Cell J
Cell Journal(Yakhteh) Homepage: https://www.celljournal.org/
Original Article Vol 25, No 1, January 2023, Pages: 62-72

In Silico Analysis of Neutralizing Antibody Epitopes on The Hepatitis C Virus Surface Glycoproteins

Raziyeh Zareh-Khoshchehreh, Ph.D.[1], Taravat Bamdad, Ph.D.[1]*, Seyed Shahriar Arab, Ph.D.[2], Mahdi Behdani, Ph.D.[3], Mahmoud Biglar, Ph.D.[4]*

1. Department of Virology, School 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, Biotechnology Research Center, Venom and Biotherapeutics Molecules Lab, Pasteur Institute of Iran, Tehran, Iran
4. Department of Pharmacy, Drug Design and Development Research Center, Tehran University of Medical Sciences, Tehran, Iran

**Corresponding Addresses: P.O.Box: 14155-4838, Department of Virology, School of Medical Sciences, Tarbiat Modares University, Tehran, Iran P.O.Box: 14155-6559, Department of Pharmacy, Drug Design and Development Research Center, Tehran University of Medical Sciences, Tehran, Iran*
Emails: BAMDAD_T@modares.ac.ir, mbiglar@tums.ac.ir

Received: 11/November/2021, Revised: 10/June/2022, Accepted: 08/August/2022

Abstract

Objective: Despite of antiviral drugs and successful treatment, an effective vaccine against hepatitis C virus (HCV) infection is still required. Recently, bioinformatic methods same as prediction algorithms, have greatly contributed to the use of peptides in the design of immunogenic vaccines. Therefore, finding more conserved sites on the surface glycoproteins (E1 and E2) of HCV, as major targets to design an effective vaccine against genetically different viruses in each genotype was the goal of the study.

Materials and Methods: In this experimental study, 100 entire sequences of E1 and E2 were retrieved from the NCBI website and analyzed in terms of mutations and critical sites by Bioedit 7.7.9, MEGA X software. Furthermore, HCV-1a samples were obtained from some infected people in Iran, and reverse transcriptase-polymerase chain reaction (RT-PCR) assay was optimized to amplify their E1 and E2 genes. Moreover, all three-dimensional structures of E1 and E2 downloaded from the PDB database were analyzed by YASARA. In the next step, three interest areas of humoral immunity in the E2 glycoprotein were evaluated. OSPREY3.0 protein design software was performed to increase the affinity to neutralizing antibodies in these areas.

Results: We found the effective in silico binding affinity of residues in three broadly neutralizing epitopes of E2 glycoprotein. First, positions that have substitution capacity were detected in these epitopes. Furthermore, residues that have high stability for substitution in these situations were indicated. Then, the mutants with the strongest affinity to neutralize antibodies were predicted. I414M, T416S, I422V, I414M-T416S, and Q412N-I414M-T416S substitutions theoretically were exhibited as mutants with the best affinity binding.

Conclusion: Using an innovative filtration strategy, the residues of E2 epitopes which have the best in silico binding affinity to neutralizing antibodies were exhibited and a distinct peptide library platform was designed.

Keywords: Broadly Neutralizing Antibody, Mutation, Peptide Library, Sequence Analysis

Citation: **Zareh-Khoshchehreh R, Bamdad T, Arab SSh, Behdani M, Biglar M. In silico analysis of neutralizing antibody epitopes on the hepatitis C virus surface glycoproteins. Cell J. 2023; 25(1): 62-72. doi: 10.22074/CELLJ.2022.253363.**

Introduction

Hepatitis C virus (HCV), a member of the genus Hepacivirus, of the *Orthornavirae* kingdom, is a positive-stranded RNA genome. The core embedded viral RNA is surrounded by a host-derived lipid envelope incorporated with two surface glycoproteins E1 and E2 (1).

These glycoproteins include a cluster of overlapping epitopes which involve in a viral attachment and a cell fusion as well as hypervariable regions (HVRs), stem regions, glycosylation sites**,** transmembrane domains (TMDs); disulfide bonds**,** heptad repeat and neutralizing antibodies response regions (2).

Among these areas, antibody binding sites are extremely immunogenic and contain a various range of interactions. For example, the virus can quickly escape with high variations in immunodominant epitopes of HVRs (3). The second group of immunogenic sites is defined as an antigenic domain B that includes conformational epitopes, which identify the majority of broadly neutralizing antibodies (bNAb) in various HCV genotypes (4, 5).

The bNAb epitopes can be targeted by a wide range of neutralizing antibodies which cause the infection inhibition. However, it is relatively difficult access to them on surface glycoproteins via the virion arrangement. So, it is possible that the virus escapes from the humoral immune system (6). A third group, an antigenic domain A, contains epitopes that lead to the non-neutralizing antibodies production (7, 8).

In the term of residue position, E1 residues, including 192-202 and 313-328, are interacted with weakly

neutralizing Ab H-111 and the bNAbs IGH505 as well as IGH526, respectively (9, 10). Furthermore, neutralizing Ab regions which are located on the E2 receptor include three areas. First, residues 412-423 are formed the Epitope I (EPI). This site encompasses the antigenic domain E related to a highly conserved linear epitope which participates in the bNAb response (11, 12). The second area is the Epitope II (EPII)/Domain B and D placed in the 434−446 residues that interact with some neutralizing monoclonal Abs (MAbs) (11, 13). The third area is the Epitope III (EPIII) in a part of Domain C (residues 538-540) that has limited neutralizing abilities and a part of Domain B (residues 523-535) which has bound by broadly neutralizing abilities (9, 13).

In general, in-silico techniques widely are used in the study of peptides and proteins (14-16). Moreover, the study of residue mutations, especially in protein critical sites has been very helpful in better understanding the function of proteins (17-19). One of the widely employed approaches in the epitope studies is to synthesize combinatorial libraries (20, 21). Totally, the peptide libraries are a powerful tool for protein-associated researches such as vaccines, drugs, and diagnostics studies (22, 23).

Different strategies can be used to produce peptide libraries. One of the most suitable methods is the "positional scanning library" that detects antibody epitopes. These types of libraries are particularly useful for peptide sequence optimization. In this strategy, the amino acid residue at a defined position is replaced by all other amino acids, one at a time. Then, activity measurement allows determining which amino-acid residue increases the peptide activity. In the Combinatorial Positional Peptide Libraries, two or more positions are changed at one time.

To aim of multi-epitope vaccines, novel antiviral drugs or new generation of detection kits design, this study focuses on the evaluating critical sites of HCV surface glycoproteins to design a distinct positional peptide library. This peptide library involves in the humoral immune responses.

Materials and Methods

This study protocol was confirmed by the Ethics Committee of Tarbiat Modares University (IR.MODARES.REC.1400.083).

Sample collection

In this experimental study, HCV-1a serum samples were obtained from fourteen Iranian patients who referred to the Keyvan Virology Medical Diagnostic Laboratory, Tehran, Iran. All of the patients had HCV-positive serological test results and no prior history of HCV treatment. We obtained a serum sample of both sexes in the age range from 20 to 70 years.

Reverse transcriptase-polymerase chain reaction and sequencing

Viral RNA was extracted from serum samples using a QIAamp viral RNA mini kit according to the manufacturer's instructions (Cat No:52904, Qiagen, Germany). Reverse transcriptase-polymerase chain reaction (RT-PCR) assay was optimized to amplify the E1 (576 bp) and E2 (1089 bp) genes of isolated viruses via four specific overlapping fragments that designed in this study. Primer pairs include:

HCV-

F1: 5′-TGTGCCCGCTTCAGCCTACCA-3′

R1: 5′-CTTCGCCCAGTTCCCCACCAT-3′

HCV-

F2: 5′-CCTGGCGGGCATAGCGTATTTCT-3′

R2: 5′-CAGCGGTGGCCTGGTGTTGTTA-3′

HCV-

F3: 5′-GTGGTGGGAACGACCGACAG-3′

R3: 5′-CCTCCGCTTGGGATATGAGTAACA-3′

HCV-

F4: 5′-AGCCTTGTCCACCGGCTCAT-3′

R4: 5′-CAGCGCCATTAACCCGACAAGAAC-3′
respectively.

Thermal conditions for all PCR assays were as follows: an initial denaturation step at 95°C for 8 minutes, 35 cycles at 95°C for 35 seconds, 60°C for 35 seconds, and 72°C for 45 seconds and a final extension step at 72°C for 10 minutes.

PCR reaction products were gel- purified, and their sequencing was determined via a bidirectional approach using the above primers. The accuracy of primary sequences was confirmed by the BLAST database (http://www.ncbi.nlm.nih.gov/BLAST/).

Sequence analysis

In addition to RT-PCR sequences, 100 sequences of HCV genotype 1a were retrieved from the GenBank of National Center for Biotechnology Information (NCBI) Databases (http://ncbi.nlm.nih.gov). The nucleotide alignment was performed for all sequences with the reference strains of the HCV genotype 1 (Accession numbers: NC_038882 and NC_004102). The Clustal W program that implemented in the Bioedit 7.7.9 and MEGA X (24) was used for the multiple alignments. Moreover, all 14 obtained sequences were deposited in the GenBank with accession numbers MW736574 to MW736578 for the E1 gene as well as the accession numbers MW736579 to MW736587 for E2.

Analysis of primary and secondary structure

All full length of E1 and E2 was evaluated to find the best neutralizing residues. At first, the active sites of the HCV-E1 and E2 structures were found from the recently published literatures (2, 25).

Then, 100 protein sequences of HCV surface glycoproteins were obtained from the UniProtKB/Swiss-Prot database (https://www.uniprot.org/statistics/Swiss-Prot). Moreover, the amino acid sequences were deduced from the nucleotide sequences of RT-PCR products. Then, missense and silent mutations were assessed in all sequences by the MEGA X software (24). Also, a selective pressure possibility was calculated to better understand of the substitution rate in the RT-PCR products. It was obtained using the methods of Nei and Gojobori in the Lanl database (https://www.hiv.lanl.gov).

Base of acquired information was selected bNAb epitopes as the appropriate sites for peptide library design that effectively in stimulation of immune response. Therefore, the online ExPASy ProtParam software and YASARA program (26) were employed for investigation and evaluation of primary and secondary properties of selected epitope structures, respectively. It was performed to acquire primitive data about peptide interaction with antibodies and presented for future *in vitro* and *vivo* studies. Moreover, the immune epitope database (IEDB) (27) was used to investigate the previous experimentally derived HCV bNAb epitope data.

The inclusion criteria for mentioning software were based on properties of amino acid sequences and exclusion criteria were based on defined criteria in software.

Three-dimensional structure analysis

All three-dimensional structures (3D-structure) of HCV-E1 and E2 glycoproteins, which were bound to antibodies were retrieved from Protein Data Bank (PDB) (https://www.rcsb.org) with Pdb Ids, including 4N0Y, 4UOI, 4DGV, 4DGY, 4G6A, 4GAG, 4GAJ, 4HS6, 4HS8, 4MWF, 4WHT, 4XVJ, 5EOC, 5FGB, 5FGC, 5VXR, 6BKB, 6BKD, 6BKC, 6BZU, 6BZY, 6BZW, 6BZV, 4WEB, 4JZN, 4JZO, 4HZL, 4Q0X, 4Z0X, 5ERW, 2KZQ, 5NPH, 5NPJ and 5NPI. All of them were investigated to find the most significant structure of the binding antigen and antibody determinants. Thus, structural coordinates for all retrieved Pdb files were analyzed and visualized by PDB viewer software YASARA and SPDBV.

Substitution frequency

In order to better understanding of most common HCV-1a, that circulating in the infected population worldwide, the amino acid substitution frequency in each position of the bNAb epitopes was determined. On average, we analyzed information about 90% of the UniProtKB/Swiss-Prot HCV-1a sequences through Homology-derived Structures of Proteins (HSSP) software (http://bioinf.modares.ac.ir/software). The reference sequence of HCV-1a (Accession numbers: NC_004102) used as our input to HSSP and the amino acid substitution percentage was considered as our output. Subsequently, these data were compared with our amplicons sequences.

This study was hypothesized that some of E2 epitopes mutants have a substantial effect on the neutralizing antibody binding affinity. Therefore, the substitution frequency step was considered to decrease the total of predicted mutants and acquire the more stable immunogenic peptides. This filtration approach was implemented to avoid confusing the immune system in the face of various immunogenic epitopes as a virus escape strategy that is very important in the success of vaccine projects.

Furthermore, to compare the mentioned area in the other genotypes of HCV, the other reference sequences of HCV were retrieved from NCBI, including NC_009823.1, NC_009824.1, NC_009825.1, NC_009826.1, NC_009827.1 and NC_030791.1, and a comparison was made.

Mutagenesis

In the step of in silico mutagenesis to predict a residue with the strongest binding affinity to antibodies in each position, OSPREY 3.0 (Open Source Protein REdesign for You) software (28) was used. The mentioned PDB files with the best interaction properties were used as the input files for OSPREY. Also, only the most frequent residues detected in the previous section were used in OSPREY and the mutants were predicted.

Furthermore, the energy landscape alteration of bound and unbound situations caused by amino acid substitutions and their conformational entropy was analyzed and calculated by OSPREY. Finally, selected mutants with the strongest binding affinity predicted as output of OSPREY and presented as PDB files. Then, a distinct positional peptide library was designed from HCV-E2 bNAb epitopes.

Results

The epitope regions and their common substitutions of the HCV-1a glycoproteins, E1 and E2, were evaluated to design a practical peptide library from the bNAb epitopes. Totally, one hundred HCV-1a isolates, of both E1 and E2, were retrieved from NCBI. Moreover, nine HCV-E2 and five HCV-E1 sequences were obtained from RT-PCR products.

Sequence analysis

Due to evaluation on the representation of complete HCV-E1 and E2 genes, multiple alignments were implemented. The missense and silent substitutions in critical epitopes of HCV-E1 and E2 regions were determined (Fig.1).

Because of more nucleotide diversity in the HCV-E2 regions, the selective pressure possibility was calculated in entire sequences obtained from patients. Its value was 0.19 which means a negative possibility in these sequences (Non-Synonymous/Synonymous <1).

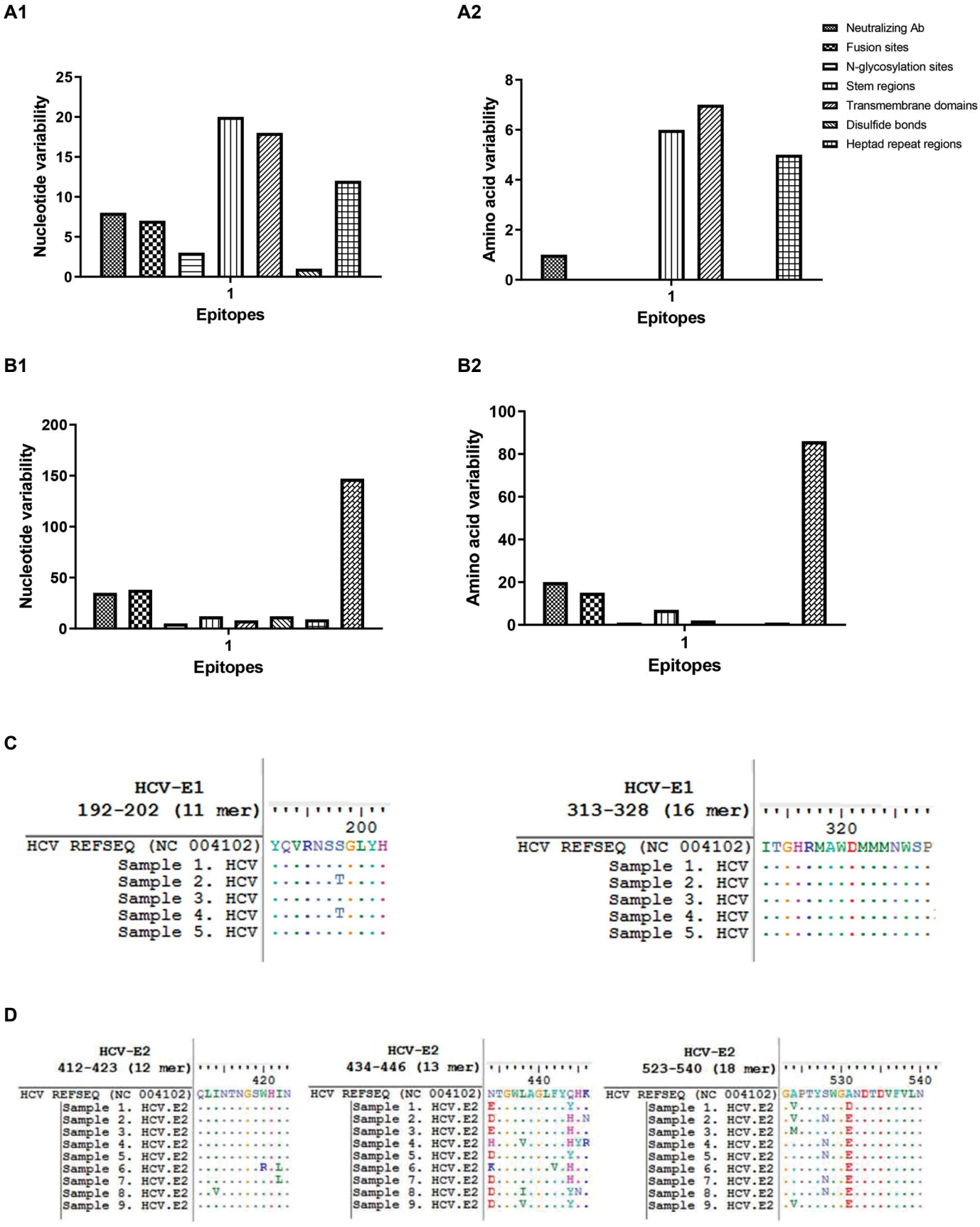

Fig.1: The sequence variability of obtained HCV-E1 and E2 genes from the PCR products. The variability of **A1.** Nucleotide and **A2.** Amino acid sequences in different epitope regions of HCV-E1. The variability of **B1.** Nucleotide and **B2.** Amino acid sequences in different epitope regions of HCV-E2. **C.** Mutations in the E1 NAb epitopes of PCR products compared to reference sequence of HCV genotype 1a (NC_004102). **D.** Mutations in the E2 NAb epitopes of PCR products in comparison with the reference sequence of HCV genotype 1a (NC_004102). HCV; Hepatitis C virus and PCR; Polymerase chain reaction.

Epitope evaluation

We selected a cluster of epitopes related to humoral immune response from the reliable article (2), IEDB, and UniProtKB/Swiss-Prot databases. Of four bNAb epitopes, three epitopes (EPI, EPII, and EPIII) belonged to the HCV-E2, and one epitope belongs to the HCV-E1.

In the next step of epitope investigation, this cluster of epitopes was analyzed in terms of physicochemical properties using the ExPASy software and their stability was confirmed (Table 1). Their high aliphatic indexes indicate these epitopes are in a wide range of thermostability, although the EPI has the highest aliphatic indexes (90.50). Also, negative GRAVY values show that these molecules are nonpolar.

In the secondary structure, evaluation was indicated a suitable condition of atoms and residues placement. Also, it was shown β-Hairpin structures to binding to the antibody by YASARA. Thus, primary and secondary structures of mentioned epitopes were exhibited suitable for binding to neutralizing antibodies.

In general, according to integrated data it looked better to select EPI in the E2 glycoprotein for the following current study.

Mutation frequency

The percentage of mutation frequency was analyzed by using the HSSP databases for about 90% of E1 and E2 sequences in the NCBI. In this filtration strategy, five positions of the EPI, nine positions of the EPII, and eleven positions of the EPIII were detected as changeable and desirable sites of mutagenesis. We detected residues that could be substituted in these situations with high stability (Table 2). The bNAb site of E1 was highly conserved and limited number of mutations was observed among all the evaluated strains. Thus, E1 was excluded from this study.

These data were in accordance with the obtained results from Sequence Analysis section. The rare substitutions contained W420R, N434E and A524M. The HCV stable variants, specifically for these epitopes, that were observed in this study were too different of our result of database analysis.

The peptide sequences of seven reference genotypes were compared to evaluate the selected epitopes in the other HCV genotypes (Table 3). It was indicated that the EPI is the most conserved epitope among all HCV genotypes, but EPII and EPIII had clear variability.

Table 1: Physicochemical properties of bNAb epitopes in the HCV-E2 glycoprotein

Physicochemical parameters	EPI (12 mer)	EPII (13 mer)	EPIII (18 mer)
Start	412	434	523
End	423	446	540
Residue sequence of ref-eeq (NC_004102)	QLINTNGSWHIN	NTGWLAGLFYQHK	GAPTYSWGANDTDVFVLN
Molecular weight	1396.53	1534.74	1927.06
Theoretical pI	6.74	8.60	3.56
Total number of atoms	192	213	260
Aliphatic indexes	90.50	67.69	65.00
Extinction coefficients	5500	6990	6990
Instability index	33.39	58.45	-3.40
Grand average of hydropathicity (GRAVY)	-0.600	-0.431	-0.122

These analyses were performed by the online ExPASy ProtParam software.

Table 2: Percentage of mutation variability in the bNAb epitopes of the HCV-E2 glycoprotein

Epitope	Amino acid positions	Percentage of amino acid substitution							
EPI	412	Q 92%	H 3%	E 1%	N 1%	S 1%	G 1%		
	413	**L 100%							
	414	I 70%	V 29%	M 1%					
	415	N 94%	R 3%	K 1%	H 1%				
	416	T 91%	S 6%	A 3%					
	417	N 99%							
	418	G 99%							
	419	S 99%							
	420	**W 100%							
	421	**H 100%							
	422	I 94%	L 5%	V 1%					
	423	**N 100%							
EPII	434	N 47%	D 19%	H 12%	Q 12%	K 2%	T 1%	S 1%	
	435	T 99%	A 1%						
	436	**G 100%							
	437	W 61%	F 39%						
	438	L 36%	I 48%	V 15%	F 1%				
	439	A 99%							
	440	G 76%	A 18%	S 6%					
	441	**L 100%							
	442	F 80%	L 11%	I 8%	V 1%				
	443	Y 98%	H 1%						
	444	Q 1%	Y 38%	H 22%	T 15%	R 9%	S 4%	V 2%	F 1%
	445	H 54%	N 27%	Y 9%	R 5%	S 2%	K 1%	Q 1%	
	446	K 73%	R 17%	N 5%	S 4%	Q 1%			
EPIII	523	**G 100%							
	524	A 59%	V 31%	E 2%	N 2%	I 2%	T 1%	L 1%	
	525	P 99%							
	526	T 96%	A 4%						
	527	Y 94%	F 5%						
	528	S 21%	N 48%	T 19%	R 8%	G 1%	K 1%		
	529	W 99%	F 1%						
	530	**G 100%							
	531	A 21%	E 57%	G 7%	S 7%	D 5%	T 1%	V 1%	C 1%
	532	N 98%	K 1%						
	533	D 45%	E 48%	V 2%					
	534	T 96%	S 3%						
	535	**D 100%							
	536	V 99%							
	537	F 81%	L 18%						
	538	V 58%	L 37%	I 4%	M 1%				
	539	L 99%							
	540	N 81%	E 7%	K 3%	R 3%	T 2%	D 1%	Q 1%	

**; Some substitutions with a special score in the table, such as 100%, are extremely close to it but not exactly.

Table 3: Comparing three peptide regions (EPI, EPII, and EPIII in genotype 1) among seven reference genotypes of the HCV in the E2 glycoprotein

Peptide region	Genotype	Accession number of ref seq	Residues
EPI	1	NC_004102	QLINTNGSWHIN
	2	NC_009823.1	QL**V**NTNGSWHIN
	3	NC_009824.1	QL**V**NTNGSWHIN
	4	NC_009825.1	QLIN**S**NGSWHIN
	5	NC_009826.1	Q**FV**N**T**NGSWHIN
	6	NC_009827.1	QLINTNGSWHIN
	7	NC_030791.1	QLINTNGSWHIN
EPII	1	NC_004102	NTGWLAGLFYQHK
	2	NC_009823.1	**H**TG**FI**A**S**LFY**T**H**S**
	3	NC_009824.1	NTG**FI**AGLFY**Y**HK
	4	NC_009825.1	NTG**F**LA**S**LFY**T**HK
	5	NC_009826.1	**Q**TG**FI**AG**IM**Y**A**HK
	6	NC_009827.1	**Q**TG**FLSA**LFY**RSN**
	7	NC_030791.1	**Q**TG**F**LAA**L**FY**T**H**R**
EPIII	1	NC_004102	GAPTYSWGANDTDVFVLN
	2	NC_009823.1	GAPTY**T**WG**E**N**E**TDVF**L**LN
	3	NC_009824.1	G**V**PTY**T**WG**E**N**EK**DVF**L**LL**K**
	4	NC_009825.1	G**V**PTY**T**WG**E**N**E**TDVF**L**LN
	5	NC_009826.1	G**C**PTY**N**WG**S**N**E**TD**ILL**LN
	6	NC_009827.1	G**N**PTY**N**WG**E**N**E**TDVF**M**L**E**
	7	NC_030791.1	G**V**PTY**T**WG**E**N**ES**DVF**X**LN

Bold letter shows the substitutions.

Desirable substitutions prediction on the bNAb regions

The structural coordinates and characteristics within the 3D-structure of the HCV-E2 in a complex with neutralizing antibodies were analyzed and visualized by the PDB viewer software, YASARA and SPDBV. Two 3D-structures, PDB IDs were selected, including 4DGV for EPI and 4MWF for EPII and EPIII. Based on the most significant situation in binding between mentioned epitope and antibodies, we considered in this selection. 4MWF is the most complete 3-D structure of E2, which has been presented and use for EPII and EPIII analysis. However, a part of EPI was deleted in this PDB file. So, 4DGV for EPI evaluation was selected. It makes the best binding structure between EPI residues and antibody determinant.

Then, to design a worthy peptide library from the bNAb epitopes of E2, substitutions with the strongest affinity to neutralize antibodies were predicted using OSPREY 3.0. The obtained residues from the previous filtration step were used in this design (Table 4).

Score alteration based on the energy landscape of possible situations and their conformational entropies were calculated (Table 4). The results revealed the residue that had the strongest binding affinity to antibodies and a minimum score.

The integrated results showed that the EPI was the best region among bNAb epitopes of E2, although the EPII and EPIII had a slightly lower score than the EPI through OSPREY. These scores can be related to differences of selected PDB IDs.

On the other hand, Q412E substitution has the lowest score among EPI residues in the PDB ID 4DGV. But it was not reproducible among other PDB IDs. Thus, I414M was acquired an excellent score in the EPI region. Moreover, T416S and I422V substitutions were exhibited pleasant scores. Also, Q412H, I414V and I422L were demonstrated more circulating among the infected population (Table 2). For this reason, they are mostly neutralized by antibodies. It may be related to the virus escape strategies.

Moreover, EPII and EPIII were evaluated in PDB ID 4MWF by the following these methods. Then, their appropriate substitution was exhibited (Table 4). Due to the greater importance of the EPI, simultaneous mutations were also investigated in this epitope. Finally, a Positional Peptide Library from a cluster of the bNAb epitope of the HCV-E2 was mapped using obtained results (Fig.2).

Table 4: The binding scores of the bNAb epitopes on the HCV-E2 glycoprotein in complex with MAbs through OSPREY3.0

Epitope of E2 glycoprotein	PDB ID	Chain	Residue number	Residue(s) in ref seq	Altered residue	Score
EPI	4DGV	A	412	Q	E	-20.25
(412-423 aa)		A	414	I	M	-14.13
		A	415	N	H	171.68
		A	416	T	S	-12.59
		A	422	I	V	-9.82
		A	414, 415	I, N	M, H	158.46
		A	414, 416	I, T	M, S	-26.05
		A	412, 414, 416	Q, I, T	N, M, S	-45.38
		A	414, 416, 422	I, T, I	M, S, V	-35.44
EPII	4MWF	C	437	W	F	-29.11
(434-446 aa)		C	440	G	S	1104.13
		C	444	Q	R	-58.02
		C	445	H	N	-39.47
		C	446	K	R	-58.69
		D	437	W	F	-39.14
		D	440	G	A	211.27
		D	444	Q	R	-58.40
		D	445	H	S	-40.39
		D	446	K	R	-57.89
EPIII	4MWF	C	524	A	N	-46.69
(523-535, 536-540 aa)		C	526	T	A	-32.72
		C	527	Y	F	-37.32
		C	528	S	R	-58.53
		C	529	W	F	-42.91
		C	531	A	E	-47.91
		C	533	D	E	-45.22
		C	534	T	S	-39.75
		C	537	F	L	35.91
		C	538	V	M	-40.43
		D	524	A	N	-43.27
		D	526	T	A	-33.67
		D	527	Y	F	-36.02
		D	528	S	R	-59.46
		D	529	W	F	-44.55
		D	531	A	G	-28.16
		D	533	D	E	-44.74
		D	534	T	S	-40.50
		D	537	F	L	33.17
		D	538	V	M	-41.13

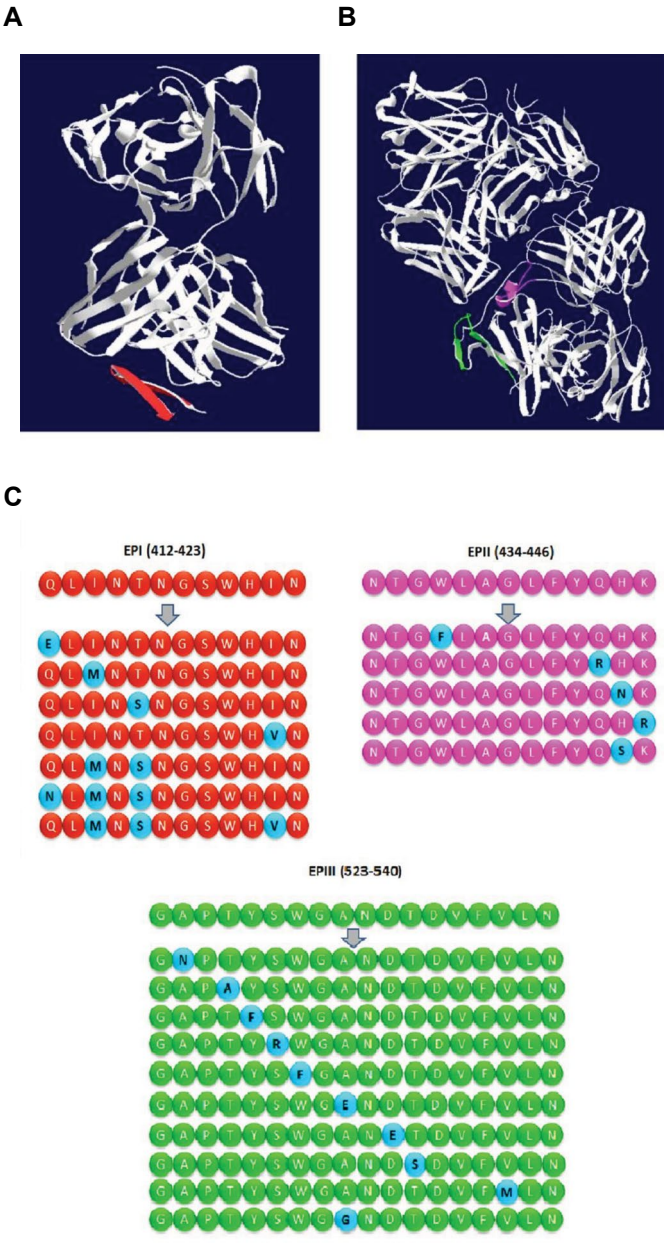

Fig.2: Designed positional peptide library from bNAb regions in the HCV-E2 glycoprotein. **A.** The crystal structure of EPI region (residues 412-423) in red (PDB ID: 4DGV). **B.** The crystal structure of the EPII region (residues 434-446) in purple and EPIII region (residues 523-540) in green (PDB ID: 4MWF). **C.** Designed positional peptide library from EPI, EPII, and EPIII of HCV-E2 glycoprotein from the most common substitutions circulating in the host HCV population using OSPREY. HCV; Hepatitis C virus and PCR; Polymerase chain reaction.

Discussion

One of the novel strategies of vaccine designing is implementing critical epitopes instead of using entire immunogenic protein. Many studies have been done on the mutagenesis of HCV epitopes and Multi-epitope vaccines until now (29, 30).

The ability of HCV in quick production of escape mutants is one of the considerable points of a vaccine production. The HCV contains a cluster of overlapping immunogenic epitopes in its surface glycoproteins that includes these escape mutants areas. This cluster is a suitable target for designing vaccines, novel drugs, or the new generation of detection kits.

To evaluate of immunogenic epitopes on the 3-D structure of HCV, PDB database was used. The results were demonstrated that the 4N0Y was the only 3-D structure of bNAb sites of E1 (residues 314-324). It was complex with the MAb IGH526 (10). Also, 4DGV and 4MWF detected as two of the best structures of E2 in investigation of binding affinity. These structures included bNAb epitopes in complex with Human neutralizing antibodies. These bNAbs named HCV1 and AR3c for 4DGV and 4MWF, respectively (31, 32).

According to integrated results, the cluster of E2 epitopes, influences B cell responses. So, E2 were picked out by focusing on bNAbs production as a key point.

In general, the results of the current study were shown that the substitution pattern of the obtained strains from Iran and the other sequences from NCBI were similar, although some mutations such as W420R are rare.

Previous studies have reported some substitution like L438V, which was visualized in the EPII of obtained sequences from PCR products, lead to the viral escape from monoclonal antibody (MAb) CBH-2 and reduce viral fitness in the cell culture by MAb HC-11 (33) or altering E2 receptor binding (34). But this substitution may be neutralized by the other neutralizing antibodies such as no-human MAbs (35). Therefore, there is a need for further studies using new strategies like what is suggested in this study.

In the current study, was shown that not only some residues are stably circulating in the host population, but also every substitution couldn't use for vaccine and immune stimulation. In fact, there was a specified pattern of substitutions, even in the protein sequences of a virus with the quasispecies ability. None of the mutagenesis studies relating to HCV has mentioned this issue in mutant production. According to previous studies, many experimentally-proven resistant mutants in peptide libraries were found rarely in the host population. In fact, 24% of them did not detect in any of the natural sequences (35, 36). Nevertheless, the residues that are great conserved in the host population mention a selection pressure to restrict mutation. Indeed, we believe that mutagenesis using common mutations in HCV strains of infected individuals can help to produce multi-epitope vaccines and development of effective immunity.

In the second step of filtration were detected the mutants with the strongest affinity to neutralizing antibodies by OSPREY. For instance, in position 414 located in EPI, a Methionine with the lowest score was detected by the most powerful connection among single mutations. Interestingly, N415 was determined inappropriate position for substitutions. However, there are several studies on mutagenesis in

this position (35). In addition to single substitutions, simultaneous mutations naturally were occurred in the special region of HCV strains. Consequently, combinatorial substitutions in unique peptide sequences are indicated. I414M-T416S with two substitutions was detected as the most stable combinatorial mutations in EPI by OSPREY.

The previous studies on conformational changes have been shown the alteration of protein function by mutation in critical binding sits (18, 19). Compared to previous HCV vaccine studies with multi-epitopes, our filtration strategy will enhance stable specific neutralizing antibody responses and reduce unwanted or non-neutralizing antibody responses. It is shown that EPI is the most conserved epitope among all HCV genotypes. Therefore, the predicted peptide library for EPI will be usable for all HCV genotypes. Hence, genotype five with three differentiations in amino acids may be an exception.

It might be predicted that a multi-epitope vaccine is reasonably type to produce a more specific induction of immune responses. Thus, it can be proposed that an approach of this study to design a peptide library and then multi-epitope vaccine for hypervariable pathogens can be useful by increasing in producing effective neutralizing antibodies.

The lack of a whole 3-D-structure for E1 and E2 glycoproteins and absence of E1 and E2 entire sequences from Iranian strains in NCBI were the limitations of the current study.

Generally, our results pointed to the design of a positional peptide library using replacement of filtered residues instead of all residues at the defined position. It theoretically to be effective against neutralizing antibodies was indicated. It was suggested to implement the mentioned strategy to design a peptide library as an innovative and cost-benefit procedure in the vaccine, drug, and diagnostic kits industry. To prove our hypothesis, the key point to design all vaccines is to induce and produce effective antibodies against the wild-type virus. This study was suggested epitopes with maximum coverage in wild-type viruses.

Conclusion

Here, we theoretically approached more stable and specific humoral immune response through an increase binding affinity of HCV mutants to neutralize antibody. Filtering many inefficient mutants of HCV that showed the best positional mutagenesis residues illustrated the significant role of positions 412, 414, 416, and 422 in the EPI of E2. Such platform may represent as a capable insight to design multi-epitope vaccines and/or novel medicine and/or diagnostics tools.

Acknowledgments

We acknowledge Fatemeh Hosseini for designing primers and helping in the RT-PCR assay as well as Zahra Sohbatzadeh. This study was performed as a part of the Ph.D. project granted by Tarbiat Modares University (Grant number: TMU-MED-4914). The authors declare that they have no conflict of interests.

Authors' Contributions

T.B.; Participated in the critical review of the manuscript and project administration. R.Z-Kh.; Conceived the study design, analyzed data and wrote the manuscript. S.Sh.A.; Introduced bioinformatics tools and contributed in silico data analysis. M.Be., M.Bi.; Were involved in the data collection and analysis. All the authors read and approved the final manuscript.

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. Freedman H, Logan MR, Law JL, 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): 749-762.
3. Shimizu YK, Hijikata M, Iwamoto A, Alter HJ, Purcell RH, Yoshikura H. Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. J Virol. 1994; 68(3): 1494-1500.
4. Law M, Maruyama T, Lewis J, Giang E, Tarr AW, Stamataki Z, et al. Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge. Nat Med. 2008; 14(1): 25-27.
5. Perotti M, Mancini N, Diotti RA, Tarr AW, Ball JK, Owsianka A, et al. Identification of a broadly cross-reacting and neutralizing human monoclonal antibody directed against the hepatitis C virus E2 protein. J Virol. 2008; 82(2): 1047-1052.
6. Keck ZY, Saha A, Xia J, Wang Y, Lau P, Krey T, et al. Mapping a region of hepatitis C virus E2 that is responsible for escape from neutralizing antibodies and a core CD81-binding region that does not tolerate neutralization escape mutations. J Virol. 2011; 85(20): 10451-10463.
7. Keck ZY, Xia J, Cai Z, Li TK, Owsianka AM, Patel AH, et al. Immunogenic and functional organization of hepatitis C virus (HCV) glycoprotein E2 on infectious HCV virions. J Virol. 2007; 81(2): 1043-1047.
8. Owsianka AM, Tarr AW, Keck ZY, Li TK, Witteveldt J, Adair R, et al. Broadly neutralizing human monoclonal antibodies to the hepatitis C virus E2 glycoprotein. J Gen Virol. 2008; 89(Pt 3): 653-659.
9. Keck ZY, 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): 7257-7263.
10. 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): 2617-2628.
11. Keck ZY, Xia J, Wang Y, Wang W, Krey T, Prentoe J, et al. Human monoclonal antibodies to a novel cluster of conformational epitopes on HCV E2 with resistance to neutralization escape in a genotype 2a isolate. PLoS Pathog. 2012; 8(4): e1002653.
12. Potter JA, Owsianka AM, Jeffery N, Matthews DJ, Keck ZY, Lau P, et al. Toward a hepatitis C virus vaccine: the structural basis of hepatitis C virus neutralization by AP33, a broadly neutralizing antibody. J Virol. 2012; 86(23): 12923-12932.
13. Sautto G, Tarr AW, Mancini N, Clementi M. Structural and antigenic definition of hepatitis C virus E2 glycoprotein epitopes targeted by monoclonal antibodies. Clin Dev Immunol. 2013; 2013: 450963.
14. Bhardwaj VK, Singh R, Sharma J, Das P, Purohit R. Structural based study to identify new potential inhibitors for dual specificity tyrosine-phosphorylation- regulated kinase. Comput Methods Programs Biomed. 2020; 194: 105494.
15. Gopalakrishnan C, Kamaraj B, Purohit R. Mutations in microRNA binding sites of CEP genes involved in cancer. Cell Biochem Biophys. 2014; 70(3): 1933-1942.
16. Singh R, Bhardwaj VK, Purohit R. Potential of turmeric-derived compounds against RNA-dependent RNA polymerase of SARS-

CoV-2: an in-silico approach. Comput Biol Med. 2021; 139: 104965.
17. Tanwar G, Purohit R. Gain of native conformation of Aurora A S155R mutant by small molecules. J Cell Biochem. 2019; 120(7): 11104-11114.
18. Rajendran V, Gopalakrishnan C, Purohit R. Impact of point mutation P29S in RAC1 on tumorigenesis. Tumour Biol. 2016; 37(11): 15293-15304.
19. Rajendran V, Gopalakrishnan C, Sethumadhavan R. Pathological role of a point mutation (T315I) in BCR-ABL1 protein-A computational insight. J Cell Biochem. 2018; 119(1): 918-925.
20. De Berardinis P, Sartorius R, Fanutti C, Perham RN, Del Pozzo G, Guardiola J. Phage display of peptide epitopes from HIV-1 elicits strong cytolytic responses. Nat Biotechnol. 2000; 18(8): 873-876.
21. Li W, Li L, Sun T, He Y, Liu G, Xiao Z, et al. Spike protein-based epitopes predicted against SARS-CoV-2 through literature mining. Med Nov Technol Devices. 2020; 8: 100048.
22. Bozovičar K, Bratkovič T. Evolving a peptide: library platforms and diversification strategies. Int J Mol Sci. 2019; 21(1): 215.
23. Mink MA, Benichou S, Madaule P, Tiollais P, Prince AM, Inchauspe G. Characterization and mapping of a B-cell immunogenic domain in hepatitis C virus E2 glycoprotein using a yeast peptide library. Virology. 1994; 200(1): 246-255.
24. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018; 35(6): 1547-1549.
25. Yechezkel I, Law M, Tzarum N. From structural studies to HCV vaccine design. Viruses. 2021; 13(5): 833.
26. Land H, Humble MS. YASARA: a tool to obtain structural guidance in biocatalytic investigations. Methods Mol Biol. 2018; 1685: 43-67.
27. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The immune epitope database (IEDB): 2018 update. Nucleic Acids Res. 2019; 47(D1): D339-D343.
28. Hallen MA, Martin JW, Ojewole A, Jou JD, Lowegard AU, Frenkel MS, et al. OSPREY 3.0: open-source protein redesign for you, with powerful new features. J Comput Chem. 2018; 39(30): 2494-2507.
29. Bazmara S, Shadmani M, Ghasemnejad A, Aghazadeh H, Pooshang Bagheri K. 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.
30. Pierce BG, Keck ZY, 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-E6954.
31. Kong L, Giang E, Nieusma T, Kadam RU, Cogburn KE, Hua Y, et al. Hepatitis C virus E2 envelope glycoprotein core structure. Science. 2013; 342(6162): 1090-1094.
32. Kong L, Giang E, Robbins JB, Stanfield RL, Burton DR, Wilson IA, et al. Structural basis of hepatitis C virus neutralization by broadly neutralizing antibody HCV1. Proc Natl Acad Sci USA. 2012; 109(24): 9499-9504.
33. Keck ZY, Olson O, Gal-Tanamy M, Xia J, Patel AH, Dreux M, et al. A point mutation leading to hepatitis C virus escape from neutralization by a monoclonal antibody to a conserved conformational epitope. J Virol. 2008; 82(12): 6067-6072.
34. El-Diwany R, Cohen VJ, Mankowski MC, Wasilewski LN, Brady JK, Snider AE, et al. Extra-epitopic hepatitis C virus polymorphisms confer resistance to broadly neutralizing antibodies by modulating binding to scavenger receptor B1. PLoS Pathog. 2017; 13(2): e1006235.
35. Rodrigo C, Walker MR, Leung P, Eltahla AA, Grebely J, Dore GJ, et al. Limited naturally occurring escape in broadly neutralizing antibody epitopes in hepatitis C glycoprotein E2 and constrained sequence usage in acute infection. Infect Genet Evol. 2017; 49: 88-96.
36. Keck Zy, Angus AG, Wang W, Lau P, Wang Y, Gatherer D, et al. Non-random escape pathways from a broadly neutralizing human monoclonal antibody map to a highly conserved region on the hepatitis C virus E2 glycoprotein encompassing amino acids 412-423. PLoS Pathogens. 2014; 10(8): e1004297.