Sequencing of E2 and NS5A regions of HCV genotype 3a in Brazilian patients with chronic hepatitis

Fernanda de Mello Malta1,*, José Eymard Moraes de Medeiros-Filho1, Raymundo Soares de Azevedo2, Luzia Gonçalves3, Luiz Caetano da Silva1, Flair José Carrilho1, João Renato Rebello Pinho1

1Departamento de Gastroenterologia, Laboratório de Gastroenterologia e Hepatologia Tropical, Instituto de Medicina Tropical, São Paulo, SP, Brasil 2Departamento de Patologia, Faculdade de Medicina, Universidade de São Paulo, Av. Dr. Enéas Carvalho de Aguiar 500, 05403-000 São Paulo, SP, Brasil 3Unidade de Epidemiologia e Bioestatística, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisboa, Portugal

Hepatitis C virus (HCV) is a major cause of liver disease throughout the world. The NS5A and E2 proteins of HCV genotype 1 were reported to inhibit the double-stranded (ds) RNA-dependent protein kinase (PKR), which is involved in the cellular antiviral response induced by interferon (IFN). The response to IFN therapy is quite different between genotypes, with response rates among patients infected with types 2 and 3 that are two-three-fold higher than in patients infected with type 1. Interestingly, a significant percentage of HCV genotype 3-infected patients do not respond to treatment at all. The aim of this paper was to analyse the sequences of fragments of the E2 and NS5A regions from 33 outpatients infected with genotype 3a, including patients that have responded (SVR) or not responded (NR) to treatment. HCV RNA was extracted and amplified with specific primers for the NS5A and E2 regions and the PCR products were then sequenced. The sequences obtained covered amino acids (aa) 636-708 in E2 and in NS5A [including the IFN sensitivity determining region (ISDR), PKR-binding domain and extended V3 region]. In the E2 and NS5A regions, we did observe aa changes among patients, but these changes were not statistically significant between the SVR and NR groups. In conclusion, our results suggest that the ISDR domain is not predictive of treatment success in patients infected with HCV genotype 3a.

Key words: E2 protein - hepatitis C virus - interferon sensitivity determining region - NS5A protein - PKR-eIF2a phosphorylation homology domain

Hepatitis C virus (HCV) is a member of the Flaviviridae family which usually causes chronic infection and can result in chronic hepatitis, liver cirrhosis and hepatocellular carcinoma. The only drug that effectively reduces viral load is interferon-α (IFN-α) (Chayama et al. 2000). Current therapy, a combination of pegylated IFN-α and ribavirin (RBV), achieves a response rate between 48-88% (Poynard et al. 2003). Individuals infected with HCV genotype 1a/b show response rates of 38-52%, whereas 66-88% of those infected with genotypes 2 and 3 achieve sustained virological response (SVR). It is likely that the sensitivity or resistance to antiviral therapy is governed by both the virus and the host itself (Goyal et al. 2007).

The ability of HCV to evade the antiviral effects of IFN-α has been previously reported. Both the HCV envelope E2 protein and the nonstructural NS5A protein of HCV-1a/b isolates have been shown to inhibit IFN-α induced double-stranded RNA-activated protein kinase (PKR) in vitro (Gale et al. 1998a, Taylor et al. 1999). Taylor et al. (1999) reported that HCV E2 protein contains a 12 amino acid (aa) domain that is highly homologous to the autophosphorylation site of PKR and initiation factor eIF2a, a target of PKR [the PKR-eIF2a phosphorylation homology domain (PePHD)]. They showed that E2 protein inhibited the kinase activity of PKR and blocked its inhibitory effect on protein synthesis and cell growth, suggesting that the interaction of E2 and PKR may be one of the mechanisms by which HCV circumvents the antiviral effects of IFN.

NS5A has also been studied as a possible mediator of IFN-α resistance (Hofmann et al. 2005). A binding domain for the IFN-α-inducible double-stranded RNA-dependent PKR was identified within the carboxy-terminal portion of NS5A in HCV 1 isolates. The PKR binding domain includes the previously described IFN sensitivity determining region (ISDR) (Enomoto et al. 1995, 1996). The binding of NS5A to PKR results in the inhibition of protein synthesis with a subsequent antiviral effect in vitro (Gale et al. 1998b). Clinical studies have addressed the question of whether the number of NS5A aa variations in HCV isolates from patients who did or did not respond to IFN-α based therapy is associated with treatment outcome. So far, the number of aa variations within the ISDR and within the entire NS5A protein from HCV genotype 1b isolates may be a predictor for SVR (Pascu et al. 2004). More recently, clinical studies have proposed that the number of aa variations within variable region V3 and its flanking regions may be associated with treatment outcome (Castelain...

In the present paper, we investigated the aa substitutions within E2 (PePHD) and NS5A from Brazilian patients chronically infected with HCV genotype 3a and their correlation with response to antiviral therapy (IFN-α plus RBV).

**PATIENTS, MATERIALS AND METHODS**

Thirty-three patients chronically infected with HCV genotype 3a from the Hepatology Branch, Department of Gastroenterology, University of São Paulo, School of Medicine, São Paulo, Brazil were enrolled. Clinical, biochemical and histological characteristics of the enrolled patients were obtained from their medical records and are summarised in Table I. Histological activity grade and fibrosis stage were evaluated according to the METAVIR scoring system (cirrhotic, n = 16; non-cirrhotic, n = 17) (Bedossa & Poynard 1996).

Response to IFN-α/RBV treatment was assessed at the end of therapy and six months after stopping therapy. HCV-RNA was determined for sera using a commercially available qualitative PCR assay (AMPLICORTM, 50 IU/mL detection limit, Roche Molecular Systems, CA, USA). Patients who achieved HCV-RNA clearance for more than six months after the end of therapy were considered to be SVR (SVR, n = 16). Patients who became HCV-RNA negative during the course of therapy but then relapsed and showed detectable HCV-RNA after stopping treatment were considered non-responders (NR) (n = 17). Informed consent to participate was obtained from each patient. This study was approved by the local Ethical Committee.

**HCV-RNA extraction** - HCV-RNA was extracted from 140 µL of pre-treatment serum using a commercially available kit (QIAamp® Viral RNA, QIAGEN, Inc, Hilden, Germany). Complementary DNA was generated using M-MLV reverse transcriptase and random hexamers (Invitrogen™, Carlsbad, CA, USA). After incubation at 70°C for 10 min, reverse transcription was carried out at 37°C for 60 min, followed by heat inactivation at 95°C for 15 min.

**Amplification of the HCV E2 and NS5A regions** - Polymerase chain reactions (PCRs) were performed using Elongase® Enzyme Mix DNA polymerases (Invitrogen™, Carlsbad, CA, USA) and a set of specific primers for each gene (Table II).

The PCRs for amplifying the C-terminal region of the E2 protein (240 bp fragment) comprising the PePHD were performed in two rounds under equivalent conditions: 30 cycles, with each cycle consisting of denaturation at 94°C for 15 sec, annealing at 55°C for 30 sec, extension at 68°C for 40 sec, followed by a final extension at 68°C for 3 min. The HCV NS5A gene (codons 1979-2414) was amplified by nested PCR in four overlapping parts with different sets of primers (Table II). The first round of PCR was performed using the external primers (NS4B_595 and NS5B_159a), after an initial denaturation step of 94°C for 2 min, 35 cycles of 94°C for 30 sec, 48°C for 30 sec and 68°C for 1 min. The second round was performed using four sets of internal primers [F1S and F1AS (551 bp), F2S and F2AS (545 bp), F3S and F3AS (439 bp) and F4S and F4AS (423 bp)] under the following conditions: an initial denaturation step of 94°C for 2 min, 30 cycles of 94°C for 15 sec, 62°C for 30 sec and 68°C for 45 sec. After cycling, a final extension period at 68°C for 3 min was added for all reactions. The amplified fragments were visualised by 2% agarose gel electrophoresis and ethidium bromide staining.

Strict procedures for nucleic acid amplification diagnostic techniques were followed during pre and post-amplification to avoid false-positive results and positive and negative controls were always included in each set of reactions.

**Sequence analysis** - The PePHD region of E2 protein and the cytoplasmic retention signal (CRS), ISDR, PKR-binding, nuclear localisation signal (NLS) and V3 domains from NS5A protein were characterised by the direct sequencing of one pre-treatment serum sample from each patient. Sequencing was performed by the PCR direct sequencing method. The nucleotide sequence was determined by the dideoxy method using the ABI Prism BigDye terminator cycle sequencing kit and the sequencing gels were run on an ABI Prism 377 DNA automated sequencer (Applied Biosystems, Foster City, CA, USA). Each amplicon was analysed in both the sense and antisense direction.

All of the sequences obtained in this study were aligned and the most frequent aa at each position was defined as the consensus sequence. The sequences of the E2 (PePHD) and NS5A domains of HCV were compared with the consensus sequence and with the prototype isolate for HCV-3a (NZL-1/D17763). Multiple aa sequence alignment was carried out using CLUSTAL X (version 1.81) and MEGA (version 2.1) (http://www.megasoftware.net/) was used to analyse aa sequences (Thompson et al. 1997, Kumar et al. 2001). The correlation between the number of mutations and the response to the treatment, in each of the region, was studied taking into account the median number of mutations in each region. In this way the populations of each region were distributed into two groups (E2 aa 636-708, > or ≤ 1; PePHD, without mutations or ≥ 1; CRS, without or ≥ 1; ISDR, without or ≥ 1; PKR-binding, > or ≤ 1; NLS, without or ≥ 1; extended-V3, > or ≤ 3; NS5A aa 1979-2414, > or ≤ 13). This approach was used by Muñoz de Rueda et al. (2008) in a similar study involving patients infected only by HCV genotype 1.

**Statistical analysis** - Statistical analysis was performed using Minitab, version 15.1.1.0 (Minitab Inc, State College, PA, USA). Descriptive statistics was carried out according to the nature of the variables and the results are presented in the Tables. The quantitative variables are ex-
pressed as means and standard deviations. Comparisons between groups (SVR and NR) were carried out using \( \chi^2 \) or Fisher’s exact test for categorical variables and the Student’s t-test for quantitative variables. A p value less than 0.05 was considered statistically significant.

Nucleotide sequence accession numbers - Nucleotide sequence data reported in this paper was submitted to GenBank (http://www.ncbi.nlm.nih.gov/) with the accession EF207987-EF208018 for the NS5A sequences and EF210487-EF210519 for the E2 sequences.

**RESULTS**

Comparison of the total number of aa changes in NS5A in SVR and NR patients - The NS5A region of HCV genotype 3a was amplified and sequenced from the sera of 32 pre-treated patients; this region was not amplified in only one sample. As we found mutations related to NZL-1 that occurred repeatedly in all of our patients, we decided to determine a consensus sequence for all patients participating in this study. We obtained results similar to those obtained with the reference strand. Mutations were always defined with respect to the consensus sequence of all NS5A sequences investigated in this study.

The aa sequences were deduced and then compared to the consensus sequence (Fig. 1). The total number of aa changes in NS5A did not differ between SVR and NR. Also, when the analyses were restricted to CRS, ISDR,
PKR-binding, NLS and the extended V3 regions, no significant difference was observed between the number of aa variations in SVR patients compared to NR patients (Table III), although the difference in the CRS domain was nearly statistically significant (p = 0.072).

Comparison of the total number of aa changes in E2 in SVR and NR patients - A partial sequence of the E2 protein was determined from the sera of all 33 pre-treated patients. This region was found to be well conserved. To determine a potential relationship between initial genetic variability and the response to antiviral treatment, the aa variations of SVR and NR patients were compared. The numbers of substitutions in PePHD and in other parts of the C-terminal region of E2 protein (aa 636-708) were not associated with a SVR to the therapy (Table III). Aa substitutions in PePHD were seen in only seven out of 33 patients (Fig. 2).

**DISCUSSION**

In the present paper, we set out to determine whether the numbers of mutations in the E2-PePHD and NSSA (CRS, ISDR, PKR-binding, NLS and extended V3) regions of HCV genotype 3a were associated with the response to therapy. Previous studies have addressed the question of whether variations in the aa sequence of the HCV E2 protein are associated with antiviral treatment, the aa variations of SVR and NR patients were compared. The numbers of substitutions in PePHD and in other parts of the C-terminal region of E2 protein (aa 636-708) were not associated with a SVR to the therapy (Table III). Aa substitutions in PePHD were seen in only seven out of 33 patients (Fig. 2).

In the present paper, we set out to determine whether the numbers of mutations in the E2-PePHD and NSSA (CRS, ISDR, PKR-binding, NLS and extended V3) regions of HCV genotype 3a were associated with the response to therapy. Previous studies have addressed the question of whether variations in the aa sequence of the HCV E2 protein are associated with patient response to IFN-α therapy. Taylor et al. (1999) described that the PePHD within E2 protein interacts with PKR in vitro and plays an important role in suppressing PKR, suggesting a possible mechanism by which HCV may evade the antiviral effects of IFN. Mutations in this region are believed to influence the response to IFN therapy; however, other studies have reported conflicting results because this region is highly conserved (Berg et al. 2000, Cochrane et al. 2000, Gerotto et al. 2000, Lo & Lin 2001, Murphy et al. 2002, Hung et al. 2003, Saito et al. 2003).

Our study showed that PePHD was highly conserved and no association was found between aa changes in this domain and the response to therapy. Only a few studies have reported a correlation between aa substitutions within PePHD and the effectiveness of IFN.

Different functions have been attributed to NS5A protein. So far, only aa variations found in the NS5A protein of HCV genotype 1b of some isolates found mainly in Japan have been correlated with a sustained virological response (Enomoto et al. 1996, Chayama et al. 2000, Pascu et al. 2004). European and American studies report conflicting results (Duverlie et al. 1998, Berg et al. 2000, Nousbaum et al. 2000, Sarrazin et al. 2000a, Murphy et al. 2002). NS5A protein has always been a research focus in the field. The interplay between NS5A and the IFN system, as well as the role of NS5A in IFN resistance has generated intense interest and has been extensively studied (Tan & Katze 2001, Gale & Foy 2005). NS5A, in particular, may play a pivotal role in the interaction between HCV and cellular signal transduction pathways.

The first studies involving NS5A protein of HCV genotype 3 and resistance to IFN were based on ISDR and PKR-binding analysis. They showed little or no changes in the aas isolated from patients (SVR and NR) compared with the NZL-1 sequence and also found no association with the response to therapy (Saiz et al. 1998, Sarrazin et al. 2000a).

In line with other studies, ISDR and PKR-binding sequences generated in our study showed little variation from the consensus sequence and no association with treatment outcome was observed. These results may indicate that, at least for HCV genotype 3a, these regions cannot be used alone as predictors of sustained virological response to the current treatment.

### TABLE III

Comparative analysis of mutations in E2 and NS5A regions between sustained virological response (SVR) and non-responders (NR) patients with chronic hepatitis C virus (HCV) infection. Data are expressed as numbers of patients with percentage in parenthesis.

<table>
<thead>
<tr>
<th>Region and mutations</th>
<th>All (n = 33)$^a$</th>
<th>SVR (n = 16)</th>
<th>NR (n = 17)$^b$</th>
<th>p value$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2 (aa 636-708)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 1</td>
<td>18 (55)</td>
<td>7 (44)</td>
<td>11 (65)</td>
<td>ns (0.227)</td>
</tr>
<tr>
<td>&gt; 1</td>
<td>15 (45)</td>
<td>9 (56)</td>
<td>6 (35)</td>
<td></td>
</tr>
<tr>
<td>E2 (PePHD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without</td>
<td>26 (79)</td>
<td>11 (69)</td>
<td>15 (88)</td>
<td>ns (0.225)</td>
</tr>
<tr>
<td>≥1</td>
<td>7 (21)</td>
<td>5 (31)</td>
<td>2 (12)</td>
<td></td>
</tr>
<tr>
<td><strong>NS5A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without</td>
<td>13 (41)</td>
<td>7 (44)</td>
<td>12 (75)</td>
<td>ns (0.072)</td>
</tr>
<tr>
<td>≥1</td>
<td>19 (59)</td>
<td>9 (56)</td>
<td>4 (25)</td>
<td></td>
</tr>
<tr>
<td>ISDR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without</td>
<td>7 (22)</td>
<td>11 (69)</td>
<td>14 (88)</td>
<td>ns (0.394)</td>
</tr>
<tr>
<td>≥1</td>
<td>25 (78)</td>
<td>5 (31)</td>
<td>2 (12)</td>
<td></td>
</tr>
<tr>
<td>PKR-binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 1</td>
<td>18 (56)</td>
<td>10 (63)</td>
<td>8 (50)</td>
<td>ns (0.476)</td>
</tr>
<tr>
<td>&gt; 1</td>
<td>14 (44)</td>
<td>6 (37)</td>
<td>8 (50)</td>
<td></td>
</tr>
<tr>
<td>NLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without</td>
<td>23 (72)</td>
<td>13 (81)</td>
<td>10 (63)</td>
<td>ns (0.433)</td>
</tr>
<tr>
<td>I</td>
<td>9 (28)</td>
<td>3 (19)</td>
<td>6 (37)</td>
<td></td>
</tr>
<tr>
<td>Extended V3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 3</td>
<td>18 (56)</td>
<td>8 (50)</td>
<td>10 (63)</td>
<td>ns (0.476)</td>
</tr>
<tr>
<td>&gt; 3</td>
<td>14 (44)</td>
<td>8 (50)</td>
<td>6 (37)</td>
<td></td>
</tr>
<tr>
<td>NS5A (aa 1979-2414)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 13</td>
<td>17 (53)</td>
<td>8 (50)</td>
<td>9 (56)</td>
<td>ns (0.723)</td>
</tr>
<tr>
<td>&gt; 13</td>
<td>15 (47)</td>
<td>8 (50)</td>
<td>7 (44)</td>
<td></td>
</tr>
</tbody>
</table>

Comparisons between groups were made by the $\chi^2$ or Fisher exact test for the categorical variables. $a$: for NS5A domains n = 32 (one sample was not amplified); aa: amino acid; $b$: for NS5A domains n = 16 (one sample was not amplified); $c$: p value of 0.05 was considered statistically significant; CRS: cytoplasmic retention signal; ISDR: interferon sensitivity determining region; NLS: nuclear localisation signal; ns: not significant; PePHD: PKR-eIF2a phosphorylation homology domain; PKR: protein kinase R.
Kohashi et al. (2006) have identified several specific sites in the ISDR that critically affect replication, including aa 2209, which were associated with a strong response to IFN in patients with genotype 1. Nevertheless, in our study we found that only one patient with SVR had a substitution in aa 2209 and two other patients (SVR and NR) showed substitutions only in aa 2216, but these findings were not correlated to therapeutic response.

Others domains of the NS5A protein were also analysed: CRS, NLS and extended V3. We may speculate that if results maintain the same trend in a larger sample of patients, statistical analysis would show a significant correlation between mutations in the CRS and virological response to treatment, as SVR patients showed a higher number of mutations in this region compared to NR. The N-terminal portion of the HCV NS5A protein (CRS) was characterised by Brass et al. (2002) as a membrane-anchoring region which forms a highly conserved amphipathic α-helix that is rich in tryptophan residues in the hydrophobic portion. This portion is necessary and sufficient to mediate the association between the NS5A protein and the membrane of the endoplasmic reticulum (ER). Meanwhile, the hydrophilic portions (located in the cytosol) support the stability of the HCV replication complex (Huang et al. 2005, Appel et al. 2006). Some authors have reported that the presence of mutations in certain codons in this region could lead to the loss of the ability of NS5A to interact with membranes, which could directly or indirectly harm viral replication, indicating that this region has an important role in the replication cycle of HCV (Brass et al. 2002, Macdonald & Harris 2004, Penin et al. 2004). In this way, the variability detected in the CRS region of some responders in our study could at least partly account for the changes in viral replication kinetics. These changes could affect the process of anchoring this protein in the ER and might contribute to the reduced efficiency of viral replication, favouring the control of viremia and the action of IFN or RBV. Studies based on mutations in the CRS region are almost nonexistent and more mutation analyses in these regions are necessary.

In the present paper, no significant differences were found upon comparing sequence variations in the NLS and extended V3 domains between SVR and NR groups. However, other studies have shown a significant association between treatment effectiveness and a high number of mutations within V3, with a significant difference between SVR and NR patients (Murphy et al. 2002, Sarrazin et al. 2002, Puig-Basagoiti et al. 2005, Goyal et al. 2007, Veillon et al. 2007). The total number of aa variations observed in NS5A of HCV genotype 3a isolates from our study was not associated with treatment response.

In summary, no significant association between sequence variations of E2/NS5A and response to IFN therapy was found. However, the lack of significant differences between the groups could be due to the small sample size. Further studies on a larger number of patients are desirable to confirm or refute this finding. However, the present paper has provided information on HCV genotype 3a within Brazil; the 65 sequences (32 sequences of NS5A and 33 sequences of E2) deposited in GenBank may contribute to evolutionary and epidemiological studies of HCV. In recent years, much has been learned about the ability of NS5A to modulate the IFN response. However, more studies are needed to clarify its effects on other aspects of immunity. With the recent development of an HCV infection model (such as the sub-genomic HCV replicon cell culture system), future work will most certainly be aimed at investigating the role of NS5A in other aspects of the HCV life cycle, including viral entry and assembly. As new in vivo models evolve, the role of NS5A in viral replication in animal models will certainly be defined. Thus, future research should provide new clues as to the various functions of this truly remarkable multifunctional regulatory protein.


