Infectious cDNA clone of the hepatitis C virus genotype 1 prototype sequence

Robert E. Lanford,¹ Helen Lee,¹ Deborah Chavez,¹ Bernadette Guerra¹ and Kathleen M. Brasky²

Department of Virology and Immunology¹ and Department of Laboratory Animal Medicine², Southwest Regional Primate Research Center, Southwest Foundation for Biomedical Research, 7620 NW Loop 410, San Antonio, TX 78227, USA

A full-length cDNA clone of the hepatitis C virus (HCV) genotype 1 prototype (subtype 1a) sequence was constructed. Synthetic RNA produced from the initial cDNA clone was not infectious following intrahepatic inoculation of a chimpanzee. A consensus clone was prepared by comparison with multiple full-length HCV sequences of genotypes 1, 2 and 3. A total of 11 non-consensus amino acid residues were altered by mutagenesis. Synthetic RNA from the repaired clone initiated a typical, acute-resolving HCV infection following intrahepatic inoculation of a chimpanzee. In addition, at least one of three chimeric cDNA clones constructed between the HCV-1 and H77 genotype 1a strains of HCV was infectious in a chimpanzee. This is the first example of an infectious chimeric HCV clone. An infectious cDNA clone of HCV-1 will be of particular value, since it is the prototype HCV sequence and many commonly used reagents are based on this sequence.

Introduction

Hepatitis C virus (HCV) is a major worldwide health problem. HCV has a high propensity for inducing life-long persistent infections that can progress to significant liver disease including cirrhosis and hepatocellular carcinoma. Approximately 3% of the world's population is chronically infected with HCV (Anon., 1997) and the recent National Health and Nutrition Exam and Survey (NHANESIII) estimates that in the United States 3–4% of the adult population between the ages of 30 and 50 are persistently infected (Alter et al., 1999). Chronic HCV infection is the leading cause for liver transplantation in the US (Hoofnagle, 1997). The current treatment for HCV infection involves 6–12 months of combination therapy with interferon-α and ribavirin (Ahmed & Keeffe, 1999). The majority of individuals infected with genotype 1 will not experience long-term viral clearance with combination therapy.

HCV is a member of the Hepacivirus genus of the Flaviviridae family. Individual isolates have considerable sequence variation and currently six major genotypes are recognized (Robertson et al., 1998). The genome is single-stranded, positive-sense RNA of approximately 9600 nt and encodes a polyprotein with a single open reading frame (ORF) of 3008–3033 aa (Houghton et al., 1994). The structural proteins (capsid, E1, E2 and potentially p7) precede the nonstructural proteins (NS2, 3, 4A, 4B, 5A and 5B) in the polyprotein. Although the precise functions of some of the HCV proteins are unknown, most have been characterized following expression in heterologous systems. NS2 forms an autoprotease with the amino terminus of NS3; NS3 is a serine protease and RNA helicase; NS4A is a cofactor for the serine protease; and NS5B is the RNA-dependent RNA polymerase (Fig. 1) (Reed & Rice, 1998). The functions of NS4B and NS5A are currently unknown. The 5' noncoding region (NCR) contains an internal ribosome entry site for translation of the polyprotein (Honda et al., 1999). The 3' untranslated region (UTR) beginning at the termination codon of the polyprotein contains a short region of significant sequence variation, a poly(U)/polypyrimidine stretch of variable length and a terminal 98 nt conserved sequence, presumably involved in RNA replication (Tanaka et al., 1995, 1996; Kolykhalov et al., 1996; Blight & Rice, 1997).

Advances in HCV research have been hampered by a number of limitations including the lack of a standardized tissue culture system and the lack of a small animal model. The development of infectious cDNA clones of the virus (Kolykhalov et al., 1997; Yanagi et al., 1997, 1998, 1999a; Hong et al., 1999; Beard et al., 1999) has provided a mechanism for performing reverse genetics, but in the absence of a tissue
culture system, the testing of mutant clones has been restricted to the intrahepatic inoculation of chimpanzees with synthetic RNA (Yanagi et al., 1999b; Kolykhalov et al., 2000). The recent description of a genotype 1b replicon system for HCV RNA (Lohmann et al., 1999) suggests that advances in the genetic analysis of HCV replication should be attained in the near future; however, the development of replicons with other strains of HCV has not met with immediate success.

This report describes the production of an infectious cDNA clone of the HCV-1 prototype strain of HCV. Although the initial clone was not infectious in chimpanzees, correction of nonconsensus residues based on alignment of multiple HCV sequences yielded an infectious cDNA. In addition, several chimeric clones with the H77 sequence were produced and at least one of the chimeras was infectious. Although several infectious cDNA clones have been described previously, infectious clones of only one other genotype 1a strain have been described. Three separate clones of the H77 strain have been constructed (Kolykhalov et al., 1997; Yanagi et al., 1997; Hong et al., 1999). In addition, two genotype 1b (Yanagi et al., 1998; Beard et al., 1999) and one genotype 2a (Yanagi et al., 1999a) infectious cDNA clones have been described. Additional infectious clones will be needed in determining whether the considerable genetic diversity of HCV is associated with differences in biological properties. In particular, an infectious clone of HCV-1 will be of value, since the HCV-1 prototype was the first sequence of HCV to be reported (Choo et al., 1991) and many investigators continue to work with reagents based on this sequence.

Methods

**Plasmids.** Development of the pHCV1-SF clone was initiated with three overlapping clones, spanning the HCV ORF, that were derived from the same starting serum as the prototype HCV-1 sequence (Choo et al., 1991; Reyes & Kim, 1991) (a gift from Jungshik Kim; Genelabs Inc.). Sequence analysis of the clones revealed several single nucleotide deletions that resulted in frame-shift mutations. Each deletion was repaired by PCR mutagenesis, and the corrections were confirmed by sequence analysis. The clones were assembled into a full-length ORF in pSP73 (Promega). The 5′ NCR and 3′ UTR were amplified by PCR and were inserted into the full-length clone as described in the text. For cloning purposes, the forward primer for the 5′ NCR contained an EcoRV site (not present within the HCV-1 sequence) directly upstream of the T7 RNA polymerase promoter. The reverse primer spanned the unique Agel site located at nt 155 in the IRES. The unique EcoRV and Agel sites were used to introduce the 5′ NCR into the ORF clone in pSP73. The introduction of the 3′ NCR was accomplished in a similar manner. The reverse primer contained an Xbal site not found within the HCV-1 sequence and the forward primer spanned the unique NotI site at nt 9220 within the ORF. The unique NotI and Xbal sites were used to introduce the 3′ NCR into the competed pHCV1-SF clone. The Xbal site was engineered such that digestion of the plasmid with Xbal and production of runoff transcripts with T7 RNA polymerase would produce either an authentic HCV 3′ terminus or a terminus with four additional nucleotides, depending on whether the single-stranded 5′ extension from the Xbal digest can be efficiently used as a template by T7 RNA polymerase.

The ORFs from the following full-length HCV sequences were aligned to produce the consensus sequence for pHCV1-SF: D50409, D00944, D19098, D17763, D28917, D49374, D63821, D90208, D11168, M58335, M84754, M61596, AF139594, M62321, M67463, D14853, and D63822. The pCV-H77 clone (AF011751) used to make chimeras with HCV-1 was obtained from Jens Bukh (Yanagi et al., 1997).

**Chimpanzees.** Chimpanzees were housed at the Southwest Regional Primate Research Center at the Southwest Foundation for Biomedical Research. Animals were cared for by members of the Department of Laboratory Animal Medicine in accordance with the Guide for the Care and Use of Laboratory Animals, and all protocols were approved by the Institutional Animal Care and Use Committee. Chimpanzees were inoculated with synthetic RNA derived from full-length HCV cDNA clones by ultrasound guided intrahepatic injection. Synthetic RNA was produced from the pHCV1/SF cDNA clone using 1 µg of Xbal linearized template per reaction in ten 20 µl transcription reactions. Transcription reactions were performed using the T7 Megascript kit as described by the manufacturer (Ambion). The quality of the RNA was verified by agarose gel electrophoresis and ethidium bromide staining. Chimpanzees were bled periodically for evaluation of increases in serum alanine transaminase (ALT), the level of viremia by quantitative
RT–PCR, and seroconversion for anti-HCV antibodies. Anti-HCV antibodies were detected using the third-generation enzyme-linked immunosorbent assay (ELISA Testing System 3.0, Ortho Diagnostic Systems, Raritan, NJ, USA).

- **TagMan quantification of HCV RNA.** HCV RNA was isolated from serum or liver tissue using RNazol (Leedo, Houston, TX, USA). HCV RNA was quantified by a real time, 5’ exoribonuclease RT–PCR (TagMan) assay using the ABI 7700 Sequence Detector (PE Biosystems). The primers and probe were derived from the conserved region of the 5’ NCR and were selected using the Primer Express software designed for this purpose (PE Biosystems). The forward primer was from nt 149 to 167 (5’ TCCGGAAACCGGTAGTACA 3’), the reverse primer was from nt 210 to 191 (5’ CGGTTTATCCAAGAAAGGA 3’) and the probe was from nt 189 to 169 (5’ CCGGTCGTCTGGCAATTCG 3’). The fluorogenic probe was labelled with FAM and TAMRA and was obtained from Synthegen (Houston, TX, USA). The primers and probe were used at 10 pmol/50 µl reaction. The reactions were performed using a TaqMan Gold RT–PCR kit (PE Biosystems) and included a 30 min 48 °C reverse transcription step, followed by 10 min at 95 °C, and then 50 cycles of amplification using the universal TagMan standardized conditions: 15 s at 95 °C for denaturation and 1 min at 60 °C for annealing and extension. Standards to establish genome equivalents (ge) were synthetic RNAs transcribed from a clone of the 5’NCR of the HCV-1 strain. Synthetic RNA was prepared using the T7 Megascript kit (Ambion) and was purified by DNase treatment, RNazol extraction and ethanol precipitation. RNA was quantified by absorbance and 10-fold serial dilutions were prepared from 10⁶ to 10 copies using tRNA as a carrier. These standards were run in duplicate in all TaqMan assays in order to calculate genome equivalents in the experimental samples. The calibration curves from one preparation of synthetic RNA to the next were essentially identical and yielded values comparable to commercially available assays.

## Results and Discussion

### Construction of a consensus cDNA clone of the HCV-1 prototype sequence

A full-length ORF of the HCV-1 prototype sequence was assembled as described in Methods. Correct polyprotein processing of the ORF was verified by expression in mammalian cells using the vaccinia system (data not shown). The 5’NCR was PCR amplified using a forward primer that included the minimal T7 promoter such that transcripts produced from the clone would initiate with the authentic 5’G of HCV. The 3’UTR was amplified using a forward primer that spanned nt 9259–9282 and a reverse primer that included a terminal Xhol site and the 3’ terminus of the 98 nt conserved element (Tanaka et al., 1995, 1996; Kolykhalov et al., 1996; Blight & Rice, 1997). A high titre human serum (> 10⁸ ge/ml) containing HCV genotype 1a was used for the amplification of the 3’ terminus, due to the lack of available high titre HCV-1 serum. A number of different clones were examined that varied in length in the poly(U/C) stretch. The clone selected for insertion into the full-length cDNA clone contained a 101 nt poly(U/C) stretch with a single G residue. The G could be an amplification error or may be present in the HCV terminus from the strain used for amplification. A full-length clone was assembled into pSP73 such that cleavage with Xhol and transcription with T7 would yield transcripts with an authentic terminus. The initial clone was not infectious following two separate intrahepatic inoculations of a chimpanzee with synthetic RNA (data not shown).

Next, attempts to produce a consensus clone were initiated, since this strategy had been successful for others (Kolykhalov et al., 1997; Yanagi et al., 1997, 1998, 1999a; Hong et al., 1999; Beard et al., 1999). We chose not to sequence multiple clones from the same starting plasma, due to the effort and expense required. Instead, our ORF was aligned to multiple full-length ORFs in GenBank. Using a similar set of reference sequences, this strategy was successful for the creation of a genotype 1b infectious clone (Beard et al., 1999). A total of 17 full-length sequences was used in the evaluation and comprised the following genotypes: three genotype 3a, two genotype 1b, one genotype 1c, three genotype 2, four genotype 3, one genotype 6 and three of undesignated genotype (see Methods). Any residue that was non-consensus in our clone in a position that was invariant for all other clones was deemed to be in error and was changed to the invariant residue. In addition, a few of the positions where our clone differed from a nearly invariant residue were altered to represent the consensus. Eleven residues were altered by PCR mutagenesis. The amino acid changes are indicated in Table 1. In each case, a DNA fragment flanked by restriction endonuclease sites unique to the full-length clone was subcloned. PCR mutagenesis was performed on a smaller fragment of the subclone that was flanked by restriction endonuclease sites unique to the subclone but not unique to the full-length clone. The modified fragment was inserted into the subclone, the domain modified by PCR was sequence confirmed, and the fragment was returned to the parental full-length clone. The resulting clone

### Table 1. Amino acid changes required to make consensus sequences

The amino acid number in the HCV-1 polyprotein is indicated. The amino acid in the initial clone is shown in single-letter code with an arrow indicating the amino acid change. The protein domain of the polyprotein within which the change was made is indicated.

<table>
<thead>
<tr>
<th>aa no.</th>
<th>Change</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>K → R</td>
<td>Core</td>
</tr>
<tr>
<td>11</td>
<td>N → T</td>
<td>Core</td>
</tr>
<tr>
<td>20</td>
<td>R → Q</td>
<td>Core</td>
</tr>
<tr>
<td>460</td>
<td>Q → R</td>
<td>E2</td>
</tr>
<tr>
<td>546</td>
<td>C → G</td>
<td>E2</td>
</tr>
<tr>
<td>765</td>
<td>P → L</td>
<td>p7</td>
</tr>
<tr>
<td>929</td>
<td>V → A</td>
<td>NS2</td>
</tr>
<tr>
<td>955</td>
<td>L → P</td>
<td>NS2</td>
</tr>
<tr>
<td>2211</td>
<td>F → L</td>
<td>NS5A</td>
</tr>
<tr>
<td>2541</td>
<td>P → S</td>
<td>NS5B</td>
</tr>
<tr>
<td>2857</td>
<td>N → D</td>
<td>NS5B</td>
</tr>
</tbody>
</table>
was infectious in a chimpanzee (see below) and was designated pHCV1/SF.

**Comparison of the HCV-1 sequence with the H77 infectious cDNA clones**

For discussion purposes our clone will be referred to as HCV1-SF to distinguish it from the original HCV-1 sequence (Choo et al., 1991) (M62321), and the H77 clones will be designated by H77 and the first letter of the first author’s last name [H77K (Kolykhalov et al., 1997), H77Y (Yanagi et al., 1997) and H77H (Hong et al., 1999)]. Comparison of HCV1-SF with HCV-1 and the H77 infectious clones revealed significant differences as well as striking similarity for isolates from different individuals. The 5′NCR is identical for each of the clones with the exception that the original HCV-1 sequence contains a C residue rather than the consensus A at nt 204. HCV1-SF differed from H77K, H77Y and H77H at 78, 82 and 76 aa, respectively. HCV1-SF differed from H77K and H77Y at 78 common residues and at 4 additional residues for H77Y. H77H matched HCV1-SF at 3 residues where HCV1-SF differed from H77Y and H77K, and H77H had only one additional mismatch with HCV1-SF not present in either H77K or H77Y. The 78 aa in HCV1-SF that differed from H77K and H77Y were distributed as follows: 0 in core; 7 in E1; 23 in E2 (8 of which were in HVR-1); 2 in p7; 9 in NS2; 11 in NS3; 1 in NS4A; 4 in NS4B; 15 in NS5A; and 6 in NS5B.

HCV1-SF differed from the published HCV-1 sequence at 12 residues (Choo et al., 1991). At 7 of these residues, HCV1-SF matched the H77 sequences, and at 3 HCV1-SF matched H77. Two of the 12 residues in HCV1-SF that differed from HCV-1 were at codons 9 and 11, where HCV-1 has lysine and asparagine, and HCV1-SF has arginine and threonine. Our original sequence actually matched HCV-1 at these positions. However, with the exception of HCV-1, the arginine and threonine residues at these positions are invariant, so we chose to make the consensus changes. Whether these nonconsensus residues affect the infectivity or pathogenesis of HCV-1 in comparison to other strains remains unknown.

The 3′UTR of the HCV-1 sequence could not be compared, since it was not complete. Comparison of the HCV1-SF 3′terminus with the H77 clones revealed that the variable region after the ORF termination codon preceding the poly(U/C) stretch was identical between HCV1-SF and H77K and differed from H77Y by only one residue. HCV1-SF had a 101 nt poly(U/C) stretch containing 11 Cs and 1 G, while H77K contained a 128 nt poly(U/C) stretch with 15 Cs, and H77Y contained an 81 nt poly(U/C) stretch with 14 Cs. In the few infectivity studies done with molecular clones, the length of the poly(U/C) stretch does not appear to have a major influence on infectivity. Although the number and placement of the C residues differ, they tend to cluster to the 3′half of the poly(U/C) stretch. The 98 nt conserved 3′terminus was identical for all four clones; however, the 3 nt preceding this region is TAT for HCV1-SF and AAT for the three H77 clones.

**Infectivity of pHCV1/SF RNA in chimpanzees**

Synthetic RNA was produced using T7 RNA polymerase and XhoI-linearized pHCV1/SF as the template. The full-length quality of the RNA was verified by gel electrophoresis. An HCV naïve chimpanzee (X300) was injected in the liver with synthetic RNA using ultrasound-guided inoculation. Viral RNA was first detected in the serum on week 2 at a level of 8.9 × 10^6 ge/ml using a real-time, quantitative TaqMan RT–PCR assay (Fig. 2). Peak viraemia was observed in week 8 at 1.2 × 10^7 ge/ml, and viral clearance occurred on week 16. The serum ALT levels were above the normal cut-off by week 4 (55 IU/ml). The ALT values peaked on week 10 (306 IU/ml) and returned to baseline by week 16. Seroconversion for anti-HCV antibodies was detected on week 10 using a third-generation anti-HCV ELISA. A liver needle biopsy taken on week 6 displayed moderate signs of hepatitis with hepatocellular cytoplasmic swelling throughout the section and disruption of hepatic cords and sinusoidal spaces. Additional liver biopsies were not taken.

**Evaluation of HCV chimeric genotype 1a clones for infectivity**

Recent studies have demonstrated that chimeras between infectious genotype 1a and 2a clones were not infectious in the chimpanzee (Yanagi et al., 1999). These chimeras combined the structural region of the 2a clone with the nonstructural region of the 1a clone. The lack of infectivity of these chimeric clones has significant implications for future studies directed at elucidating phenotypic properties of various HCV strains, replicons and tissue culture-adapted viruses, since chimeric viruses are often used to map sequences affecting phenotypic properties. The 1a and 2a clones differed by approximately 30% at the amino acid level. To determine whether chimeras produced from more closely related infectious clones would be infectious in the chimpanzee, we constructed three chimeric clones between two genotype 1a clones (HCV1-SF and H77Y). Since these chimeras were within the nonstructural proteins, they tested whether limited amino acid differences would prevent formation of a functional RNA replicase. In addition, successful replication of all three chimeras would provide an opportunity to evaluate sequence evolution in closely related genomes, and response to replication competition during acute infection and immunological selection pressures. Four conserved, unique restriction endonuclease sites between the two clones were employed to construct the chimeras. For each chimera, a segment of the HCV1-SF clone was inserted into the H77C backbone (Fig. 1). SFCV4 contained the HCV1-SF domain from aa 836 to 1101 (nt 2850–3643). SFCV5 contained the HCV1-SF domain from aa 1101 to 1648 (nt 3643–5286). SFCV6 contained the HCV1-SF domain from aa 1648 to 2959 (nt 5286–9220). The chimeras resulted in 10, 11 and 26 aa changes in the H77Y clone, respectively.

Each clone was linearized at the XhoI site and synthetic
RNA was produced using T7 RNA polymerase. The quantity of the RNAs was estimated by absorbance, and the quality of the RNAs was verified by gel electrophoresis. A mixture of the three RNAs was injected into the liver of an HCV naive chimpanzee (X246). HCV viral RNA was detected in the serum at week 2 at $1.3 \times 10^4$ ge/ml, and the viral titre peaked at week 15 at $4.0 \times 10^5$ ge/ml. The viral titre declined to $1.2 \times 10^2$ ge/ml in week 22 and remained at very low levels through week 38 (Fig. 3). Viral clearance occurred by week 43 and X246 remained RT–PCR negative through week 69.

To determine which of the chimeric clones induced infection in chimpanzee X246, RT–PCR assays were developed that were specific for each chimera. For each chimera, two assays were utilized that employed a common forward primer and a chimera-specific cDNA primer. The chimera-specific cDNA primers were chosen such that the 3’ end of the primer would match either the HCV1-SF fragment or the parental H77Y sequence. The specificity of the assays for amplification of HCV1 or H77 was confirmed using chimpanzee sera derived from animals infected with either the HCV-1 or H77 strains. Amplification of RNA derived from the chimpanzee sera in both assays confirmed that only the appropriate sequence was amplified with each set of strain specific primers. Using this approach, only the SFCV6 chimera could be detected in the serum of X246 at multiple early time-points (weeks 2, 8 and 12). In addition, using the last positive bleed date (week 38), a variable region between the two clones was amplified, cloned and eleven clones sequenced to confirm that only the SFCV6 clone could be detected. These data demonstrate that at least one chimera produced between different 1a genotype strains was infectious in the chimpanzee, and thus represents the first example of an infectious chimeric HCV clone. Only a single chimeric virus was detected at both the first and last RT–PCR positive bleeds (weeks 2 and 38). Both the strain-specific RT–PCR assays and the sequencing of multiple clones suggested that the other chimeras if present were present at a level below 10% that of SFCV6. The SFCV6 chimera introduced 26 aa changes into the H77Y clone that spanned from the carboxy terminus of NS3 (aa 1648) to the carboxy terminus of NS5B (aa 2959) (Fig. 1). The amino acid changes were located in the following domains of the HCV polyprotein: 1 in NS4A, 5 in NS4B, 15 in NS5A and 5 in NS5B.
The failure to detect replication of the other two chimeric viruses could be due to technical reasons. However, the RNAs were produced at the same time, and the quality of the RNAs was verified by gel electrophoresis. In addition, since the three RNAs were mixed at equal ratios prior to injection and were injected into three sites, it is unlikely that one RNA gained access to hepatocytes while the other two did not. The other two chimeras may have been replication competent, but the SFCV6 chimera exhibited a selective advantage over the other two chimeras early in infection. Although serum samples taken as early as 2 weeks post-inoculation contained only SFCV6 within the limits of the assay, the other chimeras could have been present at levels 10-fold below the SFCV6 without detection. Rapid replacement of one molecular clone with another was previously observed when a chimpanzee infected with a genotype 2a clone was super-infected with the H77Y genotype 1a clone. The genotype 2a clone was no longer detectable 3 weeks after super-infection (Yanagi et al., 1999a).

Alternatively, the SFCV4 and SFCV5 chimeras may not be replication competent despite the limited changes introduced by the HCV1-SF sequences. SFCV4 contained an HCV1-SF fragment that introduced 10 aa changes, 9 of which were within NS2. SFCV5 contained an HCV1-SF fragment that introduced 11 aa changes, all of which were within NS3. The amino acid changes introduced by the HCV1-SF sequences may have been incompatible with the remainder of the nonstructural proteins in the H77Y backbone. Considerable protein–protein interactions presumably occur within the viral RNA replicase, and some of these interactions may be strain specific in that a change in one protein domain is accompanied by a compensatory change in an interactive domain. Required interactions between the inserted regions and the structural proteins or RNA sequence cannot be excluded either. Unfortunately, the limited availability of chimpanzees and the expense of conducting studies in this animal model do not permit further exploration of these clones by independent inoculation of SFCV4 and SFCV5 to determine whether they possess limited replication competence. Future studies of this nature will benefit from further development of the recently described replicon system for HCV (Lohmann et al., 1999).

The authors thank Jungsuh Kim for the three overlapping HCV-1 fragments used to construct the HCV-1 ORF; Jens Bukh for the pCV-H77 clone used to make chimeras with HCV-1, and Camille Sureau for his contributions to the earlier versions of the clone. This work was supported by grants U19 AI40035 and P51 RR13986 from the National Institutes of Health.

References


Robertson, B., Myers, G., Howard, C., Brettin, T., Bukh, J., Gaschen, B., Gojobori, T., Maertens, G., Mizokami, M., Nainan, O., Netesov, S.,


Received 18 January 2001; Accepted 8 March 2001