



Bioinformatics Analysis of Domain 1 of HCV-Core Protein: Iran

Behzad Dehghani¹ · Tayebeh Hashempour¹ · Zahra Hasanshahi¹ · Javad Moayedi¹

Accepted: 8 March 2019 / Published online: 20 April 2019
© Springer Nature B.V. 2019

Abstract

Hepatitis C virus (HCV) infection is a serious global health problem and a cause of chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma (HCC). Bioinformatics software has been an effective tool to study the HCV genome as well as core domains. Our research was based on employing several bioinformatics software applications to find important mutations in domain 1 of core protein in Iranian HCV infected samples from 2006 to 2017, and an investigation of general properties, B-cell and T-cell epitopes, modification sites, and structure of domain 1. Domain 1 sequences of 188 HCV samples isolated from 2006 to 2017, Iran, were retrieved from NCBI gene bank. Using several tools, all sequences were analyzed for determination of mutations, physicochemical analysis, B-cell epitopes prediction, T-cell and CTL epitopes prediction, post modification, secondary and tertiary structure prediction. Our analysis determined several mutations in some special positions (70, 90, 91, and 110) that are associated with HCC and hepatocarcinogenesis, efficacy of triple therapy and sustained virological response, and interaction between core and CCR6. Several B-cell, T-cell, and CTL epitopes were recognized. Secondary and tertiary structures were mapped for domain 1 and core proteins. Our study, as a first report, offered inclusive data about frequent mutation in HCV-core gene domain 1 in Iranian sequences that can provide helpful analysis on structure and function of domain 1 of the core gene.

Keywords HCV · Core · Domain1 · Bioinformatics

Introduction

HCV infection is a serious global health problem and causes chronic hepatitis, liver cirrhosis, and HCC (Akuta et al. 2007; Ajourloo et al. 2015; Alborzi et al. 2017, 2015; Moayedi et al. 2018; Hashempour et al. 2018). It is estimated that 160 million people are infected worldwide (Lavanchy 2011). The prevalence rate of HCV infection is from 0.2 up to 40% in different countries, and this prevalence in Iran is 0.16% (Sefidi et al. 2013).

Lack of an effective vaccine and therapeutic choices has led to the rapid growth of HCV infection (Lauer and Walker 2001).

HCV has six large different genotypes (1–6) and 70 distinct subtypes (a, b, c, etc.) globally (Martro et al. 2011). Genotyping analysis showed 30–33% difference in each genotype, and in subtypes around 20–25% (Sefidi et al. 2013).

According to the current investigations in Iran, the predominant HCV subtype is 1a, followed by 3a and 1b (Sefidi et al. 2013).

HCV is a positive-strand RNA virus encoding three structural components, the core protein, and two E1 and E2 envelope glycoproteins (Ajourloo et al. 2015).

The core protein has many confirmed roles: core binds RNA and DNA and has an important function in RNA packing. It has been determined that HCV-core is a nucleic acid chaperone similar to retroviral nucleocapsid (NC) proteins in act acting to rearrange HCV 3'UTR, resulting in RNA dimerization in vitro (Caval et al. 2011; Cristofari G; Ivanyi-Nagy et al. 2004; Steinmann et al. 2008).

HCV-core is a highly basic protein that forms the viral NC and has interactions with cellular proteins and signal transduction pathways. As a result of HCV-core and host cell interactions, core may have a function in persistent infection and the pathogenesis of HCV liver disease (Polyak et al. 2006; Hashempour et al. 2015).

Core protein consisted of three predicted domains: aa1-aa117 domain 1 (domain D1), aa117-aa177 domain 2

✉ Tayebeh Hashempour
thashem@sums.ac.ir

¹ Shiraz HIV/AIDS Research Center, Institute of Health, Shiraz University of Medical Sciences, Shiraz 71937, Iran

(domain D2), and 177–191 domain 3 (domain D3) (Strosberg et al. 2010).

Domain 1 contains frequent positively charged amino acids, and is involved in RNA binding, promotes dimerization of the viral RNA, and has a significant role in NC formation and core envelopment by endosomal membranes (Ivanyi-Nagy et al. 2006).

Several identified mutations in domain 1 are involved in the development of HCC and hepatocarcinogenesis, the efficacy of triple therapy, and interaction between core and CXCL6 (Akuta et al. 2007, 2010, 2011; Fishman et al. 2009; Ogata et al. 2002; Takahashi et al. 2001; Idrees and Ashfaq 2013).

Humoral and cellular immune responses against HCV infections are inefficient and there is no convincing explanation to understand HCV immune pathogenesis (Gremion and Cerny 2005). However, HCV-specific IgM and IgG together were detected in acute infection. There are some outstanding proofs supporting a role for Abs in control of HCV infection and especially in reinfection (Cashman et al. 2014).

Some researchers have described a rise in anti-HCV humeral immune after immunization with core protein (Aghasadeghi et al. 2006). Other researchers have claimed that in acute HCV infection the titer of antibodies is very low, and delay in neutralizing antibody production is responsible for ineffective ability to prevent HCV infection (Netski et al. 2005). Other sources have introduced a large diversity of epitopes in HCV proteins as a possible way to escape the humoral response (Pavio and Lai 2003).

Cellular immune responses have an important role in the clearance of HCV infection. Patients with cellular immune dysfunction like human immunodeficiency virus (HIV) have rapid HCV progression. Some researchers have claimed that this system promotes liver injury by cytolysis activity of infected cells. In spite of cellular immune responses, HCV often evades recognition and has the ability to persist (Ward et al. 2002).

Several studies have described a number of pathways for HCV to escape from cell responses, including impaired oligo-/mono-specific or no virus-specific CD4⁺ and CD8⁺, mutation of epitopes, weakness of proliferative capacity, and cytotoxicity and ability to secrete TNF- α and IFN- γ by CD8⁺ T cells (Neumann-Haefelin et al. 2005). In addition, regulatory T cells (Tregs) induced by HCV infection have a significant role in the impaired activity of cellular immune response (Hashempour et al. 2015; Hashempour et al. 2010).

Bioinformatics tools are efficient means to study viruses and different parts of the HCV genome like core domains (Idrees and Ashfaq 2013; Moattari et al. 2015; Dehghani et al. 2017; Nezafat et al. 2018; Atapour et al. 2018; Sarvari et al. 2014; Behzad Dehghani and Zahra Hasanshahi 2019). Bioinformatics tools are efficient means to study viruses and different parts of the HCV genome like core domains. Many

programs have been developed to analyze function, structures, and modification of core protein, providing a large amount of information about core domains and important mutation sites (Akuta et al. 2007, 2010, 2011).

Current data are useful for prediction of HCV disease development and treatment response. In this study, we employed several bioinformatics tools to find important mutations in domain 1 of the core protein, general properties of B-cell and T-cell epitopes, modification sites, and structure of domain 1 in Iranian HCV infected samples from 2006 to 2017.

Materials and Methods

Sequence Alignment and Phylogenetic Tree

Domain 1 sequences of 188 Iranian HCV samples and reference sequences of HCV genotypes that were registered in NCBI gene bank (<http://www.ncbi.nlm.nih.gov/>) from 2006 to 2017 were downloaded. Homology among sequences was determined using multiple sequence alignment available in CLC- sequence viewer software under the following parameters: gap open cost, 10; gap extension cost, 1.0; and very accurate progressive alignment algorithm. Also, phylogenetic trees were analyzed through CLUSTAL X software, version 1.81, by neighbor-joining times to confirm the reliability of phylogenetic trees. The accession numbers of all sequences are displayed in Table 1.

Determination of Mutations

By considering previous studies, several significant mutations that are involved in: 1-hepatocellular carcinoma, 2-viral response to triple therapy, and 3-the interaction between core and CXCL6, were determined. All sequences were compared to find mentioned mutations (Akuta et al. 2007; Fishman et al. 2009; Ogata et al. 2002).

Physico-chemical Analysis

General properties of domain 1 (genotype 1a) were determined by employing “Expasy’sProtParam” (<http://expasy.org/tools/protparam.html>) and ProtScale at (<http://web.expasy.org/protscale/>) (Gasteiger et al. 2005).

B-Cell Epitopes Prediction

Chou and Fasman, Karplus and Schulz, Kolaskar & Tongaonkar, Emini, Parker, and BepiPred methods at <http://www.immuneepitope.org> (http://tools.immuneepitope.org/tools/bcell/iedb_input) were run for prediction of B-Cell epitopes positions (Chou and Fasman 2009; Karplus and

2/html) (Kelley and Sternberg 2009), (PS)²-v2 Server at (<http://ps2v2.life.nctu.edu.tw>) (Chen et al. 2006) were employed. To evaluate the stereochemistry and quality of 3D structures Qmean at (<http://swissmodel.expasy.org/qmean/cgi/index.cgi>) (Benkert et al. 2008) was used, and the Ramachandran plot was mapped by Rampage (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>).

The Signal Peptide Prediction

The Signal peptide was predicted by “Signal-BLAST”, and “SignalP 4.1 Server”.

Prediction of Epitopes Digestion: Peptide Cutter was Used to Determine Potential Cleavage Sites Cleaved by Proteases

Research Ethics

All data were collected anonymously in accordance with legal requirements regarding data protection and medical confidentiality. Approval from the Faculty Human Research Ethics Committee (Shiraz University of Medical Sciences) was obtained before the commencement of the study.

Results

Mutation and Phylogenetic Tree

By considering all submitted sequences in NCBI GenBank we could not find any sequences related to 2017.

Phylogenetic tree for all sequences was shown in Fig. 1. All 2006 sequences were placed in a cluster at the bottom of the tree, and a sequence of 2012 has a high similarity to KF218585.1 (2014). The majority of sequences were closer to 1a and 3a than other reference sequences.

All important mutation positions were listed in Table 2; the majority of mutations happened in 2013 and 2016 samples. No mutation was detected in 12, 23, 25, 39, 45 positions.

Protparam analysis

Protparam results for domain 1 are listed in Table 3. Because of the high percentage of basic amino acids, domain 1 is a highly basic peptide (Theoretical pI: 12). The instability index, an estimate of the stability of a protein in a test tube, showed that domain 1 is an unstable peptide. Aliphatic index, a positive factor for the increase of thermostability of proteins, indicated that this peptide is a thermostable peptide. GRAVY is a hydropathicity index and increasing

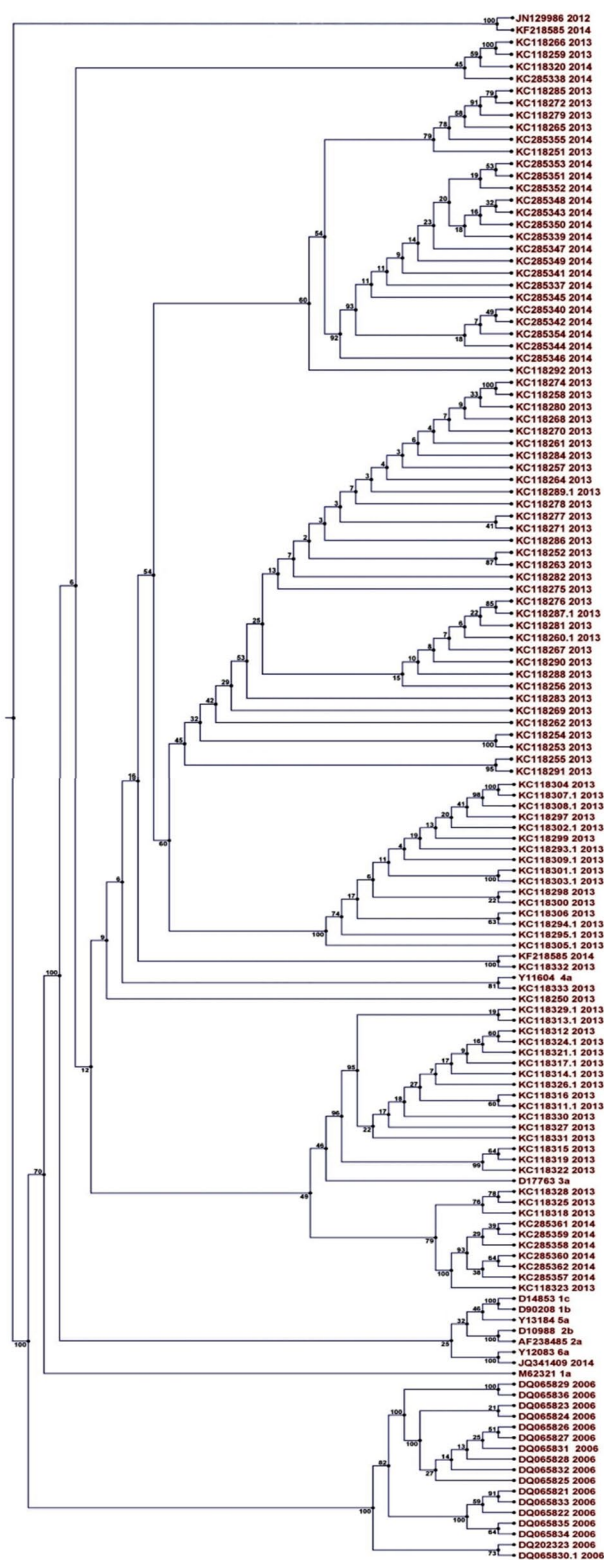


Fig. 1 Phylogenetic tree based on domain1 sequences and by using neighbor joining method. The phylogenetic tree was constructed by the NJ method. The numbers at the forks show the numbers of occurrences of the repetitive groups to the right out of 100 bootstrap samples. All used reference sequences were showed after accession numbers (1a, 1b, and etc.). Sequences were categorized in five major clusters

Table 2 Frequency of all identified mutations in HCV-core domain1 sequences

	2006	2012	2013	2014	2016	Ref
11(T)	–	–	1(P)1.2%	–	–	1a(N)
12(K)	–	–	–	–	–	–
23(K)	–	–	–	–	–	–
25(P)	–	–	–	–	–	–
39(R)	–	–	–	–	–	–
45(G)	–	–	–	–	–	–
49(T)	2(V)11.7%	–	2(P) 2.4%, 5(A) 6.1%, 1(V) 1.2%	1(V) 3.5%	1(p)1.6%	–
69(R)	–	–	–	–	–	–
70(R)	2(Q)11.7%	–	3(Q) 3.7%, 2(H) 2.4%, 1(P) 1.2%	4(Q)14.3%	25(Q)42.3%,1(H)1.6%	3a,6a (Q)
74(R)	–	–	–	–	–	–
78(Q)	–	–	–	1 (R) 3.5%	–	2a,2b (K)
90(G)	–	–	1(L) 1.2%, 1(S)1.2%	–	–	–
91(C)	2(M)11.7%	–	13(M)16.4%, 3(L) 3.7%, 1(F) 1.2%	–	43(M)72.8%, 7(L)11.8	1b (M); 2a and 5a (L)
102(G)	–	–	–	–	–	5a (S)
110	8(T)47%, 9(N)53%	–	23(N)28.4%, 5(S)6.1%, 53(T)65.4%	18(T)64.3%, 8(N)28.57%	N (10)17%	1b, 1a, 2b(T); 1c(S); 4a, 3a, 6a, 2a, 5a(N)

The majority of mutations happened in amino acid residues 49, 70, 91, 110

positive score indicates a greater hydrophobicity, so this peptide is a hydrophilic peptide.

ProtScale Analysis

Hydropathicity analysis by Kyte J. and Doolittle R.F. method showed that the major part of the peptide had a negative score; the maximum hydropath city score was on aa34 (valin) and the minimum hydropath city score was on aa 14 (asparagine).

Amino acids flexibility predicted by Bhaskaran R Pon-swamy P.K method indicated that the maximum flexibility was around amino acid 58 (proline) and the minimum was around aa 95 (glycine).

Transmembrane (TM) tendency calculated by Zhao, G., London E. method, showed that the major part of peptide had a negative score, and the maximum transmembrane tendency was on aa34 (valine) and the minimum was on aa10 (lysine).

Peptide polarity predicted by Grantham R method showed that the maximum polarity was on amino acid 12 and 13 (lysine and arginine) and the minimum polarity was on aa 34 (valine).

B-Cell Epitopes Prediction

<http://www.immuneepitope.org> online software:

Chou and Fasman Beta-Turn Prediction, which is based on the rationale for predicting turns to predict antibody epitopes, showed one high score region, 99–112.

Emini Surface Accessibility Prediction, which is based on surface accessibility scale, showed two high score positions (4–20 and 49–61).

Karplus and Schulz flexibility scale was used for B-cell prediction; this method is based on mobility of protein segments on the basis of the known temperature B factors of the α -carbons of 31 proteins of known structure. Results demonstrated two positions with the highest score (50–60 and 5–12).

Kolaskar & Tongaonkar Antigenicity method is based on physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes to predict antigenic determinants on protein. Results showed five positions of 20–39, 43–49, 63–69, 78–85, and 93–101.

Parker Hydrophilicity Prediction method is based on peptide retention times during high-performance liquid chromatography (HPLC) on a reversed-phase column. By this method, two regions were found: 9–16, 51–57.

Linear B-cell epitopes were determined using BepiPred. This method is based on a combination of a hidden Markov model and a propensity scale method. Three regions (1–25, 51–84, and 102–116) were founded by BepiPred analysis.

Bcepred Results

For a combination of all physicochemical properties (hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surface, and turns) for linear B-cell epitope prediction based on physicochemical properties on a non-redundant dataset:

Table 3 Domain1 physicochemical properties computed by “Protparam”

Amino acid composition	Number	Percentage (%)
Number of amino acids: 117		
Molecular weight: 13320.2		
Theoretical pI: 12.05		
Ala (A)	4	3.40
Arg (R)	21	17.90
Asn (N)	5	4.30
Asp (D)	2	1.70
Cys (C)	1	0.90
Gln (Q)	6	5.10
Glu (E)	3	2.60
Gly (G)	16	13.70
Ile (I)	2	1.70
Leu (L)	6	5.10
Lys (K)	7	6.00
Met (M)	1	0.90
Phe (F)	1	0.90
Pro (P)	17	14.50
Ser (S)	7	6.00
Thr (T)	6	5.10
Trp (W)	5	4.30
Tyr (Y)	3	2.60
Val(V)4		3.40
Total number of negatively charged residues (Asp + Glu)	5	
Total number of positively charged residues (Arg + Lys)	28	
The estimated half-life is		
Mammalian reticulocytes, in vitro	30 h	
Yeast, in vivo	> 20 h	
<i>Escherichia coli</i> , in vivo	> 10 h	
The instability index (II)	77.69 unstable	
Aliphatic index	40	
Grand average of hydropathicity (GRAVY)	-1.424	

Table 4 16 meric conserved B-cell epitopes regions in HCV-core domain1, predicted by ABCpred online software

Rank	Sequence	Position	Score
1	QPGYPWPLYGNEGCGW	78–93	0.92
2	TRKTSERSQPRGRRQP	49–64	0.9
3	MSTNPKPQKKNKRNTN	1–16	0.89
4	PIPKARRPEGRTWAQP	64–79	0.86
4	GSRPSWGPTDPRRRSR	102–117	0.86

Using bcepred online software, three regions (4–22, 47–59, and 109–115) with the highest combined score were found.

Five 16 meric conserved regions (78, 49, 1, 64 and 102) were found by ABCpred prediction Server (Table 4).

VaxiJen Prediction

According to a predefined cutoff of VaxiJen program, domain 1 was confirmed as a probable antigen (model: virus and threshold: 0.4).

IgE Epitopes

The prediction of allergenic proteins by mapping of IgE epitope, SVM, and hybrid methods showed that domain 1 was not an allergen protein.

T Cell-Epitopes

Regarding T-cell responses against HCV, previous researches found some hosts' human leukocyte antigen

Table 5 HLA predicted epitopes in HCV-core domain1 sequence (genotype 1a) for HLA's that were determined by previous researches in Iranian patients

HLA types	Epitopes positions
HLA-A1	1_10
HLA-A2	29–44, 77–85
HLA-A3	2–10,29–59, 96–104
HLA-B*3501	6–14, 27–35, 41–49, 57–65,78–91, 99–116
HLA-B*3801	29–37, 77–85, 89–97
HLA-Cw*0401	23–37, 58–65, 78–93
HLA-Cw*0702	27–35, 73–86
HLA-DQA1*05:01	2–16,13–27
HLA-DQA1*0201	NO
HLA-DQB1*0602	2–16,17–31, 102–116
HLA-DQB1*0301	96–104
HLA-G: HLA-G*01:01, HLA-G*01:02	4–16,8–21, 22–33, 66–79, 100–112
HLA-DRB1*0301	96–104
HLA-DRB1*0305	96–104
HLA-DRB1*0309	96–104
HLA-DRB1*0701	No
HLA-DRB1*11	34–42

Table 6 CTL epitopes in HCV-core domain 1; the high score epitopes are displayed

Peptide Rank	Position	Sequence	Score	Prediction
1	48–56	ATRKTSERS	1	Epitope
2	80–88	GYPWPLYGN	1	Epitope
3	2–10	STNPKPQKK	0.99	Epitope

(HLA) alleles associations with HCV infection in Iranian patients. We found several epitopes of HLA's shown in Table 5.

Some studies found both CD4 helper and CD8 CTL responses against HCV infection. ctpred found several epitopes for CTL (Table 6).

Postmodification

Prediction of serine, threonine, and tyrosine phosphorylation sites by DISPHOS showed one position (116) in domain1.

By “NetPhos” software we found 10 phosphorylation sites (Fig. 2), 6 sites for serine (53, 56, 99, 103, 106, and 116) 3 sites for threonine (15, 49, and 52) and one site for tyrosine (86).

NetPhosK results determined four phosphorylation sites, three threonine amino acids (3, 15, and 49) for protein kinase C and one serine (116) for protein kinase A. No glycosylation site was found by NetNGlyc and GlycoEP.

Secondary and Tertiary Structure Prediction

Secondary structure prediction for core and domain1 by using SOMPA software was summarized in Table 7 and Fig. 3.

SOMPA showed there was no alpha helix structure in domain1 and the major part of it was the random coil. But the combination of (PS)²-v2 and PHYRE² showed there was an alpha helix structure in 8–15 region (Figs. 4, 5). All programs displayed extended strand in the 29–36 region.

3D structures were determined by all three online software but only structures predicted by I-TASSER were

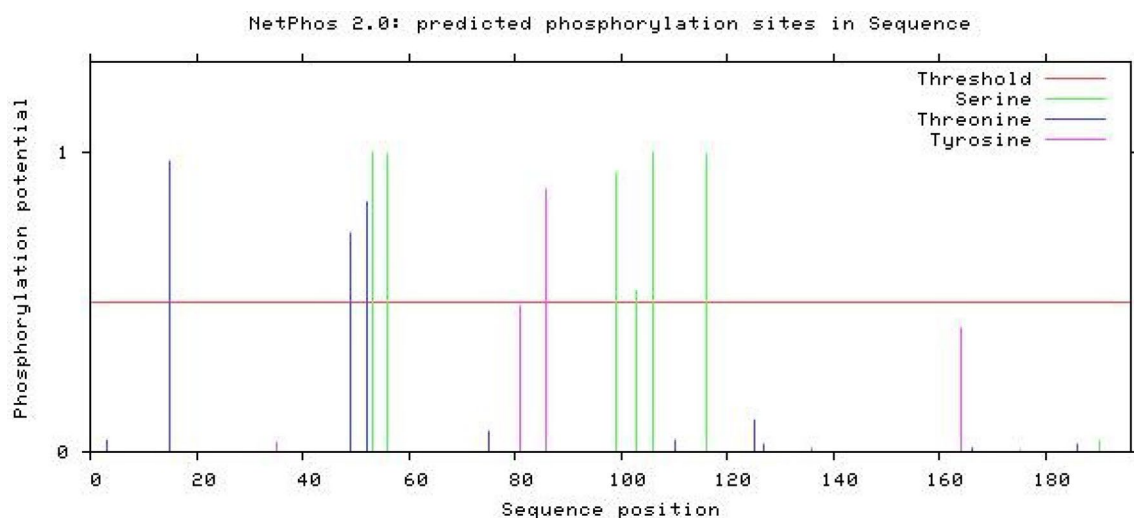
**Fig. 2** Phosphorylation sites prediction for domain 1 using “NetPhos” online software. Green lines indicate 6 sites for serine, blues lines show 3 sites for threonine, and one purple line shows tyrosine. All sites with scores above the threshold of 0.5 were considered as phosphorylation sites

Table 7 Percentage of secondary structures in core and domain1

	Alpha helix (Hh)	Extended strand (Ee)	Beta turn (Tt)	Random coil (Cc)
Core	39, 20.42%	25, 13.09%	14, 7.33%	113, 59.16%
Domain 1	NO	19, 16.24%	10, 8.55%	88, 75.21%

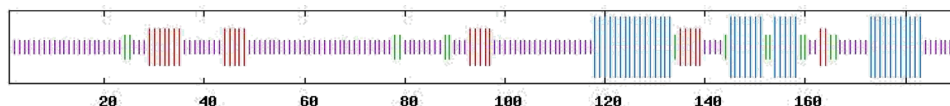


Fig. 3 Secondary structure prediction using SOMPA. Red region is extended strand, blue is the alpha helix, green is beta turn, and purple is the random coil. The majority of core structure belongs to random coil

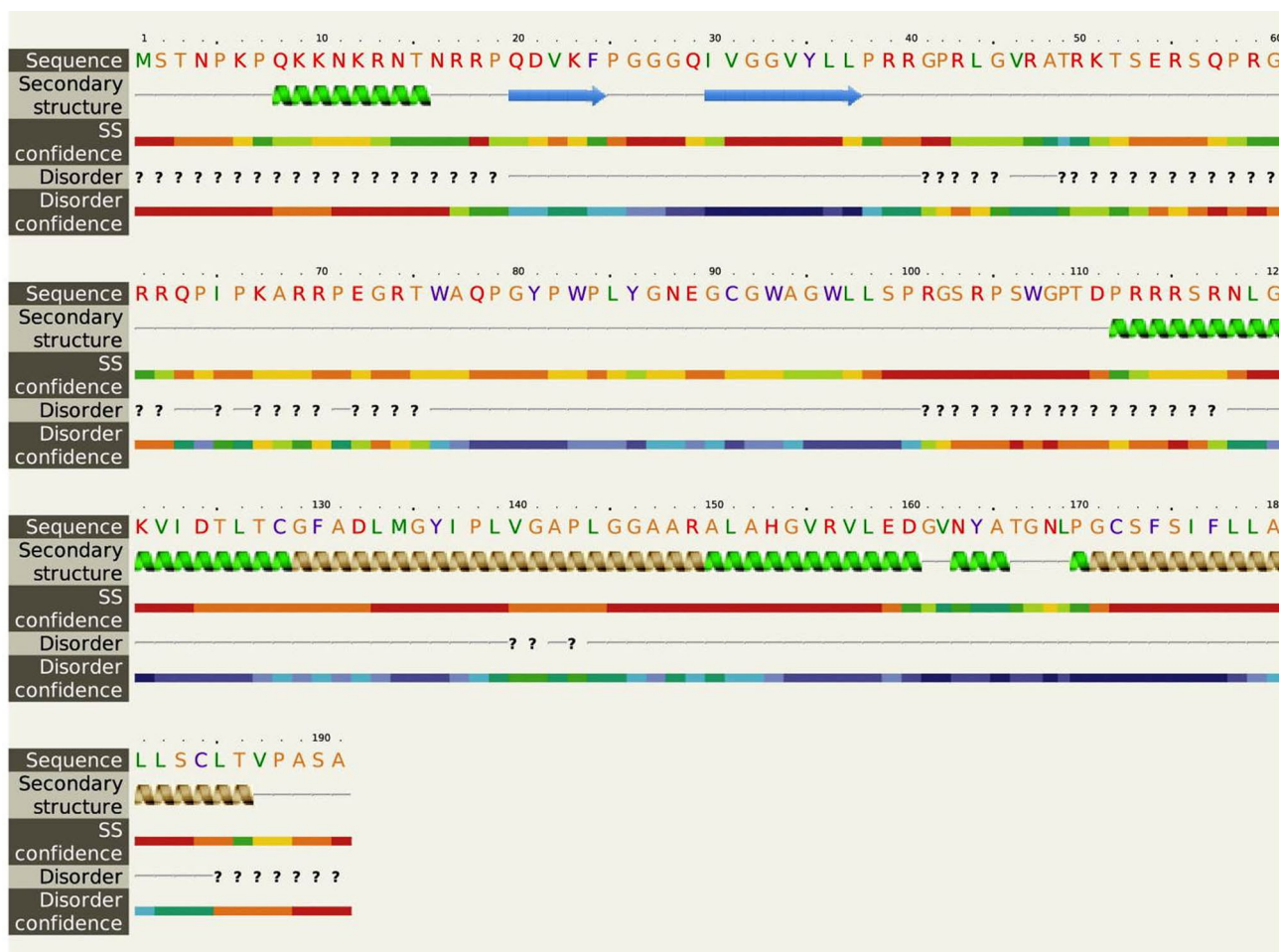


Fig. 4 Secondary structure prediction using PHYRE2. The result of this tool shows that the majority of the core structure (40%) is alpha helix which is indicated with green helix, also the confidence keys of the predicted structure for these regions are high

reliable. Final structures (Figs. 6, 7) were validated by Qmean. QMEANscore and Z-score for calculated for core were 0.242 and -5.43 . The scores were not satisfactory but at least provided an overview of the core protein structure. QMEANscore and Z-score for domain 1 were 0.61

and -1.24 confirming the quality and reliability of the predicted structure.

Ramachandran plot was assessed by RAMPAGE, and percentages of the favoured region, allowed region, and outlier region for core were 63.0%, 27.5%, and 9.5% respectively

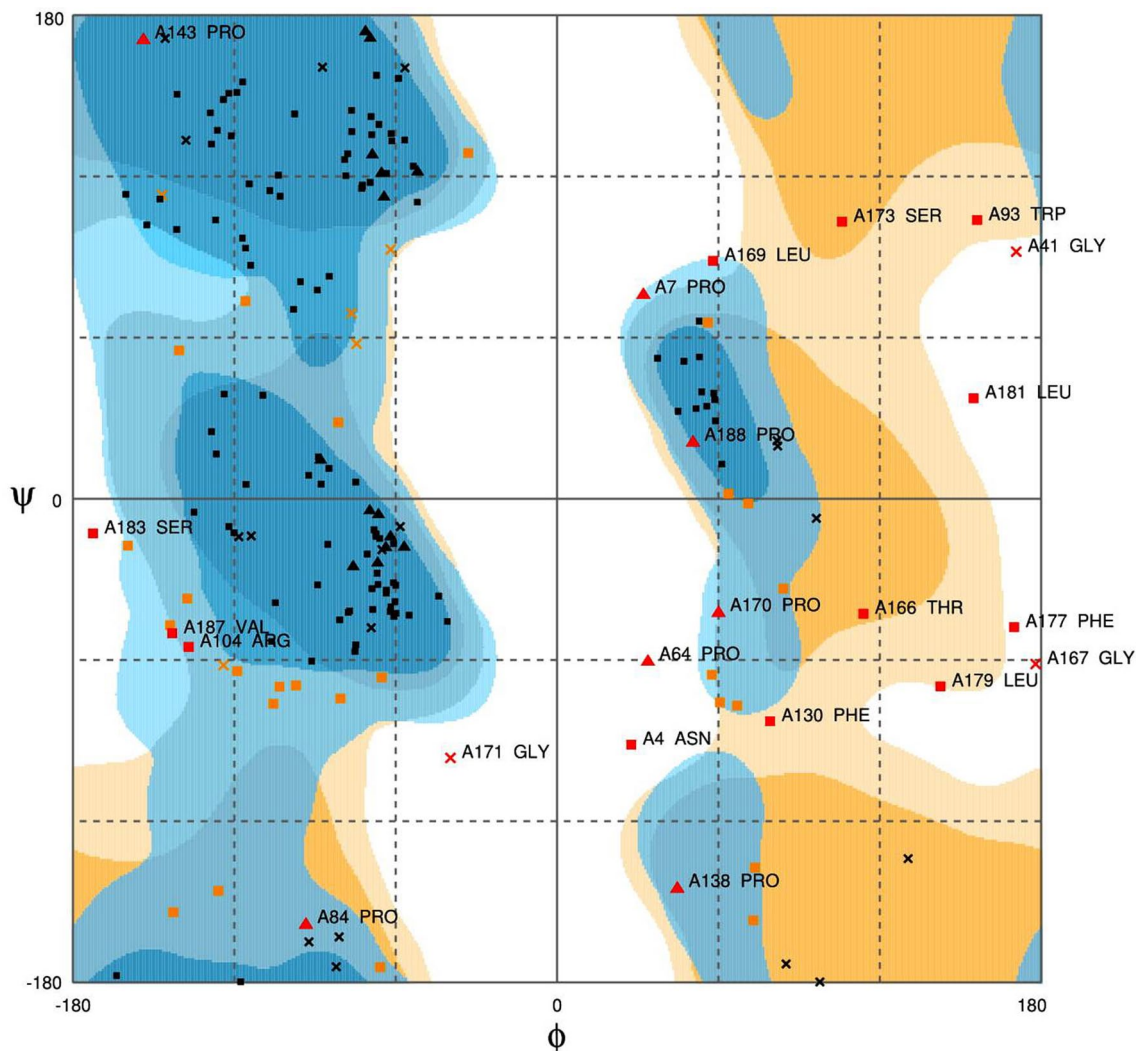


Fig. 8 Ramachandran plot was used to visualize energetically allowed regions for backbone dihedral angles ψ against ϕ of amino acid residues in modeled protein structure (LCC model) for tertiary structure

of core protein by RAMPAGE; the majority of amino acids residues were in favored region (119 amino acids) and allowed region (52 amino acids)

In selected sequences, we found that amino acid 11 was T in all sequences except one sequence in 2013 (P) and in 1a (N).

Positions 12(K) did not show any change; amino acid 25 was P and amino acid 69 was R in all sequences. Amino acid 70 in nearly all sequences was R except in 2 sequences in 2006 (Q), 5 sequences in 2013 (3(Q), 2(H), and 1(P)), 4 sequences in 2014 (Q), in 2016 [25(Q), 1(H) and 2 reference sequences (Q)].

Amino acid 91 in most sequences was C except 2 sequences in 2006(M), 17 sequences in 2013 (13(M), 3(L), 1(F)), 50 sequences in 2016(43(M), 7(L)) and three ref sequences, (1b(M), 2a and 5a (L)). In amino acid 102 we did not find any change but in one of the reference sequences, we found one change (5a (G to S)).

Akuta et al. (2007) employed PCR for detecting substitutions of aa 70 and aa 91 in HCV-core gene of genotype 1b by using the mutation-specific primer as an important predictor of hepatocarcinogenesis. For wild samples, aa 70 was arginine (R) and aa 91 was leucine (L) but for mutant aa 70 was glutamine(Q)/histidine (H) and aa 91 was methionine (M) (Akuta et al. 2007).

Also, Furui et al. (2011) identified aa 70 and aa 91 substitutions among Japanese volunteer blood donors (Furui et al. 2011).

In terms of aa 70 substitutions, we recognized 2 glutamine substitutions in 2006 sequences and 3 glutamine, 2 histidines, one proline in 2013 sequences; also 4 glutamine substitutions in 2014 sequences, and 25 glutamine, and 1

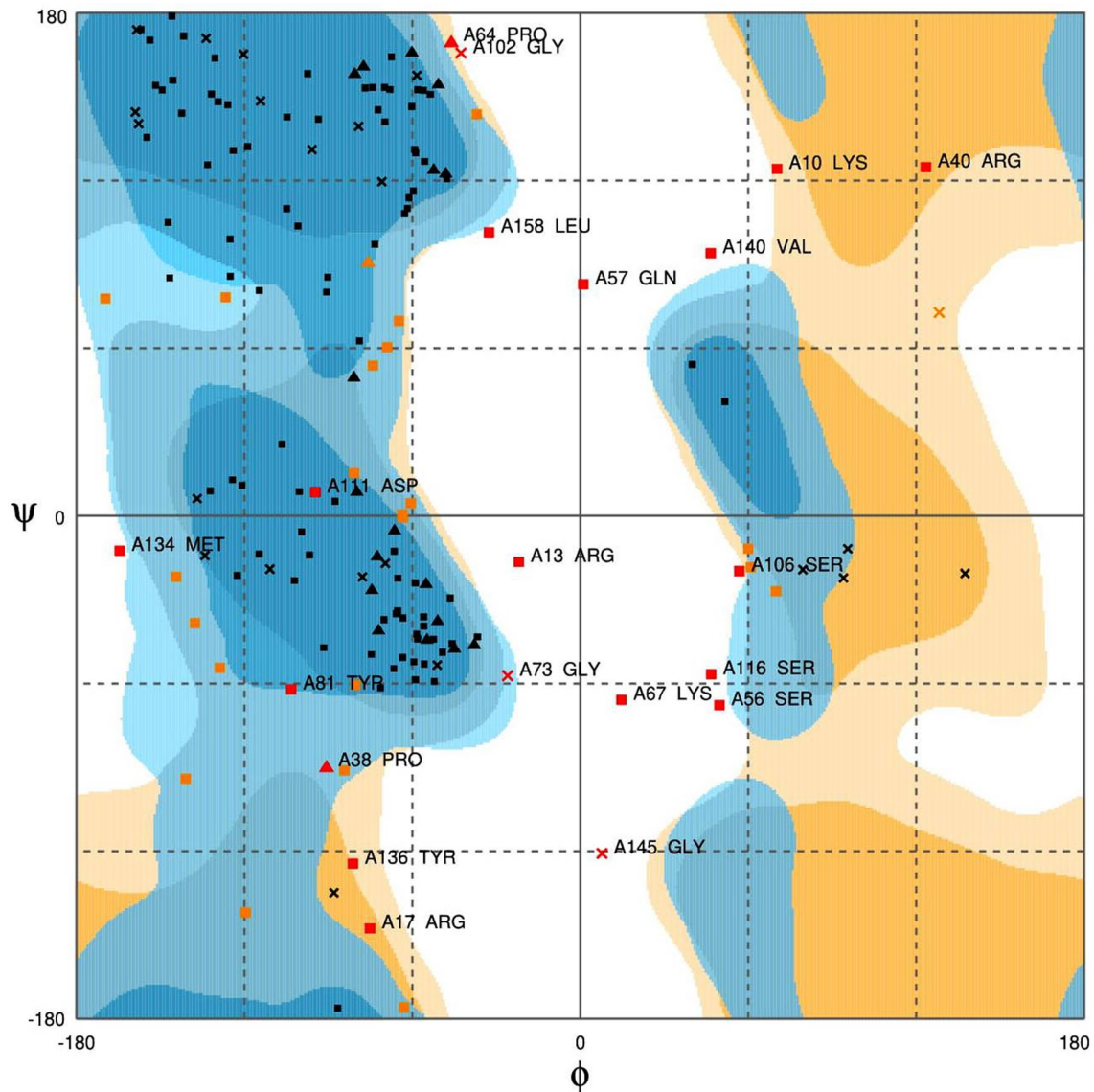


Fig. 9 Ramachandran plot was used to visualize energetically allowed regions for backbone dihedral angles ψ against ϕ of amino acid residues in modeled protein structure (LCC model) for tertiary structure

of domain1 by RAMPAGE; majority of amino acids residues were in favored region (71 amino acids) and allowed region (28 amino acids)

histidine in 2016. We found several methionine residues in aa 91 in 2006, 2014, and 2016.

Ogata et al. (2002) compared sequences of the core protein of Subtype 1b HCV strains obtained from patients with and without HCC and found some amino acid mutation sites (Ogata et al. 2002). K23Q, Q70R, and T110M substitutions were found by Ogata et al. (2002). In comparison with our results, in all sequences, aa 23 was K, in the majority of sequences aa 70 was R and in 34 sequences it was Q. In aa 110 we did not find any methionine (Ogata et al. 2002).

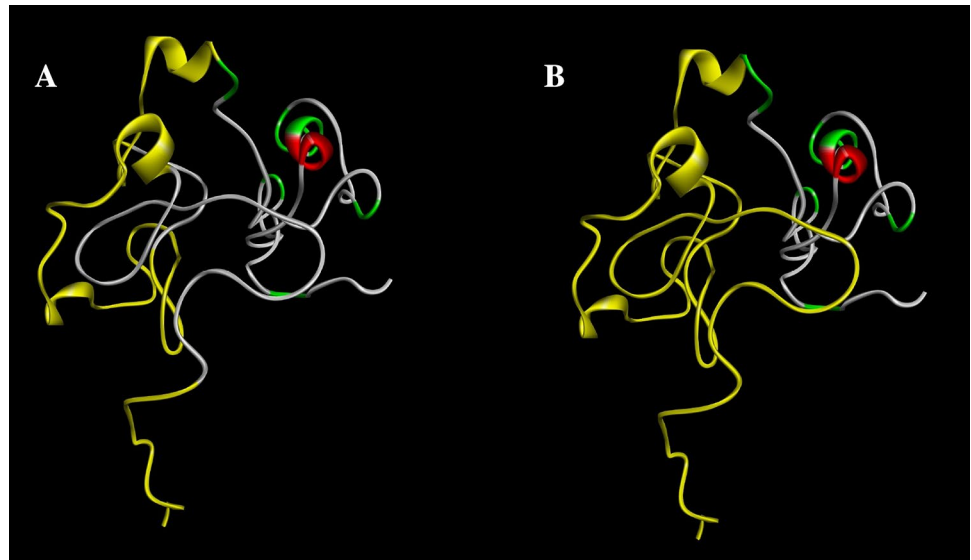
Akuta et al. (2010) confirmed the role of Gln70 (or His70) in the efficacy of triple therapy and sustained a virological response, the patient with both genotype non-TT and

Gln70 (His70) had the worst sustained virological response. Also, Akuta et al. (2011) by following up twenty-six patients determined the role of Gln70 (His70) substitution in the development of HCC. They suggested detection of aa substitutions in the core region before antiviral therapy (Akuta et al. 2010, 2011).

Alestig et al. (2011) showed substitution in aa 70 of the core was related to treatment response, but that was less important than IL28B polymorphism. In our study, 37 sequences had Gln70 (or His70) substitution (Alestig et al. 2011).

Tokita et al. (2000) approved the role of HCV-core region (Thr49Pro) to reduce the fluorescence enzyme immunoassay

Fig. 10 A: the position of the B-Cell epitopes (yellow region) on core tertiary structure and B: the T-cell epitopes (yellow region) on core 3D structure



(FEIA) sensitivity. We found 2 sequences in 2013 and one in 2016 with T to P substitutions (Tokita et al. 2000).

Findings of Horie et al. (1999) indicated that alteration from glycine to serine at core codon 45 was dominant in noncancerous liver portions rather than in cancerous liver portions and sera from HCC patients (Horie et al. 1999). Our results did not show any glycine to serine mutation in aa 45 and in all sequences, it was glycine.

Idrees and Ashfaq (2013) by using molecular docking software, reported interactions of amino acid residues arginine 149, arginine 39, arginine 74, and arginine 78 in HCV-core protein and Leu44, Ala71, Ser76 and Pro97 in CXCL6. This finding clues to understanding HCV pathogenesis. Any change in these positions can relate to HCV Infection and HCC. Our results showed no alteration in aa 39 and 74, and in aa 78 just one sequence was glutamine to arginine (Idrees and Ashfaq 2013).

Using a combination of predicted B-cell antibody epitopes by all methods on the immune epitope website and also considering bcepred and ABCpred prediction, we could define three major epitopes (4–20, 50–60, and 100–112) for domain 1.

Ferroni et al. (1993) by using the algorithm of Jameson and Wolf identified four epitopes in HCV-core protein (7–21, 31–45, 49–63, and 99–113).

Harase et al. (1995) analyzed the response to HCV-core protein in mice and found a major B cell epitope (21–40).

Pirisi et al. (1995) analyzed Sera from 97 HCV infected patients and found three 15-mer peptides as antigens in an enzyme immunoassay. They concluded that anti-R15P (50–64, RKTSEERSQPRGRRQP) as a potent antigen might help to identify a subgroup at higher risk to develop HCC (Pirisi et al. 1995).

Also, a study on HCV positive blood donor by Lechmann et al. (1996) determined a region (aa 1–24) of domain I that bound the antibodies from the sera of all patients and showed a great potential for detection of HCV infection by using serological B-Cell responses tests.

Comparison of our results with previous studies revealed that all predicted epitopes in our research have a good potential for future studies of the immune response against HCV infection, and are useful for recognition of all kinds of HCV infections.

Gededzha et al. (2014) used several bioinformatics tools to predict HLA class I and HLA class II in HCV genotype 5a. They found three T-cell epitopes of NS3, NS4B, and NS5B.

Some HLA class II alleles were found in Iranian patients by Samimi-Rad et al. (2015); DRB1*0301, DQA1*0501, DQB1*0201, DRB1*1101, and DQB1*0301 were demonstrated in patients with HCV clearance, and DRB1*0701, DQA1*0201, DQB1*0602, DRB1*0301, DRB1*11, and DQB1*0201 occurred more frequently in chronic patients (Samimi-Rad et al. 2015).

Khorrani et al. (2015) found a relationship between HLA-G, IL-10, and response to combined therapy in HCV positive patients. They concluded that HLA-G, and IL-10 have a significant role in response to therapy with IFN- α 2 α and ribavirin.

Also, HLA-A01 and HLA-B38 were determined as important alleles associated with Peg-IFN plus ribavirin therapy in Egyptian patients by Farag et al. (2013).

Pourhassan et al. (2014) found several HLA alleles associated with HCV in Iranian patients (2011–2013). A2, A3, B35, B38, BW4, CW4 and CW7 were the most frequent alleles found by this group.

Table 8 The results of predicted cleavage positions for 12 common proteases: B-cell, T-cell, and CTL predicted epitopes

Positions	Caspase	Chymotrypsin	Clostripain	Elastase	Pepsin	Proteinase K	Asp-N	Staphylococcal peptidase I	Thrombin	Proline	Trypsin
B-cell epitopes											
78–93	3(8,9,16)		5(2, 7, 11, 13, 14)		5(4, 7, 9,15, 16)	6(4, 6, 8, 9, 12, 16)	1(11)	1(12)			6(2, 3, 7, 11, 13, 14)
49–64			1(13)			3(1, 4, 6)	1(5)	1(6)			4(9, 10, 12, 13)
1_16	1(1)		3(6, 7, 11)	2(5,14)	1(12)	2(3,15)	1(8)	1(9)	1(7)	1(8)	3(4, 6, 11)
64–79	1(13)					6(2, 5, 9, 12, 13, 14)					
102–117	1(6)		5(3, 12, 13, 14, 16)			2(6,9)	1(9)		1(4)		3(12, 14, 16)
CTL epitopes											
48–56			2(3,8)	1(1)		4(1, 2, 5, 7)	1(6)	1(7)			3(3, 4, 8)
80–88	2(6,7)				3(2, 5, 7)	4(2, 4, 6, 7)					
2_10						1(2)			1(6)		2(8,9)
T-cell epitopes											
HLA-A1											
1_10	1(1)								1(7)		2(9,10)
HLA-A2											
29–44	3(6, 7, 15)	3(10, 11, 14)		2(2, 5)	4(5, 6, 8, 15)	7(1, 2, 5, 6, 7, 8, 15)			1(10)		3(10, 11, 14)
HLA-A2											
77–85	1(9)			1(1)	2(5,8)	4(1, 5, 7, 9)					
HLA-A3											
2_10											
HLA-A3											
29–59	1(25)	3(7,8,16)	7(11, 12, 15, 19, 22, 27, 31)	4(3, 6, 18, 20)	4(6 7 9 16)	12(2, 3, 6, 7, 8, 9, 16, 18, 20, 21, 24, 26,)		1(26)	1(11)		8(11 12 15 19 22 23 27 31)
HLA-A3											
96–104											
HLA-B*3501											
6_14			1(8)						1(2)		2(8,9)
HLA-B*3501											
27_35	1(9)			2(5,8)	2(8,9)	4(4, 5, 8, 9)					
HLA-B*3501											
41–49	1(4)		2(3,7)	2(6,8)	1(4)	4(4 6 8 9)					2(3,7)
HLA-B*3501											
57–65			3(3 5 6)								4(3, 5, 6, 11)

Table 8 (continued)

Positions	Caspase	Chymotrypsin	Clostripain	Elastase	Pepsin	Proteinase K	Asp-N	Staphylococcal peptidase I	Thrombin	Proline	Trypsin
HLA-B*3501											
78–91	2(8,9)				3(4, 7, 9)	5(4, 6, 8, 9, 12)	1(11)	1(12)			
HLA-B*3501											
99–116	1(9)	5(3, 6, 15, 16, 17)				2(9, 12)	1(12)		1(7)	3(3, 15, 17)	
HLA-B*3801											
29–37	3(7, 8, 9)	2(3, 6)		4(6, 7, 8, 9)		6(2, 3, 6, 7, 8, 9)					
HLA-B*3801											
77–85	1(9)			1(1)	2(5,8)	4(1, 5, 7, 9)					
HLA-B*3801											
89–97	3(5, 8, 9)			1(6)	5(4, 5, 7, 8, 9)	5(1, 5, 6, 8, 9)		1(1)			
HLA-Cw*0401											
23–37	3(7, 8, 9)			2(3, 6)	4(6, 7, 8, 9)	6(2, 3, 6, 7, 8, 9)					
HLA-Cw*0401											
58–65			3(2,4,5)								3(2,4,5)
HLA-Cw*0401											
78–93	3(8, 9, 16)				5(4, 7, 9, 15, 16)	6(4, 6, 8, 9, 12, 16)	1(11)	1(12)			
HLA-Cw*0702											
27–35	1(9)			2(5, 8)	2(8,9)	4(4, 5, 8, 9)					
HLA-Cw*0702											
73–86	3(4, 13, 14)	1(2)		1(5)	4(3, 9, 12, 14)	7(3, 4, 5, 9, 11, 13, 14)					1(2)
HLA-DQA1*05:01											
2_16		1(12)				2(2,14)			1(6)	4(8, 9, 11, 12)	
HLA-DQA1*05:01											
13–27		3(1,5,6)		1(10)	1(12)	3(3,10,12)	1(8)		1(7)	3(1,5,11)	
HLA-DQB1*0602											
2_16		1(12)				2(2,14)			1(6)	4(8, 9, 11, 12)	
HLA-DQB1*0602											
17–31								1(4)			
HLA-DQB1*0602											
102–116		2(1,2)		2(6,15)	1(8)	4(6, 8, 14, 15)			1(3)	2(1,7)	
HLA-DQB1*0301		3(1, 2, 3)	2(6, 9)	2(1, 2)		3(1, 2, 3)				2(6, 9)	
HLA-G: HLA-G*01:01, HLA-G*01:02											
4_16		1(10)				1(12)			1(4)	4(6, 7, 9, 10)	

Table 8 (continued)

Positions	Caspase	Chymotrypsin	Clostripain	Elastase	Pepsin	Proteinase K	Asp-N	Staphylococcal peptidase I	Thrombin	Proline	Trypsin
8_21			3(6,10,11)			1(8)	1(13)		1(12)		5(2, 3, 5, 6, 10)
22–33				2(1,10)	1(3)	4(1,3,9,10)					1(2)
66–79	1(11)		3(4,5,9)	2(3,12)	1(10)	5(3, 7, 10, 11, 12)	1(6)	1(7)	1(6)		3(2,4,9)
100–112	1(8)		2(2,5)			2(8,11)	1(11)		1(6)		1(2)
HLA-DRB1*0301											
96–104		3(1,2,3)	2(6,9)		2(1,2)	3(1,2,3)					2(6,9)
HLA-DRB1*0305											
96–104		3(1,2,3)	2(6,9)		2(1,2)	3(1,2,3)					2(6,9)
HLA-DRB1*0309											
96–104		3(1,2,3)	2(6,9)		2(1,2)	3(1,2,3)					2(6,9)
HLA-DRB1*11											
34–42	2(3,4)		2(7,8)	1(2)	3(2,3,5)	4(2,3,4,5)			1(7)		2(7,8)

The position of each epitope, the number of cleave sites in each epitope as well as the position of the cleave sites in each epitope are mentioned in the table

In accordance with the above-mentioned studies, we collected HLA alleles associated with HCV infection, and by employing in-silico analysis we established numerous T-cell epitopes for domain1 that can be helpful for future studies to design effective vaccine against HCV genotype 1a, and can provide benefit data for better understanding the role of domain1 in immune response.

Many researchers have proved broader CTL responses to HCV infection and the usefulness of CTL epitopes mapping to develop therapeutic interventions or vaccines (Sabet et al. 2014; Saeedi et al. 2014; Jazayeri and Carman 2005; Arashkia et al. 2011). In our research, we utilized reliable software to predict CTL epitopes and extracted data that can be useful for vaccine development studies.

By considering results of phosphorylation sites prediction by 3 programs, we concluded that 3 sites (15, 49, and 116) were the main phosphorylation sites in domain 1. Amino acid 116 is a serine that located in Arg-Arg-Arg-Ser-Arg region; this region was similar to the usual target sequence for protein kinase A [Arg-Arg-X-Ser/Thr-X]. Two threonine amino acid residues (15, 49) were calculated as protein kinase c target sites where this kinase acts through the phosphorylation of hydroxyl groups of amino acid residue. Previous studies indicated that core protein is phosphorylated by PKA and PKC.

Core phosphorylation regulates the suppressive activity of HCV-core protein on HBV gene replication and expression. Also, it has been shown that phosphorylation in core relates to the nuclear localization of the core protein. They demonstrated three serine residues (Ser-53, Ser-116, and Ser-99) as the potential phosphorylated sites in core protein, that were similar to NetPhos software results in our research (Yassin 2001; Shih et al. 1995; Lu and Ou 2002).

Secondary structure prediction indicated that the majority of domain1 was the random coil, and all B-cell epitopes and important mutations placed on random coil structure.

Tertiary structures were designed by three significant and reliable online programs, but just one of them provided a reliable and high-quality protein structure model for domain1. The quality and reliability of models were confirmed by QMEAN and RAMPAGE software.

By examining all the predictions core epitopes with and without signal peptide, we found out that there is no difference between these two different strategies of analysis (Pene et al. 2009; Targett-Adams et al. 2008; Ma et al. 2007; Okamoto et al. 2008; Oehler et al. 2012).

Based on previous studies core protein has a C terminal signal peptide (170–191) and because the focus of our study was on domain 1, it was expected that the deletion of this region could not affect the epitope prediction results.

Digestion analysis to predict possible proteases was shown that each epitope can be digested by at least 5

selected proteases which can have a significant effect on the reduction of the half-life of epitopes.

Conclusion

Finally our investigations in this research provided comprehensive data about frequent mutations in domain1, and as a first report can be useful for future study about significant mutations and their role in therapeutic pathway and response to antiviral therapy for Iranian patients. Also, identification of domain1 properties provides practical information for domain1 cloning and more researches.

Acknowledgements The authors would like to acknowledge Shiraz University of Medical Sciences for financial support (Grant Number: 95-01-59-13086) and Jean Krugman, retired professor of General Studies, DeVry University, Decatur, Georgia, USA for help in revising the English of the manuscript.

Funding This study was funded by Shiraz University of Medical Sciences (Grant Number: 95-01-59-13086).

Compliance with Ethical Standards

Conflict of interest Authors declare there is no conflict of interest.

Research Involving Human and/or Animal Participants This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Aghasadeghi M, Sadat S, Budkowska A, Khabiri A, Amini S, Bahramali G et al (2006) Evaluation of a native preparation of HCV core protein (2-122) for potential applications in immunization, diagnosis and mAb production. *Iran J Public Health* 35(1):1–10
- Ajorloo M, Bamdad T, Hashempour T, Alborzi AM, Mozghani SHR, Asadi R et al (2015) Detection of specific antibodies to HCV-ARF/CORE + 1 protein in cirrhotic and non-cirrhotic patients with hepatitis C: a possible association with progressive fibrosis. *Arch Iran Med AIM* 18(5):304–307
- Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y et al (2007) Amino acid substitutions in the hepatitis C virus core region are the important predictor of hepatocarcinogenesis. *Hepatology* 46(5):1357–1364
- Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, Sezaki H et al (2010) Amino acid substitution in hepatitis C virus core region and genetic variation near the interleukin 28B gene predict viral response to telaprevir with peginterferon and ribavirin. *Hepatology* 52(2):421–429
- Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Sezaki H, Suzuki Y et al (2011) Amino acid substitutions in hepatitis C virus core region predict hepatocarcinogenesis following eradication of HCV RNA by antiviral therapy. *J Med Virol* 83(6):1016–1022
- Alborzi AM, Bamdad T, Davoodian P, Hashempour T, Nejatizadeh AA, Moayedi J (2015) Insights into the role of HCV Plus-/Minus strand RNA, IFN- γ and IL-29 in relapse outcome in patients infected with HCV. *Asian Pac J Allergy Immunol* 33(3):173–181
- Alborzi A, Hashempour T, Moayedi J, Musavi Z, Pouladfar G, Merat S (2017) Role of serum level and genetic variation of IL-28B in interferon responsiveness and advanced liver disease in chronic hepatitis C patients. *Med Microbiol Immunol* 206(2):165–174
- Alestig E, Arnholm B, Eilard A, Lagging M, Nilsson S, Norkrans G et al (2011) Core mutations, IL28B polymorphisms and response to peginterferon/ribavirin treatment in Swedish patients with hepatitis C virus genotype 1 infection. *BMC Infect Dis* 11(1):124
- Arashkia A, Rouhvand F, Memarnejadian A, Alizadeh S, Motevalli F, Ebrahimi M (2011) Immunoinformatics modeling, construction of DNA plasmids Carrying CTL epitopes of hepatitis C virus and their preliminary immunological analysis. *Iran J Med Microbiol* 4(4):30–40
- Atapour A, Mokarram P, MostafaviPour Z, Hosseini SY, Ghasemi Y, Mohammadi S et al (2018) Designing a fusion protein vaccine against HCV: an in silico approach. *Int J Peptide Res Ther*. <https://doi.org/10.1007/s10989-018-9735-4>
- Benkert P, Tosatto SC, Schomburg D (2008) QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* 71(1):261–277
- Bhasin M, Raghava G (2004) Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine* 22(23–24):3195–3204
- Blom N, Gammeltoft S, Brunak S (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites I. *J Mol Biol* 294(5):1351–1362
- Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4(6):1633–1649
- Cashman SB, Marsden BD, Dustin LB (2014) The humoral immune response to HCV: understanding is key to vaccine development. *Front Immunol* 5:550
- Caval V, Piver E, Ivanyi-Nagy R, Darlix J-L, Pagès J-C (2011) Packaging of HCV-RNA into lentiviral vector. *Biochem Biophys Res Commun* 414(4):808–813
- Chauhan JS, Rao A, Raghava GP (2013) In silico platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences. *PLoS ONE* 8(6):e67008
- Chen C-C, Hwang J-K, Yang J-M (2006) 2: Protein structure prediction server. *Nucleic Acids Res* 34(suppl_2):W152–W157
- Chou P, Fasman GD (2009) Amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* 47:45
- Cristofari G, Ivanyi-Nagy R, Gabus C, Boulant S, Lavergne JP, Penin F et al (2004) The hepatitis C virus Core protein is a potent nucleic acid chaperone that directs dimerization of the viral (+) strand RNA in vitro. *Nucleic Acids Res* 32(8):2623–2631
- Dehghani B, Rasooli I, Gargari SLM, Nadooshan MRJ, Owlia P, Nazarian S (2013) Immunogenicity of *Salmonella enterica* serovar Enteritidis virulence protein, InvH, and cross-reactivity of its antisera with *Salmonella* strains. *Microbiol Res* 168(2):84–90
- Dehghani B, Rasooli I, Jalali-Nadoushan M, Owlia P, Rasooli Z (2014) Immunoprotectivity of *Salmonella enterica* serovar Enteritidis virulence protein, InvH, against *Salmonella typhi*. *Iran J Basic Med Sci* 17(8):560
- Dehghani B, Ghasabi F, Hashempour T, Joulaei H, Hasanshahi Z, Halaji M et al (2017) Functional and structural characterization of Ebola virus glycoprotein (1976–2015)—An in silico study. *Int J Biomath* 10(08):1750108
- Dehghani B, Hashempour T, Zahra H (2019) Using immunoinformatics and structural approaches to design a novel HHV8 vaccine. *Int J Peptide Res Ther*; In Press
- Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform* 8(1):4

- Emini EA, Hughes JV, Perlow D, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol* 55(3):836–839
- Farag RE, Arafa MM, El-Etreby S, Saoudy NS, Eldeek BS, El-Alfy HA et al (2013) Human leukocyte antigen class I alleles can predict response to pegylated interferon/ribavirin therapy in chronic hepatitis C Egyptian patients. *Arch Iran Med* 16(2):68
- Ferroni P, Mascolo G, Zaninetti M, Colzani D, Pregliasco F, Pirisi M et al (1993) Identification of four epitopes in hepatitis C virus core protein. *J Clin Microbiol* 31(6):1586–1591
- Fishman SL, Factor SH, Balestrieri C, Fan X, DiBisceglie AM, Desai SM et al (2009) Mutations in the hepatitis C virus core gene are associated with advanced liver disease and hepatocellular carcinoma. *Clin Cancer Res* 15(9):3205–3213
- Furui Y, Hoshi Y, Murata K, Ito K, Suzuki K, Uchida S et al (2011) Prevalence of amino acid mutation in hepatitis C virus core region among Japanese volunteer blood donors. *J Med Virol* 83(11):1924–1929
- Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) *The proteomics protocols handbook*. Springer, New York, pp. 571–607
- Gedezha MP, Mphahlele MJ, Selabe SG (2014) Prediction of T-cell epitopes of hepatitis C virus genotype 5a. *Virol J* 11(1):187
- Geourjon C, Deleage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics* 11(6):681–684
- Gremion C, Cerny A (2005) Hepatitis C virus and the immune system: a concise review. *Rev Med Virol* 15(4):235–268
- Gupta R, Brunak S (2001) Prediction of glycosylation across the human proteome and the correlation to protein function. In: 2002 *Biocomputing*, World Scientific. pp. 310–322
- Harase I, Moriyama T, Kaneko T, Kita H, Nomura M, Suzuki G et al (1995) Immune response to hepatitis C virus core protein in mice. *Immunol Cell Biol* 73(4):346
- Hashempour T, Bamdad T, Merat S, Janzamin E, Nemati L, Jabbari H et al (2010) Expansion of CD4+ CD25+ FoxP3+ regulatory T cells in chronic hepatitis C virus infection. *Iran J Immunol* 7(3):177–185
- Hashempour T, Alborzi AM, Moayedi J, Ajorloo M, Bamdad T, Sharifi AH et al (2018) A decline in anti-core + 1 antibody titer occurs in successful treatment of patients infected with hepatitis C virus. *Jundishapur J Microbiol* 11(2):e58294
- Hashempour T, Bamdad T, Bergamini A, Lavergne JP, Haj-Sheykholeslami A, Brakier-Gingras L et al (2015) F protein increases CD4+ CD25+ T cell population in patients with chronic hepatitis C. *Pathog Dis* 73(4):ftv022
- Horie T, Shimizu I, Horie C, Yogita S, Tashiro S, Ito S (1999) Mutations of the core gene sequence of hepatitis C virus isolated from liver tissues with hepatocellular carcinoma. *Hepatol Res* 13(3):240–251
- Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z et al (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* 32(3):1037–1049
- Idrees S, Ashfaq UA (2013) HCV infection and NS-3 serine protease inhibitors. *Virol Mycol* 2(112):2160–2161
- Ivanyi-Nagy R, Kanevsky I, Gabus C, Lavergne J-P, Ficheux D, Penin F et al (2006) Analysis of hepatitis C virus RNA dimerization and core–RNA interactions. *Nucleic Acids Res* 34(9):2618–2633
- Jazayeri SM, Carman WF (2005) Virus escape CTL or B cell epitopes? *Hapat Mon* 5(4):133–136
- Karplus P, Schulz G (1985) Prediction of chain flexibility in proteins. *Naturwissenschaften* 72(4):212–213
- Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4(3):363
- Khorrani S, Mohammadpour H, Shahzamani K, Zarif MN, Sharifi AH, Merat S et al (2015) The relationship between HLA-G and viral loads in non-responder HCV-infected patients after combined therapy with IFN- α 2 α and ribavirin. *Hum Immunol* 76(2–3):181–186
- Larsen JEP, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. *Immunome Re* 2(1):2
- Lauer GM, Walker BD (2001) Hepatitis C virus infection. *N Engl J Med* 345(1):41–52
- Lavanchy D (2011) Evolving epidemiology of hepatitis C virus. *Clin Microbiol Infect* 17(2):107–115
- Lechmann M, Ihlenfeldt HG, Braunschweiger I, Giers G, Jung G, Matz B et al (1996) T- and B-cell responses to different hepatitis C virus antigens in patients with chronic hepatitis C infection and in healthy anti-hepatitis C virus—positive blood donors without viremia. *Hepatology* 24(4):790–795
- Lu W, Ou J-h (2002) Phosphorylation of hepatitis C virus core protein by protein kinase A and protein kinase C. *Virology* 300(1):20–30
- Ma H-C, Ku Y-Y, Hsieh Y-C, Lo S-Y (2007) Characterization of the cleavage of signal peptide at the C-terminus of hepatitis C virus core protein by signal peptide peptidase. *J Biomed Sci* 14(1):31–41
- Martro E, Valero A, Jordana-Lluch E, Saludes V, Planas R, González-Candelas F et al (2011) Hepatitis C virus sequences from different patients confirm the existence and transmissibility of subtype 2q, a rare subtype circulating in the metropolitan area of Barcelona, Spain. *J Med Virol* 83(5):820–826
- Moattari A, Dehghani B, Khodadad N, Tavakoli F (2015) In silico functional and structural characterization of H1N1 influenza A viruses hemagglutinin, 2010–2013, Shiraz, Iran. *Acta Biotheor* 63(2):183–202
- Moayedi J, Hashempour T, Musavi Z, Hallaji M, Ghasabi F, Haj-Sheykholeslami A et al (2018) Comparison of IL-28B favorable genotype frequency between healthy and patients infected with HCV. *Sci J Ilam Univ Med Sci* 26(2):28–36
- Netski DM, Mosbruger T, Depla E, Maertens G, Ray SC, Hamilton RG et al (2005) Humoral immune response in acute hepatitis C virus infection. *Clin Infect Dis* 41(5):667–675
- Neumann-Haefelin C, Blum HE, Chisari F, Thimme R (2005) T cell response in hepatitis C virus infection. *J Clin Virol* 32(2):75–85
- Nezafat N, Dorosti H, Zarei M, Ghasemi Y (2018) Exploring dengue proteome to design an effective epitope-based vaccine against dengue virus AU—Sabetian, Soudabeh. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2018.1491890>
- Oehler V, Filipe A, Montserret R, Da Costa D, Brown G, Penin F et al (2012) Structural analysis of hepatitis C virus core-E1 signal peptide and requirements for cleavage of the genotype 3a signal sequence by signal peptide peptidase. *J Virol* 86(15):7818–7828
- Ogata S, Nagano-Fujii M, Ku Y, Yoon S, Hotta H (2002) Comparative sequence analysis of the core protein and its frameshift product, the F protein, of hepatitis C virus subtype 1b strains obtained from patients with and without hepatocellular carcinoma. *J Clin Microbiol* 40(10):3625–3630
- Okamoto K, Mori Y, Komoda Y, Okamoto T, Okochi M, Takeda M et al (2008) Intramembrane processing by signal peptide peptidase regulates the membrane localization of hepatitis C virus core protein and viral propagation. *J Virol* 82(17):8349–8361
- Parker J, Guo D, Hodges R (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 25(19):5425–5432
- Pavio N, Lai MM (2003) The hepatitis C virus persistence: how to evade the immune system? *J Biosci* 28(3):287–304
- Pene V, Hernandez C, Vauloup-Fellous C, Garaud-Aunis J, Rosenberg A (2009) Sequential processing of hepatitis C virus core protein

- by host cell signal peptidase and signal peptide peptidase: a reassessment. *J Viral Hepat* 16(10):705–715
- Pirisi M, Fabris C, Toniutto P, Vitulli D, Soardo G, Falletti E et al (1995) Reactivity to B cell epitopes within hepatitis C virus core protein and hepatocellular carcinoma. *Cancer Res* 55(1):111–114
- Polyak SJ, Klein KC, Shoji I, Miyamura T, Lingappa JR (2006) Assemble and interact: pleiotropic functions of the HCV core protein. *Hepatitis C viruses: genomes and molecular biology*. Horizon Bioscience, Norwich, pp. 89–119
- Pourhassan A. Hepatitis C (2014) The first report in Azeri patients. *Pak J Biol Sci* 17(6):872–875
- Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5(4):725
- Sabet LP, Taheri T, Memarnejadian A, Azad TM, Asgari F, Rahimnia R et al (2014) Immunogenicity of multi-epitope DNA and peptide vaccine candidates based on core, E2, NS3 and NS5B HCV epitopes in BALB/c mice. *Hepat Mon* 14(10):e22215
- Saeedi A, Naderi M, Tabarraie A, Kelishdi M, Ghaemi A (2014) Enhanced immune responses of a hepatitis C virus core DNA vaccine by co-inoculating interleukin-12 expressing vector in mice. *Vaccine Res* 1(2):29–33
- Saha S, Raghava GP (2004) BcePr(ed): prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. In: Saha S, Raghava GP (eds) *International conference on artificial immune systems*. Springer
- Saha S, Raghava G (2006a) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65(1):40–48
- Saha S, Raghava G (2006b) AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 34(suppl_2):W202–W209
- Samimi-Rad K, Sadeghi F, Amirzargar A, Eshraghian MR, Alavian SM, Rahimnia R (2015) Association of HLA class II alleles with hepatitis C virus clearance and persistence in thalassemia patients from Iran. *J Med Virol* 87(9):1565–1572
- Sarvari J, Mojtahedi Z, Kuramitsu Y, Fattahi MR, Ghaderi A, Nakamura K et al (2014) Comparative proteomics of sera from HCC patients with different origins. *Hepat Mon* 14(1):e13103
- Sefidi FJ, Keyvani H, Monavari SH, Alavian SM, Fakhim S, Bokharai-Salim F (2013) Distribution of hepatitis C virus genotypes in Iranian chronic infected patients. *Hepat Mon* 13(1):e7991
- Shih C-M, Chen C-M, Chen S-Y, Lee Y (1995) Modulation of the trans-suppression activity of hepatitis C virus core protein by phosphorylation. *J Virol* 69(2):1160–1171
- Singh H, Raghava G (2001) ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 17(12):1236–1237
- Singh H, Raghava G (2003) ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* 19(8):1009–1014
- Steinmann E, Brohm C, Kallis S, Bartenschlager R, Pietschmann T (2008) Efficient trans-encapsidation of hepatitis C virus RNAs into infectious virus-like particles. *J Virol* 82(14):7034–7046
- Strosberg AD, Kota S, Takahashi V, Snyder JK, Mousseau G (2010) Core as a novel viral target for hepatitis C drugs. *Viruses* 2(8):1734–1751
- Takahashi K, Iwata K, Matsumoto M, Matsumoto H, Nakao K, Hatahara T et al (2001) Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the ‘progression score’ revisited. *Hepatol Res* 20(2):161–171
- Targett-Adams P, Hope G, Boulant S, McLauchlan J (2008) Maturation of hepatitis C virus core protein by signal peptide peptidase is required for virus production. *J Biol Chem* 283(24):16850–16859
- Tokita H, Kaufmann GR, Matsubayashi M, Okuda I, Tanaka T, Harada H et al (2000) Hepatitis C virus core mutations reduce the sensitivity of a fluorescence enzyme immunoassay. *J Clin Microbiol* 38(9):3450–3452
- Ward S, Lauer G, Isba R, Walker B, Klenerman P (2002) Cellular immune responses against hepatitis C virus: the evidence base 2002. *Clin Exp Immunol* 128(2):195–203
- Yassin K (2001) *Unraveling the mystery of liver diseases in Egypt*. Jacobs Verlag, Lage

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