Differences between hepatitis C virus 5′ untranslated region quasispecies in serum and liver

Sook Jin Jang,1,2 Lian-Fu Wang,1 Marek Radkowski,1 Jorge Rakela3 and Tomasz Laskus1

1 Division of Transplantation Medicine, Thomas E. Starzl Transplantation Institute, University of Pittsburgh, BST 1513E., 200 Lothrop Street, Pittsburgh, PA 15213, USA
2 Department of Clinical Pathology, Chosun University Hospital, Chosun, Republic of Korea
3 Division of Gastroenterology and Hepatology, University of Pittsburgh Medical Center, Pittsburgh, USA

It is unclear whether the sequence populations of hepatitis C virus (HCV) quasispecies in the liver and in serum are different, as a variety of studies on this subject provide conflicting results. In the current study, the populations of HCV 5′ untranslated region (5′ UTR) sequences in paired serum and liver samples from six patients with chronic hepatitis were analysed. Liver-derived, negative-strand viral RNA was amplified with a highly strand-specific Tth-based assay, and extensive measures, including accounting for template copy number, were undertaken to lower the risk of sporadic artefactual polymorphism. Amplified sequences were compared by single-strand conformation polymorphism analysis and by direct sequencing of identified differences. In four patients, liver samples were found to contain variants within the quasispecies which were not found in serum or negative-strand viral RNA, while in the remaining two patients, low virus titre prevented a reliable quasispecies analysis. These results suggest the presence in the same individual of HCV variants differing in the 5′ UTR and possibly replicating with different kinetics.

Introduction

Hepatitis C virus (HCV) is the major aetiological agent of parenterally transmitted non-A, non-B hepatitis (Choo et al., 1989; Kuo et al., 1989). In the majority of cases, HCV infection persists indefinitely, leading to chronic hepatitis, cirrhosis and hepatoma (Alter et al., 1992; Kiyosawa et al., 1990; Simonetti et al., 1992). Like many other RNA viruses, HCV has a very high mutation rate and circulates as a population of closely related genomes, referred to as quasispecies (Domingo et al., 1985; Martell et al., 1992; Steinhauer & Holland, 1987). This dynamic and highly plastic nature of HCV has been postulated to play an important role in evading eradication by the host immune system (Shimizu et al., 1994; Taniguchi et al., 1993). The nature of quasispecies is still poorly understood and is the subject of intensive investigation, as some studies have suggested that changes in quasispecies correlate with the progression of liver disease and treatment outcome (Enomoto et al., 1994; Honda et al., 1994; Koizumi et al., 1995).

Several recent studies, in which serum- and liver-derived viral sequences were cloned for analysis, have suggested that the structure of quasispecies in the liver occasionally may be different from that circulating in the serum (Cabot et al., 1997; Navas et al., 1998; Shimizu et al., 1997). However, all these studies were conducted on the E2 hypervariable region 1, in which the number of variants within quasispecies is expected to be especially high (Hijikata et al., 1991; Weiner et al., 1991). Comparison of PCR products amplified from such divergent and rapidly changing regions is technically challenging; for example, as the template copy number in these studies was not adjusted and a relatively small number of clones was studied, part of the observed quasispecies heterogeneity could be artificial and represent sampling variations compounded by Taq polymerase and cloning errors. The results of two other studies, in which HCV quasispecies were compared by single-strand conformation polymorphism (SSCP) analysis, contradict each other (Maggi et al., 1997; Sakamoto et al., 1995). Again, both studies were conducted on the E2 hypervariable region and the template copy number was not adjusted for individual reactions. Methodological differences and the lack of accounting for template copy number are also likely to be
corresponded to 100 case of serum samples, the amount of RNA used in the reaction artefacts than cloning of PCR products and the subsequent and direct sequencing, as this approach may be less prone to morphism related to sampling error becomes a major problem.

The amplified sequences were compared by SSCP analysis and direct sequencing, as this approach may be less prone to artefacts than cloning of PCR products and the subsequent study of individual clones (Smith et al., 1997). Moreover, SSCP analysis identifies all major variants within the quasispecies comprising at least 3% of the virus population (Laskus et al., 1996). Some of these variants are likely to be missed by the typical strategy of sequencing 10–20 clones (Cabot et al., 1997; Navas et al., 1998; Shimizu et al., 1997).

Methods

**Biological samples.** Paired serum and liver samples were collected from six patients who underwent liver transplantation for end-stage, HCV-related chronic liver disease. All were hepatitis B virus surface antigen-negative and none had received any antiviral treatment prior to the study. RNA was extracted from homogenized liver tissue or serum by means of a modified guanidinium thiocyanate–phenol–chloroform technique with commercially available kits (Ultraspex 2 and Ultraspec 3; Biotecx). One μg of RNA (as determined by spectrophotometry) extracted from liver was reverse transcribed as described below; in the case of serum samples, the amount of RNA used in the reaction corresponded to 100 μl.

**RT–PCR.** MMLV RT-based detection of HCV RNA has been described in detail elsewhere (Laskus et al., 1997a, b). This assay, although totally strand-non-specific, was capable of detecting 10 genome equivalents of the template.

**Strand-specific RT–PCR.** Tth-based RT–PCR detection of HCV negative-strand RNA was performed as described elsewhere (Laskus et al., 1997a). As previously reported (Laskus et al., 1997a, b), this assay was capable of detecting about 100 genome equivalents of the correct strand, while detecting non-specifically about 10^8 genome equivalents of the incorrect strand.

All titres were determined by analysing 10-fold serial dilutions of the RNA template, since at this serial dilution the results were reproducible from run to run. Titres were calculated by assuming that the end-point dilution contained 10 genome equivalents when tested by the MMLV RT-based assay and 100 genome equivalents when tested by the Tth-based, strand-specific assay. When compared to a commercial quantitative assay (AmpliCor; Roche), the results were accurate to within one order of magnitude.

Extensive measures, outlined elsewhere (Laskus et al., 1996), were employed to prevent and to detect carry-over contamination. All RT–PCRs included positive controls, consisting of end-point dilutions of respective RNA strands, and negative controls included uninfected livers and sera.

**Analysis of HCV quasispecies.** For the purpose of sequence comparison, ‘nested’ protocols were used to maximize the yield of PCR product. Amplification of the 5’UTR region was conducted by RT–PCR as described previously (Laskus et al., 1997a); the internal primers were 5’ ACTGCTTTCAAGGAGGCATCT 3’ (nt 57–79) and 5’ CAAGCACCCTATCAGGCAGTACC 3’ (nt 307–285). HCV sequences were compared by SSCP analysis. PCR products were purified with a DNA-binding resin system (Wizard PCR, Promega). A sample of 0.1 μg of the purified product in 9 μl alkaline denaturing buffer (50 mM NaOH, 1 mM EDTA) was mixed with 1 μl SSCP loading buffer (95% formamide, 0.5% bromophenol blue, 0.5% xylene cyanol) and was subjected to non-denaturing 8% PAGE (acrylamide:bis-acrylamide, 19:1; Sigma) in 1 × Tris–borate–EDTA buffer at 400 V for 4–6 h at a constant temperature of 25 °C. The bands were visualized by silver staining (Silver Stain, Promega). When necessary, individual bands were extracted from the gel, reamplified and then sequenced directly, as described previously (Laskus et al., 1998).

Subsequently, all PCR products were sequenced directly in both directions by the Sanger dideoxy chain-termination method with a modified T7 DNA polymerase (Sequenase version 2.0; USB). Sequences were read manually with the help of an EasyReader digitizer (Hitchiti) and were analysed with the DNASIS 2.1 for Windows software (Hitachi).

HCV genotypes were determined by direct sequencing of the NS5 region, as described elsewhere (Laskus et al., 1996). RNA secondary structures were predicted by using the program RNAstructure version 2.52 (Walter et al., 1994; Zuker, 1989).

The following measures were implemented to lower the risk of artefactual polymorphism. (i) When SSCP analysis revealed differences between sequences, the analysis was repeated in an independent experiment with a new RNA template. Both of these reactions had to be identical when run next to each other on the same SSCP gel before the difference was considered to be genuine. (ii) Because of the risk of sampling errors associated with low template copy number, the titre of all samples studied was measured. Whenever possible, the analysis was repeated after adjusting the amount of virus template to be the same in each reaction compared. (iii) To lower the risk of artefacts related to high template copy number (Wang et al., 1997), the products of the first round of amplification were diluted 1:100 before they were used as the template for the second round of PCR. (iv) All samples were sequenced directly in both directions. To rule out incorporation errors by Taq polymerase, all direct sequencing was routinely repeated from a new amplification reaction.

Results and Discussion

Fragments derived from the HCV 5’UTR were amplified successfully from all paired liver and serum samples. However, the actual virus titre ranged from 10^1 to 10^4 genome
Serum sequences (Laskus et al., 1992). The presence of genotype 2b was confirmed in all samples from this patient and reactions with primers specific for other genotypes, including genotype 1a, were persistently negative (not shown). This assay, when tested on a mixture of known genotype samples, allowed for the detection of a minor strain present at a ratio of 1:10^4 with respect to the major strain. Taking into account the amount of viral template in samples from patient 4 (Table 1), this is probably close to the practical limit of sensitivity in this case.

### Table 1. Titres of positive and negative strands of HCV RNA in liver tissue and sera from six patients with end-stage chronic hepatitis

<table>
<thead>
<tr>
<th>Patient</th>
<th>Genotype</th>
<th>Liver</th>
<th>Serum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+RNA</td>
<td>−RNA</td>
<td>+RNA</td>
</tr>
<tr>
<td>1</td>
<td>1a</td>
<td>10^6</td>
<td>10^3</td>
</tr>
<tr>
<td>2</td>
<td>1a</td>
<td>10^6</td>
<td>10^3</td>
</tr>
<tr>
<td>3</td>
<td>1b</td>
<td>10^3</td>
<td>10^3</td>
</tr>
<tr>
<td>4</td>
<td>2b</td>
<td>10^3</td>
<td>10^3</td>
</tr>
<tr>
<td>5</td>
<td>1b</td>
<td>10^3</td>
<td>10^2</td>
</tr>
<tr>
<td>6</td>
<td>1a</td>
<td>10^3</td>
<td>10^2</td>
</tr>
</tbody>
</table>

Equivalents per 100 µl for serum and from 10^3 to 10^6 genome equivalents per µg RNA for liver. HCV negative-strand RNA was detected in all liver samples, at titres which were 1–2 orders of magnitude lower than the titres of the positive strand (Table 1).

SSCP analysis of PCR products amplified from serum, liver and from liver-derived negative-strand RNA revealed in-distinguishable major band patterns, compatible with the presence of identical ‘master’ viral sequences, in four of the six patients studied (patients 1–4 in Fig. 1). In these four cases, all SSCP bands were detected reproducibly from run to run. The presence of identical consensus sequences was confirmed by direct sequencing of PCR products.

In the remaining two patients (patients 5 and 6 in Fig. 1), however, differences in the ‘master’ sequences were evident: in patient 6 the negative-strand RNA sequences were not compatible with the rest of the samples and in patient 5 all three ‘master’ sequences, serum, liver and liver-derived negative strand, were different. These differences were subsequently verified by direct sequencing of PCR products (not shown). However, as shown in Table 1, samples from both of these patients contained very small amounts of viral template and, more importantly, on repeated testing, the SSCP analysis results were not reproduced invariably, thus failing the major requirement for verification of the results. This inconsistency of the results is not surprising, as we have recently shown that end-point dilution of viral template leads to amplification of diverse sequences (Laskus et al., 1998). However, increasing the amount of template was not feasible, as we found that high concentrations of RNA (≥ 5–6 µg per reaction) were commonly inhibitory to RT–PCR.

In order to exclude the possibility that the low virus titre in patients 5 and 6 was due to the use of suboptimal 5’ UTR primers, HCV RNA titres were analysed by using primers specific for the NS5 region, and were found to be similar. Thus, the low virus titre was more likely to be related directly to the biological material studied (cirrhotic livers from patients with end-stage liver disease), in which few intact hepatocytes capable of supporting HCV replication are expected to remain.
Fig. 1. Analysis by SSCP of HCV 5' UTR sequences amplified from serum and liver from six patients with chronic hepatitis C. Lanes S and L represent viral sequences recovered from serum and liver, and lanes L(–) represent viral negative-strand RNA sequences amplified from liver tissue by strand-specific Tth-based RT–PCR. Patients 1–4 show identical major band patterns and patients 5 and 6 show different major band patterns. However, in the latter two patients virus titres were low and the results were not observed consistently. In patients 1–4, SSCP bands which were unique to the liver (indicated by arrowheads) were extracted from the gel and sequenced; these sequences are shown in Fig. 2.

Fig. 2. Nucleotide sequence alignment of the 5' UTR fragments of HCV recovered from patients 1–4. The sequences are compared to the prototype sequence, HCV-1, published by Choo et al. (1991). Dots indicate identity with HCV-1; dashes indicate gaps introduced to maximize the alignment. In all four patients the consensus viral sequences (identified by the suffix C) amplified from serum, liver and liver negative-strand HCV RNA were identical. However, sequences unique to the liver (identified by the suffix Q), which were derived from SSCP bands indicated by arrowheads in Fig. 1, differed by 2–17 nucleotide substitutions from the respective ‘master’ sequences.

As the source of HCV infection in patient 4 was unknown, it is unclear whether he could have been exposed to multiple or heterogeneous infectious sources. Alternatively, the infecting strain could represent an as yet unknown sequence.

When compared with previous reports (Cabot et al., 1997; Maggi et al., 1997; Navas et al., 1998; Sakamoto et al., 1995; Shimizu et al., 1997), the present study has several unique features. Firstly, it was conducted on the most conserved part of the viral genome, which lowers the chance of artefactual polymorphism being detected. Secondly, we analysed replicating viral sequences amplified with a truly strand-specific assay. Third, the ‘effective’ template copy number was determined and, when necessary, adjusted for each sample. Not unexpectedly, gross sequence differences were limited to cases where the initial template copy number was found to be small. However, all liver samples contained variants that were not mirrored by sequences amplified from liver-derived viral negative-strand RNA, and which were not present in serum.

These differences between serum and liver quasispecies suggest that different quasispecies may replicate with different dynamics. Thus, the quasispecies replicating at a ‘high’ level would be exemplified in our study by liver negative-strand and serum sequences as well as by the bulk of liver-derived positive-strand sequences, whereas the quasispecies replicating at a ‘low’ level would be detectable only as part of the liver-derived positive-strand sequences. The existence of virus variants replicating at a low level could be a way of avoiding immune surveillance, and the biological basis for such low replication could be provided by changes in the 5' UTR. However, it is also possible that the liver-specific variants
observed represent commonly occurring defective genomes that are unable to be packaged. Interestingly, several sequence changes found in the liver-specific variants, like the presence of four A residues at positions 204–207, are unprecedented among HCV 5’ UTR sequences and several changes showed high convergence between different samples (all four Q sequences had the same C to A substitution at position 204 and three had an insertion of A after position 205).

The 5’ UTR, which is the most conserved part of the whole HCV genome, contains an internal ribosomal entry site allowing translation in a cap-independent manner (Honda et al., 1996; Tsukiyama-Kohara et al., 1992). Several studies have indicated that variability in this region is constrained by the requirement for specific secondary structures in viral RNA (Brown et al., 1992; Simmons et al., 1993; Smith et al., 1995). Interestingly, this structure appeared to be generally well-preserved in the liver-unique variants; however, even minor changes in such a crucial region might possibly affect replication.

Alternatively, 5’ UTR variations could represent adaptations for different liver-cell populations. In picornaviruses, which replicate in a similar manner, 5’ UTR structures have been shown to interact with a variety of cellular proteins (Jang & Wimmer, 1990; Pelletier & Sonenberg, 1988); moreover, these interactions seem to be important in determining the host range of individual viruses (Agol, 1991). Taking this into consideration, it could be speculated that changes in the 5’ UTR may represent cell-specific adjustment. According to this scenario, quasispecies variants that are not reflected in HCV negative-strand RNA and in serum could inhabit cells capable of sustaining a low level of virus replication (e.g. liver lymphoid cells).

The presence of two distinct functional compartments characterized by ‘high’- and ‘low’-level virus replication is currently unclear. The 5’-untranslated region quasispecies of hepatitis C virus (HCV) quasispecies in the liver may not be reflected by analysis of circulating HCV virions. Structure of replicating hepatitis C virus (HCV) quasispecies in the liver may not be reflected by analysis of circulating HCV virions. Journal of Virology 71, 1732–1734.


Received 3 August 1998; Accepted 29 October 1998