Mutational Analysis of Various Sub-Genomic Regions of HCV and Their Role in Interferon Therapy Response

Abida Raza¹, Maryam Zafer², Binish Gull Arshad³, Samina N. Shakeel⁴

¹ Nuclear Medicine, Oncology and Radiotherapy Institute, Islamabad, Pakistan
² COMSATS Institute of Information Technology, Islamabad, Pakistan
³ Quaid-i-Azam University, Islamabad, Pakistan

Abstract: Combination therapy with interferon (IFN) and ribavirin for 24 to 48 weeks is the current standard treatment for chronic hepatitis C. Success rate is 67-80%. There are four subgenomic regions of HCV in which distinct mutations have been observed in response to IFN therapy; E2-PePHD, NS5A- ISDR, NS5A-PKRBD and NS5A-V3 domains. Current meta analysis gives comprehensive details of mutations in subgenomic regions and their role in non responsivenes to IFN treatment. It has been observed that E2-PePHD region is well conserved region and show high sensitivity to IFN-ribavirin therapy. The substitutions within the NS5A-ISDR to the treatment outcome are conflicting. Mutations within the NS5A-PKRBD of HCV type 1 are associated with a long-term sustained response to IFN-ribavirin therapy. There is a considerable association between efficacy of treatment and a high mutation rate within NS5A-V3 domain has been observed. There are also other domains of HCV i.e. IRRDR, CRS and NLS, but no significant mutations been studied in these regions, more research is to be required.

1. Introduction

The hepatitis C virus (HCV) is one of the most frequent causes of chronic viral hepatitis, liver cirrhosis, hepatocellular carcinoma. The virus particle consists of a central hub of genetic material (RNA) which is enclosed by an icosahedral protecting shell of protein, and is further sheathed in a lipid envelope [1] Fig 1.

Figure 1: Simplified diagram of the structure of the Hepatitis C virus particle (Wikipedia)

The viral genome consists of a single open reading frame specifically 9600 nucleotide bases long which then translated to produce a single protein product, which is further processed to produce smaller active proteins [2] Fig 2.

Hepatitis C virus RNA

5’ NTR

Gene encoding precursor polyprotein

9500 nt bases

Structural proteins

non-structural proteins

p22

p27

p27

p17

p7

p7

p17

C

E1

E2

NS2

NS3

NS4A

NS4B

NS5A

NS5B

nucleocapsid

RNA helicase

transmembrane protein

proteases

co-factors

interferon resisting protein

RNA polymerase

2. Role of Structural Proteins

The hepatitis C virus E1 and E2 envelop proteins are the key players in entire actions which are required for viral entry into target cells. Collectively with the capsid protein, the core components of HCV virion are envelop proteins E1 and E2. E1 and E2 proteins are located in the N-terminal region of poly protein precursor with the non-structural proteins. E1 and E2 glycoproteins interrelate through their

Figure 2: Genome organization of Hepatitis C virus

On the 5’ and 3’ ends of the Ribonucleic acid are the NTR which are not translated into proteins but they are important towards translation and replication of the viral RNA. The 5’ NTR has a ribosome binding site [IREs – Internal Ribosome Entry Site] that initiates the translation of a very long protein containing about 3,000 amino acids [3]. This large pre-protein has major role in viral replication and assembly [4]. Virus has both structural proteins (Core protein, E1 and E2) and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5, NS5A, and NS5B).
transmembrane domain to make non-covalent heterodimers that represent the putative viral spike [5]. While in the transmembrane fix region of both E1 and E2 is concerned in ER retention and E1E2 association [6]. The soluble part plays main role in the folding process of glycoproteins, in virus entry, and in the modulation of immune responses [4], [7].

3. Role of Non-structural Proteins

The NS2 region makes part of a protease which also includes the N terminus of NS3 region. This facilitates autoproteolytic cleavage at the NS2/NS3 intersection of the polyprotein. The NS3 also translates serine protease, nucleotide triphosphatase and RNA helicase activities. The serine protease activity has predominantly well described and possibly plays a vital role in HCV processing, making it an enviable target for antiviral drugs. The NS4 region translates NS4A and NS4B. The NS4 protein functions as a cofactor in NS3 activity. It is necessary for the NS3/4A and NS4B/5A cleavage trials and enhances cleavage at the NS4A/4B sites. The function of NS4B which is a very hydrophobic protein is unidentified though it is assumed to characterize the viral RNA polymerase [8].

4. Functions of Sub-genomic Regions of HCV in Response to IFN Therapy

Combination therapy with interferon (IFN) and ribavirin for 24 to 48 weeks is the current standard treatment for chronic hepatitis C, but it does not eliminate the virus completely in 50-80% patients. Many viral and host factors are concerned in response to interferon treatment. Viral factors that support sustained virologic response to Interferon therapy include HCV genotypes other than genotype 1 and low viral load. There are four subgenomic regions of HCV in which distinct mutations have been observed in response to IFN therapy. The regions include E2-PePHD (PKR/eIF2α phosphorylation homology domain); NS5A-ISDR (IFN sensitivity-determining region), PKRBD (PKR binding domain), and V3 (Variable region 3). A PKR/eIF2α phosphorylation homology domain (PePHD) within the E2 protein has been found to interact with PKR (double-stranded RNA-dependent protein kinase) and inhibit PKR in vitro, suggesting a possible mechanism of HCV to elude the antiviral effects of IFN therapy [9]. Mutations in this region are thought to manipulate the response to IFN therapy [10], [11], however, the results of different studies are conflicting because this region is highly conserved [12]-[14]. Genetic heterogeneity of a specific domain of NS5A region, termed the IFN sensitivity-determining region (ISDR), is closely related to the response in Japanese patients with HCV genotype 1b, so that patients with at least 4 mutations within ISDR achieved a sustained virologic response (SVR) to IFN-a monotherapy. However, the reported results in Europe and the United States concerning the correlation of the substitutions within the ISDR to the treatment outcome are conflicting. It is well known that the ISDR is essential but not adequate for the interaction between NS5A and PKR enzyme (which is very important for the activation of IFN) [20].

This region mediate disruption of PKR dimerization resulted in the repression of PKR function and the inhibition of PKR-mediated eIF2α phosphorylation. The introduction of multiple mutations within the PKR-binding region, including those within the ISDR, abrogated the ability of NS5A to bind to PKR. Mutations within the PKRBD of HCV type I are associated with a long-term sustained response to IFN-α and IFN-α/RBV therapy [21], [22]. Recent, clinical studies have projected that the number of amino acid variations within variable region V3 may be associated with the treatment outcome [23], [24].

5. Analysis of Mutations in Sub Genomic Regions of HCV

Different studies have been proposed to check the mutations in different regions so far. Here we will describe some of them briefly.

5.1. E2-PePHD

Some of the researchers reported a link between amino acid switching in PePHD domain of HCV genotype 1, 2 and 3 strains [25]-[27] and treatment responses while it has been revealed in some others found that PePHD domain is a conserved domain and there is no link exists between amino acid substitution and treatment response [12]-[13], [17]-[21]. In this research we aligned amino acid sequences of PePHD domain from 31 variants of six HCV genotype 3 strains together with consensus sequences. The region was found to be conserved with amino acid switching at only two positions; 4 and 5. At position 4 Glutamine (Q) was switched by Leucine (L) in variants of one of the breakthrough sample and at position 5 Histidine (H) was switched...
by Q in variants of one of the rapid responder. Other amino acid locations were more or less conserved with either no switching or very rare switching in any one of the variants [28] Fig 3. As replacements were found in both rapid responders and breakthrough responders, therefore this result is reliable with previous reported replacements in PePHD which are not linked to treatment response.

In another analysis [29], the pretreatment amino acid sequence of E2-PePHD, resulted by direct sequencing by PCR, was found highly conserved among the HCV-1b isolates. In another analysis [29], the pretreatment amino acid sequence of E2-PePHD, resulted by direct sequencing by PCR, was found highly conserved among the HCV-1b isolates. A small number of mutations were spotted in the flanking regions of PePHD domain (29 amino acids upstream and 31 amino acids downstream), but these mutations were not linked with a specific pattern of response to treatment. The amino acid sequence of E2-PePHD before therapy was contrasted with that after therapy by direct sequencing. The major profiles spotted after M6 (for non-responders) illustrated a variety of amino acid substitutions in the clones from 4 of the 9 patients studied: substitution of an S by an A or T at position 2, substitution of a P by an S at position 6, and substitution of a T by an A at position 12. Fluctuations in amino acids were also observed in the clones from 2 responders between the beginning of treatment and M2 (for responders to dual therapy): substitution of an S by an A at position 2, substitution of a P by an S at position 6, and substitution of a T by an A at position 12. For these 6 selected patients, we carried out pretreatment PePHD quasi species evaluation for 20 clones per patient. All but 2 clones had a PePHD motif which is identical to that of the HCV-1b prototype; these mutations (substitution of a P by a T at position 664) were found in 2 clones from a single non-responder contrasted from that found in the same patient at M6 (substitution of an S by an A at position at 660) [29] Fig 4b.

**Figure 3:** PePHD amino acids multiple alignment of 31 variants of six HCV genotype 3a baseline samples subjected to IFN alpha and Ribavirin combination therapy.

In another analysis [29], the pretreatment amino acid sequence of E2-PePHD, resulted by direct sequencing by PCR, was found highly conserved among the HCV-1b isolates. In another analysis [29], the pretreatment amino acid sequence of E2-PePHD, resulted by direct sequencing by PCR, was found highly conserved among the HCV-1b isolates. A small number of mutations were spotted in the flanking regions of PePHD domain (29 amino acids upstream and 31 amino acids downstream), but these mutations were not linked with a specific pattern of response to treatment. The amino acid sequence of E2-PePHD before therapy was contrasted with that after therapy by direct sequencing. The major profiles spotted after M6 (for non-responders) illustrated a variety of amino acid substitutions in the clones from 4 of the 9 patients studied: substitution of an S by an A or T at position 2, substitution of a P by an S at position 6, and substitution of a T by an A at position 12. Fluctuations in amino acids were also observed in the clones from 2 responders between the beginning of treatment and M2 (for responders to dual therapy): substitution of an S by an A at position 2, substitution of a P by an S at position 6, and substitution of a T by an A at position 12. For these 6 selected patients, we carried out pretreatment PePHD quasi species evaluation for 20 clones per patient. All but 2 clones had a PePHD motif which is identical to that of the HCV-1b prototype; these mutations (substitution of a P by a T at position 664) were found in 2 clones from a single non-responder contrasted from that found in the same patient at M6 (substitution of an S by an A at position at 660) [29] Fig 4b.

**Figure 4a:** Alignment of the sequences of amino acid residues 630 to 701 of the HCV E2 protein from pretreatment samples from 25 HCV-1b-infected patients. Amino acid sequences derived by direct PCR sequencing are shown. The sequence of HCV-J is taken as the reference sequence for the HCV-1b subtype. Dashes indicate residues identical to those in this prototype. The vertical lines indicate the PePHD domain. sr1 to sr10, patients with no virologic response; sr1,sr4, sr6, sr7 and sr10, patients with sustained virologic responses to IFN-α alone; sr2*, sr3*, sr5*, sr8*, sr9*, sr11*, sr12*, sr13*, sr14* and sr15*, patients with sustained virologic responses to IFN-α plus ribavirin combination therapy.
5.2. NS5A

In agreement with previous studies, Kurosaki et al [30] found that amino acid mutations are mostly at codon 2213 and 2218. However, they detected a specific amino acid substitution pattern in responder (R) patients at codon 2213. Of the 18 non-responder patients (NR), seven had an Alanine to Glutamic acid change at codon 2213, whereas this variation was not observed in any of the R patients. In addition, the mutation at codon 2218 was determined more in NR patients at codon 2218 (10/21; 48%) than in R patients (2/18; 11%). Positive correlation between mutations in the region of the NS5A gene of HCV genotype 1b (NS5A 2209-2248) and sensitivity to IFN-α was reported in several studies from Japan. Accordingly, genotype 1b patients with 4-11 amino acid changes in the region of the NS5A gene of HCV genotype 1b (NS5A 2209-2248) and sensitivity to IFN-α were reported in several studies from Japan. In addition, patients infected with intermediate type sequence (11/18) also had complete response (63.2%). In the NR patients (21/39), there were only eight patients (8/21) infected with the wild type virus and the rate of intermediate type virus (13/21) was nearly two times higher than the wild type virus. Thus, we found no correlation between the sequence of the ISDR and sensitivity to IFN [30], [31]. Data demonstrates that no significant association exists between the number of amino acid changes in ISDR and response to IFN-α treatment. Amino acid sequences in ISDR are well conserved in the majority of patients infected with genotype 1b and 1a.

Figure 4b: Variation of the sequence of the HCV E2-PePHD domain in clones from 25 HCV-1b-infected patients during treatment. PePHD sequences obtained by direct sequencing (major profile) are shown. After M2, 5 patients tested negative for HCV RNA and remained negative thereafter; 20 patients tested positive for HCV RNA and benefited from the dual therapy, but 10 of them subsequently developed an SVR. Amplification and sequencing were performed with the M2 sample (amplification was a failure in four cases). 10 of the patients remained HCV RNA positive, despite dual therapy. Amplification and sequencing were performed with the M6 samples (amplification was a failure in one case). The sequences of HCV-J, HCV-1, HCV-J6, HCV-J8, and HCV-NZL1 are taken as references for the HCV-1b, HCV-1a, HCV-2a, HCV-2b and HCV-3a subtypes, respectively.

In NS5A, mutations within ISDR, PKR-BD, V3, and the interferon/ribavirin resistance-determining region (RRDR) have been correlated with the IFN-based therapy response. In Tunisia, where a high prevalence of HCV-1b has been found, no data regarding the implication of NS5A in treatment response were available. The current study examined the relationship between the pre-treatment mutation number within ISDR, PKR-BD, V3, RRDR, as well as in the entire ISDR-V3 region of NS5A (amino acid 2209-2379) and the response to the 48-week course of combined IFN plus ribavirin therapy in 15 HCV-1b-infected Tunisian patients. Referring to HCV-J sequence, a significant high genetic variability was observed within PKR-BD in the sustained virologic responder patients compared to non-responders (P=0.040). More importantly, when considering the entire region from ISDR to V3, referred to as NS5A-ISDR-V3, a clear difference in the mutation number was observed between sustained
virologic responders (19.6±3.16) and non-responders (15.0±1.41) (P=0.002). Additionally, a more detailed analysis of NS5A ISDR-V3 region revealed an elevated degree of mutation rate within the region located between amino acids 2282 and 2308 (P=0.0006). Interestingly, an analysis of specific amino acid variations defined Proline and Serine at position 2300 as signature patterns for sensitive and resistant strains, respectively. The genetic variability within the NS5A region of HCV-1b strains was associated with the response to the combined IFN plus ribavirin therapy in Tunisia [32].

In another study, the non-structural NS5A protein of hepatitis C virus has been controversially implicated in the resistance to Interferon therapy in clinical studies. In Japan, mutations in the Interferon Sensitivity-Determining Region (ISDR) in the NS5A gene were associated with response to Interferon therapy in patients infected with genotype 1b. In contrast, studies from Europe did not confirm such association. More recently, it has been suggested that the V3 domain outside the putative ISDR might also have amino acids changes that may be involved in the resistance to IFN. In this study, the relationship between NS5A mutations in ISDR and V3 domains and virologic response to therapy were investigated.

In the ISDR domain, no significant differences have been observed in amino acids changes between responders (1.7 ± 1.8, n= 19, range 0-6) and non-responders (1.1 ± 0.8, n= 14, range: 0-3), (P= 0.483), to therapy before the beginning of treatment. In the V3 domain, we found more mutations in responders (6.5 ± 1.9, range: 2-11) than in non-responders (4.7 ± 1.2, range: 3-8) (P= 0.0013), before the beginning of treatment [33]. From data analysis it is confirmed that, in Europe, the ISDR domain is not predictive for treatment success but suggests that the V3 domain have greater variability in responders than non-responders.

5.4. NLS and V3

In this research, no noteworthy differences were found over comparing sequence variations in the NLS and extended V3 domains among SVR and NR groups. In spite of this, other researches have shown an important relationship between treatment efficacy and a high number of mutations within V3, with a considerable difference among SVR and NR patients [39]. The total number of amino acid variations observed in NS5A protein of HCV genotype 3a isolates from this study was not associated with treatment response.

5.5. IRRDR

The entire NS5A region of the HCV genome was amplified from the pretreated sera and the amino acid sequences deduced. We compared each NS5A sequence with a consensus sequence inferred from aligning the NS5A-1b sequences [40] Fig 5.
In this connection, the consensus sequence for IRRDR differs from the corresponding sequence of a prototype strain of IFN-resistant HCV-1b by a single residue at position 2367 (alanine instead of glycine). The numbers along the sequence indicate amino acid positions. Dots indicate residues identical to those of the Cons sequence. Ala^{2360} and Thr^{2378} are written in boldface.

As described, there was no difference in the mean number of mutations in ISDR or PKR-BD between SVR and non-SVR. Only four patients had HCV with four or more mutations in ISDR (data not shown), the criterion for IFN-sensitive HCV strains according to Enomoto et al. Although there appeared to be a trend for patients with HCV having four or more mutations in ISDR toward SVR (3 of 4), the difference was not statistically significant. Also, the prevalence of HCV with four or more mutations in ISDR was not significantly different between SVR (3 of 21; 14.3%) and non-SVR (1 of 24; 4.2%). It would be interesting to note, however, that all three HCV strains with four or more mutations in ISDR obtained from SVR had HCV of IRRDR of 6 or greater, whereas the only strain with four or more mutations in ISDR from non-SVR had three mutations in IRRDR (data not shown) [41]. It is thus possible that the IRRDR sequence variation is associated with PEG-IFN/RBV responsiveness more closely than is the ISDR variation.

From the meta analysis, we conclude that there is a link between amino acid substitution in PePHD domain of HCV genotype 1, 2 and 3 strains with treatment responses whereas E2-PePHD domain is a conserved one and there is no correlation exists between amino acid substitution and the IFN treatment response. A positive correlation between mutations in the region of the NS5A gene of HCV genotype 1b and sensitivity to IFN-α is reported in several studies from Japan. In line with other studies from Europe, we found no difference in the number of amino acid mutations in NS5A-ISDR between patients who responded to an IFN treatment regime and those who did not. Likewise in Pakistan, no correlation was found between amino acid substitution and treatment response. Data demonstrates no significant association exists between number of amino acid changes in ISDR and response to IFN-α treatment. A significant high genetic variability is present within PKRBD in sustained virologic responder patients compared to non-responders. The genetic variability within the NS5A region of HCV-1b strains is associated with the response to the combined. The ISDR domain is not predictive for treatment success but suggests that the V3 domain have greater variability in responders than non-responders.

References


