



DIVERSITY OF HEPATITIS C VIRUS; ANALYSIS OF GLYCOPROTEINS E1 AND E2

Dr. Hamid Mahmood¹, Dr. Nasir Zulfiqar², Dr. Sitara Hassan³, Ammara Waqar⁴, Naeem Shahzad⁵, Irfan Farouqi⁶.

1. Professor of Biochemistry
Continental Medical College,
Lahore
2. Assistant Professor
Sir Syed Institute of Medical Sciences,
Karachi
4. PIQC, Lahore
5. PIQC, Lahore
6. PIQC, Lahore

Correspondence Address:
Dr. Hamid Mahmood
Professor of Biochemistry
Continental Medical College, Lahore
drhamidmahmood373@gmail.com

Article received on:
07/03/2014
Accepted for Publication:
28/05/2014
Received after proof reading:
16/08/2014

ABSTRACT... Objective: (1)To understand the diversity in HCV glycoprotein E1 and E2 sequences at different stages of infection. (2) Method widely used in evolutionary studies of HIV, as a new gold standard in HCV research. **Setting:** Department of Biochemistry Sir Syed Institute of Medical Sciences Karachi. **Period.** 1st Jan 2012 to 31 Dec 2012. **Design:** Experimental analysis. **Methods:** The samples derived from a xenomouse model of transmission and several samples from naturally occurring transmission as well as sequences from acute stage of infection. we utilized single genome amplification (SGA) technique to recover full-length E1 sequences. SGA of full-length E1 glycoprotein sequences, followed by direct sequencing minimizes in vitro generated artifacts and experimental biases associated with the standard bulk amplification and cloning approach, giving an accurate representation of investigated intrahost population. **Analysis and results:** Comparative analyses of sequences derived from three different settings were performed, in conjunction with a range of phylogenetic tests. As full-length E1 sequences were utilised. **Conclusions:** The role of glycoprotein E1 and E2 during the infection of HCV was Known. The new gold standard in HCV research prove to be of great importance.

Key words: HCV glycoproteins, PCR

Article Citation: Mahmood H, Zulfiqar N, Hassan S, Waqar A, Shahzad N, Farouqi I. Diversity of hepatitis c virus; Analysis of glycoproteins E1 and E2 professional Med J 2014;21 (3): 723-727.

INTRODUCTION

HCV viral particle contains E1 glycoproteins in its bilayer envelope. These surface proteins are essential for virus life cycle as they mediate entry into the host cell and have also been shown to possess epitopes targeted by host immune system. HCV glycoproteins E1 is type 1 transmembrane proteins with N-terminal ectodomains and C-terminal membrane anchors. E1 and E2 are cleaved from the HCV polyprotein as soon as translation of their coding region occurs. Cleavage site between E2 and p7 seems to be recognized less efficiently by signal peptidase giving rise to different pools of E2: fully cleaved E2 and E2p7. Folding of E1 is a relatively slow process which is dependent on E2. Cysteine cross-linking that supports the conformation in E2 occurs in a time sufficient for its cleavage from the rest of the precursor protein. Both E1 and E2

possess ER retention signals and undergo maturation in this cellular organelle. The C-termini of E1 and E2 are responsible for translocation of glycoprotein ectodomains into the ER lumen. HCV glycoproteins associate in ER with chaperones like calreticulin, Binding Immunoglobulin Protein (BiP) and calnexin. Trafficking of E1 and E2 between these chaperone proteins depends on proper processing of N-linked glycans that, among other roles, has an impact on folding of these proteins. Calreticulin and BiP interact preferentially with misfolded aggregates, whereas calnexin associates with properly folded heterodimers.

Domains responsible for heterodimerisation of HCV glycoproteins have not been characterized accurately. Inconsistent with most Flaviviruses, HCV envelope proteins undergo extensive glycosylation as part of the maturation processes.

Mapping of glycosylation sites showed that E1 has up to 6 and E2 - 11 potential N-linked glycosylation sites. Glycans present on HCV envelope proteins nearly double their mass. Digested by Peptide-N-Glycosylase F (PNGase F) E1 and E2 are about 37 and 17 kDa, whereas glycosylated proteins migrate at about 65 and 30 kDa, respectively. The characterisation of HCV glycoproteins has shown that glycosylation patterns differ between various truncated and full length proteins indicating their different properties. Absence of E2 or E1, during expression of the other glycoprotein, leads to production of different glycoforms. Analyses with conformation- dependent antibodies show that folding of HCV glycoproteins is a slow process. In addition it has also been shown that although some degree of folding can be observed in E2 expressed alone, both glycoproteins need to be co-expressed to acquire functional properties.

MATERIALS AND METHODS

Source of samples

To investigate genetic and phenotypic variability of HCV envelope glycoproteins, sequences derived from three novel clinical or experimental settings were utilised. For the first sequence dataset, HCV RNA samples were obtained as described in a previously published study⁷. Eleven human liver-chimeric Alb-uPA/SCID mice were inoculated intrajugularly with 100 µl genotype 1a HCV infected serum KP (2.3×10^6 IU/ml) to simulate a natural exposure to virus. Patient KP acquired HCV via an unknown route and at the time of sampling was in the chronic phase of infection. Experimental infections were monitored over time by tail bleed sampling. HCV RNA in mouse serum was quantified by a real-time 24 TaqMan PCR assay⁷. RNA was recovered from serum aliquots with use of commercially available RNA extraction kit (Qiagen) and resuspended in 20 µl of H₂O. The sequences obtained from these RNA samples were designated the “mouse model” dataset.

The second sequence dataset was generated from a proposed hospital acquired transmission event where the donor and recipient were suspected¹. The index case patient had been

receiving haemodialysis three times a week at the same unit for 7 years. Blood samples were screened every three months and were HCV negative until September 2012 when routine (unknown) HCV antibody test reached equivocal concentration. Retrospective analysis of stored samples revealed HCV PCR +/-antibody - results for a sample taken 3 months earlier, suggesting acute HCV infection.

Investigation of possible risk factors leading to index case infection pinpointed treatment received in March/April time for diabetic ulcer at the renal inpatient clinic. Four haemodialyses were performed of which one was directly after a HCV positive patient. A transmission event between these two haemodialysis patients was suspected and subsequent genotyping identified the infecting HCV genotype as 3a in both cases. The putative date of the transmission event was established to be 26.03.2012. Sequence data was generated from serum samples obtained from these two patients, along with epidemiologically unrelated genotype 3a sequences from various sources including nine US and eight Pakistani cases. The sequences utilised in this part of the investigation were termed the “transmission” dataset. RNA was extracted from this set of samples utilising two different method.

The third set of sequences was derived directly from patient samples previously enrolled in the Trent Study cohort². Serum samples utilised were collected exclusively from patients in the acute phase of infection of unknown source. Patients were infected with genotype 1a HCV. Window pre-seroconversion period of disease was determined by HCV RNA presence in serum and lack of detectable anti-HCV antibodies in second and third generation enzyme linked immunosorbent assays (unknown manufacturer). The presence of HCV RNA in samples was determined initially by an inhouse RT- PCR assay and, post 1995, by a commercially available reaction (Amplicor; Roche Diagnostics, East Sussex, UK)². HCV glycoprotein sequences generated in this part of project were designated as the “acute” dataset (see Table-I).

	Patient initials	Sample time	Date of collection	Special notes
Acute 1	N/A	1A20	04/01/2012	1 months after first sample collected, month after HCV RNA detected, Anti-HCV negative
Acute 2	MA	1A30.3	07/10/2012	2 months after first sample collected, month after HCV RNA detected, Anti-HCV negative
Chronic		1A30.12	04/02/2012	16 months post infection, RNA+, Anti-HCV+, patient deceased
Acute 3	KB	1A57.2	02/09/2012	5 months after first sample collected, first time HCV RNA detected, no information about Anti-HCV, later sample still Anti-HCV negative
Chronic		1A57.14	01/02/2012	42 months post infection, RNA+, Anti-HCV+
Acute 4	RW	1A65.1	05/04/2012	First sample taken, positive RNA, negative Anti-HCV, first Ab+ sample collected 5 months post PCR+
Acute 5		1A83.1	14/12/2012	2 months post infection, RNA+, Anti-HCV+

Table-I. Acute samples details

Amplification of E1E2

Two PCR based methods were used in sampling viral quasi species. The first method involved generating full length E1E2 glycoprotein sequences using unknown amount of cDNA molecules as a template, followed by direct sequencing of PCR products. This is referred to as “bulk amplification” method as aliquots of non-diluted cDNA samples were used in PCR amplification. It was mainly used to amplify E1E2 fragments from low titre samples or to serve as a comparison to single genome amplification method described below.

In order to avoid in vitro generated recombinants, cloning induced errors and also to ensure accurate and representative sampling of quasi species populations, a single genome amplification (SGA) approach was utilised. The SGA method, followed by direct sequencing, has been previously applied to characterise viral quasi species populations in HIV-1³⁻⁵ and HCV infection⁸. In this approach viral cDNAs were serially diluted and aliquots of 2-fold dilutions (1 µl) were used as template in the first round of a nested full-length E1E2 PCR. Two µl of first-round product were subsequently used as template in second-round

reactions. For each sample, end-point titration PCRs revealed the dilution at which the concentration of viral cDNA was <1 molecule per µl. Subsequently multiple E1E2 amplicons were obtained for each sample at the end-point dilution ($\leq 3/10$ PCR positives). According to Poisson distribution.

When no more than 30% of PCR reactions are positive, more than 80% of amplicons will be derived from single molecule templates⁶. These amplicons were sequenced directly and further referred to as “single molecule” derived quasi species sequences. Chromatograms were checked by eye for presence of double peaks that would disqualify an amplicon from further analysis due to being derived from multiple templates. Both bulk and single-molecule methods utilised a nested PCR comprised of two round of amplification primed by two sets of primers. A second pair of primers was designed to attach to sequences positioned internally on the template in relation to the first pair of primers. For every PCR carried out negative controls were used to ensure lack of carryover contamination. Amplification reactions were set up in 20 µl volumes containing 4 pmol of sense and anti-sense outer primers, 0.25

mM dNTPs, 0.5 U of Platinum® Taq High Fidelity polymerase (Invitrogen), 1× High Fidelity polymerase buffer and 2 mM MgSO₄. One µl of neat or diluted cDNA was used as reaction template. The PCR-cycling parameters were 35 cycles of 94°C for 15 seconds, 50°C for 30 seconds and 68°C for 3 minutes. Two micro litres of the first-round product was subsequently used in second-round reactions with inner primers using identical amplification and cycling conditions to the first round but increasing the cycle number to 45.

RESULTS

Having assessed the dynamics of HCV experimental transmission in the chimeric mouse model, we then investigated a case of HCV transmission acquired through sharing haemodialysis machine. Patients undergoing haemodialysis are at increased risk of HCV infection. Horizontal transmission between haemodialysis patients has been well documented. Standard infection control procedures have been shown to drastically reduce nosocomial spread of HCV within dedicated haemodialysis units. Transmission of HCV to multiple patients using the same haemodialysis machine is not well documented. Consequently dedicated hemodialysis machines for HCV infected patients are not routinely used. Using phylogenetic analysis we confirm a case of nosocomial HCV transmission between patients receiving haemodialysis in the same unit, after sharing the same machine. These analyses have implications for health care management guidelines.

DISCUSSION

In the first part of this study, the xenomouse model was utilised to investigate the evolutionary mechanisms underlying HCV transmission. Using partial E1E2 sequences also precludes meaningful subsequent phenotypic investigation. Contrastingly, our investigation is the first to describe HCV full-length E1E2 glycoprotein variants distribution from a defined source at transmission. E1E2 sequences were derived from single genome templates, minimizing in vitro generated artefacts, which may sequence variants

distribution and misrepresent evolutionary processes underlying observed sequence diversity. In addition, Alb uPA/SCID xenomice were experimentally infected with a known source inoculum whose sequence population was also interrogated. The full-length nature of the E1E2 sequences retrieved also enabled phenotypic characterization of donor and recipient glycoproteins at transmission. These data suggest the different xenomouse hosts constitute highly similar selective environments, despite possessing human liver grafts from different donors.

In this respect, the xenomouse Alb uPA/SCID model appear to mimic the HVR1 evolutionary stasis observed at transmission between human hosts and experimental transmission from human to chimpanzee. However, expanding the sequence coverage to encompass full-length E1E2 sequences revealed contrasting results to previous transmission studies, that advocated phylogenetic grouping of envelope sequences based on HVR1 sequence analysis⁹.

Our data revealed three key amino acid sites (198, 448 and 474) and subtle variation at a fourth position (570) played a role in establishment of a productive infection in a new host in the xenomouse system. The major donor variant SNHV appeared unsuccessful in establishing infection post transmission and was outcompeted by the TDYD variant, which became the dominant component in all recipient xenomice. All post-transmission E1E2s harbour an N?D PNGS knockout, which presumably confers a selective advantage in each xenomouse host. A PNGS knockout at position 448 (corresponding to glycan E2N4) has previously been shown to abrogate H77 E1E2 pseudoparticle infectivity [347]. However, in the JFH-1 HCVcc system PNGS 448, has not affected infectivity although has been demonstrated to render envelopes more susceptible to neutralisation by a panel of anti-E2 monoclonal antibodies, indicating PNGS 448 is critical in shielding neutralisation sensitive epitopes from immunological targeting [71]. This would suggest that TDYV variant is selectively

amplified due to the fitness it harbours in the new host in the absence of neutralising antibodies.

CONCLUSIONS:

The role of glycoprotein E1 and E2 during the infection of HCV was known. Horizontal transmission between the patients was not found. Nosocomial HCV transmission between the patients receiving hemodialysis in the same unit or on the same machine is important for health care management. The SGA method, followed by direct sequencing, has been previously applied to characterise viral quasi species populations in HIV-1 and HCV infection. The new gold standard in HCV research prove to be of great importance.

Copyright© 12 May, 2014.

REFERENCES

1. Thomson, P.C., et al., **A case of hepatitis C virus transmission acquired through sharing a haemodialysis machine.** Nephrology Dialysis Transplantation Plus, 2011. 4(1): p. 32-35.
2. Mohsen, A.H. and T.H.S. Group, **The epidemiology of hepatitis C in a UK health regional population of 5.12 million.** Gut, 2001. 48(5): p. 707-713.
3. Boom, R., et al., **Rapid and simple method for purification of nucleic acids.** Journal of Clinical Microbiology, 1990. 28(3): p. 495-503.
4. Keele, B.F., et al., **Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection.** Proceedings of the National Academy of Sciences, 2008. 105(21): p. 7552-7557.
5. Simmonds, P., et al., **Human Immunodeficiency Virus-Infected Individuals Contain Provirus In Small Numbers Of Peripheral Mononuclear-Cells And At Low Copy Numbers.** Journal of Virology, 1990. 64(2): p. 864-872.
6. Bar, K.J., et al., **Wide Variation in the Multiplicity of HIV-1 Infection among Injection Drug Users.** Journal of Virology, 2010. 84(12): p. 6241-6247.
7. Law, M., et al., **Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge.** Nature Medicine, 2008. 14(1): p. 25-27.
8. Ramachandran, S., et al., **End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: Performance under optimal and suboptimal conditions.** Journal of Virological Methods, 2008. 151(2): p. 217-224.
9. Fafi-Kremer, S., et al., **Viral entry and escape from antibody-mediated neutralization influence hepatitis C virus reinfection in liver transplantation.** Journal of Experimental Medicine, 2010. 207(9): p. 2019-31.
10. Fissell, R.B., et al., **Patterns of hepatitis C prevalence and seroconversion in hemodialysis units from three continents: The DOPPS.** Kidney International, 2004. 65(6): p. 2335-2342.
11. Sambrook and Russell, **Molecular Cloning: A Laboratory Manual.** 3 ed. Vol. 3. 2001: Cold Spring Harbor Laboratory Press.
12. Sanger, F. and A.R. Coulson, **A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.** Journal of Molecular Biology, 1975. 94(3): p. 441-446.