**Iranian Journal of Virology 2022;16 (1):7-15**
**©2022, Iranian Society of Virology**

## Original Article

# Protein Sequence Analysis of E1 Surface Glycoprotein of Hepatitis C Virus

**Zareh-Khoshchehreh R[1], Bamdad T[1*], Arab SS[2], Behdani M[3], Biglar M[4]**

1. Department of Virology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran.
2. Department of Biophysics, Faculty of Biological Sciences, Tarbiat Modares University-TMU, Tehran, Iran.
3. Department of Biotechnology, Venom and Biotherapeutics Molecules Laboratory, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran.
4. Department of Pharmacy, Drug Design and Development Research Center, Tehran University of Medical Sciences, Tehran, Iran.

## Abstract

**Background and Aims:** Hepatitis C Virus infects more than 170 million people globally despite highly effective direct acting antiviral drugs that greatly improved treatment. The Hepatitis C virus envelope glycoproteins E1 and E2 are the major target to induce immune responses. Since, the different aspects of E1 such as its function and structure are still discussed and require further study, in current study critical regions of E1 were evaluated.

**Materials and Methods:** Mutation diversity in these areas was determined using strains that were available in online databanks and authentic software. Furthermore, RT-PCR for E1 was done on HCV-1a positive samples and the sequences were analyzed. The percentage of substitutions, desired and stable residues for mutation in each position were indicated.

**Results:** The integrated results exhibited bNAb epitope (residues 313-328) which is the most conserved epitope in E1 glycoprotein sequence among all genotypes of HCV.

**Conclusion:** These kinds of studies may shed light on identification more binding sites of virus and broadly cross-neutralization of antibodies. Moreover, it may facilitate the modeling of peptides to new antiviral design or boosting the immune response in multi-epitope vaccine studies.

**Keywords:** E1 envelope glycoprotein, Mutation, Epitope, Neutralization, Vaccine design

## Introduction

The hepatitis C virus (HCV) is a single-stranded, positive-sense RNA virus and has been classified in the Flaviviridae family. The 9.6-kb virus genome encodes a polyprotein that is processed into three structural and seven nonstructural proteins (1).The virion is made of a nucleocapsid enclosed within a host-derived lipid envelope.

The structure of virion envelops composed of heterodimer glycoproteins E1 and E2. These glycoproteins are the HCV major determinants for immune response, viral entry, fusion and release (2).

The immunogenicity property of these proteins has been led to design different subunit and peptide vaccine studies (3).

Generally, more studies have been done on E2 glycoprotein than E1. However, due to the critical role of E1 in structure and function of E2 as well as it's interaction with immune system elements, its use in designing vaccine alongside E2 can be very effective (4).

E1 structures consists of a cluster of overlapping epitopes that investigating them help to design vaccine and new drug of HCV.

These regions include N-glycosylation sites (N196, N206, N234, N250, N305 and N325 in H77 strain), Fusion sites (residues 264-290), Stem region (residues 309-349), Transmembrane Domains or TMDs (residues 350-383), Disulfide bonds (C207, C226, C229, C238, C272, C281, C304, and C306), Heptad

[*]**Corresponding author:**
Taravat Bamdad, Email: Bamdad_t@modares.ac.ir.

Repeat regions (residues 330-347) and neutralizing antibodies response sites (5).

Among these regions the antibody binding sites are notable for vaccine studies. These area in E1 contain residues 192−202 that interact with the weakly neutralizing Ab such as monoclonal antibody (mAb) H-111 (17) as well as residues 313−328 that interact with the Broadly Neutralizing Antibodies (bNAbs) such as mAbs IGH505 and IGH526 (6, 7).

Globally, HCV infects more than 170 million people (8) in spite of highly effective direct acting antiviral drugs that greatly improved treatment (9). Moreover, according to the global viral hepatitis elimination program of the World Health Organization (WHO), the lack of an effective vaccine has made this route difficult (10).

This study focuses on general evaluating remarkable regions on whole HCV-E1 glycoprotein in obtained sequences from GenBank.

It determines conserved positions, potential substitutions and desired residues for substitutions in these critical sites.

This type of evaluation could help to find novel treatment pathway and multi-epitope vaccine discovery by increasing knowledge about suitable mutagenesis. The mutation variability is very important item in silico analysis particularly for investigating of highly variable viruses such as HCV. Moreover, it could help to improve the WHO strategies in order to HCV elimination program.

.

## Methods

**Sequence collection:** Reference sequence of seven HCV genotypes with accession numbers NC_004102, NC_009823 to NC_009827 and NC_030791 in addition to 100 sequences of E1 glycoprotein of HCV genotype 1a retrieved from GenBank of National Center for Biotechnology Information (NCBI) Databases (http:// ncbi.nlm.nih. gov).

Moreover, seven critical regions in E1 glycoprotein structure of HCV were selected from the recently published literatures to investigate HCV-E1 as a candidate of HCV vaccine in future studies. These seven regions included Neutralizing Antibody Sites, N-glycosylation sites, Stem regions, Fusion sites, Transmembrane Domains, Disulfide Bonds and Heptad Repeat Regions. Therefore, totally 113 residue substitutions were evaluated.

**RT-PCR and Sequencing**: To evaluate of substitutions in hot spots of Iranian HCV-E1 glycoprotein E1 along with the sequences that retrieved from GenBank, RT-PCR was done on HCV-1a positive samples.

They were received from five patients referred to "Keyvan Virology Medical Diagnosis Laboratory" in Tehran, Iran. The study protocol was confirmed by Ethics Committee of Tarbiat Modares University (no: IR.MODARES.REC. 1400.083).

Viral RNA was extracted from serum samples using a QIAamp viral RNA mini kit according to the manufacturer's instructions (Qiagen, Germany). RT-PCR test was performed to synthesis of E1 (576 bp) gene via two overlapping primer pairs consist of HCV-F1: 5'-TGTGCCCGCTTCAGCCTACCA-3'/ HCV -R1: 5'-CTTCGCCCAGTTCCCCACCA-T-3', HCV-F2: 5'-CCTGGCGGGCATAGCGTATT-TCT-3'/ and HCV-R2: 5'-CAGCGGTGGCCT-GGTGTTGTTA-3'.

The cycling conditions were: 8 min at 95°C as an initial denaturation, 35 cycles at 95°C for 35 sec, 60°C for 35 sec, and 72°C for 45 sec and a final extension step at 72°C for 10 min.

The reaction products were purified and were sequenced using above primers with bidirectional approach.

**Sequence Analysis:** Accuracy of nucleotide sequences was confirmed by BLAST database (http://www.ncbi.nlm.nih.gov/BLAST/).

Afterward, the 100 nucleotide sequences in addition to five Iranian E1 sequences alignment were performed with the reference strains of HCV genotype 1 (Accession numbers: NC_ 004102).

Clustal W program that implemented in Bioedit 7.7.9 and MEGA X were used for the multiple alignments.

On the other hand, the amino acid sequences of E1-HCV were deduced from the nucleotide sequences as well as were obtained from UniProtKB/Swiss-Prot database (https://www. uniprot.org/statistics/Swiss-Prot).

| **Table 1.** Mutation Variability in Neutralizing Ab (NAb) Epitopes of HCV-E1 glycoprotein | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Epitope** | **Amino Acid Positions** | **Percentage of Amino Acid Substitution** | | | | | | | **Substitution in** Iranian E**1 sequences** |
| **NAb (a.a 192-202)** | 192 | Y 72% | L 16% | V 6% | I 3% | F 1% | A 1% | H 1% | |
| | 193 | Q 59% | E 35% | N 2% | H 2% | T 1% | | | |
| | 194 | V 80% | W 14% | Y 5% | | | | | |
| | 195 | R 96% | K 2% | A 1% | G 1% | | | | |
| | 196 | N 100% | | | | | | | |
| | 197 | S 61% | V 14% | T 12% | A 7% | I 4% | | | |
| | 198 | S 62% | T 38% | | | | | | S198T |
| | 199 | G 95% | S 3% | T 1% | | | | | |
| | 200 | L 70% | V 13% | I 9% | S 4% | M 1% | A 1% | G 1% | |
| | 201 | Y 100% | | | | | | | |
| | 202 | H 79% | V 8% | I 7% | Y 3% | M 2% | | | |
| **NAb (a.a.313-328)** | 313 | I 67% | V 19% | L 15% | | | | | |
| | 314 | T 78% | S 22% | | | | | | |
| | 315 | G 100% | | | | | | | |
| | 316 | H 99% | | | | | | | |
| | 317 | R 99% | | | | | | | |
| | 318 | M 100% | | | | | | | |
| | 319 | A 100% | | | | | | | |
| | 320 | W 100% | | | | | | | |
| | 321 | D 100% | | | | | | | |
| | 322 | M 100% | | | | | | | |
| | 323 | M 100% | | | | | | | |
| | 324 | M 96% | L 3% | | | | | | |
| | 325 | N 99% | | | | | | | |
| | 326 | W 100% | | | | | | | |
| | 327 | S 100% | | | | | | | |
| | 328 | P100% | | | | | | | |

Mutations less than one percent are not shown

**Mutation Evaluation:** In order to achieve a comprehensive investigation of conserved residues on the circulating and HCV-E1 sequences in GeneBank, the amino acid substitution frequency in mentioned critical sites were analyzed by using Homology-derived Structures of Proteins (HSSP) (http:// bioinf.modares.ac.ir/software/osprey).

In this step, the most conserved and suitable epitope was detected. In addition, the desired positions and residues for mutagenesis were exhibited.

**Physicochemical Analysis:**

The most conserved epitope site in HCV-E1 glycoprotein was selected and the online ExPASy ProtParam (11) software was used for evaluation of its physicochemical properties. This step will be useful for future in vitro and in vivo studies.

From this point of view, the computation of different physical and chemical characteristics such as molecular weight, extinction coefficient, theoretical isoelectric point (pI), instability and aliphatic indexes and grand average hydropathy (GRAVY) for the target peptide was done.

## Results

To investigate HCV-E1 beside HCV-E2 as a desired candidate to help for designing the multi-epitope HCV vaccine in future studies, seven critical regions in primary structure of E1 glycoprotein were selected. The nucleotide and protein HCV-E1 sequences retrieved from NCBI and UniProtKB. Moreover, five sequences of Iranian E1 gene from HCV-1a isolates were obtained by RT-PCR assay.

**Table 2.** Mutation Variability in Fusion Sites of HCV-E1 glycoprotein

| Epitope | Amino Acid Positions | Percentage of Amino Acid Substitution | | | | | Substitution in Iranian E1 sequences |
|---|---|---|---|---|---|---|---|
| **Fusion Sites (a.a 264-290)** | 264 | L 94% | M 4% | V 1% | | | |
| | 265 | L 92% | I 4% | M 3% | V 1% | | |
| | 266 | V 98% | A 2% | | | | |
| | 267 | G 95% | M 4% | | | | |
| | 268 | S 57% | A 34% | G 5% | T 3% | | |
| | 269 | A 100% | | | | | |
| | 270 | T 82% | A 16% | V 2% | | | |
| | 271 | L 63% | F 15% | M 14% | V 5% | I 2% | |
| | 272 | C 100% | | | | | |
| | 273 | S 100% | | | | | |
| | 274 | A 98% | S 1% | | | | |
| | 275 | L 83% | M 17% | | | | |
| | 276 | Y 100% | | | | | |
| | 277 | V 96% | I 3% | | | | |
| | 278 | G 100% | | | | | |
| | 279 | D 100% | | | | | |
| | 280 | L 81% | M 9% | V 8% | A 3% | I 1% | |
| | 281 | C 100% | | | | | |
| | 282 | G 100% | | | | | |
| | 283 | S 76% | A 20% | G 4% | | | |
| | 284 | V 94% | L 3% | I 2% | A 1% | | |
| | 285 | F 94% | M 5% | L 1% | | | |
| | 286 | L 96% | I 3% | | | | |
| | 287 | V 90% | I 5% | A 4% | | | |
| | 288 | G 76% | S 22% | A 2% | | | |
| | 289 | Q 100% | | | | | |
| | 290 | L 75% | A 18% | M 5% | | | |

Mutations less than one percent are not shown

Then, these sequences were investigated in the mentioned E1 regions. The most conserved and variable residues in hot spot areas were detected. Then, the percentage of the substitution of the various residues at a specific position was indicated.

**Mutation Variability:** Mutation variability in mentioned seven regions of E1was investigated. It was shown which amino acids out of twenty amino acids can be placed in a specific position in E1 critical epitopes. Moreover, the percentage of the presence of each residue in these positions was demonstrated (Tables 1-4).

Whereas, a conserved epitope is a remarkable region for vaccine design, the most conserved epitope among seven sites was determined. Its position was found in residues 313-328 that interact with broadly neutralizing antibodies in E1 (Table 1). Furthermore, it was exhibited which position in this conserved epitope was desired for mutagenesis. Base on the obtained results from Table 1, positions 313 and 314 were shown variable. The other residues in this epitope were great conserved. The appropriate residues for substitution were detected I, V and L in position 313 as well as T and S in position 314.

No substitution was detected in PCR products in this area of E1. Another region of NAbs in E1 structure was residues 192-202 that has diversity in its amino acids.

Indeed, among 11 amino acids were detected seven positions convertible. S198T mutation was detected in sequences which obtained from RT-PCR products (Table 1).

Based on integrated results, eight positions in Fusion area were variable with probability less than 90% for each residue (Table 2). However, most residues with the ability to participate in substitution were limited to 3-5 residues.

**Table 3.** Mutation Variability in Stem Regions and Heptad Repeat Regions of HCV-E1 glycoprotein

| Epitope | Amino Acid Positions | Percentage of Amino Acid Substitution | | | | | | Substitution in Iranian E1 sequences |
|---|---|---|---|---|---|---|---|---|
| **Stem Regions (a.a 309-349)** | 309 | Y 100% | | | | | | |
| | 310 | P 92% | Q 3% | T 3% | S 1% | | | |
| | 311 | G 100% | | | | | | |
| | 312 | H 98% | T 1% | R 1% | | | | |
| | 313 | I 67% | V 19% | L 15% | | | | |
| | 314 | T 78% | S 22% | | | | | |
| | 315 | G 100% | | | | | | |
| | 316 | H 99% | | | | | | |
| | 317 | R 99% | | | | | | |
| | 318 | M 100% | | | | | | |
| | 319 | A 100% | | | | | | |
| | 320 | W 100% | | | | | | |
| | 321 | D 100% | | | | | | |
| | 322 | M 100% | | | | | | |
| | 323 | M 100% | | | | | | |
| | 324 | M 96% | L 3% | | | | | |
| | 325 | N 99% | | | | | | |
| | 326 | W 100% | | | | | | |
| | 327 | S 100% | | | | | | |
| | 328 | P 100% | | | | | | |
| | 329 | T 84% | A 16% | | | | | T329P |
| | **330** | **A 26%** | **T 51%** | **V 18%** | **L 3%** | **I 1%** | | **A330S** |
| | **331** | **A 76%** | **G 15%** | **T 8%** | | | | |
| | **332** | **L 80%** | **M 18%** | **Y 1%** | | | | |
| | **333** | **V 86%** | **I 8%** | **A 4%** | **L 1%** | | | |
| | **334** | **V 74%** | **M 15%** | **L 8%** | **I 2%** | **T 1%** | | **V334M** |
| | **335** | **A 78%** | **S 22%** | | | | | **A335S** |
| | **336** | **Q 77%** | **H 15%** | **Y 5%** | **S 2%** | | | |
| | **337** | **L 72%** | **V 16%** | **M 5%** | **A 4%** | **I 3%** | | |
| | **338** | **L 92%** | **M 4%** | **A 3%** | | | | |
| | **339** | **R 100%** | | | | | | |
| | **340** | **I 67%** | **V 18%** | **L 15%** | | | | **I340V** |
| | **341** | **P 100%** | | | | | | |
| | **342** | **Q 92%** | **E 6%** | **S 1%** | | | | |
| | **343** | **A 75%** | **T 17%** | **L 4%** | **V 2%** | **I 1%** | | |
| | **344** | **I 60%** | **V 21%** | **L 16%** | **F 1%** | **A 1%** | **C 1%** | |
| | **345** | **M 12%** | **L 47%** | **V 22%** | **F 15%** | **I 2%** | **A 1%** | **M345L** |
| | **346** | **D 95%** | **E 4%** | **S 1%** | | | | |
| | **347** | **M 73%** | **I 19%** | **V 5%** | **L 3%** | | | |
| | 348 | I 75% | V 21% | L 3% | F 2% | | | |
| | 349 | A 90% | F 4% | V 2% | T 2% | L 1% | S 1% | |

Bold residues 330−347 are Heptad Repeat Regions and overlap with Stem Regions.
Mutations less than one percent are not shown.

**Table 4.** Mutation Variability in Transmembrane Domains of HCV-E1 glycoprotein

| Epitope | Amino Acid Positions | Percentage of Amino Acid Substitution | | | | | | | Substitution in Iranian E1 sequences |
|---|---|---|---|---|---|---|---|---|---|
| **Transmembrane Domains (TMDs) (a.a 350-383)** | 350 | G 99% | E 1% | | | | | | |
| | 351 | A 91% | G 8% | | | | | | |
| | 352 | H 100% | | | | | | | |
| | 353 | W 100% | | | | | | | |
| | 354 | G 100% | | | | | | | |
| | 355 | V 83% | I 17% | | | | | | |
| | 356 | L 93% | M 3% | V 2% | I 1% | | | | |
| | 357 | A 91% | F 5% | V 1% | L 1% | G 1% | | | |
| | 358 | G 97% | A 3% | | | | | | |
| | 359 | I 49% | L 40% | M 8% | V 3% | | | | |
| | 360 | A 98% | L 1% | | | | | | |
| | 361 | Y 99% | | | | | | | |
| | 362 | F 68% | Y 32% | | | | | | |
| | 363 | S 97% | G 2% | A 1% | | | | | |
| | 364 | M 99% | | | | | | | M364I |
| | 365 | V 75% | K 22% | A 3% | | | | | V365A |
| | 366 | G 96% | A 3% | | | | | | |
| | 367 | N 95% | A 4% | | | | | | |
| | 368 | W 100% | | | | | | | |
| | 369 | A 98% | L 1% | | | | | | |
| | 370 | K 100% | | | | | | | |
| | 371 | V 99% | | | | | | | |
| | 372 | L 74% | A 14% | I 5% | V 3% | F 1% | M 1% | C 1% | L372V |
| | 373 | V 57% | I 29% | L 7% | A 6% | | | | V373L, V373I, V373F |
| | 374 | V 80% | I 20% | | | | | | |
| | 375 | L 71% | M 27% | V 1% | I 1% | | | | |
| | 376 | L 81% | V 10% | I 5% | F 3% | | | | |
| | 377 | L 85% | M 15% | | | | | | |
| | 378 | F 95% | V 3% | A 1% | T 1% | | | | |
| | 379 | A 84% | S 16% | | | | | | |
| | 380 | G 98% | S 2% | | | | | | |
| | 381 | V 99% | | | | | | | |
| | 382 | D 98% | E 2% | | | | | | D382E |
| | 383 | A 84% | G 16% | | | | | | |

Mutations less than one percent are not show.

**Table 5.** Physiochemical Properties of Neutralizing Abs Epitopes in HCV-E1 glycoprotein

| | bNAb Epitope (16 mer) |
|---|---|
| Start | 313 |
| End | 328 |
| Residue Sequence of Ref-Seq (NC_004102) | ITGHRMAWDMMMNWSP |
| Molecular Weight | 1964.32 |
| Theoretical pI | 6.74 |
| Total number of atoms | 261 |
| Aliphatic indexes | 30.63 |
| Extinction Coefficients | 11000 |
| Instability Index | 29.16 |
| Grand average of hydropathicity (GRAVY) | -0.381 |

Heptad Repeat Regions (residues 330-347) was located inside Stem Regions (residues 309-349) that were demonstrated extremely variable (Table 3). While, it were found residues 315-328 great conserved (overlap with bNAb epitope). In TMD, 12 positions out of 34 were detected with higher mutation rate (Table 4). But, all asparagines in N-glycosylation sites and Cysteines in disulfide bonds were detected 100% conserved.

Generally, E1 mutation in Iranian sequences followed the mutation pattern of sequences that obtained from GeneBank. But, the substitutions T329P, A330S, M364I and V373F were exception which included rare mutations.

**Epitope Analysis:** The residues 313-328 (bNAb epitope) was selected the best area for vaccine studies because of it was the most conserved area in HCV-E1 structure.

Therefore, its physicochemical properties were analyzed by the online ExPASy ProtParam software (Table 5). According to this, the average molecular weight of this epitope was 1964.32 Da. Also, that's theoretical isoelectric pH values were 6.74 that it means a negatively charged peptide. The aliphatic index of E1 bNAb epitope demonstrates this peptide is stable. The negative average hydropathicity value indicates the mentioned peptide is non-polar.

**Epitope Investigation in different HCV Genotypes:** Moreover, bNAb region (ITGHRMAWDMMMNWSP) was investigated among seven different HCV genotypes by comparison of their reference sequences.

Isoleucine in position 313 was converted to leucine on genotype 3 and valine on genotypes 6 and 7. Further, threonine in position 314, methionine in position 324 and asparagine in position 325 were replaced to serine in genotype 3, lysine genotype 7 and glutamine genotype 5, respectively.

## Discussion

The high capacity of HCV in antigenic diversity and evasion from immune system is one of the notable points in vaccine design. This strategy particularly related to the surface glycoproteins of the virus that included a cluster of overlapping immunogenic epitopes (5). The molecular investigation of these immunogenic epitopes could be useful in to new drug production and subunit or multi-epitope vaccine design.

Moreover, novel methods in vaccine researches widely use critical epitopes instead of entire proteins (14, 15).

Commonly, glycoprotein E2 has been studied until now despite the importance structural and functional of E1 (12, 13), in current study investigation of the overlapping immunogenic epitopes in whole E1 structure was performed.

Based on obtained results, residues 315-328 in bNAb epitope (is an overlap site with Stem region) and residues 272-286 (expect positions 283,280 and 275) in fusion site were conserved at least 94%. The other positions were exhibitted variable.

Furthermore, all asparagines in N-glycosylation sites and cysteines in disulfide bonds were detected 100% conserved. Since, the protein sequences that used in this study retrieved from GeneBank and UniProtKB/Swiss-Prot database, the results confirmed the importance of glycosylation sites and disulfide bonds in stability of the protein and viral infectivity. Furthermore, the mutation pattern of PCR products was similar to substitutions in mentioned sequences that retrieved from GeneBank. However, T329P, A330S, M364I and V373F substitutions in Stem region and TDM area were rare mutations.

Based on integrated results, bNAb epitope that included residues 313-328 detected as the most conserved epitope in E1 glycoprotein. However, in this epitope, two positions 313 and 314 were indicated almost variable.

Furthermore, the appropriate residues for substitution in these variable positions were determined. So, three nonpolar amino acids Valine, Leucine and Isoleucine in position 313 as well as two polar amino acids Threonine and Serine in position 314 were exhibited suitable for mutagenesis.

Methionine in position 324 was indicated almost 96% conserved and the other residues in 313-328 regions were 99-100% conserved.

Moreover, the obtained results from compareson of seven HCV reference genotypes that

retrieved from GenBank demonstrated bNAb region of E1 is approximately universally conserved across all HCV strains despite high sequence variability of E1. Therefore, using this epitope in vaccine design could be lead to cross-neutralize several HCV genotypes that is very important item in efficiency of vaccine.

Also, according to physichochemical characteristic of E1 bNAb epitope obtained from ExPASy software, we predicted that bNAb peptide had great stability for future vitro and vivo studies. Therefore, it could be as a good immunogen selection in a subunit or multi-epitope vaccine.

In confirmation of these results, previous studies had shown that MAb IGH526 recognizes a linear and helical epitope on HCV-E1 site 313-327. The linear component of the IGH526 epitope was exhibited highly flexible by molecular dynamics investigations. Also, it had shown IGH526 neutralized the prototypic HCV strain H77 (6, 7).

It should be mentioned that mutagenesis is a common procedure for peptide library and multi-epitope vaccine design.

It could be increases the binding affinity that leads to produce stable and effective neutrallizing antibodies (16). Moreover, previous study had indicated the single mutations did not disrupt the epitope on the virus (6).

However, our study suggests that mutagenesis studies be performed on positions that are naturally variable in the structure of the virus instead of conserved sites. For example in current study, 313 and 314 positions in bNAb epitope of HCV-E1 were exhibited variable in circulating HCV strains.

## Conclusion

In conclusion, the heterodimer E1/E2 is the major target of the humoral response during HCV infection and E1 glycoprotein has an important role in this process beside E2. Based on many structural details of the E1 are still incomplete, the in silico studies could be useful for better understanding about different aspects of E1 glycoprotein.

These types of studies could lead to raise recognition and production of a wider range of virus neutralizing antibodies that are helpful in new antiviral design or vaccine studies. In current study, seven critical part of E1 in HCV strains that were available in online authentic databanks were investigated. Furthermore, some E1 sequences that obtained from PCR products were investigated in these regions.

The results of evaluation for mutation variability indicated the percentage of the presence of each amino acid in specific positions of E1. Moreover, it exhibited which residues are appropriate for mutagenesis in these positions. The mutation evaluation demonstrated bNAb epitope in residues 313-328 is the most conserved epitope in E1.

## Acknowledgment

We acknowledge Fatemeh Hosseinii for designing primers and helping in RT-PCR assay.

## Conflict of interest

No conflict of interest is declared.

## Funding

This study was performed as a part of PhD project granted by Tarbiat Modares University (Grant number: TMU-MED-4914).

## References

1. Lindenbach BD, Rice CM. The ins and outs of hepatitis C virus entry and assembly. Nat Rev Microbiol. 2013;11(10):688-700.

2. Ball JK, Tarr AW, McKeating JA. The past, present and future of neutralizing antibodies for hepatitis C virus. Antiviral Res. 2014;105:100-11.

3. Houghton M. Prospects for prophylactic and therapeutic vaccines against the hepatitis C viruses. Immunol Rev. 2011;239(1):99-108.

4. Drummer HE. Challenges to the development of vaccines to hepatitis C virus that elicit neutralizing antibodies. Front Microbiol. 2014;5:329.

5. Freedman H, Logan MR, Law JLM, Houghton M. Structure and function of the hepatitis C virus envelope glycoproteins E1 and E2: antiviral and vaccine targets. ACS Infect Dis. 2016;2(11):749-62.

6. Meunier J-C, Russell RS, Goossens V, Priem S, Walter H, Depla E, et al. Isolation and characterization

of broadly neutralizing human monoclonal antibodies to the e1 glycoprotein of hepatitis C virus. J Virol. 2008;82 (2):966-73.
7. Kong L, Kadam RU, Giang E, Ruwona TB, Nieusma T, Culhane JC, et al. Structure of hepatitis C virus envelope glycoprotein E1 antigenic site 314–324 in complex with antibody IGH526. J Mol Biol. 2015;427 (16):2617-28.
8. Li D, Huang Z, Zhong J. Hepatitis C virus vaccine development: old challenges and new opportunities. Natl Sci Rev. 2015;2(3):285-95.
9. Gane E, Agarwal K. Directly acting antivirals (DAAs) for the treatment of chronic hepatitis C virus infection in liver transplant patients:“a flood of opportunity”. Am J Transplant. 2014;14(5):994-1002.
10. Pourkarim MR, Razavi H, Lemey P, Van Ranst M. Iran’s hepatitis elimination programme is under threat. Lancet. 2018;392(10152):1009.
11. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook. 2005:571-607.
12. Samreen B, Khaliq S, Ashfaq UA, Khan M, Afzal N, Shahzad MA, et al. Hepatitis C virus entry: role of host and viral factors. Infect Genet Evol. 2012;12(8):1699-709.
13. Ghasemi F, Ghayour-Mobarhan M, Gouklani H, Meshkat Z. Development of preventive vaccines for hepatitis C virus E1/E2 protein. Iran J Pathol. 2018;13 (2):113.
14. Bazmara S, Shadmani M, Ghasemnejad A, Aghazadeh H, Bagheri KP. In silico rational design of a novel tetra-epitope tetanus vaccine with complete population coverage using developed immunoinformatics and surface epitope mapping approaches. Med Hypotheses. 2019;130:109267.
15. Pierce BG, Keck Z-Y, Lau P, Fauvelle C, Gowthaman R, Baumert TF, et al. Global mapping of antibody recognition of the hepatitis C virus E2 glycoprotein: implications for vaccine design. Proc Natl Acad Sci USA. 2016;113(45):E6946-E54.
16. Keck Z-Y, Sung VM, Perkins S, Rowe J, Paul S, Liang TJ, et al. Human monoclonal antibody to hepatitis C virus E1 glycoprotein that blocks virus attachment and viral infectivity. J Virol. 2004;78(13):7257-63.