Conservation of hepatitis C virus nonstructural protein 3 amino acid sequence in viral isolates during liver transplantation


Instituto Butantan, Laboratório de Imunologia Viral, São Paulo, Brazil; Inserm U 748, Strasbourg, France; Université Louis Pasteur, Strasbourg, France; Departamento de Gastroenterologia, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil; Departamento de Física e Informática, Instituto de Física de São Carlos, São Carlos, SP, Brazil; Departamento de Patologia, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil; and Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil

Received December 2008; accepted for publication January 2009

SUMMARY. As a consequence of selective pressure exerted by the immune response during hepatitis C virus (HCV) infection, a high rate of nucleotide mutations in the viral genome is observed which leads to the emergence of viral escape mutants. The aim of this study was to evaluate the evolution of the amino acid (aa) sequence of the HCV nonstructural protein 3 (NS3) in viral isolates after liver transplantation. Six patients with HCV-induced liver disease undergoing liver transplantation (LT) were followed up for sequence analysis. Hepatitis C recurrence was observed in all patients after LT. The rate of synonymous (dS) nucleotide substitutions was much higher than that of nonsynonymous (dN) ones in the NS3 encoding region. The high values of the dS/dN ratios suggest no sustained adaptive evolution selection pressure and, therefore, absence of specific NS3 viral populations. Clinical genotype assignments were supported by phylogenetic analysis. Serial samples from each patient showed lower mean nucleotide genetic distance when compared with samples of the same HCV genotype and subtype. The NS3 samples studied had an N-terminal aa sequence with several differences as compared with reference ones, mainly in genotype 1b-infected patients. After LT, as compared with the sequences before, a few reverted aa substitutions and several established aa substitutions were observed at the N-terminal of NS3. Sites described to be involved in important functions of NS3, notably those of the catalytic triad and zinc binding, remained unaltered in terms of aa sequence. Rare or frequent aa substitutions occurred indiscriminately in different positions. Several cytotoxic T lymphocyte epitopes described for HCV were present in our 1b samples. Nevertheless, the deduced secondary structure of the NS3 protease showed a few alterations in samples from genotype 3a patients, but none were seen in 1b cases. Our data, obtained from patients under important selective pressure during LT, show that the NS3 protease remains well conserved, mainly in HCV 3a patients. It reinforces its potential use as an antigenic candidate for further studies aiming at the development of a protective immune response.

Keywords: HCV, hepatitis, liver transplantation, NS3.

INTRODUCTION

As a consequence of selective immune pressure during hepatitis C virus (HCV) infection, a high rate of nucleotide mutations in the HCV genome leads to the generation and selection of escape mutants [1,2], which in turn contribute to the persistence of infection. Nucleotide mutations leading to changes in amino acid (aa) sequence and possibly in protein structure take place mainly in the structural proteins of the virion envelope that include the envelope glycoproteins 1 and 2 (E1, E2), which are the main molecules involved in virus entry into target cells and in the generation of immune responses. Therefore, the development of HCV-antigenic proteins based on the natural candidates, such as E1 and E2, is hampered by the high rate of nucleotide mutations leading to aa substitutions that characterize the quasispecies variants in the host [3].
During HCV replication, a single polyprotein encoded by the HCV genome is processed by cell and viral proteases giving rise to structural and nonstructural proteins (NS) respectively. In addition to E1 and E2, the structural region yields also the core protein, and all the three proteins constitute the virus particles. The NS proteins play several roles during viral replication [4]. The NS3 has been described as a viral-encoded protease, containing a serine protease (N-terminal portion) and NTPase/helicase (C-terminal portion) activity [4]. Besides its key role in HCV processing and replication [5,6], NS3 has been shown to contribute to activation of genes for inducible nitric oxide synthase (iNOS) and production of reactive oxygen species (ROS), which lead to double-stranded DNA breaks [7]. NS3 was also shown to transform cells [8,9], to block transforming growth factor-α/Smad3-mediated apoptosis [10], to suppress beta-interferon induction, thus helping to establish persistent HCV infection [11–13], and to inhibit the function of the tumour suppressor p53 [14]. Therefore, there is evidence for a critical role of NS3 in HCV multiplication and oncogenesis [15], although correlations between NS3 primary and secondary structures with hepatocellular carcinoma (HCC) are still unconvincing [16,17].

As NS3 is essential for viral polyprotein processing and, hence, productive HCV replication, it has been considered as a potential target molecule for the development of specific inhibitors [18] as well as a candidate for vaccine development. Isolated HCV-specific cytotoxic T lymphocytes (CTL) are thought to be effective in limiting viral spread and in clearing virus during infection. When isolated from patients with acute hepatitis C, such cells were shown to recognize epitopes in the NS3 region [19].

Further support that the NS3 protein may constitute a target to control HCV replication relies on the fact that the NS proteins are more conserved entities than E1, and more so than E2 [20]. Nevertheless, single mutations in the NS3 region occur naturally or under selective pressure from NS3 protease inhibitors [18]. Therefore, further studies on the frequency and localization of NS3 nucleotide and potential aa substitutions, in patients subjected to different selective pressure, are essential to establish candidate epitopes for drug treatment or to use as antigen targets.

In this study, patients undergoing liver transplantation (LT) as a result of HCV-induced cirrhosis/HCC had their NS3 encoding region sequences defined periodically, their N-terminal aa sequence determined following translation and their protease secondary structure deduced.

METHODS

Patients

Six patients undergoing liver transplantation for end-stage HCV-associated cirrhosis were included in this study. Five of them had been diagnosed also with HCC. Three of them were infected with genotype 1b and three with genotype 3a. The viral load BT varied from 0.7 to 9 × 10⁵ IU/mL and 3 months after transplantation varied from 5.9 to 7.5 × 10⁵ IU/mL. No antiviral treatment was given during the first month after LT. All patients had HCV recurrence as defined additionally by histological criteria [21].

RNA extraction, RT-PCR and N-terminal HCV NS3 amplification

The NS3 encoding region was amplified from HCV-RNA extracted from plasma samples collected before and after LT. Total RNA was extracted using the commercially available QIAamp Viral RNA Kit (Qiagen, Uniscience, SP, Brazil). RNA was reverse-transcribed into cDNA using Moloney murine leukaemia virus reverse transcriptase (InvitrogenTM Life Technologies, Carlsbad, CA, USA) and random primers (InvitrogenTM Life Technologies). For the first and second round of amplification, we used sense and anti-sense primers specific for each genotype and subtype. For the first PCR, 10 μL of cDNA was added to 40 μL of a PCR mix [20 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM of each four dNTPs, 25 pmol of 853 IU and 1302 OD primers, and 1.25 U of Taq DNA polymerase (InvitrogenTM Life Technologies)]. PCR was carried out in a thermocycler under the following conditions: 30 s at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 52 °C for first PCR and 55 °C for second PCR and 70 s for first PCR and 45 s at 72 °C for second PCR, with a final extension step of 10 min at 72 °C. For the second PCR, 5 μL of the product from the first round PCR was added to another tube containing 45 μL of PCR mix.

NS3 nucleotide sequencing and analysis

The second round PCR products were subjected to cycle sequencing using the ABI Prism BigDye terminator cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) and the nested PCR primers. Sequencing gels were run on an automated ABI Prism 377 DNA Sequencer (Applied Biosystems). All sequences obtained were analysed using the Phred-Phrap-Consed programs (www.phrap.org/phredphrapconsed.html). We analysed the quality of the sequences and aligned them in one contig representing the HCV NS3 sequence. The nucleotide sequence contigs were aligned with the Clustal X program (version 1.81: www.clustal.org) and aa sequences were obtained. Primer sequences were removed by using the BioEdit program (version 7.0.5.3: www.bioedit.software.informer.com). The whole set of NS3 nucleotide sequences obtained from the patients was used to estimate their nucleotide genetic distance (d) with Nei-Gojobori (p-distance) methods by using the MEGA program (version 4.0) [22].

The frequency of nonsynonymous substitutions per nonsynonymous site (dN) and of synonymous substitutions per
synonymous site (dS), as well as its ratio, were calculated by pairwise comparison of every single sequence using the Synonymous Nonsynonymous Analysis Program. Nucleotide sequence data have been deposited in the GeneBank database under the following accession numbers: EU847430–EU847463.

The frequency of aa substitutions in the protein and known CTL epitopes in our samples was evaluated. Analyses were carried out using the Los Alamos Data set (http://hcv.lanl.gov/content/sequence/HCV/ToolsOutline.html).

For the phylogenetic analysis, reference HCV sequences [NZL (D17763), EU315120, EF154713, D28917 for HCV 3a and HPCLYPRE (M62321), AF165064, AB426117, AB435162, AF165059, AF356827 for HCV 1b] were compared with the NS3 sequences obtained in our study. After alignment (first 994 NS3 nucleotides), molecular phylogenies were estimated using the distance approach implemented in PAUP* v4.0b [23]. The evolutionary model of DNA substitution and parameters used (TrN + I + G) were estimated by MODELTEST v3.06 [24]. For the distance tree reconstruction, the neighbour-joining algorithm [25] was performed. Robustness of phylogenetic groups was evaluated using 1000 bootstrap replicates [26,27].

RESULTS AND DISCUSSION

In a previous study [21,29], we had analysed the nucleotide sequence of the hypervariable region 1 of E2 before and after LT and observed a significant decrease in aa diversity of the quasispecies at day 7 and month 1 compared with that BT. Nucleic acid diversity was lower for genotype 1 than for genotype 3.

Among the tools we have to control viral disease progression, vaccines and antiviral drugs hold promise and are the subject of intense investigations. Nevertheless, we still do not have at our disposal effective vaccines or protease inhibitors for HCV patients which could be used in combination with interferon-alpha and ribavirin [30,31].

To understand the failure of vaccines, antiviral drugs or LT in controlling HCV progression, efforts have been made to characterize HCV particles present in the host during infection [21]. Indeed, the main factor determining the failure in controlling HCV progression is represented by the appear-

The secondary structure of the NS3 protease was predicted by the computer-assisted Garnier-Osguthorpe-Robson method as implemented in the NPSA GOR IV server [28].
ance of so-called escape mutants. Interventions, such as vaccination, viral inhibitors or LT, exert a selective pressure on HCV replication and drive the generation of HCV escape mutants. The immune response, through its humoral component, would drive mutations in the structural surface glycoproteins E1 and E2. Antiviral drugs would play their selective pressure on other regions of the virus genome, a clear example represented by mutations occurring in the NS genes coding for protease upon treatment with protease inhibitors [32]. The selective pressure driven by LT in HCV patients is much less studied and more difficult to predict as it may play a role in different pathways of virus infection.

By studying the evolution of the NS3 nucleotide sequence and deducing the NS3 N-terminal aa sequence and secondary structure of isolates from LT patients (data not shown), we could observe that the rate of dS nucleotide substitutions was much higher than that of dN ones, the high values of the dS/dN ratio suggesting no adaptive evolution sustained by LT mutation selection pressure and no variant NS3 viral population selection. By constructing a phylogenetic tree based on the NS3 sequences (Fig. 1), serial samples from each patient were found in clusters without significant bootstrap association, with the exception of those from patient 15, indicating an important similarity among these sequences for the studied region. In agreement with the dS/dN ratios, the phylogenetic tree did not show any relevant changes in NS3 after LT compared with the time BT, arguing against adaptive NS3 selection. The clinical genotype and subtype assignments (5'UTR) were confirmed by the phylogenetic study.

Several differences in the NS3 N-terminal aa sequence were observed in patient samples BT as compared with the NS3 GenBank aa reference sequences (data not shown). These differences were higher for patients carrying 1b. Several aa substitutions were observed at the N-terminus of NS3 at different times after LT, but residues of sites described to be

<table>
<thead>
<tr>
<th>HCV NS3 sequence location</th>
<th>Epitope (HCV genotype)</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1031–1039</td>
<td>AYSQQTRGL (1b)</td>
<td>61.11</td>
</tr>
<tr>
<td>1038–1047</td>
<td>GLGCITSL (1b)</td>
<td>88.89</td>
</tr>
<tr>
<td>1055–1069</td>
<td>VEGQVISTAAQTFT (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1067–1081</td>
<td>QTLATCINGVCWTV (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1071–1090</td>
<td>ATCINGVCWTVHGAGTRTI (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1073–1081</td>
<td>CINGVCWTV (1a)</td>
<td>16.67</td>
</tr>
<tr>
<td>1085–1100</td>
<td>AGTRTIASPKGVPVM (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1097–1101</td>
<td>TRITIASPKGVPQMY (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1100–1107</td>
<td>MYYNVDQD (na)</td>
<td>100</td>
</tr>
<tr>
<td>1107–1121</td>
<td>DLYGWPAPQGRSLT (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1123–1131</td>
<td>CTCGSSDLY (1a)</td>
<td>100</td>
</tr>
<tr>
<td>1131–1139</td>
<td>YLVRHADV (na)</td>
<td>100</td>
</tr>
<tr>
<td>1169–1177</td>
<td>LLCPAGHAV (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1171–1190</td>
<td>CPAGHAGIFRAVCTRGRG (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1175–1183</td>
<td>HAVGLFFRAA (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1243–1251</td>
<td>ATLGFGAYM (1b)</td>
<td>83.33</td>
</tr>
<tr>
<td>1261–1270</td>
<td>TLGFYAMS (1a)</td>
<td>83.33</td>
</tr>
<tr>
<td>1262–1270</td>
<td>LGEFGAYMSK (HCV-H)</td>
<td>83.33</td>
</tr>
<tr>
<td>1265–1274</td>
<td>GAYMSKAHGV (1b)</td>
<td>88.89</td>
</tr>
<tr>
<td>1284–1298</td>
<td>TITGAPVYSTGK (1a)</td>
<td>0</td>
</tr>
<tr>
<td>1287–1296</td>
<td>TGSPITSTDY (1b)</td>
<td>0</td>
</tr>
<tr>
<td>1292–1300</td>
<td>TSTY TGKFL (1b)</td>
<td>100</td>
</tr>
<tr>
<td>1300–1310</td>
<td>LADGCSGGAY (1a)</td>
<td>100</td>
</tr>
<tr>
<td>1318–1327</td>
<td>CHAQQATTVL (3a)</td>
<td>0</td>
</tr>
</tbody>
</table>

na, not available information.

Table 1 Frequency of known CTL epitope sequences described in the Los Alamos data set in the studied samples
involved in important functions of NS3, notably those of the catalytic triad and zinc binding, remained unaltered in terms of aa sequence (data not shown).

For both genotype 1b and 3a samples, established or reverted aa substitutions in NS3 took place randomly. We observed that most CTL epitopes described for genotype 1a were not present in our 1b or 3a samples. Most of the CTL epitopes described for genotype 1b (and some for 1a or without genotype information) were present at high frequency in our 1b samples. Some CTL epitopes were found in all samples and the only CTL epitope described for genotype 3a was not present in our samples (Table 1).

As far as we could ascertain, at the primary structure level, we did not observe a pattern of nucleotide mutation or aa substitution that could be related to HCC prognosis. Nevertheless, by analysing the NS3 protease-induced secondary structure, we observed at different times after LT a few alterations in samples of patients carrying genotype 3a but none in those of 1b (data not shown).

In conclusion, our data showed that N-terminal NS3 aa sequences of genotype 1b differed significantly from the NS3 GenBank reference aa sequences. Interestingly, the N-terminal NS3 aa sequences of genotype 3a did not differ significantly from the GenBank reference aa sequences. These aa sequences of different LT patients carrying both genotypes 1b and 3a remained quite well conserved up to 12 months after LT, particularly in their regions of well-known biological activity. N-terminal NS3 aa secondary structure analysis revealed only a few changes in genotype 3a samples. These findings support the hypothesis that, in contrast to immune responses or antiviral drug treatment, LT selective pressure does not drive important changes in the NS3 region, at least in genotype 1b patients. This may have important implications in designing an HCV NS3 candidate vaccine and its efficacy in LT patients.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the CAPES-COFECUB (463/04). C.A. Pereira is a recipient of CNPq IA senior research fellowship.

REFERENCES


© 2009 Blackwell Publishing Ltd


