



Production of Antibodies against Hepatitis C Virus Envelope Glycoprotein E2- A Potential Vaccine Against HCV Infection

Gul Afshan^{1,2,*}, Soumble Zulficar², Sumaira Mehboob², Muhammad Tahir Javed Khan¹ and Abdul Rauf Shakoori^{2*}

¹University College of Pharmacy, University of the Punjab, Allama Iqbal Campus, Lahore

²School of Biological Sciences, University of the Punjab, Quaid-i-Azam Campus, Lahore

ABSTRACT

Currently a number of vaccines are under development against hepatitis C, though an effective hepatitis C vaccine is yet to come. The present study is focused on the development of recombinant vaccine against HCV by utilizing HCV3a envelope glycoprotein E2. HCV3a E2 gene was amplified and cloned first in cloning vector pTG19 and subcloned in expression vector pET21a. E2 protein, expressed in insoluble form, was purified by repeated sonications, followed by denaturation and refolding through fractional dialysis with urea. Multiple alignment with other sequences showed nucleotide variations of HCV3a E2 compared with already reported sequences. A total of 140 (11%) point mutations were found in Pakistan's gene sequence, out of which general and specific differences were 93 (9%) and 47 (4%), respectively. The general differences usually occur in common population while the specific differences are present only in some cases. The 3D structure of E2 was determined. Protein docking with appropriate ligand revealed that V342 and L349 residues were involved in the ligand binding. Interaction of HCV3a infected human sera with HCV3a E2 purified protein was carried out through Enzyme Linked Immunosorbant Assay (ELISA) using HCV antihuman conjugated protein. It was confirmed that the HCV3a E2 purified protein bound with the human serum almost three times more efficiently than the positive control. For raising antibodies against the antigenic protein HCV3a E2, the purified protein E2 was injected subcutaneously into the rabbit (*Oryctolagus cuniculus*). The antibodies titer in rabbit after immunization in 100% pure serum sample was 0.736 which was 15 times more than that of pre immune control sera (0.049). It is concluded from the current study that purified HCV3a envelope glycoprotein E2 has significant antigenic activity and can potentially be used as recombinant vaccines against HCV infection.

Article Information

Received 19 September 2018

Revised 11 March 2019

Accepted 25 May 2019

Available online 12 September 2019

Authors' Contribution

ARS conceived and designed the study. GA performed all the lab work. GA and ARS wrote the article. SZ provided immediate guidance in experimental work and in preparation of the manuscript. SM helped in protein purification. ARS and MTJK supervised the work.

Key words

ELISA, Hepatitis C virus, HCV3a envelope glycoprotein E2, Recombinant vaccine.

INTRODUCTION

Hepatitis C is a transfusion related disease which leads to liver cirrhosis, hepatocellular carcinoma (HCC) and eventually death. HCV is a member of the family Flaviviridae (Kato, 2000). Three genotypes exist worldwide of which 1a and 1b are more commonly found. Genotype 3a is the most prevalent in Pakistan with prevalence rate of 2.2-14% (Raja and Janjua, 2008). It is followed by HCV genotypes 3b and 1a (Idrees and Riazuddin, 2008; Idrees *et al.*, 2009).

HCV is a small 55-65 nm virus containing positive sense single strand RNA genome. It consists of about 9600 nucleotides. HCV encodes a single polyprotein consisting of almost 3010 amino acids. These proteins are either structural or nonstructural in nature. The structural proteins

comprise of one core and two envelope proteins E1 and E2, whereas the nonstructural proteins are P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B. E1 and E2 are viral envelope glycoproteins that are fixed in the lipid envelope (Beeck and Dubuisson, 2003). These are responsible for viral entry in host cells via receptor binding and are prime target for the development of recombinant vaccine against HCV infection.

E2 gene is 1056bp in length, while its protein is composed of 352 amino acids with molecular weight of 70-72 kDa. It has three hypervariable regions namely, HVR1, HVR2 and HVR3. HVR1, near its amino terminus, is the main target for neutralizing antibodies. In this region the amino acid sequence varies up to 80% between different genotypes and subtypes of the same genotype (Goffard *et al.*, 2003; Weiner *et al.*, 1991).

* Corresponding author: arshakoori.sbs@pu.edu.pk; gulafshan099@gmail.com
0030-9923/2019/0006-2311 \$ 9.00/0
Copyright 2019 Zoological Society of Pakistan

Abbreviations

HCV, Hepatitis C virus; E2, Envelope glycoprotein; HCC, Hepatocellular carcinoma; PCR, Polymerase chain reaction; ELISA, Enzyme linked immunosorbant assay.

As there is no vaccine available against HCV till date, production of a prophylactic vaccine against HCV is need of the hour. Drug therapy is costly and to some extent unaffordable especially in the developing world. Effective immunization against HCV could be helpful for the host to protect against the infection through host immune response.

First attempt to produce an HCV vaccine was reported in 1994 (Choo *et al.*, 1994). This vaccine was composed of HCV envelope glycoproteins which were expressed in mammalian cells by vaccinia virus. It resulted in the induction of high antibody titer. A disadvantage was that this antibody was short lasting and ineffective to produce cross genotype protection. Induction of high titer, long lasting and cross-reactive anti-envelope antibodies and an effective cellular immune response are necessary for an effective vaccine (Lechmann and Liang, 2000).

E2 antibodies that cause cross-neutralization of diverse HCV isolates include mouse mAb AP33, rat mAb 3/11 and human mAb HCV1 (Broering *et al.*, 2009; Owsianka *et al.*, 2005). These antibodies bind to the linear epitopes present within the antigenic site of the most conserved region (residues 412-423) of E2 and in turn block the interaction of E2 and CD81 receptors present on hepatocytes. Moreover, human polyclonal antibodies after purification from the sera of HCV infected patients consisting of E2 residues 412-423 can neutralize HCV (Broering *et al.*, 2009). This conserved region of E2 antigenic site 412-423 is a prime target for vaccine development. There are other mAbs that recognize intermittent overlapping of E2 epitopes with CD81 binding sites involving the residues 395-424, 425-447 and 523-540 (Johansson *et al.*, 2007; Keck *et al.*, 2008; Law *et al.*, 2008; Schofield *et al.*, 2005). Much work has been done worldwide in this respect, but no effective prophylactic HCV vaccine is available till now.

The present study aims at developing HCV recombinant antibodies utilizing envelope glycoprotein E2 to prevent HCV infections. This will open the door for developing effective vaccine against HCV at commercial level.

MATERIALS AND METHODS

RNA isolation and cDNA synthesis

Serum sample of a local HCV3a infected patient was obtained from CMH Diagnostic Laboratory, Lahore, Pakistan. Total RNA from the sample was extracted with the help of INSTANT Virus RNA Kit (AJ Roboscreen GmbH, Cat # 0209200501) according to the manufacturer's instructions. cDNA was synthesized with 1 µg of RNA and random hexamers by using Superscript III cDNA synthesis kit (Invitrogen life technologies, CA, Cat # 18080200).

Primer designing

HCV full length genome sequences reported from various countries worldwide were selected from Los-Alamos (<https://hcv.lanl.gov/content/sequence/HCV/otherweb.html>). The detail of these sequences is given in Table I.

Table I.- List of the Hepatitis C Virus genotype 3a complete genome sequences used in the analysis.

Serial No.	Genbank accession No.	Country	Year
1.	GU294484.1	Pakistan	2009
2.	HQ108107.1	Pakistan	2010
3.	KF035127.1	India	2013
4.	KC844041.1	China	2013
5.	D17763.1	New Zealand	1993
6.	X76918.1	Germany	1995
7.	DQ437509.2	Switzerland	2006
8.	GU814263.1	Denmark	2010
9.	GQ356217.1	UK	2009
10.	GQ356200.1	UK	2009
11.	AY956467.1	USA	2006
12.	JQ803486.1	USA	2012

The sequences of structural gene *E2* were subjected to Multiple Sequence Alignment by CLUSTALW [GenomeNet site (<http://www.genome.jp/tools/clustalw/>)]. Conserved regions flanking the *E2* gene were identified from this alignment. Two sets of primers, from these conserved regions, were designed for the first and the second round of PCR (Table II).

Amplification of E2

In the first PCR, cDNA (2 µl) was used as template and amplification was carried out with the help of primers E2a-F and E2a-R. *E2* gene (1056bp) was amplified using 2 µl of amplicon of first PCR and E2b-F and E2b-R primers. In both PCRs the reaction composition (except template and primers) and reaction conditions were the same. PCR amplifications were carried out in 20 µl reaction mixture containing 2 µl DNA template, 1 µM each primer, 3mM MgCl₂, 0.125mM dNTPs and 5 units Taq polymerase. Thermal cycle reaction conditions for the first and the second round of PCRs were, initial denaturation at 95 °C for 5 min., 35 cycles of each denaturation at 95 °C for 30 sec., annealing at 58 °C for 30 sec. and polymerization at 72 °C for 1 min., and final extension of 72 °C for 10 min.

Table II.- Primers used in the analysis.

Serial No.	Primer-Id	Primer sequence	Size
1.	E2a-F (inner)	5' ATGTTCTCAGGGGTCGATGC	20 nts
2.	E2a-R (inner)	5' AGGTACCAGCCGATAACCATG	20 nts
3.	E2b-F (outer)	5' CCAAGGTCGCAATCATCATG	20 nts
4.	E2b-R(outer)	5' GTCTTTGGTCCCTAGCATTGC	21 nts
5.	E2EXP-F	5' CATATGCACACATATA CCACCG	22 nts
6.	E2EXP-R	5' TTATGCGTCTGCTA AGAGGAG	21 nts

In these primer sequences, restriction sites are in bold.

Table III.- Dosage schedule of HCV E2 purified protein injected S/C in the rabbit.

Serial No.	Purified protein (E2)	Dose
1.	0 dose	E2 200µg in 500µl PBS (pH7.4) + Freund's Complete Adjuvant 500µl
2.	1 st booster (after 4 weeks)	E2 100µg in 500µl PBS (pH7.4) + Freund's Incomplete Adjuvant 500µl
3.	2 nd booster (after 7 weeks)	E2 100µg in 500µl PBS (pH7.4) + Freund's Incomplete Adjuvant 500µl

Cloning of E2

E2 was cloned by ligating the amplified gene with pTG19. DH5α strain of *E.coli* (cloning host) was transformed with the recombinant vector pTG19-E2. The transformation was confirmed through single restriction analysis using *EcoRI* and double restriction analysis using *BamHI* as its restriction site is present before as well as after the insert in the MCS region of the vector pTG19. The transformation was also confirmed through sequencing of cloned E2 gene.

In order to ligate E2 gene at right position in pET21a (expression vector), the gene was again amplified by using forward primer (E2EXP-F) containing *NdeI* site and reverse primer (E2EXP-R) containing *HindIII* site (Table II) using pTG19-E2 recombinant vector as template. The PCR reaction composition and conditions remained the same as described earlier. The amplicon was cloned using pTG19 vector and DH5α host strain. E2 was subcloned through *NdeI* and *HindIII* sites using pET21a (expression vector) first in DH5α and then in BL21 C+ strain (expression host) of *E. coli*.

Expression of E2

For E2 gene expression, the transformed cell culture was grown in LB medium (10 ml, containing 100µg/ml ampicillin) at 37 °C for 16 h. Next day, from this culture 1% secondary inoculum was given to ampicillin containing fresh LB medium (1L) under sterile conditions. It was incubated at 37 °C with constant shaking until culture O.D.₆₀₀ reached upto 0.5. For protein expression this culture was induced with 0.1mM IPTG while the

control culture was kept uninduced. Both cultures were further grown at 37 °C with shaking at 100 rpm for 4 h followed by centrifugation at 3700 x g for 5 min. at 4 °C to harvest the cell pellet.

Each cell pellet was resuspended in 20mM Tris-Cl, pH 8.0 by diluting them upto O.D.₆₀₀ 10. Both induced and uninduced cells pellet were lysed by ultra sonication technique. Total 20 sonication cycles with 30 second burst and 1 min. rest at 60 Hz intensity using Vibra Cell Sonicator were performed. The samples were centrifuged at 14810 x g for 10 min. at 4 °C to get separate soluble and insoluble fractions. The protein production was examined on 12% SDS-PAGE (Sambrook and Russel, 2001). The insoluble fraction was further resuspended in same volume of 20 mM Tris-Cl buffer.

Purification of E2 protein

As E2 protein was expressed in inclusion bodies, it was further purified by suspending in wash buffer (20mM Tris-Cl pH 8.0, 5mM NaCl and 0.01% Triton). The mixture was then sonicated for 20 cycles, each for 30 sec. burst with 1 min. rest at 60 Hz intensity. The sonicated sample was centrifuged again at 3700 x g for 5 min. at 4 °C and the same procedure was repeated 4 times to get rid of cell's membranous proteins and other debris from inclusion bodies. All samples were further examined on 12% SDS-PAGE. Protein concentration was estimated by measuring its absorbance at 260 and 280 nm. For this purpose, 100µl sonicated sample was centrifuged and resuspended in 5% SDS (100µl).

Solubilization of the inclusion bodies

For refolding of E2, 9 mg of inclusion bodies were solubilized in 18 ml solubilization buffer (20mM Tris-Cl pH 8.0, 8M urea, 10% glycerol, 6mM DTT). The inclusion bodies were resuspended in the solubilization buffer by pipetting and vortexing well in the buffer with final protein concentration of 0.5 mg/ml. It was incubated at 37 °C for half an hour in a tightly closed tube to allow complete denaturation of insoluble aggregates. The contents were vortexed well for complete resuspension of inclusion bodies. Sample was subjected to sonication to get rid of any suspended particle followed by centrifugation at 3700 x g at 4 °C for 10 min. The denatured protein sample was further saved at 4 °C for overnight.

Refolding of E2

The denatured protein was poured in a 12kDa cut off value dialysis tube and dialysed against 1L refolding buffer (20mM Tris-Cl pH 8.0, 6M urea, 0.5M EDTA) at 4 °C for 4-6 h (during which buffer was changed twice). The urea concentration was gradually decreased by replacing with fresh refolding buffer of 4M, 3M, 2M, 1M and 0M urea. Dialysis was completed in 7 days.

A clear dialysate having protein concentration 0.4 mg/ml was obtained. It was centrifuged at 3700 x g for 5 min. at 4 °C to remove any micro debris or particles. Protein sample was further concentrated by passing through a membrane filter (10kDa MWCO) having the final concentration (0.5 mg/ml). The refolded sample was analyzed on 12% SDS-PAGE.

ELISA of HCV3a E2 purified protein with HCV3a infected human sera

Enzyme Linked Immunosorbant Assay (ELISA) was performed to verify the HCV3a E2 purified protein. The kit used for this purpose was HCV ELISA 3rd generation kit (Wiener lab. Cat No. 1483258). The OD 450/360 was recorded using Humareader plus (Human GmbH).

Production of antibodies against E2 purified protein in rabbit

For the preparation of immunogen, 200µg of filtered antigen (purified E2) in 0.5 ml phosphate buffer saline (PBS) (pH7.4) was mixed with 0.5 ml of Freund's complete Adjuvant (Sigma Aldrich Cat # F5881). The immunogen was administered subcutaneously into the rabbit to raise antibodies against E2. Subsequent 1st booster and 2nd booster doses were given as half the 1st dose antigen (100µg) in 0.5 ml PBS (pH7.4) mixed with 0.5 ml Freund's Incomplete Adjuvant (Sigma Aldrich Cat # F5506) (Table III). The rabbit was kept for one week more and the blood was collected through cardiac puncture. Aliquots of serum

were made and saved at -20 °C.

Antibodies confirmation through ELISA

Antibodies raised in response to HCV purified protein E2 were confirmed through ELISA. The 96 well ELISA plate was coated with purified antigenic protein E2 in borate buffer pH 8.0 (200ng /100µl) in each well overnight at 4 °C in humid conditions. Dilution (100 X) of rabbit serum was prepared in buffer (25mM PBS+0.05% BSA) and 100µl of the serum was added in duplicate for each test antisera and control sera and incubated at 4 °C overnight. HRP conjugated anti rabbit IgG (10,000 fold diluted) was used further for incubation and OD_{450/360} was recorded using Humareader plus (Human GmbH).

Bioinformatics study of HCV E2 protein

The sequence of HCV 3a E2 gene obtained in this study was submitted in DNA Data Bank of Japan (DDBJ) under the accession numbers LC217312.1. It was subjected to Multiple Sequence Alignment by CLUSTALW with other already taken reported E2 gene sequences. Phylogenetic tree was developed based on the multiple alignment.

The primary sequence of E2 protein was deduced from nucleotide sequence. Its analysis revealed that E2 is a cysteine rich protein (18 cysteine residues out of 352). Further PI of protein was measured using ExPasy online server which was 7.12.

Molecular protein assimilation and docking

The structure of HCV3a E2 protein was produced by I-TASSER (zhanglab.ccmb.med.umich.edu/I-TASSER) on the basis of amino acid sequence. The structure in PDB format was visualized in PyMOL (<https://www.schrodinger.com/pymol>). The results obtained are shown in Figure 8. The structure obtained was submitted to COACH docking tool for protein docking with appropriate ligands.

Statistical analysis

Data are presented as mean \pm SD. Numerical data were analyzed using Student's t-test and ANOVA. P value < 0.05 was considered statistically significant.

RESULTS

E2 gene

The amplified E2 appeared as 1056bp band (Fig. 1a), which was later cloned in DH5 α using pTG19. Restriction of recombinant vector pTG19-E2 with *EcoRI* resulted in a single band at ~3.856kb, while restriction with *BamHI* appeared as two bands, one at ~ 1077bp for E2 and

second at 2.8kb for cloning vector pTG19 (Fig. 1b). Both restrictions confirmed the presence of right size of insert in recombinant plasmid. Total point mutations found in HCV E2 were 140 (11%), out of which general differences were 93 (9%) while specific differences were 47 (4%). General differences were almost double as compared to specific differences (Fig. 2).

Phylogenetic analysis of HCV E2 of Pakistan revealed it as the most closely related to D17763.1 of New Zealand. The HCV sequence of X76918.1 of Germany is allied to GU814263.1 of Denmark which are closely related to HCV3a E2 of Pakistan and D17763.1 of New Zealand. HCV E2 of Pakistan is also related to KC844041 of China and JQ803486.1 of USA. The nucleotide sequence of HCV E2 is then related to GQ356217.1 of UK and AY956467.1 of USA. It is then related to DQ437509.2 of Switzerland and GQ356200.1 of UK which are closely related to each other. The nucleotide sequence of HCV E2 is far apart to HQ108107.1 and GU294484.1 of Pakistan and KF035127.1 of India (Fig. 3).

The 3D structural determination of E2 showed that there were 15% β -sheets at positions 3-5, 29-31, 60-61, 90-91, 114-115, 132-134, 154-157, 167-168, 180-181, 216-218, 230-232, 244-256, 296-297, 313-315 and 320-326 and 17% α -helices at positions 7-19, 54-59, 71-72, 79-82, 207-209, 259-264, 282-284, 288-290 and 330-349.

Protein docking with proper ligands indicated that V342 and L349 residues are involved in the ligand binding (Fig. 4).

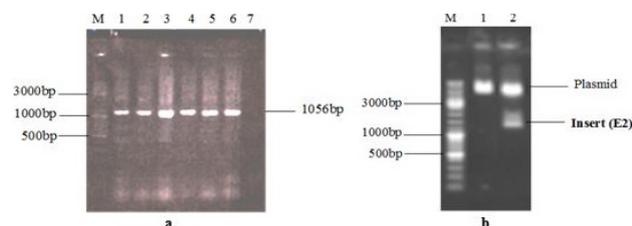


Fig. 1. Agarose gel electrophoresis (a) PCR amplification of HCV3a E2 gene. Lanes 1-6, Amplified 1056bp E2 gene whereas lane 7 is E2 no template control. (b) Restriction analysis of pTG19-E2 with restriction enzymes *EcoRI* and *BamHI*. Lane 1, Single restriction of pTG19-E2 with restriction enzyme *EcoRI*. Lane 2, Restriction of pTG19-E2 with *BamHI* that cuts at both ends of the insert; 2.8kb digested vector releasing 1077bp band of E2 gene; Lane M, GeneRuler™DNA ladder mix (Fermentas Cat # SM0331).

Cloning of HCV3a E2 gene with *NdeI* and *HindIII* restriction sites

Figure 5a shows the PCR amplification of full length E2 gene with restriction sites *NdeI* and *HindIII*. Figure 5b

shows the recombinant plasmids (pTG19-E2). The size for isolated recombinant plasmid pTG19-E2 is 3.856kb. Double restriction showed bands for E2 at 1.056kb (1056bp) and 2.8kb for pTG19, while for pET21a band was observed at 5.4kb (Fig. 5c).

Subcloning

In Figure 6 the bands observed represented the isolated recombinant plasmid pET21a-E2 carrying E2 gene. Single restriction with *NdeI* resulted in a single band at 6.456kb. Double restriction analysis of pET21a-E2 with *NdeI* and *HindIII* restriction enzymes showed required bands at right positions *i-e*; 1056bp of E2 and 5.4kb of pET21a that confirmed the presence of desired DNA fragment in the expression vector.

Expression and purification of E2 protein

Figure 7a shows the bands of E2 protein expressed at 39kDa in pellet as inclusion bodies. Partially purified protein was illustrated after 12% SDS-PAGE analysis (Fig. 7b). Single band was observed at 39kDa that showed the >95% purified form of E2 protein (Fig. 7c).

Interaction of HCV infected human sera with HCV E2 purified protein

The purified protein was confirmed through ELISA with human sera infected with HCV3a. The positive control showed red color while the negative control showed green color. Blue color was observed in E2 protein samples. When stop solution was added in the first well positive control showed yellow color while in the second well negative control was colorless. In the third well in which E2 purified proteins was added, yellow color was observed which confirmed that the purified protein was HCV E2 protein (Fig. 8). HCV antihuman conjugated protein bound with antigen HCV E2 purified protein. The antigen antibody combination was completed.

Figure 9 shows the comparison of HCV positive control with HCV E2 purified protein. The HCV E2 protein showed the antigen antibody response in ELISA as 0.38 which was almost three-fold the positive control with value 0.156. It confirmed that the HCV E2 purified protein combined with the human serum almost three times more effectively than the positive control.

Confirmation of the antibodies against HCV3a E2 protein

The test was verified by the appearance of blue color after the addition of TMB substrate. The blue color changes to yellow after adding stop solution 12.5% H_2SO_4 that significantly confirmed that antibodies were raised against HCV E2.

HVR1

```

GQ356217.1  GGCACACACCTCAAGGTGGCAAGCAGCGCTTACTGCTGAGGCTTCCGACCTCTTTA
AY956467.1  ACCACACACCTTACGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCGCTCTTTA
HCV_E2      CACACCTACTACAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCGCTCTTTA
GUB14263.1  SAARACATATATCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
X76918.1    SAARACATATATCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
HQ108107.1  SAARACATATATCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
D17763.1    CACACATATATCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
JQ803486.1  GGCACACACCTTCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
DQ437509.2  ACCACATATATCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
GU294484.1  GTTACATATATCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
GQ356200.1  GTACACACCTTCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
KF035127.1  ACTACACCTTCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
KC844041    ACTACACCTTCAAGGTGGCAAGCAGCTCTACTGCTGCTTCCGCTACCTTTT
  **          **          *

```

HVR3 **HVR2**

```

GQ356217.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
AY956467.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
HCV_E2      TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
GU814263.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
X76918.1    TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
HQ108107.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
D17763.1    TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
JQ803486.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
DQ437509.2  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
GU294484.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
GQ356200.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
KF035127.1  TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
KC844041    TTTTCAGGTAAGGCTGGGGCCCTGACAGAGCTAACATCACCCTCCCTCTTGA
  *          *          *          *          *          *          *

```

HVR1

```

GQ356217.1  AGGCTGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
AY956467.1  ACCCTGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
HCV_E2      ACCTGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
GUB14263.1  AGTATGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
X76918.1    AGTTGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
HQ108107.1  GACATGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
D17763.1    GACATGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
JQ803486.1  ACTCCTGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
DQ437509.2  ACTCCTGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
GU294484.1  AGTCCGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
GQ356200.1  AGTCCGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
KF035127.1  AGTCCGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
KC844041    AGTCCGAGGCTCCGACAGAAATGAGCTGCTCAACACCAATGGCTCGTGGCACATCAA
  **          **          *          *          *          *          *

```

```

GQ356217.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
AY956467.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
HCV_E2      AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
GU814263.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
X76918.1    AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
HQ108107.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
D17763.1    AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
JQ803486.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
DQ437509.2  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
GU294484.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
GQ356200.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
KF035127.1  AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
KC844041    AACCCTACTGCTGGCACTACCTCCTAACCCTTGTGAACTGTCAGCTCAGAGTGC
  *          *          *          *          *          *          *

```

HVR3

```

GQ356217.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
AY956467.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
HCV_E2      AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
GUB14263.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
X76918.1    AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
HQ108107.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
D17763.1    AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
JQ803486.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
DQ437509.2  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
GU294484.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
GQ356200.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
KF035127.1  AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
KC844041    AGGACTGCCCTGAAGTGCATGATTCCTTAAACACCGGGTTATAGCTGGGTTGTTTAT
  **          **          *          *          *          *          *

```

```

GQ356217.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
AY956467.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
HCV_E2      TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
GU814263.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
X76918.1    TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
HQ108107.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
D17763.1    TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
JQ803486.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
DQ437509.2  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
GU294484.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
GQ356200.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
KF035127.1  TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
KC844041    TGC GGCC CTGT GTACT GCTT CACA CCAT CGCC AG-TGGTGTSTGGCACTACTGAT
  *          *          *          *          *          *          *

```

HVR3 **HVR2**

```

GQ356217.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
AY956467.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
HCV_E2      TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
GUB14263.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
X76918.1    TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
HQ108107.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
D17763.1    TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
JQ803486.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
DQ437509.2  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
GU294484.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
GQ356200.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
KF035127.1  TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
KC844041    TACCATAGGTTCAACTCTACTGGAATGTCCTCAAAGGCTCAGCAGCTGCAAGCCCATCACT
  **          **          *          *          *          *          *

```

```

GQ356217.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
AY956467.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
HCV_E2      GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
GU814263.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
X76918.1    GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
HQ108107.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
D17763.1    GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
JQ803486.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
DQ437509.2  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
GU294484.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
GQ356200.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
KF035127.1  GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
KC844041    GGGCGTCCCACTACTGGGGTGAATGAGACAGAGTGTCTGCTGTAAGTCCCT
  *          *          *          *          *          *          *

```

(continued on next page)

GQ356217.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTA
 AY956467.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 HCV82 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTC
 GU814263.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTC
 X76918.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 HQ108107.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTC
 D17763.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTC
 JQ803486.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 DQ437509.2 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 GU294484.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 GQ356200.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 KF035127.1 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 KC844041 GCGGCTCCCGCTG-CGCGTGCTTGGGTC--CCTGGATGAATCCACGGGGTTTTC
 * * * * *

GQ356217.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 AY956467.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 HCV82 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 GU814263.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 X76918.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 HQ108107.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 D17763.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 JQ803486.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 DQ437509.2 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 GU294484.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 GQ356200.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 KF035127.1 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 KC844041 CGGCTTGGCATTACCCATGACAGTCAATTTACNTGTTCAAGGTGAGGATGTTTGTG
 * * * * *

GQ356217.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 AY956467.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 HCV82 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 GU814263.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 X76918.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 HQ108107.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 D17763.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 JQ803486.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 DQ437509.2 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 GU294484.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 GQ356200.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 KF035127.1 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 KC844041 TCAAGAGCTGGGAG-CGCGGCTTGTAAACATCTATGGGGTGA----GGG-GAATCC
 * * * * *

GQ356217.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 AY956467.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 HCV82 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 GU814263.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 X76918.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 HQ108107.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 D17763.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 JQ803486.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 DQ437509.2 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 GU294484.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 GQ356200.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 KF035127.1 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 KC844041 GCTGGGTTGACGACGGTTTACCGCGCTTGCACCTGGACCGGGGGAGCGCTGCAT
 * * * * *

GQ356217.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 AY956467.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 HCV82 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 GU814263.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 X76918.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 HQ108107.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 D17763.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 JQ803486.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 DQ437509.2 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 GU294484.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 GQ356200.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 KF035127.1 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 KC844041 AAAAATGATTCAGACCTCTTCTGCCCCACCGACTGCTTCAGGAAACATCTCGAGGCCA
 * * * * *

GQ356217.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 AY956467.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 HCV82 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 GU814263.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 X76918.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 HQ108107.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 D17763.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 JQ803486.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 DQ437509.2 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 GU294484.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 GQ356200.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 KF035127.1 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 KC844041 ATCGAGGATCGTGACCGCAGCGAGCAACATCCGCTGCTGCATTCACCAACTGAGCTTGT
 * * * * *

GQ356217.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 AY956467.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 HCV82 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 GU814263.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 X76918.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 HQ108107.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 D17763.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 JQ803486.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 DQ437509.2 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 GU294484.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 GQ356200.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 KF035127.1 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 KC844041 TACAGCCGCTGTGGTGGGGGCGCTGTTGACACCTGGATGCTGGTGGACTATCCATAC
 * * * * *

GQ356217.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 AY956467.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 HCV82 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 GU814263.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 X76918.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 HQ108107.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 D17763.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 JQ803486.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 DQ437509.2 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 GU294484.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 GQ356200.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 KF035127.1 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 KC844041 ATAATGCTTGGCTCTTCAGGCCCATGCCGCATTTGTCAACAGGTTAATACACCTCCAC
 * * * * *

(continued on next page)

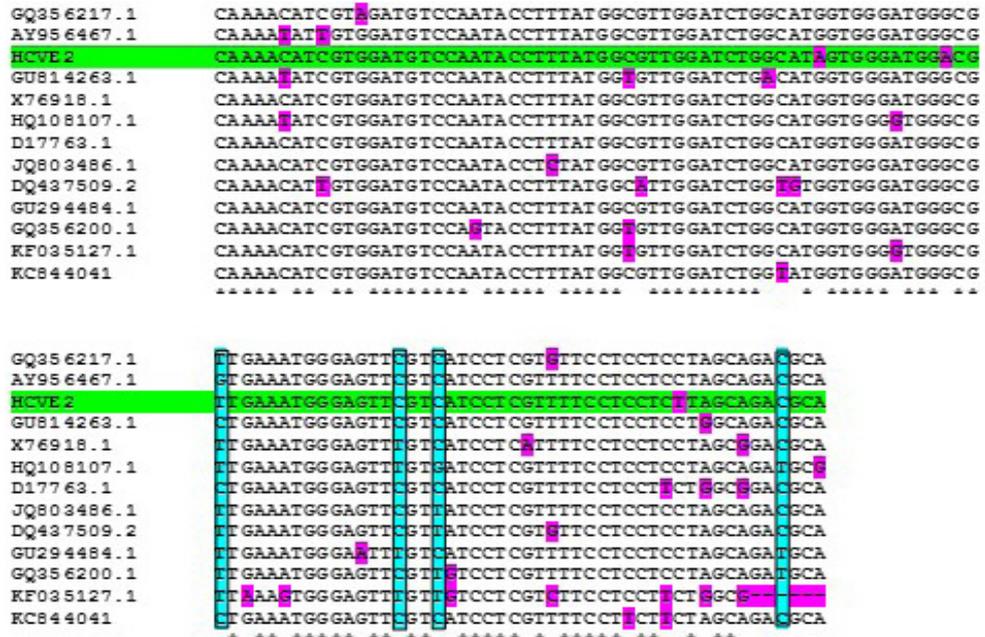


Fig. 2. Multiple sequence alignment of HCV3a E2 (Green color) with already reported HCV3a E2 genomes of various countries on the basis of nucleotide sequence. Yellow box shows the HVR1; purple box shows the HVR2 and red box shows the HVR3. Blue color boxes shows general differences; Pink color boxes show specific differences; (*) shows nucleotide similarity among all the sequences.

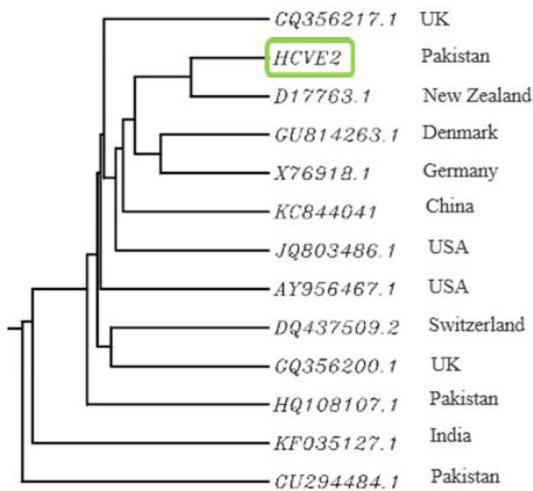


Fig. 3. Rooted phylogenetic tree with branch length (UPGMA) of hepatitis C virus (HCV) genotype 3a isolates worldwide based on E2 nucleotide sequences.

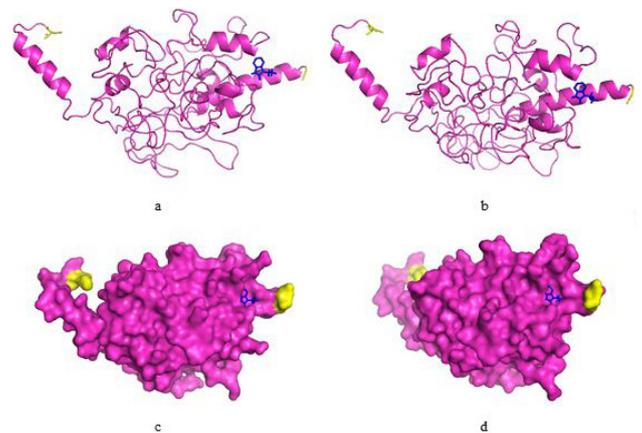


Fig. 4. 3D model of HCV3a E2 protein inferred through I-TASSER and COACH results. (a, b) Ribbon structure of E2 protein showing binding site residues in blue ball and stick. -NH2 and -COOH terminals are shown by yellow color; (c, d) Surface structure of E2 protein displaying two views.

Figure 11 illustrates the comparison of preimmune sera as control and HCV E2 serum antibodies raised in response to the E2 purified protein injected S/C into the rabbit. The figure explains the results comprising of different dilution factors ranging from 100% pure to 10 times serially diluted serum samples consequently from left to right 1-12,

respectively. The antibodies raised in 100% pure serum sample (0.736) is 15 times as compared to the control sera (0.049). In 10% diluted sample (bar 2) the antibodies response is slightly less than the pure sample as compared to the control sera. In the case of 100% dilution (bar 3), antibodies response is slightly less than the 10% diluted

sample. The next dilutions showed gradually low results as compared to the first three ones. This showed that first three dilutions can be taken into consideration and the rest ignored.

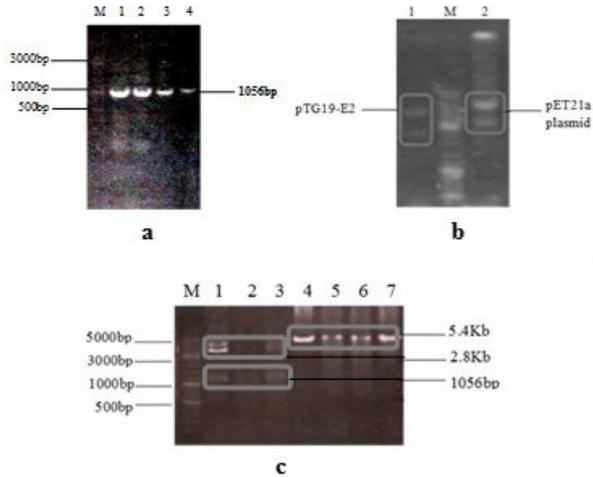


Fig. 5. Cloning of HCV3a E2 gene with *NdeI* and *HindIII* restriction sites. (a) Lanes 1-4, Amplified gene E2 with *NdeI* and *HindIII* restriction sites. (b) Lane 1, Recombinant plasmid pTG19-E2; Lane 2, pET21a. (c) Lanes 1-3, Restricted pTG19-E2; Lanes 4-7, Restricted pET21a; Lane M, GeneRulerTMDNA ladder mix (Fermentas Cat # SM0331).

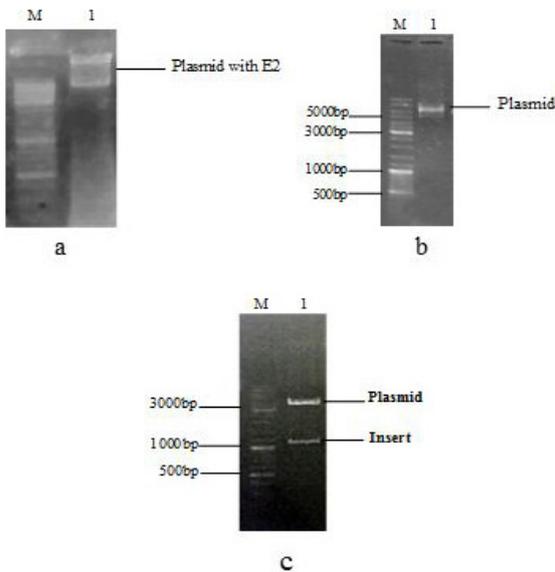


Fig. 6. Construction of pET21a-E2 recombinant vector. (a) Lane 1, Isolated pET21a-E2 gene. (b) Lane 1, Single restriction of pET21a-E2 with *NdeI*. (c) Lane 1, Double restriction of pET21a-E2 with *NdeI* and *HindIII* restriction enzymes; Lane M, GeneRulerTMDNA ladder mix (Fermentas Cat # SM0331).

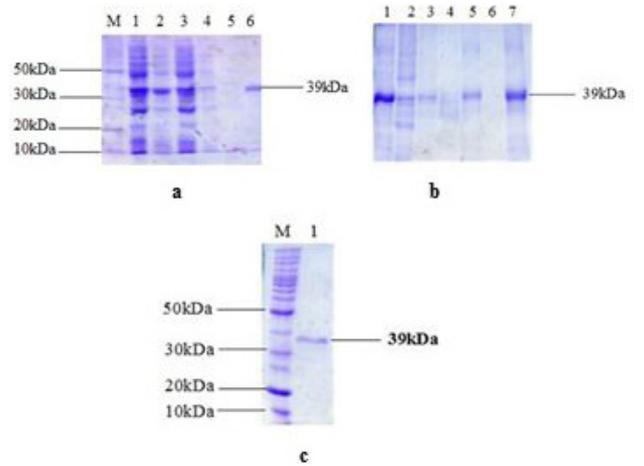


Fig. 7. SDS-PAGE analysis. (a) HCV3a E2 protein expression in inclusion bodies. Lane 1, Uninduced E2 total cell lysate; Lane 2, Induced E2 total cell lysate; Lane 3, Uninduced E2 supernatant; Lane 4, Induced E2 supernatant; Lane 5, Uninduced E2 pellet; Lane 6, Induced E2 pellet. (b) Partial protein purification of HCV3a envelope glycoprotein E2 after fractional dialysis with urea. Lane 1, E2 protein expression; Lanes 2-7, Fractional dialysis with 6M, 4M, 3M, 2M, 1M and 0M urea, respectively. (c) HCV3a E2 purified protein at 39 kDa; Lane 1, E2 purified protein; Lane M, BenchMark™ Protein ladder (Novex Cat # 10747-102).



Fig. 8. Confirmation of HCV E2 purified proteins through ELISA with HCV infected human sera. (a) Well 1 has positive, well 2 has negative and well 3 has E2 protein. (b) shows color change after adding stop solution.

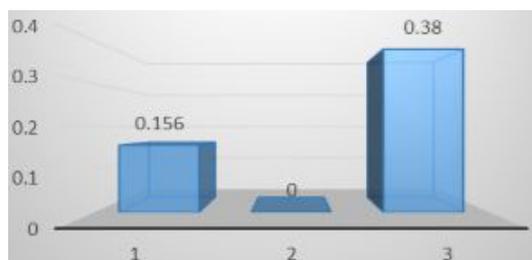


Fig. 9. Analysis of ELISA of HCV3a E2 protein. Comparison of HCV3a positive control vs HCV3a E2 purified protein. Bar 1, HCV positive control; Bar 2, HCV negative control; Bar 3, HCV E2 purified protein.

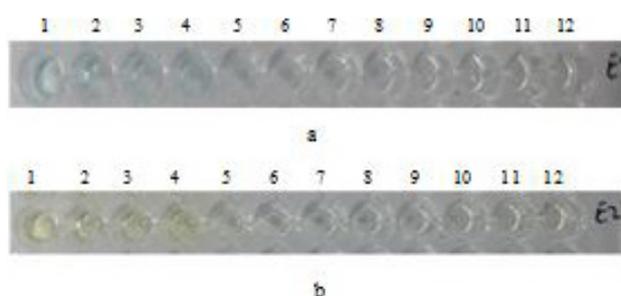


Fig. 10. HCV3a E2 antibodies confirmation after ELISA. Wells 1-12 contain different dilutions ranging from 100% pure to 10 times serially diluted E2 antibodies. (a) Blue color after adding TMB substrate. (b) Color change from blue to yellow after adding 12.5% H_2SO_4 as stop solution.

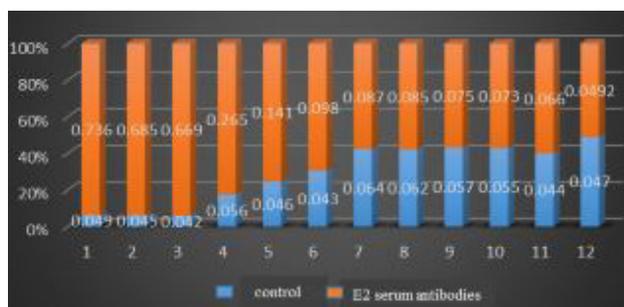


Fig. 11. Comparison of control (preimmune sera) vs HCV3a E2 serum antibodies. Bars 1-12, Comparative analysis of preimmune sera with HCV3a E2 antibodies raised in response to E2 purified protein as antigen. The dilution ranges from 100% pure to 10 times serially diluted serum sample accordingly from bar 1 through 12.

DISCUSSION

HCV infection is a critical global issue. Currently no vaccine is available to prevent HCV infection due to divergence of nucleotide sequence of various strains. There is a strong need to develop vaccine against HCV

infection because the current therapy is long, costly and has significant side effects. The immune response through recombinant vaccine could be helpful in producing effective immunization against the infection.

Currently using HCV polymerase or protease inhibitors are more effective against genotype 1 than the other genotypes because of variations among HCV genotypes. E2 protein is the major target for host immunity and vaccine development. The vaccine inhibits the viral entry into the hepatocytes and blocks the viral multiplication. Thus, prevents the viral infection that may develop in hepatitis C and liver cirrhosis (Zeisel *et al.*, 2011).

The present study is an attempt to produce antibodies using recombinant HCV protein E2. The advantage of using this method is that vaccines produced by recombinant proteins are non-pathogenic and the organism culture is not required (Swadling *et al.*, 2013). In recombinant vaccines, a number of effective viral epitopes are present to elicit immune response responsible for protective immunity. Sometimes, the immune response can be developed by recombinant proteins themselves only, while in some adjuvants are required.

The challenge to develop an efficient vaccine may need to target another sequence of HCV genome. However, it may be difficult as mutation rate in HCV genome is very high. A sequence to prevent further growth is to be found that matches the highly conserved sequences among different HCV genomes. In this study, complete genomes of HCV3a taken from various countries worldwide through Los-Alamos database were selected. The most conserved region of E2 amongst all the HCV3a genomes was taken into account. It can be effective against HCV in different countries of the world.

Keeping the reaction conditions same, it was observed that cDNA synthesized by using random hexamers rather than gene specific primer proved to be better for downstream amplification of gene. An enhanced sensitivity was observed in the synthesis of cDNA using random hexamers than the gene specific auto-sense primers (Radhakrishnan *et al.*, 1999). In the present study, E2 protein expressed in *E. coli* was obtained in insoluble fraction in form of inclusion bodies. E2 was expressed as insoluble form of protein.

The inclusion bodies are protein aggregates having non-native conformation. The biological functions of the gene can be studied by biologically active proteins for the synthesis of therapeutic drugs. The expressed E2 protein was purified by repeated sonications to remove membranous protein. These were later subjected to denaturation and protein refolding methods through fractional dialysis with urea to generate active soluble

recombinant protein. In this technique, the amount of protein aggregation was reduced. The protein aggregates were dissolved in high concentration of urea, used as a denaturant, to reduce the non-covalent interactions between protein molecules. Another chemical reagent dithiothreitol was added to diminish the unwanted inter and intra molecular disulfide bonds between cysteines present in the protein. The denaturant was removed by the refolding process to convert the denatured or unfolded protein to its pure active or folded form. E2 was stable, functional and antigenically active.

The purified protein E2 was confirmed with human HCV3a infected sera through ELISA. It showed that the HCV E2 purified proteins bound with the human sera almost three times more effectively than the positive control. These can be used in diagnostic purpose to test the presence of anti HCV antibodies in HCV infected human serum.

The protein was injected into the rabbit to develop antibodies. Provision of suitable animal model for vaccine trials is an important approach. In this study, rabbit model was used to assess the immunogenicity of HCV envelope glycoprotein E2. Rabbits have comparatively greater advantages over mice. They are frequently used in the testing of drugs, medical devices and vaccines because of genetic heterogeneity (Sominskaya *et al.*, 2015). E2 specific rabbit antisera were used to passively immunize the chimpanzees against HCV (Farci *et al.*, 1996). Rabbit antiserum and E2 protein may be used for the development of prophylactic, diagnostic and therapeutic effects against HCV (Liu *et al.*, 2005). Rabbit immunization with HCV E2 proteins could generate anti-E2 antibodies. Immunogenicity of E2 can be enhanced and antisera can be used for further vaccine development against HCV (Xiang *et al.*, 2006).

Freund's incomplete adjuvant was used in recombinant E2 glycoproteins to produce cellular and humoral responses against HCV in mice (Zhu *et al.*, 2013). In the present study, E2 recombinant protein was used in combination with Freund's complete adjuvant in 0 dose and Freund's incomplete adjuvant in 1st and 2nd booster dose and antibodies were raised successfully in rabbit which were confirmed through ELISA.

The rabbit antibodies produced against the E2 protein used as antigens were highly specific. *In vivo* antigenic activity of this purified recombinant protein was analyzed. The antibodies raised in 100% pure serum sample were 15 times more for E2, as compared to the control sera. Thus the purified protein E2 can be used as a vaccine against hepatitis C.

Formation of an active vaccine with less side effects is need of the hour. This study is a milestone in developing

a hepatitis free society. This may lead to the innovation of vaccine types and to cut the HCV burden worldwide.

ACKNOWLEDGEMENTS

Financial support by Higher Education Commission and all the facilities provided by School of Biological Sciences, University of the Punjab, Lahore are highly gratefully acknowledged.

Statement of conflict of interest

Authors have declared no conflict of interest.

REFERENCES

- Beeck, A. and Dubuisson, J., 2003. Topology of hepatitis C virus envelope glycoproteins. *Rev. med. Virol.*, **13**: 233-241. <https://doi.org/10.1002/rmv.391>
- Broering, T.J., Garrity, K.A., Boatright, N.K., Sloan, S.E., Sandor, F., Thomas, W.D. Jr., Szabo, G., Finberg, R.W., Ambrosino, D.M. and Babcock, G.J., 2009. Identification and characterization of broadly neutralizing human monoclonal antibodies directed against the E2 envelope glycoprotein of hepatitis C virus. *J. Virol.*, **83**: 12473-12482. <https://doi.org/10.1128/JVI.01138-09>
- Choo, Q.L., Kuo, G., Ralston, R., Weiner, A., Chien, D., Van Nest, G., Han, J., Berger, K., Thudium, K. and Kuo, C., 1994. Vaccination of chimpanzees against infection by the hepatitis C virus. *Proc. natl. Acad. Sci. U.S.A.*, **91**: 1294-1298. <https://doi.org/10.1073/pnas.91.4.1294>
- Farci, P., Shimoda, A., Wong, D., Cabezon, T. De Gioannis, D., Strazzer, A., Shimizu, Y., Shapiro, M., Alter, H.J. and Purcell, R.H., 1996. Prevention of hepatitis C virus infection in chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. *Proc. natl. Acad. Sci. U.S.A.*, **93**: 15394-15399. <https://doi.org/10.1073/pnas.93.26.15394>
- Goffard, A. and Dubuisson, J., 2003. Glycosylation of hepatitis C virus envelope glycoproteins. *Biochimie*, **85**: 295-301. [https://doi.org/10.1016/S0300-9084\(03\)00004-X](https://doi.org/10.1016/S0300-9084(03)00004-X)
- Idrees, M. and Riazuddin, S., 2008. Frequency distribution of hepatitis C virus genotypes in different geographical regions of Pakistan and their possible routes of transmission. *BMC Infect. Dis.*, **8**: 69-77. <https://doi.org/10.1186/1471-2334-8-69>
- Idrees, M., Rafique, S., Rehman, I., Akbar, H., Yousaf, M.Z., Butt, S., Awan, Z., Manzoor, S., Akram, M., Aftab, M., Khubaib, B. and Riazuddin, S.,

2009. Hepatitis C virus genotype 3a infection and hepatocellular carcinoma: Pakistan experience. *World J. Gastroenterol.*, **15**: 5080-5085. <https://doi.org/10.3748/wjg.15.5080>
- Johansson, D.X., Voisset, C., Tarr, A.W., Aung, M., Ball, J.K., Dubuisson, J. and Persson, M.A., 2007. Human combinatorial libraries yield rare antibodies that broadly neutralize hepatitis C virus. *Proc. natl. Acad. Sci. U.S.A.*, **104**: 16269–16274. <https://doi.org/10.1073/pnas.0705522104>
- Kato, N., 2000. Genome of human hepatitis C Virus (HCV): Gene organization, sequence diversity, and variation. *Microb. Comp. Genom.*, **5**: 129-151. <https://doi.org/10.1089/mcg.2000.5.129>
- Keck, Z.Y., Li, T.K., Xia, J., Gal-Tanamy, M., Olson, O., Li, S.H., Patel, A.H., Ball, J.K., Lemon, S.M. and Fong, S.K.H., 2008. Definition of a conserved immunodominant domain on hepatitis C virus E2 glycoprotein by neutralizing human monoclonal antibodies. *J. Virol.*, **82**: 6061–6066. <https://doi.org/10.1128/JVI.02475-07>
- Law, M., Maruyama, T., Lewis, J., Giang, E., Tarr, A.W., Stamatakis, Z., Gastaminza, P., Chisari, F.V., Jones, I.M., Fox, R.I., Ball, J.K., McKeating, J.A., Kneteman, N.M. and Burton, D.R., 2008. Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge. *Nat. Med.*, **14**: 25–27. <https://doi.org/10.1038/nm1698>
- Lechmann, M. and Liang, T.J., 2000. Vaccine development for hepatitis C. *Semin. Liver Dis.*, **20**: 211-226. <https://doi.org/10.1055/s-2000-9947>
- Liu, J., Zhu, L.X., Kong, Y.Y., Li, G.D. and Wang, Y., 2005. Purification and application of C-terminal truncated hepatitis C virus E1 proteins expressed in *Escherichia coli*. *World J. Gastroenterol.*, **11**: 503-507. <https://doi.org/10.3748/wjg.v11.i4.503>
- Owsianka, A., Tarr, A.W., Juttla, V.S., Lavillette, D., Bartosch, B., Cosset, F.L., Ball, J.K. and Patel, A.H., 2005. Monoclonal antibody AP33 defines a broadly neutralizing epitope on the hepatitis C virus E2 envelope glycoprotein. *J. Virol.*, **79**: 11095-11104. <https://doi.org/10.1128/JVI.79.17.11095-11104.2005>
- Radhakrishnan, S., Abraham, P. and Sridharan, G., 1999. Improved sensitivity of reverse transcriptase polymerase chain reaction of hepatitis C virus using random hexamer primers. *Diagnos. Microbiol. Infect. Dis.*, **33**: 153-156. [https://doi.org/10.1016/S0732-8893\(98\)00152-7](https://doi.org/10.1016/S0732-8893(98)00152-7)
- Raja, N. S. and Janjua, K. A., 2008. Epidemiology of hepatitis C virus infection in Pakistan. *J. Microbiol. Immunol. Infect.*, **41**: 4-8. <https://doi.org/10.1615/CritRevEukaryotGeneExpr.2017018953>
- Sambrook, J. and Russel, W.D., 2001. *Molecular cloning: A laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Schofield, D.J., Bartosch, B., Shimizu, Y.K., Allander, T., Alter, H.J., Emerson, S.U., Cosset, F.L. and Purcell, R.H., 2005. Human monoclonal antibodies that react with the E2 glycoprotein of hepatitis C virus and possess neutralizing activity. *Hepatology*, **42**: 1055–1062. <https://doi.org/10.1002/hep.20906>
- Sominskaya, I., Jansons, J., Dovbenko, A., Petrakova, N., Lieknina, I., Mihailova, M., Latyshev, O., Eliseeva, O., Stahovska, I., Akopjana, I., Petrovskis, I. and Isagulians, M., 2015. Comparative immunogenicity in rabbits of the polypeptides encoded by the 5' terminus of hepatitis C virus RNA. *J. Immunol. Res.*, **2015**: Article ID 762426. <https://doi.org/10.1155/2015/762426>
- Swadling, L., Klenerman, P. and Barnes, E., 2013. Ever closer to a prophylactic vaccine for HCV. *Expert Opin. Biol. Ther.*, **13**: 1109–1124. <https://doi.org/10.1517/14712598.2013.791277>
- Weiner, A.J., Christopherson, C., Hall, J.E., Bonino, F., Saracco, G., Brunetto, M.R., Crawford, K., Marion, C.D., Crawford, K.A. and Venkatakrishna S., 1991. Sequence variation in hepatitis C viral isolates. *J. Hepatol.*, **13**: 6–14. [https://doi.org/10.1016/0168-8278\(91\)90015-4](https://doi.org/10.1016/0168-8278(91)90015-4)
- Xiang, Z.H., Cai, W.J., Zhao, P., Kong, L.B., Ye L.B. and Wu, Z.H., 2006. Purification and application of bacterially expressed chimeric protein E1 E2 of hepatitis C virus. *Protein Expr. Purif.*, **49**: 95-101. <https://doi.org/10.1016/j.pep.2006.02.013>
- Zeisel, M.B., Fofana, I., Fafi-Kremer, S. and Baumert, T.F., 2011. Hepatitis C virus entry into hepatocytes: Molecular mechanisms and targets for antiviral therapies. *J. Hepatol.*, **54**: 566-576. <https://doi.org/10.1016/j.jhep.2010.10.014>
- Zhu, W., Fu, J., Lu, J., Deng, Y., Wang, H., Wei, Y., Deng, L., Tan, W. and Liang, G., 2013. Induction of humoral and cellular immune responses against hepatitis C virus by vaccination with replicon particles derived from Sindbis-like virus XJ-160. *Arch. Virol.*, **158**: 1013–1019. <https://doi.org/10.1007/s00705-012-1564-8>