Evolution of naturally occurring 5′ non-translated region variants of hepatitis C virus genotype 1b in selectable replicons

Hans C. van Leeuwen,[1] Chantal B. E. M. Reusken,[1]
Marko Roeten,[1] Tim J. Dalebout,[1] Jose Ignacio Riezu-Boj,[2]
Juan Ruiz[2] and Willy J. M. Spaan[1]

Correspondence
Hans C. van Leeuwen
H.C.van_Leeuwen@lumc.nl

[1]Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

[2]Division of Hepatology and Gene Therapy, Clínica Universitaria/Department of Medicine, Fundación para la Investigación Medica Aplicada (FIMA), University of Navarra, Pamplona, Spain

Quasispecies shifts are essential for the development of persistent hepatitis C virus (HCV) infection. Naturally occurring sequence variations in the 5′ non-translated region (NTR) of the virus could lead to changes in protein expression levels, reflecting selective forces on the virus. The extreme 5′ end of the virus' genome, containing signals essential for replication, is followed by an internal ribosomal entry site (IRES) essential for protein translation as well as replication. The 5′ NTR is highly conserved and has a complex RNA secondary structure consisting of several stem–loops. This report analyses the quasispecies distribution of the 5′ NTR of an HCV genotype 1b clinical isolate and found a number of sequences differing from the consensus sequence. The consensus sequence, as well as a major variant located in stem–loop IIIa of the IRES, was investigated using self-replicating HCV RNA molecules in human hepatoma cells. The stem–loop IIIa mutation, which is predicted to disrupt the stem structure, showed slightly lower translation efficiency but was severely impaired in the colony formation of selectable HCV replicons. Interestingly, during selection of colonies supporting autonomous replication, mutations emerged that restored the base pairing in the stem–loop. Recloning of these altered IRESs confirmed that these second site revertants were more efficient in colony formation. In conclusion, naturally occurring variants in the HCV 5′ NTR can lead to changes in their replication ability. Furthermore, IRES quasispecies evolution was observed *in vitro* under the selective pressure of the replicon system.

INTRODUCTION

A major complication of hepatitis C virus (HCV) infection is its asymptomatic progression into chronic hepatitis, which can lead to cirrhosis of the liver and ultimately hepatocellular carcinoma and liver failure (Farci & Purcell, 2000). Therapy with interferon alpha and ribavirin, a guanoside analogue, provides the highest sustained virological response rates pivotal to the treatment for chronic HCV liver disease (Pawlotsky, 2003). A contributory factor for development of persistent infection of HCV is that its genome shows considerable genetic heterogeneity as a result of accumulation of mutations during viral RNA replication. This variability, which is characteristic of RNA viruses, results from the fact that the viral RNA polymerase lacks a proofreading activity needed for repair of mis-incorporated nucleotides during RNA synthesis. Mutations that occur during virus replication might confer selective advantages or disadvantages to the progeny virus, depending on varying local environment conditions including the presence of innate and adaptive immune response, antivirus treatment and availability of permissive cells. The existence of such a population of variants within a single individual is called a quasispecies (Farci & Purcell, 2000).

HCV, which is a distinct member of the family *Flaviviridae*, has a positive-stranded RNA genome of ∼9600 nt that encodes a single long open reading frame of ∼3000 aa. This polyprotein is co-translationally and post-translationally cleaved into ten structural and non-structural proteins (Bartenschlager & Lohmann, 2000). Flanking the open reading frame at each end are a 5′ and a 3′ non-translated region (NTR), approximately 342 and 225 nt in length, which contain signals required for replication (Friebe *et al.*, 2001; Friebe & Bartenschlager, 2002). Translation of the HCV open reading frame is mediated via the 5′ NTR that

includes a structural RNA element identified as an internal ribosomal entry site (IRES) (Tsukiyama-Kohara et al., 1992; Wang et al., 1993). The whole 5′ NTR folds into a complex RNA structure containing four distinct domains, with domains II, III and part of domain IV (up to the AUG) forming the IRES (Fig. 1a). Through a direct interaction with this IRES element, the ribosomal 40S subunit is recruited to the vicinity of the start codon, directly placing the AUG in the ribosomal P site, thereby eliminating the requirement for a 5′ cap structure and ribosomal scanning (Hellen & Sarnow, 2001).

The 5′ NTR, even though the most highly conserved part of the virus genome, shows a quasispecies distribution with minor variants in its sequence (Lu et al., 2000). Because virus particles in serum are likely to be released from the liver but also from compartments such as lymphocytes or dendritic cells, the sequence diversity found in the IRESs of these assorted particles was suggested to reflect their translational activity and tropism for these compartments (Laporte et al., 2000, 2003; Lerat et al., 2000).

In this work, we studied the 5′ NTR sequence heterogeneity from serum of an HCV genotype 1b (HCV1b)-infected human containing extremely high virus titres. Several minor and major variants were found in its sequence. One major variant, which had lost base pairing in stem–loop IIIa of the HCV IRES, and therefore altered the predicted secondary RNA structure fold, was tested in colony formation of subgenomic replicon HCV RNAs. We find that under selection pressure there is a strong constraint on the HCV IRES to conserve its structure and function.

METHODS

Patient. Serum/EDTA blood was retrieved from a 70-year-old male patient. Transmission route and date of transmission are not known. No known risk factors for HCV positivity were identified. HCV antibodies were present as seen by ELISA (Cobas Core anti-HCV EIA II; Roche Diagnostics) and genomic HCV RNA was detected in peripheral blood by an in-house qualitative RT-nested PCR using two pairs of primers from the 5′ NTR: outer sense primer 5′-CCTGTGAGGAACTACTGTCT-3′, outer antisense primer 5′-CTATCAGGCAGTACCACAAG-3′, inner sense primer (JL3) 5′-ACTGTCTTCACGCAGAAAGC-3′ and inner antisense primer (JL4) 5′-GACCCAACACTACTCGGCTA-3′. HCV RNA was found to be repeatedly positive indicating a chronic infection. In June 2000, the HCV load was $7 \cdot 9 \times 10^8$ IU ml^{-1} as measured by real-time RT-PCR assay, using the LightCyler System (Roche Diagnostics). HCV quantification was performed in a total volume of 10 μl containing 7 μl of the PCR mixture (1× LightCycler-FastStart DNA Master SYBR Green I; Roche Diagnostics), 6 mM MgCl$_2$, 8·35 pmol of each primer (JL3 and JL4) and 3 μl cDNA. The following parameters were used: initial denaturation at 95 °C for 15 s and 45 cycles of 10 s at 95 °C for denaturing, 5 s at 56 °C for annealing and 10 s at 72 °C for extension. Fluorescence was measured at the end of each extension period and the amount of HCV RNA was calculated by interpolation, using a standard curve generated by amplification under the same conditions of known amounts of HCV RNA (NAP HCV RNA Quantification Panel kit; Acrometrix). Evidence of inflammation and fibrosis was found on liver biopsy, with a Knodell score of six. The patient was otherwise in good condition. He never received antivirus treatment. An informed consent was signed. All HCV sequences derived from this strain were marked with the letters EU.

HCV 5′ NTR amplification. The 5′ NTR was identified by cRACE (Maruyama et al., 1995). In short, first strand cDNA was obtained from the isolated HCV RNA using a 5′-phosphorylated oligonucleotide AS2648, 5′-ACCAGGCAGCACAGAAGAACAC-3′ and Thermoscript reverse transcriptase (Invitrogen). This cDNA starts at position 2648 and runs to the 5′ end of the genome. After removal of the template RNA with 0·5 M NaOH, the cDNA was circularized by T4 RNA ligase (Invitrogen). This circularized single-stranded DNA was subsequently used as a template for amplification using primers S2293, 5′-TGCAATTGGACTCGAGGAGAGAGCG-3′, and H179, 5′-ACTCGGCTAGCAGTC-3′. The resulting 600 bp band was cloned into the pCR2.1-TOPO vector (Invitrogen). These clones, which included the junctional region between the 5′ end and the primer used for cDNA synthesis, were then sequenced in both directions by using an ABI Prism 310 genetic analyser (Applied Biosystems) and the Dye Dideoxy terminator sequencing kit (Applied Biosystems).

Plasmid constructions. The replicon vector used to test IRES mutations was pFK-I389neo/NS3-5/5.1 [pFK5.1, see Krieger et al. (2001)]. Mutant IRES replicon constructs were generated using the QuickChange Site-Directed Mutagenesis Kit (Stratagene). Because additional unwanted mutations could be introduced during the QuickChange procedure, a large NruI–SpeI fragment from the original pFK5.1 construct was placed back into the mutated replicon. The region corresponding to nt 1–269 containing the mutated IRES was sequenced.

Dual luciferase reporter constructs containing mutated IRESs A53G and C168G were generated by a fusion PCR-based strategy from pDualLuc-HCVwt/ΔC that has been described previously (Reusken et al., 2003). For pDualLuc-A53G we used primers (A) 5′-ACTCAC-CGGTTCCGCAGACCACTATG-3′ and (B) 5′-CCATAGATCACT-CCCCTGTGAGGGAC-3′ on replicon pFK5.1 A53G. The resulting PCR product was used as primer together with (C) 5′-AAATA-CAAAGGATATCAGGTGGCCCCCGCTGAATTGG-3′ on pDualLuc-HCVwt/ΔC. This fusion PCR product was cloned into pDualLuc-HCVwt/ΔC using EcoRV and AgeI. For pDualLuc-C168G we used primers (D) 5′-TATCAGGCAGTSWCCACAAGGCCTTTCGCG-3′ and (E) 5′-GGTCTGCGGAACCGGTGAGTACAGCGG-3′ on replicon pFK5.1 C168G. The resulting PCR product was used as primer together with (F) 5′-ATTCATGCATACGCGTGCCAGCCCCCGAT-TGGGG-3′ on pDualLuc-HCVwt/ΔC. This fusion PCR product was cloned into pDualLuc-HCVwt/ΔC using NsiI and StuI. All fragments generated by PCR were sequenced upon cloning. pDualLuc-A53G + C168G was generated by exchanging the 1681 bp AgeI–ApaI fragment from pDualLuc-A53G with the corresponding fragment from pDualLuc-C168G.

Cell culture. Huh-7 hepatocellular carcinoma cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5 % heat-inactivated FCS, 2 mM L-glutamine and 1 % non-essential amino acids. For cell lines carrying HCV replicons, 500 μg G418 ml^{-1} was added (Geneticin; Life Technologies).

Transfection of cultured cells. Electroporation and selection of G418-resistant cell lines were done as described by Lohmann et al. (1999).

Luciferase assays. Tissue culture expression studies were performed as described previously (Reusken et al., 2003). In short, Huh-7 cells were infected with recombinant vaccinia virus, expressing T7 RNA polymerase, vTF7-3 (Fuerst et al., 1986), at an m.o.i. of 10. After 1 h incubation at 37 °C, cells were washed in PBS and transfected with plasmid using LipofectACE reagent OptiMEM (Life Technologies). Cells were washed with PBS 7 h post-transfection,

and lysed in passive lysis buffer (Dual-luciferase Reporter Assay; Promega). Cells were stored at $-80\,^{\circ}C$ and thawed before reporter activities were measured with the dual-luciferase reporter assay using a luminometer (TD-20/20; Turner BioSystems). The experiments were performed in triplicate.

RESULTS

Quasispecies distribution in the 5′ NTR

Characterization of the heterogeneity within the HCV 5′ NTR was performed on two separate RNA extractions from a 200 µl serum sample of an untreated patient (designated EU) with chronic hepatitis related to an HCV1b infection. After reverse transcription followed by circularization of the cDNA formed (see Methods) and PCR amplification using high-fidelity DNA polymerase a total of 26 cDNA clones of the 5′ NTR was generated (12 and 14 cDNA clones, respectively). As expected from the genotype determination (see Methods), all sequenced variants had the characteristic group of nucleotide substitutions reported for genotype 1b (Honda et al., 1999). Furthermore, when comparing the consensus sequence of the 26 cDNA clones to other HCV1b IRES sequences, it differed only in the loop of stem–loop I (see Fig. 1b). The extreme 5′ end of the genome including domain I is not part of the IRES but is essential for replication (Rijnbrand et al., 1995; Friebe et al., 2001). Nine cloned variants corresponded to the consensus sequence and make up the prevalent quasispecies. The other variants contained one to three nucleotide substitutions in the IRES as compared to the consensus (see Fig. 1 and legend). In total 20 different nucleotide changes were found, including one insertion (C insertion at position 119) and two nucleotides deleted (63 and 277, see Fig. 1a). Most mutations occurred in the single-stranded structures, or the base changes resulted in preservation of base pairing when they arose in the double-stranded structures (i.e. $A:U\leftrightarrow G:U\leftrightarrow G:C$). No compensatory base changes that maintain base pairing were found in any double-stranded stem. Major variants, representing more than 10 % of the quasispecies, contained substitutions at position A53 to G (6/26 = 23 %) and at position C168 to G (5/26 = 19 %) and an insertion of a cytidine in a stretch of C residues between nt 119 and 127 (4/26 = 15 %). These three major variants were found in RT-PCRs derived from the two separate RNA isolations, making it unlikely that these were due to errors introduced by the reverse transcriptase and polymerase. The A53G mutation and the C119 insertion were also observed in other IRES quasispecies studies and are present in the EMBL database (Lu et al., 2000; Laporte et al., 2000). A natural variant at position C168 has also been described where this position was changed to an A (AF041307; Quarleri et al., 2000). The C168G variant was always found in combination with the A53G mutation but one cloned A53G variant was found without the C168G substitution (Fig. 1 legend), suggesting that A53G is the founder of C168G. The C insertion at position 119 was not found in combination with the other

major variants and probably arose separately during quasispecies selection.

In the secondary RNA structure fold, the alteration of cytidine 168, located in stem–loop IIIa, to a guanidine (see Fig. 1a) results in a loss of pairing with base G159. Because stem–loop IIIa, as part of the IIIabc four-way junction, is a functionally important element (Kieft et al., 2002), which is directly contacted by both the 40S ribosomal subunit and the eukaryotic initiation factor 3 (eIF3) (Kieft et al., 2001), we subjected this major variant to further analysis.

Effect of stem–loop IIIa mutations on translation

To establish whether the major quasispecies mutations had an effect on translational activity we used the bicistronic dual-luciferase reporter system (Collier et al., 1998), which provides a way to correct for possible differences in transfection efficiencies (Reusken et al., 2003). This dual luciferase reporter construct contains an upstream cap-dependent firefly luciferase gene (Fluc), serving as a control for transfection efficiency, and downstream an internal Renilla luciferase gene (Rluc), which is driven by the HCV IRES. A T7 promoter drives cytoplasmic transcription of this bicistronic mRNA. Huh-7 hepatocellular carcinoma cells, infected with recombinant vaccinia virus expressing T7 RNA polymerase, were transfected with dual luciferase plasmids containing the consensus IRES or the IRES quasispecies variants. Cells were lysed 7 h post-transfection and both luciferase reporter activities measured (see Methods). IRES relative translation efficiency was calculated as the ratio of the two luciferase activities (Rluc/Fluc). Relative activities of the IRES variants were compared to that of the consensus IRES, which was set at 1. The results, which represent values obtained from three independent experiments and luminometric measurements done in duplicate, are summarized in Fig. 2. The consensus IRES was the most efficient at initiating translation. The variant containing the A53G and C168G mutations was less efficient showing approximately 70 % activity levels. To analyse the contribution of the individual nucleotide changes to this reduced activity, we introduced the A53G and C168G mutations separately into the consensus IRES. Introducing A53G had very little effect on translation efficiency but replacing position C168 with a G resulted in an almost 20 % reduction in IRES activity (Fig. 2a). This shows that although the combination of the two mutants had the highest effect on IRES activity, the C168G mutation contributed most to the reduction in translation. When considering that stem–loop IIIa is a major determinant of eIF3 and 40S ribosomal subunit binding (Hellen & Sarnow, 2001; Kieft et al., 2001), this decrease of ∼20 % in translation activity seems small for a mutation that is predicted to disrupt the stem–loop (see also Laporte et al., 2000, 2003).

(a)

II

5′ NTR

IIIb

IIIa

19 %

IIIc

IIId
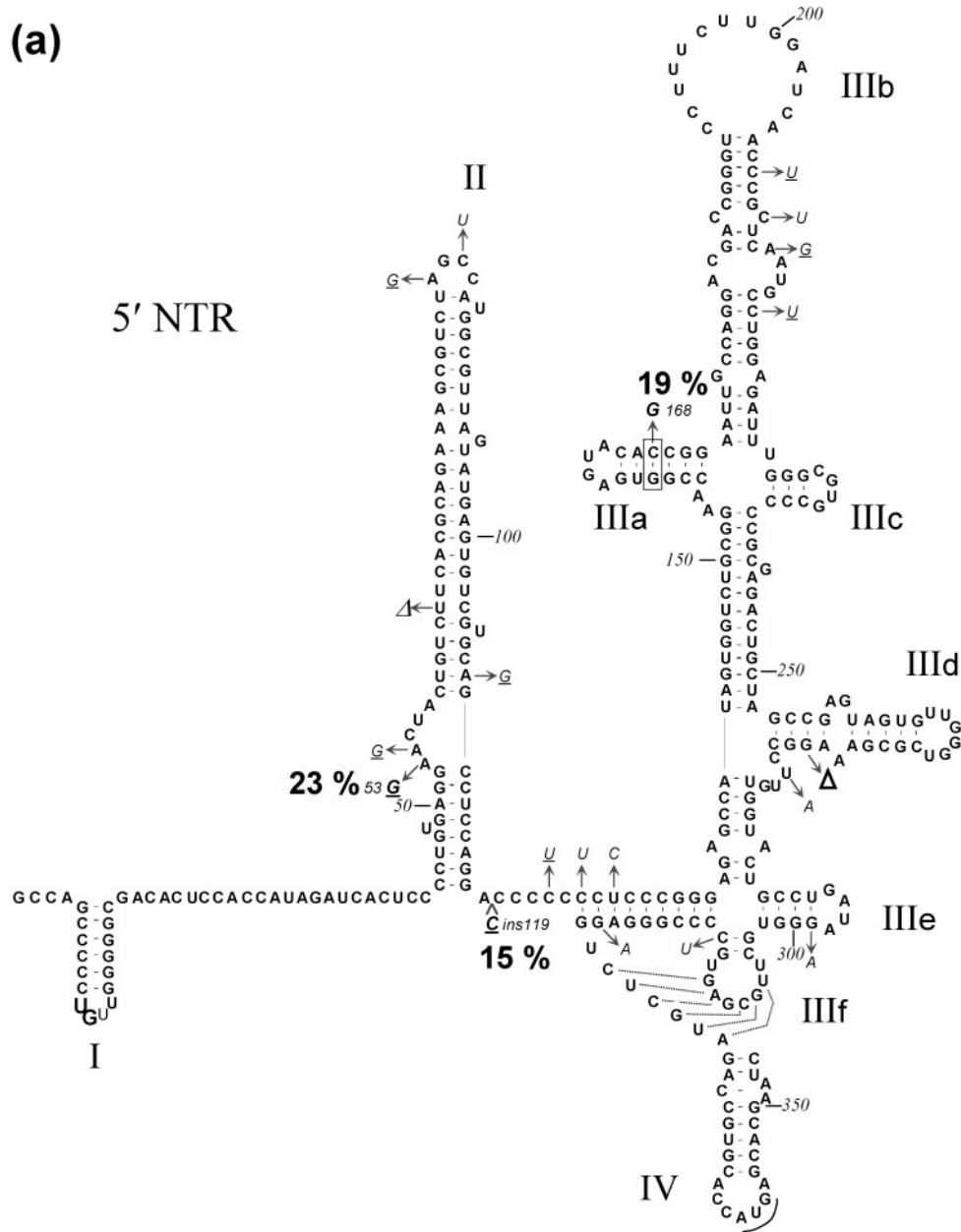
23 %

IIIe

15 %

IIIf

I

IV

Fig. 1. (a) Scheme of the predicted secondary structure of the HCV1b IRES, showing the locations of the nucleotide mutations found in sequenced quasispecies. Major structural domains are labelled I, II, IIIa, IIIb, IIIc, IIId, IIIe, IIIf, IV. Arrows followed by the nucleotide substitution depict mutations found; Δ, base deleted; *ins*, base inserted. Nucleotide substitutions reported previously in the literature or in the HCVDB database are underlined. Note that the secondary structure of stem–loop II is as determined using NMR (Lukavsky *et al.*, 2003) and differs from the predicted secondary structure. The occurrence of major variants is shown as percentages next to the substitutions. Double mutants: A53G+C168G (3×), A53G+G299A, U127C+C211U, Cins119+G322A, C81U+Cins119, A109G+C314U. Triple mutants: A53G+Δ63+C168G, A53G+C168G+C208U, G54A+C123U+U281A. (b) Comparison of nucleotides forming stem–loop I found in genotypes 1a and 1b.

(b)

```
GCCAGCCCCCUGUUGGGGGCGA   1b (EU)
GCCAGCCCCCGAUUGGGGGCGA   1b (con1)
GCCAGCCCCCUGAUGGGGGCGA   1a
```
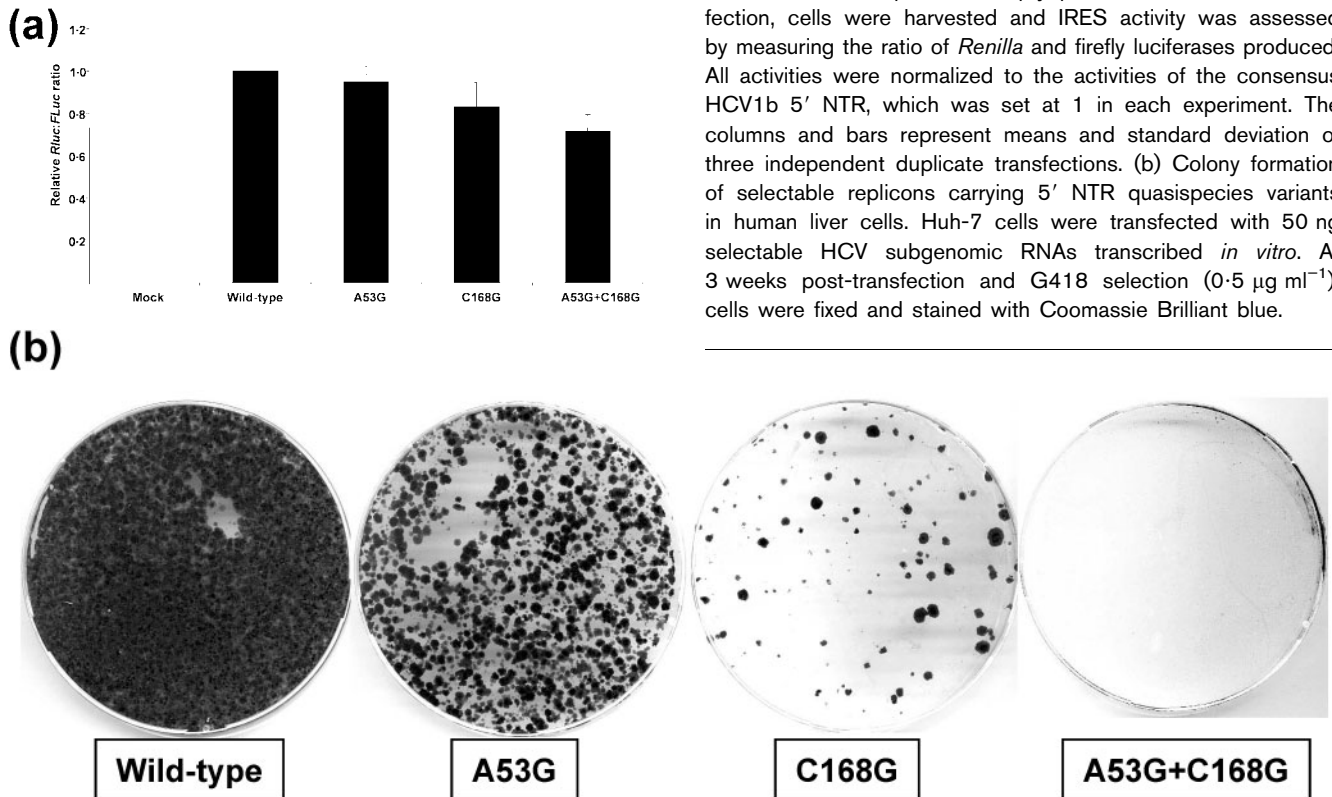
(a)

Fig. 2. Effects of IRES quasispecies variants on translation and selectable replicon colony formation. (a) Relative translational activities of HCV 5′ NTR quasispecies variants in human liver cells (ratio *Rluc* : *Fluc*). vTF7-3-infected Huh-7 cells were transfected with plasmids expressing dicistronic RNAs containing the consensus HCV1b IRES or the IRES quasispecies variants indicated. Mock represents empty plasmid. At 7 h after transfection, cells were harvested and IRES activity was assessed by measuring the ratio of *Renilla* and firefly luciferases produced. All activities were normalized to the activities of the consensus HCV1b 5′ NTR, which was set at 1 in each experiment. The columns and bars represent means and standard deviation of three independent duplicate transfections. (b) Colony formation of selectable replicons carrying 5′ NTR quasispecies variants in human liver cells. Huh-7 cells were transfected with 50 ng selectable HCV subgenomic RNAs transcribed *in vitro*. At 3 weeks post-transfection and G418 selection ($0.5\ \mu g\ ml^{-1}$), cells were fixed and stained with Coomassie Brilliant blue.

(b)

Effect of stem–loop IIIa mutations on replication

Since the HCV IRES contains *cis*-acting elements for translation as well as replication (Friebe *et al.*, 2001; Reusken *et al.*, 2003), we tested the stem–loop IIIa variant in selectable HCV replicons, which enable the study of viral RNA replication in cell culture (Lohmann *et al.*, 1999). In these replicons the region encoding the structural proteins is replaced by a neomycin phosphotransferase gene, which upon expression confers resistance to G418. After transfection of human hepatoma Huh-7 cells, those which sustain high levels of self-replicating HCV RNAs can be selected by adding G418 to the growth medium. The number of G418-resistant colonies that develop directly reflects the efficiency with which a replicon multiplies in cells. For our studies we made use of the HCV1b con1 replicon (pFK–I389neo5.1) (Krieger *et al.*, 2001) containing three cell culture-adaptive mutations in the non-structural genes that enhance replication (Krieger *et al.*, 2001) and thereby enables the detection of 5′ NTR mutants with lower levels of replication.

Because sequences upstream of the IRES are essential for RNA replication, we first engineered selectable replicons in which the con1–1b 5′ NTR was replaced with the consensus 5′ NTR of our EU strain, effectively replacing nt 11 and 12 of stem–loop I (see Fig. 1b). We observed that both replicons were equally efficient in G418-resistant colony formation (data not shown), and therefore used the EU strain consensus IRES in further experiments. Subsequently, the IRES quasispecies variants were engineered in this replicon. Fig. 2(b) shows the effect of these variants in Huh-7 cells using the selectable replicon system. Unexpectedly, the double mutation rendered the replicon non-viable as no colonies were obtained. To analyse why this variant failed to generate stable colonies, we cloned the separate mutations, A53G and C168G, in the replicon. When the individual mutants were assayed in the replicon system, colonies were obtained (Fig. 2b). Compared with the consensus 5′ NTR, the efficiency of colony formation was ∼5-fold lower for the A53G mutation and ∼50- to 100-fold lower for the C168G mutation (see Fig. 2b). This shows that although each mutation did yield G418-resistant colonies, the additive effect of both mutations is apparently

deleterious in the A53G–C168G mutant. When the selection pressure was lowered to 250 µg ml^{-1} G418, a concentration that increases the efficiency of colony formation (Lohmann et al., 1999), eight resistant colonies were found for the double mutation. However, none of them survived the initial trypsinization after picking from the plate. The replicon system might be more stringent in selecting replication-competent RNA molecules, compared to circulating virus genomes (Bartenschlager et al., 2003), explaining the inability of the naturally occurring variant A53G–C168G to form replicon colonies.

Revertants in stem–loop IIIa enhance replication

The lower efficiency of colony formation of the C168G mutant could be due to lower levels of replication or because early on during selection vital changes need to be generated within the replicon in order to enhance replication to sufficient levels. To examine the latter possibility, two independent G418-resistant colonies, isolated after electroporation of the C168G replicon, were expanded, RNA was isolated, and replicon RNA was amplified by RT-PCR and cloned as described above. Ten cloned variants from each G418-resistant colony were sequenced. In both colonies the point mutation C168G was found to be present. None of the cloned variants had reverted to the parental sequence but one or two other mutations accompanied most 5′ NTR variants. Interestingly, in both colonies the nucleotide originally opposite mutation G168, i.e. guanosine 159, had mutated either to a cytidine in one clone

(G159C, frequency 4/10) or to a uridine in the other clone (G159U, frequency 6/10). Both alterations restore the predicted base pairing in stem–loop IIIa (see Fig. 3a). To assess whether these base changes represent mutations that restore efficient replication, we cloned and tested these revertant base pairs, 159C.G168 and 159U.G168, in the replicon. The number of colonies obtained after selection with G418 clearly increased for both revertants compared to the original C168G mutant (Fig. 3b). Levels for the 159U.G168 revertant were five to 10 times higher, while almost wild-type levels were reached for the 159C.G168 revertant (Fig. 3b). This shows that stem–loop IIIa mutations that restore the base pairing between positions 168 and 159 enhance replication and that the poor replication activity observed for the stem–loop IIIa C168G mutant is attributable to disruption of the base pairing.

DISCUSSION

This study provides a detailed analysis of 5′ NTR quasi-species variants present in the serum of an HCV1b-infected individual. The isolated consensus 5′ NTR was typical for HCV1b but differed at two positions in stem–loop I (Fig. 1b). This short stem–loop formed by residues 5–20, upstream of the IRES, is not required for translation (Rijnbrand et al., 1995). RNA replication, in contrast, is completely blocked when stem–loop I is deleted (Friebe et al., 2001). The negative strand of stem–loop I folds into its mirror image (Schuster et al., 2002; Smith et al., 2002), and probably acts as an element for initiation of new
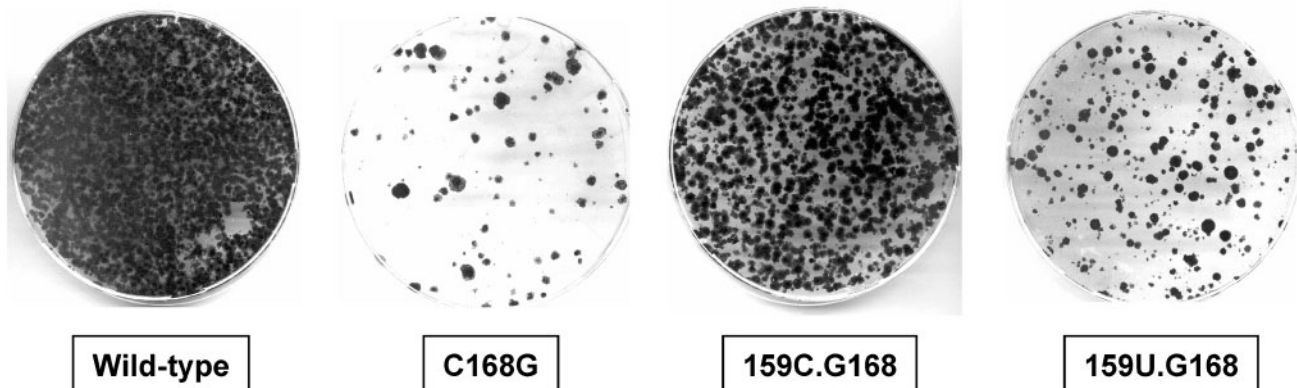

(a)
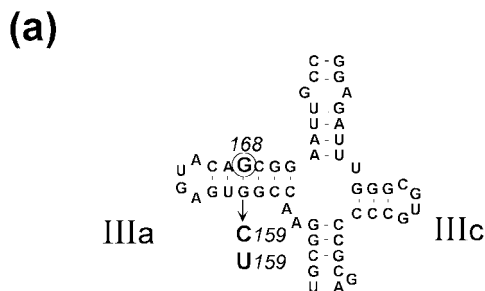
Fig. 3. Revertants of the stem–loop IIIa mutation restore base pairing and enhance colony formation. (a) Secondary structure of the stem–loop IIIabc junction, including the C168G quasi-species variant (circled). The two second site revertants found at position 159 in two independent G418-resistant selectable replicon colonies are indicated. (b) Colony formation of revertants of the 5′ NTR C168G replicon in human liver cells. Huh-7 cells were transfected with selectable HCV subgenomic RNAs containing 5′ NTR revertants 159C.G168 and 159U.G168. At 3 weeks post-transfection and G418 selection (0·5 µg ml^{-1}), cells were fixed and stained with Coomassie Brilliant blue.

(b)

genomic HCV plus-strand RNA (Bartenschlager & Lohmann, 2000). The stem of stem–loop I, on the negative strand, is critical for interaction with the HCV NS3-helicase, whereas the loop does not contribute to NS3 binding (Banerjee & Dasgupta, 2001). In the subgenomic replicon system, we observed that the HCV1b stem–loop I variant, containing UG at positions 11 and 12 in the loop (see Fig. 1), replicated as efficiently as the HCV1b con1 strain.

In addition to the differences in stem–loop I, a major variant was found in stem–loop IIIa, located at position 168, which is predicted to alter the secondary RNA structure due to loss of base-pairing with nucleotide 159 (see Fig. 1a). This mutation, although resulting in only a small reduction in translation activity (Fig. 2a), showed a dramatic reduction in replication (Fig. 2b). Mutations that disrupt the IIIabc four-way junction result in a pronounced loss of translation initiation activity (Kieft et al., 1999, 2002). The minor reduction in IRES activity observed for the C168G variant therefore suggests that the cruciform structure for IIIabc is still intact. Stem–loop IIIa is conserved between pestiviruses and HCV (Fletcher & Jackson, 2002) and is involved in direct interaction with both eIF3 and 40S ribosomal subunits (Kieft et al., 2001). NMR and crystallography have been used to solve detailed structures of HCV IRES stem–loops IIId, IIIe and the IIIabc junction (Lukavsky et al., 2000; Kieft et al., 2002; Collier et al., 2002). In the crystal of the IIIabc junction, part of loop IIIa, containing the base paired region 159–161, was missing due to disorder (Kieft et al., 2002), thereby making structural analysis of the C168G mutation impossible.

Interestingly, when RNA was recovered from colonies originating from replicons containing the stem–loop IIIa C168G variant, several molecules in the population were found to carry an additional mutation that restored base pairing with the mutant G168 position (Fig. 3a). When these clones were inserted back in the original replicon, higher levels of colony formation were observed (Fig. 3b), confirming that they are true revertants and base pairing of stem–loop IIIa is required for optimal replication. The observation that the compensatory mutations failed to rescue replication to wild-type levels suggests that sequence-specific contacts with the initiation complex at this position might be essential. In addition, as colony formation by the 159C.G168 revertant reached higher levels than the 159U.G168 revertant, stability of the base pairing appears important [C.G = −2·4 kcal mol^{-1} (−10·0 kJ mol^{-1}) versus U.G = −1·4 kcal mol^{-1} (−5·9 kJ mol^{-1})].

These emerging revertants could be due to their existence prior to selection in the replicon system, i.e. generated during in vitro transcription by the T7 polymerase, or by mutation from the template during RNA replication. Since the two selected colonies arose from the same in vitro transcription reaction but show different solutions for restoration of the base pairing (either G.C or G.U base pair), we consider the latter mechanism more likely.

Owing to the error-prone nature of the viral polymerase, mutations are expected to occur randomly distributed over the 5′ NTR region. Clearly, only mutations compatible with replication and translation can be propagated. Whether the stem–loop II/IIIa mutations we observed confer a survival advantage or disadvantage remains unclear. We suggest that, although the virus adaptation led to a low replication, it could well be that the genetic defect reflects genetic flexibility.

The con1 replicon we used contains the non-structural proteins of a genotype 1b strain different to our strain from which we obtained IRES quasispecies variants. It is possible that the 5′ NTR variants found are tailored to the virus coding region and interactions may be suboptimal. Further work with subgenomic replicons containing the non-structural genes of our isolate should help to clarify this point.

The unique structure of the HCV IRES makes it an attractive target for the development of antivirus agents directed against this RNA element such as antisense oligonucleotides, transacting ribozymes and siRNAs (Kurreck, 2003). Since stem–loop IIIa is highly conserved and involved in both replication and translation, it could present a good candidate for antivirus targeting. This work, however, shows that revertants can arise in the IRES during selective pressure of the subgenomic replicon system. In vivo, drug-resistant variants in HCV-infected patients have been reported in the virus RdRp during ribavirin monotherapy, where a single amino acid change conferred increased resistance (Young et al., 2003). Strategies directed against HCV RNA elements should therefore target multiple regions in order to reduce the probability of emerging escape mutants.

ACKNOWLEDGEMENTS

REFERENCES

Banerjee, R. & Dasgupta, A. (2001). Specific interaction of hepatitis C virus protease/helicase NS3 with the 3′-terminal sequences of viral positive- and negative-strand RNA. J Virol 75, 1708–1721.

Bartenschlager, R. & Lohmann, V. (2000). Replication of hepatitis C virus. J Gen Virol 81, 1631–1648.

Bartenschlager, R., Kaul, A. & Sparacio, S. (2003). Replication of the hepatitis C virus in cell culture. Antiviral Res 60, 91–102.

Collier, A. J., Tang, S. X. & Elliott, R. M. (1998). Translation efficiencies of the 5′ untranslated region from representatives of the six major genotypes of hepatitis C virus using a novel bicistronic reporter assay system. J Gen Virol 79, 2359–2366.

Collier, A. J., Gallego, J., Klinck, R., Cole, P. T., Harris, S. J., Harrison, G. P., Aboul-ela, F., Varani, G. & Walker, S. (2002). A conserved

RNA structure within the HCV IRES eIF3-binding site. *Nat Struct Biol* **9**, 375–380.

Farci, P. & Purcell, R. H. (2000). Clinical significance of hepatitis C virus genotypes and quasispecies. *Semin Liver Dis* **20**, 103–126.

Fletcher, S. P. & Jackson, R. J. (2002). Pestivirus Internal Ribosome Entry Site (IRES) structure and function: elements in the 5′ untranslated region important for IRES function. *J Virol* **76**, 5024–5033.

Friebe, P. & Bartenschlager, R. (2002). Genetic analysis of sequences in the 3′ nontranslated region of hepatitis C virus that are important for RNA replication. *J Virol* **76**, 5326–5338.

Friebe, P., Lohmann, V., Krieger, N. & Bartenschlager, R. (2001). Sequences in the 5′ nontranslated region of hepatitis C virus required for RNA replication. *J Virol* **75**, 12047–12057.

Fuerst, T. R., Niles, E. G., Studier, F. W. & Moss, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci U S A* **83**, 8122–8126.

Hellen, C. U. T. & Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15**, 1593–1612.

Honda, M., Rijnbrand, R., Abell, G., Kim, D. S. & Lemon, S. M. (1999). Natural variation in translational activities of the 5′ nontranslated RNAs of hepatitis C virus genotypes 1a and 1b: evidence for a long-range RNA–RNA interaction outside of the internal ribosomal entry site. *J Virol* **73**, 4941–4951.

Kieft, J. S., Zhou, K. H., Jubin, R., Murray, M. G., Lau, J. Y. N. & Doudna, J. A. (1999). The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J Mol Biol* **292**, 513–529.

Kieft, J. S., Zhou, K. H., Jubin, R. & Doudna, J. A. (2001). Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* **7**, 194–206.

Kieft, J. S., Zhou, K. H., Grech, A., Jubin, R. & Doudna, J. A. (2002). Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nat Struct Biol* **9**, 370–374.

Krieger, N., Lohmann, V. & Bartenschlager, R. (2001). Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations. *J Virol* **75**, 4614–4624.

Kurreck, J. (2003). Antisense technologies. Improvement through novel chemical modifications. *Eur J Biochem* **270**, 1628–1644.

Laporte, J., Malet, I., Andrieu, T. & 7 other authors (2000). Comparative analysis of translation efficiencies of hepatitis C virus 5′ untranslated regions among intraindividual quasispecies present in chronic infection: opposite behaviors depending on cell type. *J Virol* **74**, 10827–10833.

Laporte, J., Bain, C., Maurel, P., Inchauspe, G., Agut, H. & Cahour, A. (2003). Differential distribution and internal translation efficiency of hepatitis C virus quasispecies present in dendritic and liver cells. *Blood* **101**, 52–57.

Lerat, H., Shimizu, Y. K. & Lemon, S. M. (2000). Cell type-specific enhancement of hepatitis C virus internal ribosome entry site-directed translation due to 5′ nontranslated region substitutions

selected during passage of virus in lymphoblastoid cells. *J Virol* **74**, 7024–7031.

Lohmann, V., Korner, F., Koch, J. O., Herian, U., Theilmann, L. & Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **285**, 110–113.

Lu, M., Kruppenbacher, J. & Roggendorf, M. (2000). The importance of the quasispecies nature of hepatitis C virus (HCV) for the evolution of HCV populations in patients: study on a single source outbreak of HCV infection. *Arch Virol* **145**, 2201–2210.

Lukavsky, P. J., Otto, G. A., Lancaster, A. M., Sarnow, P. & Puglisi, J. D. (2000). Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. *Nat Struct Biol* **7**, 1105–1110.

Lukavsky, P. J., Kim, I., Otto, G. A. & Puglisi, J. D. (2003). Structure of HCV IRES domain II determined by NMR. *Nat Struct Biol* **10**, 1033–1038.

Maruyama, I. N., Rakow, T. L. & Maruyama, H. I. (1995). Crace – a simple method for identification of the 5′-end of messenger-RNAs. *Nucleic Acids Res* **23**, 3796–3797.

Pawlotsky, J. M. (2003). The nature of interferon-alpha resistance in hepatitis C virus infection. *Curr Opin Infect Dis* **16**, 587–592.

Quarleri, J. F., Robertson, B. H., Mathet, V. L., Feld, M., Espinola, L., Requeijo, M. P., Mando, O., Carballal, G. & Oubina, J. R. (2000). Genomic and phylogenetic analysis of hepatitis C virus isolates from Argentine patients: a six-year retrospective study. *J Clin Microbiol* **38**, 4560–4568.

Reusken, C. B. E. M., Dalebout, T. J., Eerligh, P., Bredenbeek, P. J. & Spaan, W. J. (2003). Analysis of hepatitis C virus/classical swine fever virus chimeric 5′NTRs: sequences within the hepatitis C virus IRES are required for viral RNA replication. *J Gen Virol* **84**, 1761–1769.

Rijnbrand, R., Bredenbeek, P., van der Straaten, T., Whetter, L., Inchauspe, G., Lemon, S. & Spaan, W. (1995). Almost the entire 5′ non-translated region of hepatitis C virus is required for cap-independent translation. *FEBS Lett* **365**, 115–119.

Schuster, C., Isel, C., Imbert, I., Ehresmann, C., Marquet, R. & Kieny, M. P. (2002). Secondary structure of the 3′ terminus of hepatitis C virus minus-strand RNA. *J Virol* **76**, 8058–8068.

Smith, R. M., Walton, C. M., Wu, C. H. & Wu, G. Y. (2002). Secondary structure and hybridization accessibility of hepatitis C virus 3′-terminal sequences. *J Virol* **76**, 9563–9574.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. *J Virol* **66**, 1476–1483.

Wang, C., Sarnow, P. & Siddiqui, A. (1993). Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *J Virol* **67**, 3338–3344.

Young, K. C., Lindsay, K. L., Lee, K. J., Liu, W. C., He, J. W., Milstein, S. L. & Lai, M. M. C. (2003). Identification of a ribavirin-resistant NS5B mutation of hepatitis C virus during ribavirin monotherapy. *Hepatology* **38**, 869–878.