 2021;17(6):2359-2360.
8. Jafarpour R, Pashangzadeh S, Dowran R. Host factors: Implications in immunopathogenesis of COVID-19. Pathol Res Pract. 2021;228:153647.
9. Mariano G, Farthing RJ, Lale-Farjat SLM, Bergeron JRC. Structural Characterization of SARS-CoV-2: Where We Are, and Where We Need to Be. Front Mol Biosci. 2020;7:605236.
10. Neitthoffer B, Alvarez F, Larrous F, Caillet-Saguy C, Etienne-Manneville S, Boëda B. A short sequence in the tail of SARS-CoV-2 envelope protein controls accessibility of its PDZ-binding motif to the cytoplasm. J Biol Chem. 2024;300(1).
11. Zhu Y, Alvarez F, Wolff N, Mechaly A, Brûlé S, Neitthoffer B, et al. Interactions of SARS-CoV-2 protein E with cell junctions and polarity PDZ-containing proteins. bioRxiv. 2021:2021.12.04.471219.
12. Park SH, Siddiqi H, Castro DV, De Angelis AA, Oom AL, Stoneham CA, Let al. Interactions of SARS-CoV-2 envelope protein with amilorides correlate with antiviral activity. PLoS Pathog. 2021;17(5):e1009519.
13. Javorsky A, Humbert PO, Kvansakul M. Structural basis of coronavirus E protein interactions with human PALS1 PDZ domain. Commun Biol. 2021;4(1):724.
14. Lo Cascio E, Toto A, Babini G, De Maio F, Sanguinetti M, Mordente A, et al. Structural determinants driving the binding process between PDZ domain of wild type human PALS1 protein and SLiM sequences of SARS-CoV E proteins. Comput Struct Biotechnol J. 2021;19:1838-1847.
15. Faraji SN, Raee MJ, Hashemi SMA, Daryabor G, Tabrizi R, Dashti FS, et al. Human interaction targets of SARS-CoV-2 spike protein: A systematic review. Eur J Inflamm. 2022; 20:1721727X221095382.
16. Khan MI, Khan ZA, Baig MH, Ahmad I, Farouk AE, Song YG, Dong JJ. Comparative genome analysis of novel coronavirus (SARS-CoV-2) from different geographical locations and the effect of mutations on major target proteins: An in silico insight. PLoS One. 2020;15(9):e0238344.
17. Rahman MS, Hoque MN, Islam MR, Islam I, Mishu ID, Rahaman MM, et al. Mutational insights into the envelope protein of SARS-CoV-2. Gene Rep. 2021;22:100997.
18. Mou K, Abdalla M, Wei DQ, Khan MT, Lodhi MS, Darwish DB, et al. Emerging mutations in envelope protein of SARS-CoV-2 and their effect on thermodynamic properties. Inform Med Unlocked. 2021;25:100675.
19. Ayatollahi AA, Aghcheli B, Amini A, Nikbakht H, Ghassemzadehpirsala P, Behboudi E, Rajabi A, Tahamtan A. Association between

[DOI: 10.61186/rbmb.13.1.124]

blood groups and COVID-19 outcome in Iranian patients. Future Virol. 2021;16(10):657-65.
20. Hassan SS, Choudhury PP, Roy B. SARS-CoV2 envelope protein: non-synonymous mutations and its consequences. Genomics. 2020;112(6):3890-3892.
21. DeDiego ML, Nieto-Torres JL, Jimenez-Guardeño JM, Regla-Nava JA, Castaño-Rodriguez C, Fernandez-Delgado R, et al. Coronavirus virulence genes with main focus on SARS-CoV envelope gene. Virus Res. 2014;194:124-37.
22. Rahmanzade R, Rahmanzadeh R, Hashemian SM, Tabarsi P. Iran's Approach to COVID-19: Evolving Treatment Protocols and Ongoing Clinical Trials. Front Public Health. 2020;8:551889.
23. Rafieepoor M, Mohebbi SR, Hosseini SM, Tanhaei M, Saeedi Niasar M, Kazemian S, et al. Detection of SARS-CoV-2 RNA in selected agricultural and food retail environments in Tehran, Iran. Front Public Health. 2022;10:823061.
24. Soltani A, Jamalidoust M, Hosseinpour A, Vahedi M, Ashraf H, Yousefinejad S. First molecular-based detection of SARS-CoV-2 virus in the field-collected houseflies. Sci Rep. 2021;11(1):13884.
25. Ko J, Park H, Seok C. GalaxyTBM: template-based modeling by building a reliable core and refining unreliable local regions. BMC Bioinformatics. 2012;13:198.
26. Heo L, Park H, Seok C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic Acids Res. 2013;41(Web Server issue):W384-8.
27. Zhang Q, Wang P, Kim Y, Haste-Andersen P, Beaver J, Bourne PE, et al. Immune epitope database analysis resource (IEDB-AR). Nucleic Acids Res. 2008;36(Web Server issue):W513-8.
28. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. Nat Protoc. 2017;12(2):255-278.
29. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. Bioinformatics. 2016;32(23):3676-3678.
30. Bianchi M, Benvenuto D, Giovanetti M, Angeletti S, Ciccozzi M, Pascarella S. Sars-CoV-2 Envelope and Membrane Proteins: Structural Differences Linked to Virus Characteristics? Biomed Res Int. 2020;2020:4389089.
31. Suryawanshi RK, Koganti R, Agelidis A, Patil CD, Shukla D. Dysregulation of Cell Signaling by SARS-CoV-2. Trends Microbiol. 2021;29(3):224-237.
32. Rizwan T, Kothidar A, Meghwani H, Sharma V, Shobhawat R, Saini R, et al. Comparative analysis of SARS-CoV-2 envelope viroporin mutations from COVID-19 deceased and surviving patients revealed implications on its ion-channel activities and correlation with patient mortality. J Biomol Struct Dyn. 2022;40(20):10454-10469.
33. Alam I, Kamau AA, Kulmanov M, Jaremko Ł, Arold ST, Pain A, et al. Functional Pangenome Analysis Shows Key Features of E Protein Are Preserved in SARS and SARS-CoV-2. Front Cell Infect Microbiol. 2020;10:405.
34. Mandala VS, McKay MJ, Shcherbakov AA, Dregni AJ, Kolocouris A, Hong M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. Nat Struct Mol Biol. 2020;27(12):1202-1208.
35. Chan KK, Tan TJC, Narayanan KK, Procko E. An engineered decoy receptor for SARS-CoV-2 broadly binds protein S sequence variants. Sci Adv. 2021;7(8):eabf1738.
36. Bandoy DJDR, Weimer BC. Analysis of SARS-CoV-2 genomic epidemiology reveals disease transmission coupled to variant emergence and allelic variation. Sci Rep. 2021;11(1):7380.
37. Caohuy H, Eidelman O, Chen T, Mungunsukh O, Yang Q, Walton NI, et al. Inflammation in the COVID-19 airway is due to inhibition of CFTR signaling by the SARS-CoV-2 Spike protein. Sci Rep. 2024;14(1):16895.
38. Lee HJ, Zheng JJ. PDZ domains and their binding partners: structure, specificity, and modification. Cell Commun Signal. 2010;8:8.
39. Nardella C, Visconti L, Malagrinò F, Pagano L, Bufano M, Nalli M, et al. Targeting PDZ domains as potential treatment for viral infections, neurodegeneration and cancer. Biol Direct. 2021;16(1):15.

[DOI: 10.61186/rbmb.13.1.124]

40. Troyano-Hernáez P, Reinosa R, Holguín Á. Evolution of SARS-CoV-2 Envelope, Membrane, Nucleocapsid, and Spike Structural Proteins from the Beginning of the Pandemic to September 2020: A Global and Regional Approach by Epidemiological Week. Viruses. 2021;13(2):243.
41. Li S, Wang J, Dai X, Li C, Li T, Chen L. The PDZ Domain of the E Protein in SARS-CoV Induces Carcinogenesis and Poor Prognosis in LUAD. Microbes Infect. 2024; 22:105381.
42. Schoeman D, Cloete R, Fielding BC. Comparative studies of the seven human coronavirus envelope proteins using topology prediction and molecular modelling to understand their pathogenicity. bioRxiv. 2021. 2021.03.08.434384;
43. Toto A, Ma S, Malagrinò F, Visconti L, Pagano L, Stromgaard K, Gianni S. Comparing the binding properties of peptides mimicking the Envelope protein of SARS-CoV and SARS-CoV-2 to the PDZ domain of the tight junction-associated PALS1 protein. Protein Sci. 2020;29(10):2038-2042.
44. Chai J, Cai Y, Pang C, Wang L, McSweeney S, Shanklin J, Liu Q. Structural basis for SARS-CoV-2 envelope protein recognition of human cell junction protein PALS1. Nat Commun. 2021;12(1):3433.
45. Castaño-Rodriguez C, Honrubia JM, Gutiérrez-Álvarez J, DeDiego ML, Nieto-Torres JL, Jimenez-Guardeño JM, et al. Role of severe acute respiratory syndrome coronavirus viroporins E, 3a, and 8a in replication and pathogenesis. MBio. 2018; 14(6):e0248523.

[DOI: 10.61186/rbmb.13.1.124]