Hepat Mon. 2014 November; 14(11): e23660. DOI: 10.5812/hepatmon.23660

Published online 2014 November 22. **Research Article**

Characterization of HCV Genotype 5a Envelope Proteins: Implications for Vaccine Development and Therapeutic Entry Target

Maemu Petronella Gededzha [1,*]; Maphahlanganye Jeffrey Mphahlele [1]; Selokela Gloria Selabe [1]

[1]Department of Virology, HIV and Hepatitis Research Unit, University of Limpopo, Medunsa Campus/National Health Laboratory Service, Pretoria, South Africa

[*]*Corresponding Author*: Maemu Petronella Gededzha, Department of Virology, HIV and Hepatitis Research Unit, University of Limpopo, Medunsa Campus/National Health Laboratory Service, Pretoria, South Africa. Tel: +27-125213631, Fax: +27-125215794, E-mail: maemup@yahoo.com

Received: September 15, 2014; **Revised:** October 8, 2014; **Accepted:** October 26, 2014

Background: Hepatitis C virus (HCV) is one of the major causes of cirrhosis and hepatocellular carcinoma with an estimation of 185 million people with infection. The E2 is the main target for neutralizing antibody responses and the variation of this region is related to maintenance of persistent infection by emerging escape variants and subsequent development of chronic infection. While both E1 and E2 are hypervariable in nature, it is difficult to design vaccines or therapeutic drugs against them.
Objectives: The objective of this study was to characterize genotype 5a E1 and E2 sequences to determine possible glycosylation sites, conserved B-cell epitopes and peptides in HCV that could be useful targets in design of vaccine and entry inhibitors.
Patients and Methods: This study was conducted through PCR amplification of E1 and E2 regions, sequencing, prediction of B-cell epitopes, analysis of N-linked glycosylation and peptide design in 18 samples of HCV genotype 5a from South African.
Results: Differences in the probability of glycosylation in E1 and E2 regions were observed in this study. Three conserved antigenic B-cell epitopes were predicted in the E2 regions and also 11 short peptides were designed from the highly conserved residues.
Conclusions: This study provided conserved B-cell epitopes and peptides that can be useful for designing entry inhibitors and vaccines able to cover a global population, especially where genotype 5a is common.

Keywords: Hepatitis C Virus; Genotype; Epitopes; Peptides

1. Background

Globally, an estimated 185 million people have been infected with hepatitis C virus (HCV) as one of the major causes of cirrhosis and hepatocellular carcinoma (1). HCV genome consists of approximately 9.6 kilobases, positive-sense single-stranded RNA, which encodes three structural (C, E1 and E2) and 7 non-structural (p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) proteins flanked by 5' and 3' untranslated regions (UTR) (2). E1 and E2 proteins are type I transmembrane proteins with both N-terminal ectodomain and a C-terminal domain (3) and contain 6 and 11 glycosylation sites, respectively (4, 5). These proteins are involved in viral entry by interacting with CD81 and Scavenger receptor class B member 1 (SRB1) (6-8). HCV glycosylation sites play an essential role in envelope proteins to ensure correct conformation for virus entry (5, 9) and antigenic variation (10). HCV E2 glycosylation sites interact with cell surface receptors directly allowing the virus to enter the cell (11, 12). Glycosylation sites may mask important epitopes from host antibody responses (13, 14). B-cell epitopes are essential in increasing the preferred immune responses (15, 16) and number of epitopes and modulation of immune recognition of antigens can be influenced by deglycosylation of E1 proteins (17). The E1 derived peptide p35 (amino acid (aa) 315–323) (18), E2-conserved synthetic peptides p37 (aa 517–531) and p38 (aa 412–419) have been reported to neutralize HCV particles, as important components of a candidate peptide vaccine (19). The molecular targets for current HCV Direct-acting antiviral (DAA) in development are mainly focused on non-structural proteins such as the NS3 protease, NS5A and the NS5B RdRp (20). Recently, considerable progress has been made to understand HCV entry (21, 22) and development of entry inhibitors (20, 21, 23, 24). Many patients do not respond to the current available therapy, therefore, there is an urgent need to develop effective HCV vaccines and specific therapeutic drugs. While both E1 and E2 are hypervariable in nature, it is difficult to design vaccines or therapeutic drugs against them. Genotype 5a accounts for over 50% of HCV infections in South Africa (25).

2. Objectives

This study aimed to characterize genotype 5a E1 and E2 sequences to determine possible glycosylation sites, conserved B-cell epitopes and peptides in HCV that could be useful targets in the design of vaccine and entry inhibitors.

3. Patients and Methods

3.1. Study Population

This study included 18 genotype 5a samples collected from treatment-naive HCV infected patients at Dr. George Mukhari Academic Hospital (DGMAH), north-west of Pretoria, South Africa, from 2007 to 2011. Patients' demographics and genotyping based on 5'UTR were previously described in detail (25). Six of 18 samples were sequenced as part of the genotype 5a near-full length analysis previously described (26). DGMAH is an academic hospital serving a population of around 4 million from both rural and urban areas. It is a referral hospital for patients from the North West, Mpumalanga, Limpopo and the north-west part of Pretoria, Gauteng. The Medunsa Research and Ethics Committee approved the study.

3.2. PCR and Sequencing

Viral RNA was extracted from 140 μL of serum using the QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. HCV RNA was converted into cDNA using the enzyme RevertAid TM RT-PCR (Fermentas, Vilnius, Lithuiana). The cDNA was amplified in three overlapping fragments (Table 1) covering complete E1 and E2 regions. Direct sequencing was performed with ABI 3500XL (Inqaba Biotechnological Industry, PTY, Ltd, Pretoria, South Africa) using second round PCR primers. Sequence fragment assembly was performed using Chromas Pro1.5 (www.technelysium.com.au/chromas.html). All sequences were aligned by Mafft (mafft.cbrc.jp/alignment/server/) and translated into amino acids using BioEdit (27).

3.3. Analysis of N-Linked Glycosylation Sites

The N glycosylation sites were predicted using the online prediction server NetNGlyc version 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/), which predicts N glycosylation sites in proteins by artificial neural networks that examine the sequence context of Asn-Xaa-Ser/Thr sequins. The networks can identify 86% of the glycosylated and 61% of the non-glycosylated sequins, with an overall accuracy of 76%.

3.4. Prediction of B-Cell Epitopes

For identification of B-cell epitopes, 16-mer B-cell epitopes was predicted using the program ABCpred (http://www.imtech.res.in/raghava/abcpred/) at a 0.51 default threshold using a consensus sequence from 18 genotype 5a sequences created using Bioedit. ABCpred server predicts B-cell epitopes using artificial neural network using fixed length patterns (28). Antigenicity of all predicted epitopes was analyzed using VaxiJen v2.0 online antigen prediction (www.ddg-pharmfac.net/vaxijen/). Proteins having antigenic score more than 0.4 were selected as antigenic. VaxiJen v2.0 allows antigen classification based on physicochemical properties of proteins without recourse to sequence alignment. All predicted epitopes were analyzed for conservation using the IEDB database (http://tools.immuneepitope.org/tools/conservancy/iedb_input) at a threshold of 100% conservation compared to 406, 221, 98, 33, 45, 45 randomly selected sequences from each of the HCV genotypes 1a, 1b, 2, 3, 4 and 6, respectively.

3.5. Peptide Design

Structure analysis of sequence was performed using the Protparam online tool (29). Protparam computed different parameters including the molecular weight, theoretical pI, AA composition, atomic composition, extinction coefficient, instability index, aliphatic index and grand average of hydropathicity (GRAVY). To check post-translational modifications, predicted peptides were predicted for N-linked glycosylation as described above and for N-linked phosphorylation using the NetPhos 2.0 (30) program. The NetPhos 2.0 produces neural network predictions for serine, threonine and tyrosine phosphorylation sites in sequences. Only those motifs with NetPhos score of 0.7 or greater were considered.

3.6. GenBank Accession Numbers

Sequences were submitted to GenBank under the accession numbers KC7678835 - KC767846.

Table 1. Sequences of the HCV Primers Used in This Study

Sequences		Primers	Reference
A			this study
F1A	1088	GAC CAT TTC ATC ATC ATG TCC CA	
R1A	1425	TGT ATG CGG CGG CGA ACA AGA CC	
F2A	1113	CTT CGG AGG GCC GTT GAC TAC TTA GCG	
R2A	1413	CGA ACA AGA CCC CCC AGT GGG	
B			
M105	1292	ATG GCA TGG GAC ATG ATG ATG	(27)
R1B	2061	TAG GCC CTA AGT TGC AGG GTG GA	this study
M106	1298	TGG GAC ATG ATG ATG AAT TGG	(27)
R2B	2022	CAA ACC CTG TGG AAT TCA TCC AG	this study
C			this study
F1C	1743	GGC TGG GGA ACT ATC AGC TAT	
R1C	2636	AAA CCC ATG AGT CCC CGC AGC C	
F2C	1773	TCG GGC CCC AGT GAT GAC AAG	
R2C	2612	AGC CGC GTT TAG GAC AAT GAC GTT CT	

4. Results

4.1. Sequence Alignment and Genetic Distances

Sequence alignment of 18 genotype 5a sequences with a reference sequence from the GenBank showed that most regions in the genotype 5a E1 and E2 proteins were conserved except hypervariable 1 (HVR1), which was highly variable as expected. Comparison of genetic distances between sequences in this study showed intragroup genetic distances ranging from 8% to 17%, with an average distance of 13% (Table 2).

4.2. Analysis of E1 and E2 N-Linked Glycosylation

E1 and E2 proteins of 18 sequences were analyzed for possible glycosylation sites. Differences in the probability of glycosylation in E1 and E2 were observed in most sequences. Whereas other studies reported five N-linked glycosylation sites in the E1 region, all strains in the current study showed three or four glycosylation sites, except for ZADGM2088, which showed 2 glycosylation sites, with N325 site not predicted as glycosylation sites from all sequences. In the E2 region, three sequences (ZADGM1104, ZADGM1707 and ZADGM3013) showed nine glycosylation sites, while the remaining had variations in the number of glycosylation sites. In ZADGM308, position N430 was replaced by H, while in ZADGM6544, N448 was replaced by D. Site N476 was found in only 6 of analyzed 18 sequences. The E2 sites N423 and N576 were not predicted as glycosylation sites in all genotype 5a sequences in this study (Table 3).

4.3. B-Cell Epitopes Prediction

Three conserved antigenic B-cell epitopes were predicted for genotype 5a sequences in the E2 region. Epitope $E2^{504-609}$ (GPVYCFTPSPVVVGTT) had the highest antigenic score of 1.1613, while $E2^{675-690}$ (LPCSFTPTPALSTGLI) and $E2^{685-700}$ (LSTGLIHLHQNIVDTQ) had antigenic scores of 0.5340 and 0.6639, respectively. For conservancy analysis, epitope $E^{2504-609}$ was highly conserved among other genotypes, while epitope $E^{2675-690}$ and $E^{2685-700}$ were variable (Table 4).

Table 2. Genetic Distances in E1 and E2 Sequences of Genotype 5a in This Study

		Genetic Distances [a, b]																	
	Sequence	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1	ZADGM7890																		
2	ZADGM6544	0.12																	
3	ZADGM4227	0.13	0.12																
4	ZADGM1908	0.14	0.12	0.13															
5	ZADGM1707	0.10	0.12	0.11	0.12														
6	ZADGM651	0.13	0.13	0.14	0.13	0.12													
7	ZADGM308	0.12	0.12	0.12	0.12	0.10	0.11												
8	ZADGM6485	0.13	0.10	0.13	0.12	0.12	0.12	0.10											
9	ZADGM4124	0.13	0.13	0.13	0.13	0.12	0.14	0.13	0.12										
10	ZADGM2439	0.14	0.15	0.15	0.15	0.12	0.14	0.12	0.13	0.14									
11	ZADGM2352	0.13	0.14	0.14	0.14	0.13	0.12	0.11	0.12	0.14	0.13								
12	ZADGM525gp	0.14	0.14	0.15	0.14	0.12	0.15	0.14	0.14	0.14	0.16	0.16							
13	ZADGM869	0.14	0.14	0.15	0.13	0.13	0.10	0.13	0.13	0.14	0.14	0.14	0.17						
14	ZADGM3013	0.15	0.14	0.14	0.15	0.12	0.15	0.13	0.12	0.15	0.15	0.15	0.17	0.15					
15	ZADGM0518	0.9	0.12	0.11	0.12	0.08	0.13	0.10	0.11	0.11	0.12	0.13	0.14	0.13	0.11				
16	ZADGM2582	0.14	0.14	0.13	0.14	0.13	0.13	0.12	0.13	0.14	0.13	0.13	0.16	0.15	0.15	0.13			
17	ZADGM2088	0.14	0.13	0.13	0.13	0.12	0.12	0.11	0.12	0.13	0.14	0.12	0.13	0.14	0.14	0.12	0.12		
18	ZADGM1104	0.13	0.14	0.13	0.14	0.12	0.13	0.09	0.12	0.14	0.14	0.14	0.13	0.15	0.15	0.12	0.14	0.13	

[a] The values range between 0 (0%) and 1 (100%) substitutions per nucleotide site.
[b] The numbers 1-18 corresponds to the sequence number on the vertical side.

Table 3. Probability of Glycosylation in E1 and E2 Sequences

		Probability at Glycosylation Site [a, b]															
	Sequence	E1					No of Sites	E2									No of Sites
		196	209	234	305	325		417	430	448	476	533	541	557	623	645	
1	ZADGM7890	+	++	++	+	-	4	+	++	-	-	+	++	+	+	+	7
2	ZADGM6544	+	++	+	-	-	3	+	++	-	-	-	++	+	+	+	6
3	ZADGM4227	+	++	++	+	-	4	++	++	-	-	+	-	+	+	+	6
4	ZADGM1908	+	++	+	-	-	3	++	+	-	+	++	+	+	+	-	7
5	ZADGM1707	+	++	+	+	-	4	++	+	+	+	+	++	+	+	+	9
6	ZADGM651	+	++	++	+	-	4	++	+	-	-	+	+	+	+	-	6
7	ZADGM308	+	++	++	-	-	3	++	-	+	-	+	++	+	+	-	6
8	ZADGM6485	+	++	++	-	-	3	++	++	-	+	+	++	+	+	-	7
9	ZADGM4124	+	++	+	+	-	4	++	+	-	-	+	+	+	+	+	7
10	ZADGM2439	+	++	+	+	-	4	++	+	-	+	+	+	+	+	-	7
11	ZADGM2352	+	++	++	-	-	3	++	++	-	-	+	++	+	+	-	6
12	ZADGM525gp	-	++	++	+	-	3	++	+	-	-	-	+	+	+	+	6
13	ZADGM869	+	++	+	-	-	3	+	++	-	-	+	+	+	+	+	7
14	ZADGM3013	+	++	+	+	-	4	++	++	++	+	+	++	+	+	+	9
15	ZADGM0518	+	++	+	-	-	3	++	++	++	-	-	++	+	+	-	6
16	ZADGM2582	+	++	+	-	-	3	++	+	-	+	+	++	+	+	+	8
17	ZADGM2088	-	++	++	-	-	2	++	+	-	-	+	+	+	+	-	6
18	ZADGM1104	+	++	++	-	-	3	++	++	+	+	+	+	+	+	+	9

[a] Numbering is based on the M62321 full-length sequence.
[b] Glycosylation probability is shown by +++ (probability > 70%), ++ (probability between 60 and 70%), + (probability between 50 and 60%), and - (not predicted).

Table 4. Predicted B-Cell Epitopes of HCV Genotype 5a and Their Antigenicity Score, Number of Allele and Conservancy (Percentage) in Different Genotypes

Position	Predicted Epitopes	Antigen score	Genotype 1a	Genotype 1b	Genotype 2	Genotype 3	Genotype 4	Genotype 6
504	GPVYCFTPSPVVVGTT	1.1613	92	97	88	94	89	73
675	LPCSFTPTPALSTGLI	0.5340	0	0	0	0	0	0
685	LSTGLIHLHQNIVDTQ	0.6639	0	0	0	0	0	2

Table 5. Predicted Peptides for HCV E1 and E2 Conserved in Genotype 5a

Position [a]	Peptides	Length	Molecular Weight	Theoretical PI	Extinction Coefficient (/cm M)	Instability Index	Aliphatic index	GRAVY	Composition of Hydrophobic AA's, % [b]	N-linked Glycosylation C [c]	N-linked Phosphorylation
201	YHTNDCPNSSI	14	1611.7	5.08	2980	17.36	69.29	-0.343	21.4	+	-
262	VDYLAGGAA	9	835.9	3.80	1490	-3.53	108.89	0.867	22.2	-	-
304	CNCSIYSGH	9	983	6.72	1615	5.69	43.33	-0.056	11.1	++	-
314	TGHRMAWDMMMNWSPT	16	1952.2	6.41	11000	29.16	-0.706	6.25	37.5	-	-
352	HWGVLFAAAY	10	1134.3	6.74	6990	-4.25	98	1.040	40	-	-
562	VKTCGAPPC	9	875	8.03	125	30.68	43.33	0.311	11.1%	-	-
585	TDCFRKHP	8	1003.1	7.92	0	5.15	0	-1.512	12.5	-	-
645	ACNWTRGERCDL	12	1423.5	6.1	5625	27.31	40.83	-0.908	16.7	+	-
664	LSPLLHTTTQ	10	1110.2	6.74	-	37.86	117	0.020	30%	-	-
675	AILPCSFTPTPALSTGLIHL-HQNIVDTQ	28	2988.4	5.97	0	28.19	115	0.421	32.1	-	-
725	FLLLADAR	8	918.1	5.84	-	-1.86	171.25	1.225	50	-	-

[a] Numbering is based on the M62321 full-length sequence.
[b] list of hydrophobic amino acids (Leu, Val, Ile, Met, Phe and Trp).
[c] Glycosylation probability is shown by +++ (probability > 70%), ++ (probability between 60 and 70%), + (probability between 50 and 60%), and - (not present).

4.4. Peptide Design

From the consensus sequences of genotype 5a E1 and E2, eleven short peptides of 8-28 amino acids were designed from the highly conserved residues. Five peptides of 9-16 amino acids in length were derived in the E1 region, while six peptides of 8-26 amino acids were derived in the E2. Three of the peptides had post-translation modification, which is the N-linked glycosylation, although at a low probability. None of the peptides has either serine, threonine and tyrosine phosphorylation sites predicted. Most peptides were found to be the best predicted peptides useful for designing entry inhibitors (Table 5).

5. Discussion

Genotype 5 is the most conserved HCV genotype classified into only one subtype (5a) (26). This study was designed to identify conserved sequences of these proteins to predict antigenic epitopes and peptides that could serve as best targets for vaccine design and potential entry inhibitors. Using different structural and sequence analyses tools helped with in-silico analysis for E1 and E2 regions. HCV genotype 5a sequences were found to be conserved in most regions of E1 and E2 proteins. The most variable region within the study sequences was the HVR1 and these HVR1 differed by up to 80% between HCV genotypes and subtypes (31). Although highly variable, the HVR1 is the only region that contains neutralization determinant, which is the target for immune response (32). As expected due to HVR variability, comparison of genetic distances between sequences in this study showed high genetic distances ranging from 8% to 17%, with an average distance of 13%. Variability within the HVR1 is one of the reasons describing why human antibodies raised against HCV E2 epitopes do not provide protection against multiple viral infections (19). In this study, analysis of N-linked glycosylation sites revealed that genotype 5a sequences were not conserved at glycosylation sites as compared to other genotypes. Site N476 with a level of 75% conservation among different genotypes was absent from the sequences of genotype 5a (5) and was found in six of the 18 analyzed sequences. As reported previously, E2 sites N423 and N576 were absent in all genotype 5a sequences including the 18 sequences from this study, which is notable because these two sites were reported to be 99-100% conserved across all genotypes (5). The glycosylation sites were reported to be highly conserved among different genotypes (9). These sequence variations in genotype 5a glycosylation sites could be useful to design efficient vaccine to help host to produce good antibody response. E2 is the main target for neutralizing antibody responses and variation of this region is thought to be related to maintenance of persistent infection by emerging escape variants and subsequent development of chronic infection (33, 34). Recently, a linear region of E2 encompassing amino acids 434 to 446 has been reported to elicit non-neutralizing antibodies that can inhibit neutralizing activity of antibodies targeting amino acids 412 to 423 (35). However, a study by Tarr et al. reported conflicting results showing that human antibodies that target the region encompassing amino acids 434 to 446, are not inhibitory but capable of neutralizing HCVpp and HCVcc entry (36). All B-cell epitopes included in this study were found to be antigenic ally effective, and it can be implied that these epitopes may be important for inducing the desired immune response. The $E^{2504-609}$ epitope was found to be the most conserved among other genotypes. Recently a study by Ikram et al. reported conserved epitopes among genotype 3a that was also conserved among other genotypes (37). Highly conserved epitopes might influence the immunogenic potential since variability within the epitopes can increase the chance of immune escape (38). Short polypeptides derived from viral envelope sequences of other viruses have been used to investigate protein interactions involved in viral entry and some antiviral agents have been successfully developed (39). Envelope protein peptide inhibitors for other viruses in the same family with HCV like Dengue and West Nile were shown to inhibit viral entry (40, 41). In HCV, the post-binding entry step was prevented using peptides derived from the C terminal region of E2, which plays an important role in the HCV entry process (42). For this study, conserved peptides were derived that can be used as targets for therapeutic purposes. In this study, only three peptides had glycosylation sites at low probability and no phosphorylation sites were predicted. Post translational modifications such as glycosylation and phosphorylation affect the stability of therapeutic peptides (43). Using HCV glycoproteins in therapeutic strategies may offer protection against HCV infection (44). In conclusion, genotype 5a sequences are conserved and can be used to design epitopes and peptides. The results showed that antigenic conserved predicted B-cell epitopes and stable peptides with few post-translational modifications. These epitopes and peptides are potential candidates to design entry inhibitors and vaccines able to cover a global population, especially where genotype 5a is common. Further investigations would analyze these peptides to better understand their involvement in blocking HCV entry.

Acknowledgements

We would like to acknowledge all members of the HIV and Hepatitis Research Unit, Department of Virology, Pretoria, South Africa for their technical help and writing assistance.

Authors' Contributions

Maemu Petronella Gededzha designed the study, performed lab work, analyzed data and drafted the manuscript. Selokela Gloria Selabe and Maphahlanganye Jeffrey Mphahlele supervised the project and critically reviewed the manuscript. All authors read and approved the final manuscript.

Funding/Support

This work was supported by grants from Medical Research Council, National Research Foundation, National Health Laboratory Service Research Trust, and The Stella & Paul Loewestein Charitable and Educational Trust in South Africa.

References

1. World health Organization.. *Guidelines for the screening care and treatment of persons with hepatitis C infection.*: WHO; 2014. Available from: http://www.who.int/hiv/pub/hepatitis/hepatitis-c-guidelines/en/.
2. Lemon SM, Walker C, Alter MJ, Yi M. Hepatitis C virus. In: Knipe DM, Howley PM editors. *Fields Virology.*: Lippincott Williams and Wilkins; 2007. pp. 1253–304.
3. Chevaliez S, Pawlotsky JM. HCV Genome and Life Cycle. In: Tan SL editor. *Hepatitis C Viruses: Genomes and Molecular Biology.*. Norfolk UK: Horizon Bioscience; 2006. pp. 45–7.
4. Drummer HE, Maerz A, Poumbourios P. Cell surface expression of functional hepatitis C virus E1 and E2 glycoproteins. *FEBS Lett.* 2003;**546**(2-3):385–90.
5. Goffard A, Dubuisson J. Glycosylation of hepatitis C virus envelope proteins. *Biochimie.* 2003;**85**(3-4):295–301.
6. Higginbottom A, Quinn ER, Kuo CC, Flint M, Wilson LH, Bianchi E, et al. Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J Virol.* 2000;**74**(8):3642–9.
7. Scarselli E, Ansuini H, Cerino R, Roccasecca RM, Acali S, Filocamo G, et al. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J.* 2002;**21**(19):5017–25.
8. Ashfaq UA, Qasim M, Yousaf MZ, Awan MT, Jahan S. Inhibition of HCV 3a genotype entry through host CD81 and HCV E2 antibodies. *J Transl Med.* 2011;**9**:194.
9. Brown RJ, Tarr AW, McClure CP, Juttla VS, Tagiuri N, Irving WL, et al. Cross-genotype characterization of genetic diversity and molecular adaptation in hepatitis C virus envelope glycoprotein genes. *J Gen Virol.* 2007;**88**(Pt 2):458–69.
10. Slater-Handshy T, Droll DA, Fan X, Di Bisceglie AM, Chambers TJ. HCV E2 glycoprotein: mutagenesis of N-linked glycosylation sites and its effects on E2 expression and processing. *Virology.* 2004;**319**(1):36–48.
11. Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, Petracca R, et al. Binding of hepatitis C virus to CD81. *Science.* 1998;**282**(5390):938–41.
12. Helle F, Dubuisson J. Hepatitis C virus entry into host cells. *Cell Mol Life Sci.* 2008;**65**(1):100–12.
13. Schonning K, Jansson B, Olofsson S, Hansen JE. Rapid selection for an N-linked oligosaccharide by monoclonal antibodies directed against the V3 loop of human immunodeficiency virus type 1. *J Gen Virol.* 1996;**77 (Pt 4)**:753–8.
14. Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, et al. Antibody neutralization and escape by HIV-1. *Nature.* 2003;**422**(6929):307–12.
15. Bugli F, Mancini N, Kang CY, Di Campli C, Grieco A, Manzin A, et al. Mapping B-cell epitopes of hepatitis C virus E2 glycoprotein using human monoclonal antibodies from phage display libraries. *J Virol.* 2001;**75**(20):9986–90.
16. Giang E, Dorner M, Prentoe JC, Dreux M, Evans MJ, Bukh J, et al. Human broadly neutralizing antibodies to the envelope glycoprotein complex of hepatitis C virus. *Proc Natl Acad Sci U S A.* 2012;**109**(16):6205–10.
17. Fournillier A, Wychowski C, Boucreux D, Baumert TF, Meunier JC, Jacobs D, et al. Induction of hepatitis C virus E1 envelope protein-specific immune response can be enhanced by mutation of N-glycosylation sites. *J Virol.* 2001;**75**(24):12088–97.
18. El-Awady MK, Tabll AA, Atef K, Yousef SS, Omran MH, El-Abd Y, et al. Antibody to E1 peptide of hepatitis C virus genotype 4 inhibits virus binding and entry to HepG2 cells in vitro. *World J Gastroenterol.* 2006;**12**(16):2530–5.
19. El-Awady MK, Tabll AA, El-Abd YS, Yousif H, Hegab M, Reda M, et al. Conserved peptides within the E2 region of Hepatitis C virus induce humoral and cellular responses in goats. *Virol J.* 2009;**6**:66.
20. Baldick CJ, Wichroski MJ, Pendri A, Walsh AW, Fang J, Mazzucco CE, et al. A novel small molecule inhibitor of hepatitis C virus entry. *PLoS Pathog.* 2010;**6**(9).
21. Dorner M, Horwitz JA, Robbins JB, Barry WT, Feng Q , Mu K, et al. A genetically humanized mouse model for hepatitis C virus infection. *Nature.* 2011;**474**(7350):208–11.
22. Sainz BJ, Barretto N, Martin DN, Hiraga N, Imamura M, Hussain S, et al. Identification of the Niemann-Pick C1-like 1 cholesterol absorption receptor as a new hepatitis C virus entry factor. *Nat Med.* 2012;**18**(2):281–5.
23. Mittapalli GK, Zhao F, Jackson A, Gao H, Lee H, Chow S, et al. Discovery of ITX 4520: a highly potent orally bioavailable hepatitis C virus entry inhibitor. *Bioorg Med Chem Lett.* 2012;**22**(15):4955–61.
24. Zhu H, Wong-Staal F, Lee H, Syder A, McKelvy J, Schooley RT, et al. Evaluation of ITX 5061, a scavenger receptor B1 antagonist: resistance selection and activity in combination with other hepatitis C virus antivirals. *J Infect Dis.* 2012;**205**(4):656–62.
25. Gededzha MP, Selabe SG, Kyaw T, Rakgole JN, Blackard JT, Mphahlele MJ. Introduction of new subtypes and variants of hepatitis C virus genotype 4 in South Africa. *J Med Virol.* 2012;**84**(4):601–7.
26. Gededzha MP, Selabe SG, Blackard JT, Kyaw T, Mphahlele MJ. Near full-length genome analysis of HCV genotype 5 strains from South Africa. *Infect Genet Evol.* 2014;**21**:118–23.
27. Koletzki D, Dumont S, Vermeiren H, Peixe P, Nina J, Camacho RJ, et al. Full genome sequence of three isolates of hepatitis C virus subtype 4b from Portugal. *Arch Virol.* 2009;**154**(1):127-32.
28. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM editor. *The Proteomics Protocols Handbook.*. Humana Press; 2005. pp. 571–607.
29. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins.* 2006;**65**(1):40–8.
30. Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol.* 1999;**294**(5):1351–62.
31. Penin F, Dubuisson J, Rey FA, Moradpour D, Pawlotsky JM. Structural biology of hepatitis C virus. *Hepatology.* 2004;**39**(1):5-19
32. Genovese D, Dettori S, Argentini C, Villano U, Chionne P, Angelico M, et al. Molecular epidemiology of hepatitis C virus genotype 4 isolates in Egypt and analysis of the variability of envelope proteins E1 and E2 in patients with chronic hepatitis. *J Clin Microbiol.* 2005;**43**(4):1902-9.
33. Zhang P, Zhong L, Struble EB, Watanabe H, Kachko A, Mihalik K, et al. Depletion of interfering antibodies in chronic hepatitis C patients and vaccinated chimpanzees reveals broad cross-genotype neutralizing activity. *Proc Natl Acad Sci U S A.* 2009;**106**(18):7537–41.
34. Tarr AW, Urbanowicz RA, Jayaraj D, Brown RJ, McKeating JA, Irving WL, et al. Naturally occurring antibodies that recognize linear epitopes in the amino terminus of the hepatitis C virus E2 protein confer noninterfering, additive neutralization. *J Virol.* 2012;**86**(5):2739–49.
35. Manzin A, Solforosi L, Petrelli E, Macarri G, Tosone G, Piazza M, et al. Evolution of hypervariable region 1 of hepatitis C virus in primary infection. *J Virol.* 1998;**72**(7):6271–6.
36. Farci P, Alter HJ, Wong DC, Miller RH, Govindarajan S, Engle R, et al. Prevention of hepatitis C virus infection in chimpanzees after antibody-mediated in vitro neutralization. *Proc Natl Acad Sci U S A.* 1994;**91**(16):7792–6.
37. Ikram A, Anjum S, Tahir M. In Silico Identification and Conservation Analysis of B-cell and T-Cell Epitopes of Hepatitis C Virus 3a Genotype Enveloped Glycoprotein 2 From Pakistan: A Step Towards Heterologous Vaccine Design. *Hepat Mon.* 2014;**14**(6).
38. Yusim K, Kesmir C, Gaschen B, Addo MM, Altfeld M, Brunak S, et al. Clustering patterns of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1 (HIV-1) proteins reveal imprints of immune evasion on HIV-1 global variation. *J Virol.* 2002;**76**(17):8757–68.

39. Greenberg ML, Cammack N. Resistance to enfuvirtide, the first HIV fusion inhibitor. *J Antimicrob Chemother.* 2004;**54**(2):333–40.
40. Hrobowski YM, Garry RF, Michael SF. Peptide inhibitors of dengue virus and West Nile virus infectivity. *Virol J.* 2005;**2**:49.
41. Bai F, Town T, Pradhan D, Cox J, Ledizet M, Ashish., et al. Antiviral peptides targeting the west nile virus envelope protein. *J Virol.* 2007;**81**(4):2047–55.
42. Liu R, Tewari M, Kong R, Zhang R, Ingravallo P, Ralston R. A peptide derived from hepatitis C virus E2 envelope protein inhibits a post-binding step in HCV entry. *Antiviral Res.* 2010;**86**(2):172–9.
43. Wang G. Post-translational Modifications of Natural Antimicrobial Peptides and Strategies for Peptide Engineering. *Curr Biotechnol.* 2012;**1**(1):72–9.
44. Beyene A, Basu A, Meyer K, Ray R. Hepatitis C virus envelope glycoproteins and potential for vaccine development. *Vox Sang.* 2002;**83 Suppl 1**:27–32.