JOURNAL OF VIROLOGY, June 1996, p. 3363–3371 Vol. 70, No. 6
0022-538X/96/$04.00+0
Copyright © 1996, American Society for Microbiology

Identification of a Highly Conserved Sequence Element at the 3′ Terminus of Hepatitis C Virus Genome RNA

ALEXANDER A. KOLYKHALOV,[1] STEPHEN M. FEINSTONE,[2] AND CHARLES M. RICE[1]*

Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110-1093,[1] and Division of Virology, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland 20892[2]

Received 12 December 1995/Accepted 21 February 1996

Previous reports suggest that the hepatitis C virus (HCV) genome RNA terminates with homopolymer tracts of either poly(U) or poly(A). By ligation of synthetic oligonucleotides followed by reverse transcription-PCR, cDNA cloning, and sequence analysis, we determined the 3′-terminal sequence of HCV genome RNA. Our results show that the HCV 3′ nontranslated region consists of four elements (positive sense, 5′ to 3′): (i) a short sequence with significant variability among genotypes, (ii) a homopolymeric poly(U) tract, (iii) a polypyrimidine stretch consisting of mainly U with interspersed C residues, (iv) a novel sequence of 98 bases. This latter nucleotide sequence is not present in human genomic DNA and is highly conserved among HCV genotypes. The 3′-terminal 46 bases are predicted to form a stable stem-loop structure. Using a quantitative-competitive reverse transcription-PCR assay, we show that a substantial fraction of HCV genome RNAs from a high-specific-infectivity inoculum contain this 3′-terminal sequence element. These results indicate that the HCV genome RNA terminates with a highly conserved RNA element which is likely to be required for authentic HCV replication and recovery of infectious RNA from cDNA.

Throughout the world, hepatitis C virus (HCV) continues to be a major cause of human liver disease (see reference 17 for a review). Chronic infections with varied clinical outcomes, including chronic hepatitis, cirrhosis, and hepatocellular carcinoma, are common. Protective vaccines for HCV are not yet available, and long-term remission after interferon therapy is observed in only a minority of treated patients. This situation has focused considerable attention on obtaining a better understanding of HCV replication and pathogenesis.

Since the molecular cloning of HCV was reported in 1989 (11), rapid advances have been made in the areas of diagnostics, epidemiology, and molecular virology (17). Sequence analysis of HCV isolates has uncovered enormous genetic diversity, leading to the classification of multiple genotypes and subtypes (see reference 8 for a review). This group of viruses has been classified as a separate genus in the family *Flaviviridae*, which contains two other genera, the flaviviruses (type virus, yellow fever virus) and the pestiviruses (e.g., bovine viral diarrhea virus) (13). All of these viruses are enveloped positive-strand RNA viruses which express their encoded proteins via translation of a single long open reading frame (ORF) (reviewed in reference 25). In the case of HCV, the genome RNA is ~9.4 kb and contains a long and highly conserved 5′ nontranslated region (NTR) which functions as an internal ribosome entry site (36). This element drives expression of a polyprotein of ~3,000 amino acid residues which is proteolytically processed by host signal peptidase and two viral proteinases to produce at least 10 structural and nonstructural (NS) proteins (25, 27). Following the long ORF, most reports suggest that HCV genome RNA contains a short 3′ NTR followed by a poly(U) homopolymer tract (10, 16, 20, 22, 32). In contrast, the genome RNA of HCV-1 (genotype 1a) has been reported to bind to oligo(dT) cellulose (15) and contain a 3′-terminal poly(A) tract.

These reported differences in the HCV 3′ NTR structure are surprising since 5′- and 3′-terminal sequences and structures of positive-strand RNA viruses often function as *cis* elements important for RNA replication and/or packaging; such elements are typically highly conserved. Correct terminal sequences can therefore be of critical importance for recovery of infectious RNA transcripts from cloned cDNA (7). Our previous work on HCV has focused on the HCV H strain which, like HCV-1, is a genotype 1a isolate. In the course of our efforts to construct a full-length functional cDNA clone for this isolate, we have determined the 3′-terminal sequence of the HCV-H genome RNA as well as 3′-terminal sequences for several other independent isolates. These studies have identified a highly conserved 98-base nonhomopolymeric sequence at the 3′ terminus of HCV genome RNA.

MATERIALS AND METHODS

Preparation of HCV RNA from human serum. Typically, total RNA was extracted from 50 μl of HCV-positive serum samples by using 450 μl of RNAzol-B as described by the manufacturer (Tel-Test, Inc.). To aid recovery, glycogen (10 μg) was added to extracted samples prior to precipitation with isopropanol or ethanol. Pellets were washed with 100% ethanol, dried, and resuspended in 10 μl of RNase-free water. H77 patient serum, containing HCV-H subtype 1a, has been described previously (see reference 26 and citations therein). Using nested primer pairs corresponding to a region in the 5′ NTR and an internal competitor RNA (30), we found the RNA titer of this serum to be ~10^7 RNA molecules per ml, which is in general agreement with previous determinations (26). Additional patient samples were obtained from David Gretch (WD), Peter Simmonds (43E), and Jay Hoofnagle (RB and WS).

3′-end determination by using T4 RNA ligase. Synthetic oligodeoxyribonucleotides for ligation and cDNA synthesis were purified by electrophoresis on 10% acrylamide–7 M urea sequencing gels (2). To prevent intramolecular and intermolecular ligation of oligonucleotides, 3′ ends were blocked by addition of ddATP. Blocking reaction mixtures (20 μl) contained 0.5 nmol of gel-purified oligonucleotide, 20 nmol of ddATP, 1 mM $CoCl_2$, 37 U of terminal deoxynucleotidyltransferase, and the buffer supplied by the manufacturer (Boehringer Mannheim, Indianapolis, Ind.). After incubation at 37°C for 1 h, $MgCl_2$ was added to 10 mM, and the blocked oligonucleotide was precipitated by the addition of 5 volumes 100% ethanol and collected by centrifugation. The blocked oligonucleotide was 5′ phosphorylated in a 50-μl reaction mixture containing 100

* Corresponding author. Mailing address: Department of Molecular Microbiology, Washington University School of Medicine, 600 South Euclid Ave., St. Louis, MO 63110-1093. Phone: (314) 362-2842. Fax: (314) 362-1232.

TABLE 1. Oligonucleotides used for ligation, cDNA synthesis, and PCR amplification

Oligonucleotide	Sequence (5′ to 3′)	Position[a]	Sense[b]
143	CGCACCCTGTCCGACTACAACATCC		
192	TTGAATTCGGCCCTGCAGGCCACAACAGTC		
206	GACTGTTGTGGCCTGCAGGGCCGAATT		
228	**CAAGTCGAC**GGGGAGACATTTATCACAGC	9271–9290	+
229	**TTGTCG**ACTCCTGCTCGCTGCAGGGGTA	9329–9350	+
230	TTGAATTCGACCCTGCAGGCCACAACA		
246	ACATGATCTGCAGAGAGGCCAG	9497–9518	−
247	**CCAAGAATT**CCCTAGTCACGGCTAGC	9437–9453	+
249	CTCACGGACCTTTCACAGC	9452–9470	−
260	CAGAATTCTTGTAGTCGGACAGGGTGCG		
284	AGTATCAGCACTCTCTGCAGTCATGCGG	9471–9498	−
285	GGTGGCTCCATCTTAGCCCTAG	9421–9442	+
286	CAGTCATGCGGCTCACGGACCTT	9459–9481	−
288	GGAAATGGCCTAAGAGGCCGGAG	9393–9415	−
289	TTATCACAGCGTGTCTCATGCC	9281–9302	+
290	TTCATCGGTTGGGGAGGAGGTAGAT	9354–9378	−
1045	AATCTTCACCGGTTGGGGAGGAGGTA	9357–9382	−

[a] In the HCV-H nucleotide sequence. The numbering excludes the poly(U)/polypyrimidine tract. In chimeric oligonucleotides, non-HCV-specific sequences are shown in boldface. Restriction sites used for cDNA cloning are underlined.

[b] The polarities of HCV-specific oligonucleotides are indicated as either the HCV genome RNA sense (+) or its complement (−).

nmol of ATP, 10 U of T4 polynucleotide kinase, and the buffer supplied by the manufacturer (New England Biolabs, Beverly, Mass.). Following incubation at 37°C for 1 h, the reaction mixture was extracted with phenol and then butanol, and the oligonucleotide was precipitated by addition of $MgCl_2$ and ethanol as described above.

To determine the HCV-H 3′ sequence, 3′-blocked and 5′-phosphorylated oligonucleotide 206 (Table 1) was ligated to HCV genome RNA by using T4 RNA ligase. Typical reaction mixtures (10 μl) contained 1.5×10^5 molecules of HCV RNA, 10 mM Tris-Cl (pH 7.5), 10 mM $MgCl_2$, 5 mM dithiothreitol, 1 mM ATP, 1 mM hexamine cobalt chloride, 30% (vol/vol) dimethyl sulfoxide, 10 pmol of oligonucleotide 206, 20 U of RNasin, and 8 U of T4 RNA ligase (Gibco-BRL, Grand Island, N.Y.). RNA was initially mixed with the dimethyl sulfoxide, after which the remaining components were added. Incubations were at 19°C for 4 to 20 h. Ligation reactions were used directly for cDNA synthesis. Reaction mixtures (15 μl) contained 5 μl of the ligation reaction, 12.5 pmol of primer 192 or 230 (complementary to oligonucleotide 206; Table 1), 0.5 mM each deoxynucleoside triphosphate (dNTP), 20 U of RNasin, 12.5 U of avian myeloblastosis virus reverse transcriptase, and the buffer supplied by the manufacturer (Promega, Madison, Wis.) and were incubated sequentially at 43°C (20 min), 48°C (20 min), and 55°C (20 min). After heating at 96°C for 5 min, 1 μl of the cDNA reaction was amplified by 40 cycles (25 s at 95°C; 30 s at 60°C; 15 s at 72°C) of PCR, using KlenTaqLA (6). Reaction mixtures (50 μl) contained 25 pmole of primer 192, 25 pmol of primer 228 (Table 1), 50 mM Tris-Cl (pH 9.2), 2.5 mM $MgCl_2$, 16 mM $(NH_4)_2SO_4$, 150 μg of bovine serum albumin per ml, and 0.2 μl of KlenTaqLA. Amplified products were extracted with phenol, ethanol precipitated, resuspended and digested with *Eco*RI and *Sal*I, and cloned into pGEM3Zf(−) which had been digested with the same two enzymes. These restriction sites were present in PCR primers 192 and 228, respectively. Plasmids containing *Eco*RI and *Sal*I inserts were identified and sequenced by the dideoxy-chain termination method, using M13 universal and reverse primers.

The sequences of additional patient samples were obtained by cloning two overlapping portions of the HCV 3′ NTRs. RNAs from patient sera were prepared as described above. To clone the internal region containing the poly(U)/polypyrimidine tract, a negative-sense primer within the novel 3′-terminal element (primer 246) was used for cDNA synthesis. First-strand cDNA was amplified using primers 228 [positive sense, upstream from the poly(U)/polypyrimidine tract] and 249 (negative sense, within the novel 101-base element; Table 1), treated with T4 DNA polymerase to produce blunt ends, and cloned into pGEM3Zf(−) which had been digested with *Sma*I and *Hin*cII. The extreme 3′-terminal sequences were cloned by the oligonucleotide ligation method essentially as described above for H77. Terminal sequences were amplified by using primers 247 and 192, digested with *Eco*RI, and cloned into *Eco*RI-digested pGEM3Zf(−). Selected clones were sequenced as described above.

Analysis of human genomic DNA. Human genomic DNA (3 ng; ~1,000 copies) was analyzed for the presence of the novel sequence by nested PCR. Primer pairs specific for the novel 98-base element at the 3′ end of HCV genome RNA were used. Initial amplification reactions (50 μl) were performed with primers 284 and 285 (Table 1); 30 cycles were performed (10 s at 95°C; 30 s at 60°C; 15 s at 72°C). One microliter of the products from these reactions was amplified under the same conditions for an additional 20 cycles, using nested primers 247 and 286 (Table 1). One-tenth of the resulting products was analyzed by electrophoresis on a 6% polyacrylamide gel. As a positive control for amplification of HCV-specific sequences, 100 molecules of a plasmid containing the novel 3′ element (pRS2/HCV-3′/Rz/T7term [20a]) were added to some reactions. As a positive control for amplification of single-copy sequences from human genomic DNA, EST 00205 primers (provided by R. Mazzarella, Washington University; DXS no. 1012E) (21) were used.

Quantitative-competitive reverse transcription (RT)-PCR. For production of competitor RNA, we constructed a plasmid (pRS1/T7HCV-3′end/comp) which contained HCV cDNA between nucleotides (nt) 9220 and the 3′ end of the genome [nt 9518; numbering excludes the poly(U)/polypyrimidine tract] positioned downstream from the T7 RNA polymerase promoter. This plasmid contains markers 5′ and 3′ to the poly(U)/polypyrimidine tract, including a 29-bp deletion between nt 9310 and 9340 and a 21-bp insertion (5′-CTAGCTCTC TCTCTCTCTCTC-3′) between nt 9448 and 9449. The poly(U)/polypyrimidine tract in pRS1/T7HCV-3′end/comp is 105 bases.

Competitor RNA was transcribed from *Eco*RI-linearized pRS1/T7HCV-3′end/comp DNA. Transcription reaction mixtures (100 μl) contained 1 mM rNTPs, 5 μCi of [^{3}H]UTP, 40 U of RNasin, 2 μg of template DNA, 100 U of T7 RNA polymerase, and the reaction buffer supplied by the manufacturer (Promega). After 1 h at 37°C, $CaCl_2$ was added to 10 mM, and template DNA was digested with 4 U of RQ1 DNase for 15 min. The RNA was extracted with phenol-chloroform, precipitated with 0.6 volume of isopropanol, resuspended in 50 μl, and digested a second time with 0.3 U of RQ1 DNase in 10 mM Tris-Cl (pH 8.0)–5 mM dithiothreitol–40 U of RNasin–10 mM $CaCl_2$–10 mM $MgCl_2$. After purification using the RNAeasy kit (Qiagen), eluted RNA was concentrated by butanol extraction and precipitated with ethanol. The RNA pellet was dissolved in 4 M urea and separated by electrophoresis on 1.5% low-melting-temperature agarose. The band of competitor RNA was visualized by staining with ethidium bromide, excised, and purified by phenol extraction after addition of 10 μg of tRNA. Gel-purified RNA was digested a third time with RQ1 DNase as described above, extracted with phenol-chloroform, precipitated and washed with ethanol, and dissolved in water. The resulting competitor RNA was checked for integrity by electrophoresis and quantified by liquid scintillation counting. Dilutions of competitor RNA (in distilled H_2O containing 1 μg of carrier tRNA) were tested for DNA contamination by nested PCR; no signal was detected for 1 μl of undiluted or 100-fold-diluted competitor RNA stock (estimated to contain ~10^7 to 10^8 molecules per μl), using primer pairs for amplifying sequences 5′ (primers 228 and 288 followed by 289 and 290) or 3′ (primers 285 and 286 followed by 247 and 287) of the poly(U)/polypyrimidine tract (data not shown).

Ten micrograms of tRNA and 1 μl of the desired dilution of competitor RNA were mixed with 50 μl of TRIzol (Gibco/BRL). In parallel, samples of HCV-containing plasma were mixed with 10 to 100 volumes of TRIzol, and a portion (typically 1 μl) was added to the competitor RNA-TRIzol mixture. RNAs were isolated as suggested by the manufacturer and resuspended in 3 μl of H_2O. RT reaction mixtures (5 μl) contained 10 mM dithiothreitol, 0.5 mM dNTPs, 1 pmol of primer, 4 U of RNasin, 20 U of Superscript II, and 1× RT buffer (Gibco/BRL). Following incubation at 43°C for 1 h, samples were heated to 96°C for 3 min. To each sample, 45 μl of a PCR mix containing the desired primers was added, after which 35 cycles of amplification were carried out as described above. For some samples, a nested-PCR step (20 cycles) was performed with an internal primer pair and 1 μl of the first amplification product (see Results).

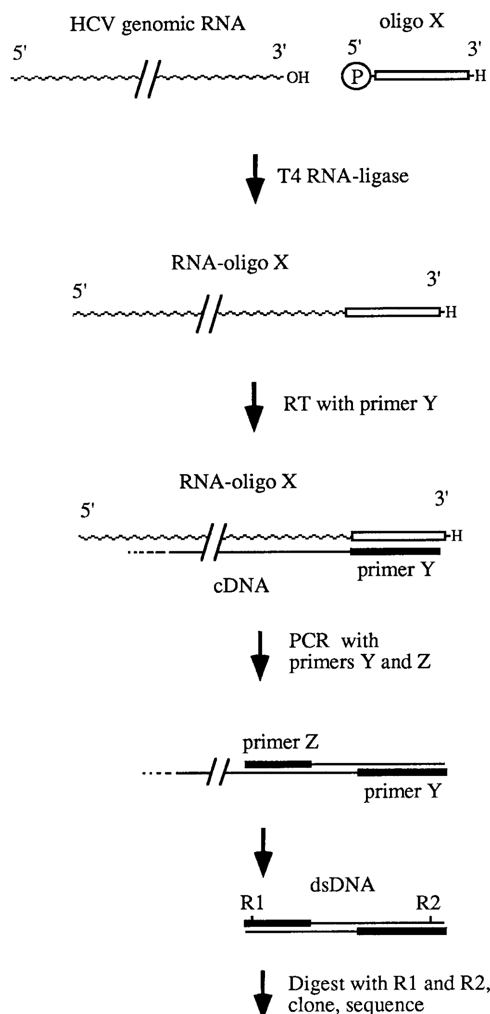


FIG. 1. Method for 3′-end determination. Shown is a schematic of the steps used to clone the extreme 3′ ends of HCV genome RNAs. RNA is indicated by wavy lines, and cDNA is indicated by straight lines. The 3′-blocked, 5′-phosphorylated oligonucleotide used for ligation with T4 RNA ligase is shown as an open box (oligonucleotide [oligo] X). Primers (thick lines) used for cDNA synthesis (oligonucleotide Y), PCR amplification (primers Y and Z), and the restriction sites (R1 and R2) used for cloning of amplified products are indicated. Specific details, including reaction conditions and primer sequences, are given in Materials and Methods and Table 1. dsDNA, double-stranded DNA.

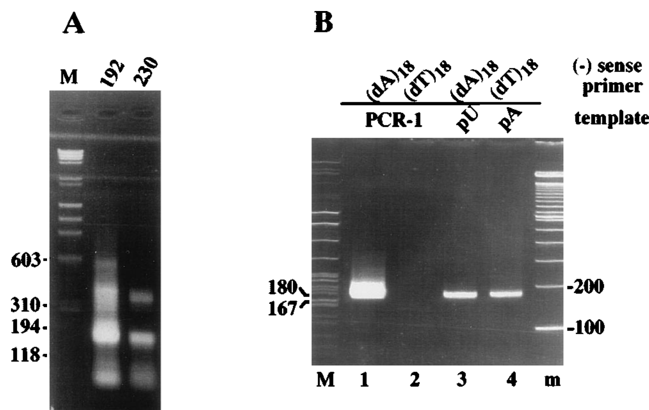


FIG. 2. HCV-specific amplification products obtained by using ligated oligonucleotides. (A) 5′-phosphorylated, 3′-blocked oligonucleotide 206 was ligated to a crude RNA preparation from H77 plasma as described in Materials and Methods and the legend to Fig. 1. First-strand cDNA synthesis was carried out with primer 192 or 230 (indicated above the respective lanes; Table 1), which are complementary to oligonucleotide 206. cDNA was amplified (with primers 192 and 228) and digested with *Eco*RI and *Sal*I, and the products were separated on a 2% low-melting-temperature agarose gel and visualized by staining with ethidium bromide. A mixture of λ*Hin*dIII and φX174 *Hae*III size markers was run in lane M. Sizes are indicated in base pairs. (B) The HCV-H 3′ NTR contains poly(U) but not poly(A). PCR-amplified cDNA, prepared essentially as described for panel A, was purified to remove excess primers. Then 1/200 of the unfractionated sample was reamplified by using sense primer 228 and either oligo$(dA)_{18}$ or oligo$(dT)_{18}$ as the antisense primer; 15 cycles (20 s at 95°C; 1 min at 40°C; 15 s at 72°C) were performed. Positive control DNAs for the amplification reaction (2 ng) were prototype full-length HCV-H plasmids (20b) containing either 3′-terminal poly$(dA)_{35}$ (pA; pTET/HCV/T75′G3′AFL) or poly $(dT)_{21}$ (pU; pTET/HCV/T75′G3′UFL). Products were separated by electrophoresis on a 6% polyacrylamide gel. Size markers are a mixture of *Bsp*RI-digested λ and *Msp*I-digested pBR322 DNAs (lane M) or a 100-bp ladder (Gibco-BRL: lane m). The expected minimal size of the 3′ NTR amplification product, including an 18-base homopolymer tract, is 174 bp.

RESULTS

A method for determining 3′-terminal sequences of rare RNAs. The general scheme used to determine the 3′-terminal sequences of HCV genome RNAs is summarized in Fig. 1. Synthetic oligodeoxyribonucleotides, blocked at their 3′ termini to prevent self-ligation, were ligated to the 3′ end of the RNA to serve as specific priming sites for cDNA synthesis and PCR. Ligation conditions were optimized by assaying the ability of T4 RNA ligase to ligate 5′-end-labeled oligonucleotides to an in vitro-transcribed acceptor RNA. Critical parameters included the batch of RNA ligase (some were heavily contaminated with RNase), the concentration of dimethyl sulfoxide (20 to 40%), and the particular oligonucleotide used for ligation (20a). Restriction sites useful for cloning 3′-terminal products were incorporated into the primers used for amplification of 3′-end sequences.

For 3′ analysis of the HCV-H genome RNA, $\sim 10^5$ molecules of RNA were obtained from high-titer H77 plasma and ligated to synthetic oligonucleotide 206, and this modified RNA was used for RT-PCR. Primers for cDNA synthesis (primer 192 or 230) and PCR amplification were complementary to the oligonucleotide used for ligation to the RNA, while a second positive-sense primer (primer 228) corresponded to a sequence near the 3′ end of the HCV ORF. As resolved by agarose gel electrophoresis, a smear of amplified products was obtained after 40 cycles of amplification (Fig. 2A). The presence of predicted internal HCV sequences was confirmed by nested PCR (using primers 229 and 1045) and restriction analysis of the resulting products (data not shown). The presence and type of homopolymer tracts were assayed by using a positive-sense primer (primer 228) and either oligo$(dA)_{18}$ or oligo$(dT)_{18}$ (Fig. 2B). A product of ~170 bp was obtained with the oligo(dA) primer (lane 1); no product was found with oligo(dT) (lane 2). Prototype full-length cDNA clones, terminating in either poly(dA) or poly(dT) (20b), served as positive and negative controls for these primer pairs and gave the expected results (Fig. 2B, lanes 3 and 4; data not shown). The size of the amplified segment obtained using oligo$(dA)_{18}$ was consistent with the 3′ NTR structures previously determined for other HCV isolates (see the introduction) and verified by direct cycle sequencing (data not shown and see below). These data strongly suggest that HCV-H does not contain poly(A) but rather, as found for most HCV isolates, contains a poly(U) tract [or at least a site for priming by oligo(dA)] at or near its 3′ terminus.

Additional sequences are present at the 3′ terminus of HCV-H genome RNA. As mentioned above, amplified products

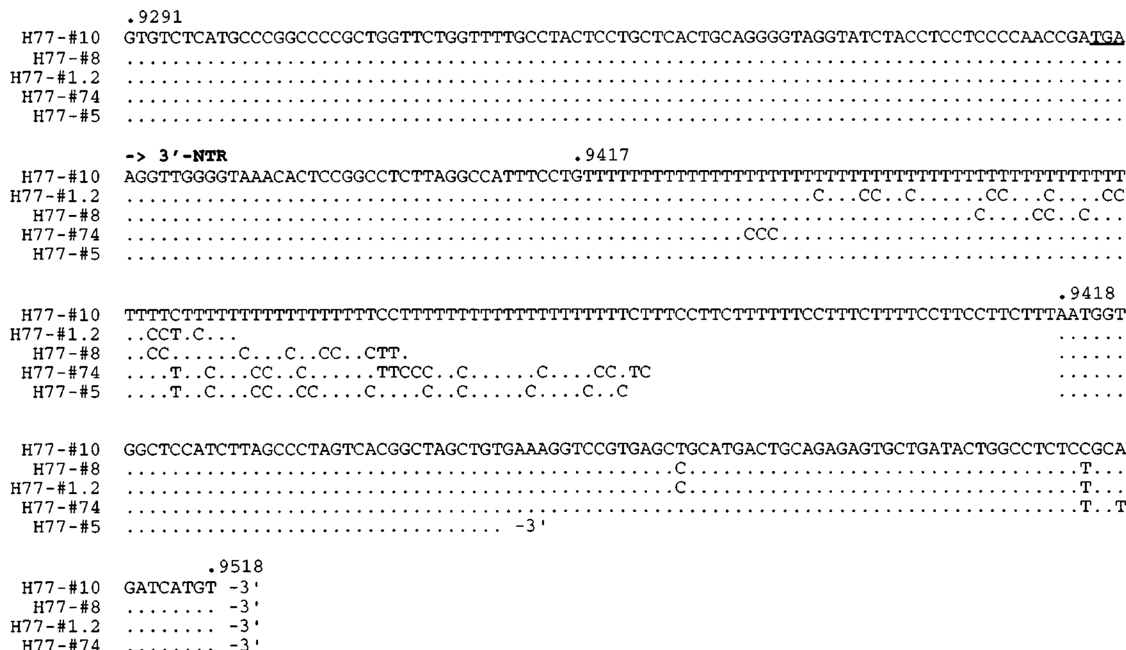


FIG. 3. Nucleotide sequences of 3′-terminal cDNA clones of HCV-H. The sequences (shown as cDNA; 5′ to 3′) of five independent HCV-H clones with extended 3′ NTR sequences are aligned. Numbering for the HCV-H genome sequence excludes the variable-length poly(U)/polypyrimidine tract (located between nt 9417 and 9418). The UGA stop codon terminating the long ORF (9375 to 9377) is underlined; identities between HCV-H clones are indicated by dots; spaces represent gaps in the aligned sequences.

obtained by the T4 RNA ligase method were heterogeneous. Most products migrated between 160 and 200 bp; additional smaller species, as well as a smear of products ranging from 250 bp to more than 400 bp, were evident (Fig. 2A). Without size fractionation, amplified products digested with *Eco*RI and *Sal*I (sites for these enzymes were present in the primers used for amplification) were cloned. Thirty clones were screened by restriction analysis, and eight clones with the longest inserts were sequenced. Following the sequence of the positive-sense primer used for amplification (primer 228), all clones contained 83 nt corresponding to the C-terminal portion of the NS5B coding region followed by the UGA stop codon at the end of the HCV ORF. Following the stop codon were 40 bases of 3′ NTR followed by a poly(U) tract. This sequence contained three substitutions relative to the previously published HCV-H sequence (18) (at positions 9380, 9386, and 9390) and extended that sequence by 21 bases. The same sequence was also found by direct cycle sequencing of uncloned material (data not shown).

Five independent clones, derived from two different PCR amplification experiments, were found to have unusual structures (Fig. 3). Following variable lengths of poly(U) and polypyrimidine stretches consisting of mainly U with occasional interspersed C residues, these clones contained additional nonhomopolymeric sequences prior to the sequence of ligated oligonucleotide 206. Beginning with an A residue, four clones contained nearly identical sequences of 101 bases (two clones differed by two substitutions each; Fig. 3). One clone (H77-5) terminated after only 39 bases of this new sequence (Fig. 3).

Novel 3′ NTR sequences are not present in human genomic DNA. An exhaustive search of the sequence databases (Blast and FastA; Genetics Computer Group package) revealed no entry with significant homology to the novel 101-base sequence or its complement. To determine if this sequence might be endogenous to the human genome, PCR primers for amplification of the new sequence were synthesized. One thousand copies of human genomic DNA were analyzed for the presence of the novel sequence by nested PCR (Fig. 4). Under these conditions, 100 molecules of the 3′-terminal element could be readily detected either in the absence (lane 2) or in the pres-

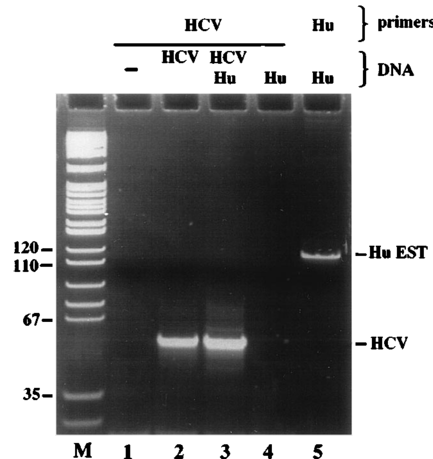


FIG. 4. Analysis of human genomic DNA. Human genomic DNA was analyzed for the presence of the novel 3′-terminal sequence by nested PCR and electrophoresis on a 6% polyacrylamide gel as described in Materials and Methods. Lanes 1 to 4 contain samples subjected to nested PCR using HCV-specific primer pairs. Lane 1, no DNA; lane 2, 100 molecules of pRS2/HCV-3′/Rz/T7term (which contains the complete HCV 3′ NTR); lane 3, 100 molecules of pRS2/HCV-3′/Rz/T7term and 1,000 genome equivalents of human (Hu) genomic DNA; lane 4, 1,000 genome equivalents of human DNA. Lane 5 shows the amplified (30 cycles) product from 1,000 genome equivalents of human DNA obtained by using the EST 00205 primer pair. HCV-specific and human (Hu EST) PCR products are indicated at the right. The sizes (in base pairs) of products from an *Msp*I digest of pBR322 (lane M) are indicated at the left.

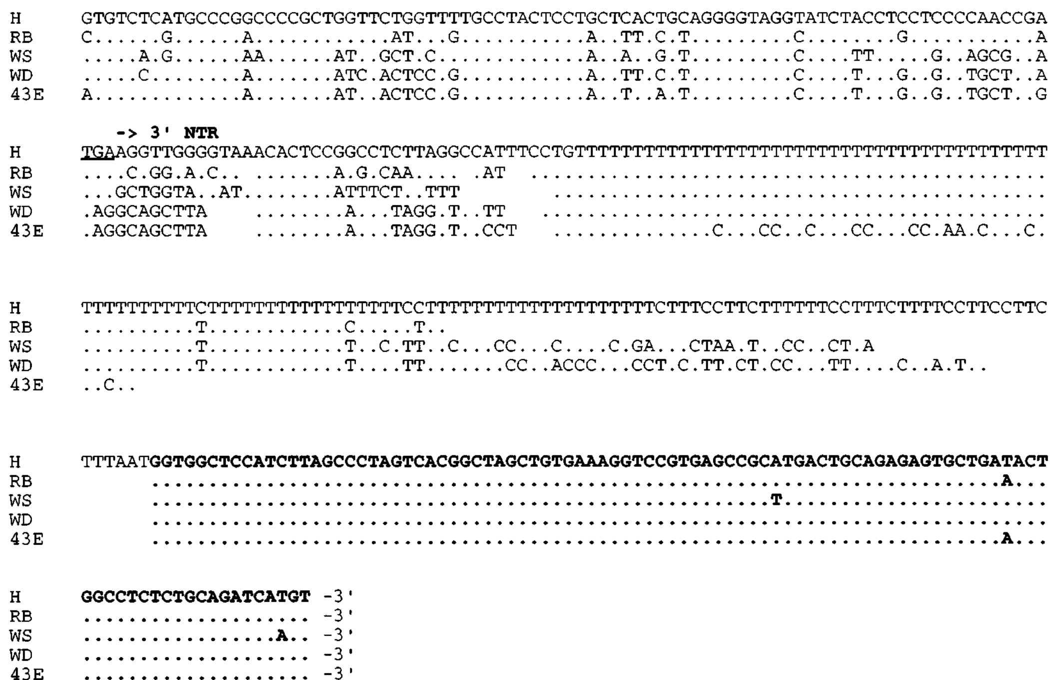
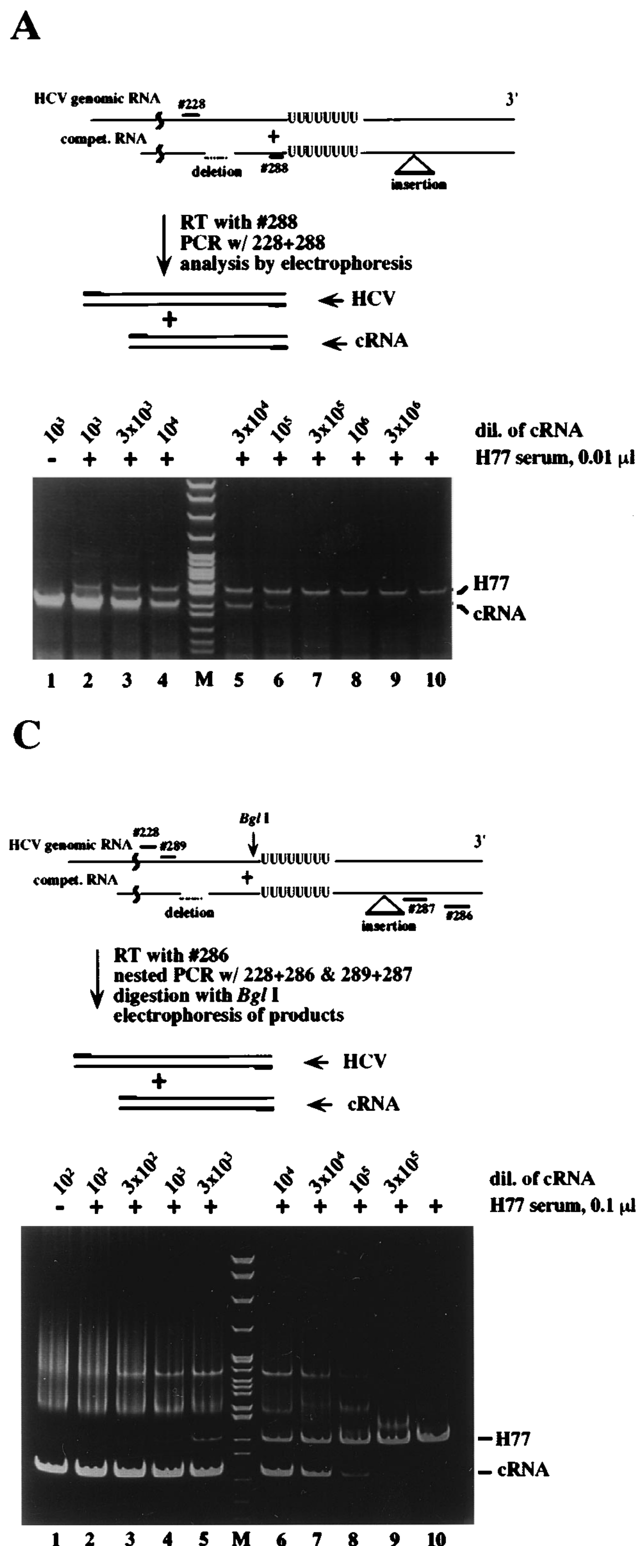


FIG. 5. Sequence conservation of the 3′-terminal element. The deduced 3′ NTR sequences (shown as cDNA; 5′ to 3′) of five independent HCV isolates (described in Materials and Methods and Results) are aligned. The top sequence represents the HCV-H consensus sequence as determined by analysis of multiple independent clones (see Fig. 3) except for the poly(U)/polypyrimidine tract, which is that of clone H77-10; the sequences for the other isolates were derived from single cDNA clones. The stop codon terminating the HCV-H long ORF is underlined; nucleotides identical to the HCV-H sequence are indicated by dots; spaces represent gaps; the highly conserved 98-base 3′-terminal element is shown in boldface type.

ence (lane 3) of human genomic DNA. No product was detected when we attempted to amplify the sequence from human genomic DNA alone (lane 4) or in the absence of added DNA (lane 1). Efficient amplification of single-copy sequences from human genomic DNA was verified by using the EST 00205 primer pair (lane 5), which produced the expected 118-bp product (an X-linked expressed sequence tag [21]) after only a single round of PCR amplification. Although an intron within the 45-bp amplified segment might preclude its detection by this method, these analyses suggest that the novel HCV-associated sequence is not present in the human genome.

Analysis of additional patient samples and HCV genotypes. Using a negative-sense primer based on the novel 3′-terminal sequence and an upstream positive-sense primer (primer 228), we examined a number of different clinical samples from patients with chronic hepatitis C which were HCV RNA positive. The expected amplification product was obtained for all of these samples (data not shown). For some of these patients, 3′ NTR amplification products, including the extreme 3′ end obtained by the oligonucleotide ligation method, were cloned and sequenced (see Materials and Methods). Figure 5 shows an alignment of four of these sequences (RB, WS, WD, and 43E) with the H77 consensus sequence. Comparison of the 83 bases of coding sequence and the 3′ NTR sequence between the stop codon and poly(U) tract suggests that these isolates correspond to HCV genotypes and subtypes different from HCV-H (data not shown). Isolate RB appears to correspond to a 1b subtype, WS appears to correspond to genotype 3, and the genotypes of WD and 43E are unclear but do not correspond to genotypes 1 to 3. As expected, the coding sequence corresponding to the C terminus of NS5B shows high conservation among these isolates, in sharp contrast to the sequence following the stop codon and preceding the beginning of the poly(U) tract. This portion of the 3′ NTR is highly variable, with only an ACACUCC motif and a UG dinucleotide immediately preceding the poly(U) start (which have been noted previously [34]) being conserved. Considerable heterogeneity, both in length and in composition, is apparent in the poly(U)/polypyrimidine tracts of these different isolates. Most sequences begin with a homopolymer poly(U) tract, with increasing numbers of C residues toward the 3′ portion of the polypyrimidine tract. For some clones, interspersed purine residues (usually A) were also found near the 3′ end of the polypyrimidine tract. It is unclear how much of the length and sequence heterogeneity actually exists in the HCV genome RNAs, since PCR amplification can lead to the expansion or contraction of internal homopolymer tracts (unpublished observation). Following the polypyrimidine tract, all of the isolates contained a sequence of 98 bases fused to the oligonucleotide used for 3′-end determination. This 98-base sequence shows remarkable conservation among these isolates (98 to 100%; see below). These results suggest that the HCV 3′ NTR organization determined for HCV-H is common to other HCV isolates and genotypes.

Evidence that the novel 3′ RNA element represents the actual 3′ end of HCV genome RNA. An additional experiment provided more compelling evidence that the 3′ novel sequence actually represents the 3′ terminus of HCV genome RNA. It could be argued that clones with the novel structure could be obtained by internal priming within the 3′ NTR if, by chance, the 3′ portion of the synthetic primer used for cDNA synthesis (or PCR) was complementary to a sequence within the HCV 3′ NTR. To address this concern, we repeated the cloning of 3′-terminal sequences by using serum from patient WD and a distinct oligonucleotide in the RNA ligation step whose sequence was not homologous to the oligonucleotide used in the



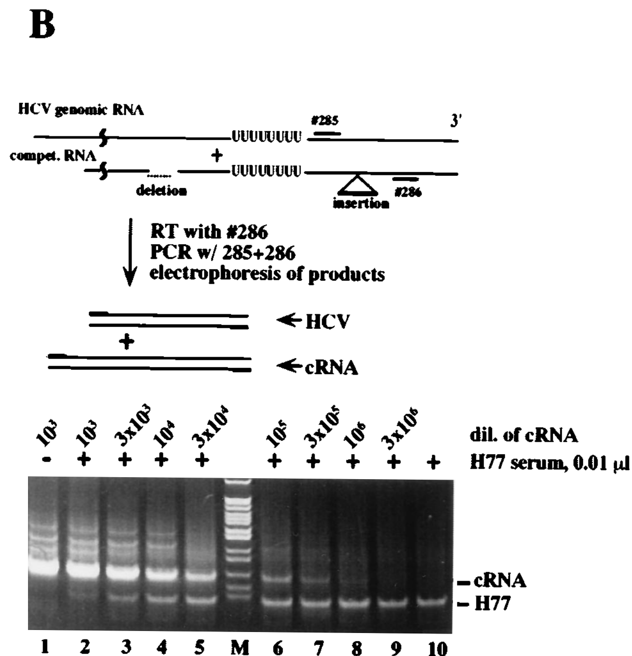


FIG. 6. The majority of HCV genome RNAs contain the 3′-terminal element. (A) Quantitation of 3′ NTR sequences upstream of the poly(U)/polypyrimidine tract in H77 plasma. A schematic of the 3′ NTR quantitative-competitive RT-PCR assay is shown (see Materials and Methods for details). Three-fold dilutions of competitor RNA (cRNA) stock were added as indicated (expressed at the reciprocal of the dilution; 10^3 indicates a 1,000-fold dilution) to RNA extracted from the equivalent of 0.01 μl of H77 plasma and subjected to RT (with primers 288, 286, and 284) followed by 35 cycles of PCR using primers 228 and 288 (Table 1). (B) Quantitation of 3′ NTR sequences downstream of the poly(U)/polypyrimidine tract. The experiment was identical to that described for panel A except that primers 286 and 284 were used for cDNA synthesis and primers 285 and 286 (Table 1) were used for 35 cycles of PCR amplification. (C) Quantitation of HCV RNA molecules which contain sequences both upstream and downstream of the poly(U)/polypyrimidine tract. cDNA synthesis and nested PCR were conducted with RNA extracted from the equivalent of 0.1 μl of H77 plasma. cDNA was synthesized by using primer 286 followed by PCR with primers 228 and 286. A portion of this amplified product was used for nested PCR with primers 289 and 287. To resolve the H77- and cRNA-derived amplification products, samples were digested with *Bgl*I prior to electrophoresis. Size markers in panels A to C are an *Msp*I digest of pBR322 (lanes M; see Fig. 4 for sizes). The first sample in each panel contained competitor RNA without H77 RNA; the last sample contained no competitor RNA.

initial experiments (oligonucleotide 143; Table 1). The complement of this oligonucleotide was used for cDNA synthesis and PCR amplification, and the products were cloned and sequenced. The same novel 3′-terminal sequence was found joined, at exactly the same breakpoint, to the sequence of the alternative synthetic oligonucleotide (data not shown). These results suggest that the novel sequence represents the authentic 3′ terminus of HCV genome RNA.

Prevalence of the novel sequence at the 3′ ends of genome RNAs obtained from the high-specific-infectivity H77 inoculum. If the newly identified 3′ NTR sequence is essential for authentic replication, then it should be present in infectious HCV genome RNAs. In the HCV-H H77 plasma, a substantial fraction of genome RNAs must be infectious since the RNA titer is similar to that in chimpanzees (26). To quantify 3′-terminal sequences present in RNA isolated from HCV-H H77 plasma, we developed quantitative-competitive RT-PCR assays for HCV-specific sequences 5′ and 3′ to the poly(U)/polypyrimidine tract (see Materials and Methods). Representative experiments are shown in Fig. 6. By comparison of Fig. 6A and B, it is apparent that the H77 inoculum contains similar levels of HCV-specific sequences 5′ and 3′ of the poly(U)/polypyrimidine tract. For the upstream 3′ NTR sequences, similar levels of H77 and competitor RNA coamplification

products were observed at a ~3 × 10^4-fold dilution of the competitor RNA stock. For the downstream element, equivalence was reached between competitor RNA dilutions of 3 × 10^4- and 10^5-fold. In Fig. C, actual linkage of these sequences was shown by nested PCR amplification across the poly(U)/polypyrimidine tract. Although this RT-PCR is significantly less efficient and required the use of more H77 RNA (data not shown), the level of H77 molecules containing both upstream and downstream sequences was found to be only slightly lower [~3- to 10-fold, which may be an underestimate because of susceptibility of the poly(U)/polypyrimidine tract to RNases during isolation or cDNA synthesis]. Although not definitive, these results are consistent with the hypothesis that infectious HCV genome RNAs contain the novel 3′ NTR element.

Secondary structure predictions. RNA folding programs mfold (19) and FoldRNA (Genetics Computer Group package) were used to predict thermodynamically favored secondary structures for the HCV 3′ NTR. RNA folding predictions were examined for a 3′-terminal nested set of RNA molecules with increasing lengths of upstream sequence. In all cases, the 3′-terminal 46 bases were predicted to form a stable stem-loop (SL) structure (3′-SL I; $\Delta G = -25.7$ kcal [1 cal = 4.184 J]/mol; Fig. 7A). As shown in Fig. 7A, folding of the entire 3′ NTR sequence with limited upstream sequence revealed additional, less stable SL structures (3′-SL II and III). Interestingly, the two substitutions in the 98-base conserved sequence found for the genotype 3 WS isolate (nt 9474 and 9516; HCV-H numbering as in Fig. 3) result in the preservation of the predicted stem structure in 3′-SL I by substituting a U-A for an A-U base pair (Fig. 7A). Another variable position in the 98-base element, which is A or U (nt 9495; HCV-H numbering), is predicted to be in the loop portion of 3′-SL I. Given that interactions between 5′- and 3′-terminal sequences of other viral RNAs can be important in replication and packaging, we also examined predicted structures for a series of RNAs with different lengths of 5′- and 3′-terminal sequences. Figure 7B shows a portion of the predicted structure for an RNA sequence composed of the 358 bases from the HCV-H 5′ terminus fused to 300 bases from the 3′ terminus. Interestingly, the 3′-terminal 3′ SL-I structure was preserved whereas 3′-SL II and III were disrupted by partial base pairing with sequences near the 5′ terminus of HCV genome RNA.

DISCUSSION

Three distinct features of the HCV 3′ NTR have emerged from previous work and the studies reported here, which define the 3′ terminus of HCV genome RNA. Between the stop codon terminating the long ORF and the internal poly(U) tract is a short sequence (40 bases in HCV-H) which is poorly conserved among HCV genotypes, aside from a few short motifs (34) (see Results). This sequence is followed by an internal poly(U)/polypyrimidine tract of variable length and composition. The 5′ portion of this tract is typically U rich, with increasing numbers of C residues found in the 3′ region. Occasional purine residues were also found, including two A residues at the 3′ boundary of HCV-H. At present, it is unclear how much of the heterogeneity observed in these cDNA clones is actually present in HCV genome RNAs. During cDNA synthesis and PCR amplification, polymerase slippage, premature termination in the homopolymer tract, and misincorporation of bases could lead to expansion or contraction of homopolymer sequences and the appearance of sporadic purine residues. Sequence analysis of cDNA clones encompassing this region which have been produced without the use of PCR should help to clarify the structure of this region of the 3′ NTR. The most striking new feature of the 3′ NTR is the identification of a 98-base nonhomopolymeric sequence which appears to represent the authentic 3′ terminus of HCV genome RNA (see below). We have shown that this sequence is not present in human genomic DNA, that it is found at the 3′ termini of several independent HCV isolates, and that it is highly conserved, with between 98 and 100% sequence identity for the isolates examined thus far. The 3′-terminal 46 bases of this sequence are predicted to form a stable SL structure. By using quantitative-competitive PCR assays, this element was shown to be present at the 3′ termini of a significant fraction of genome RNAs in the highly infectious H77 inoculum. As discussed below, these results indicate that this novel 3′-terminal RNA element is likely to be essential for authentic HCV replication.

Using independent patient sera and different oligonucleotides for ligation and RT-PCR, we found the same novel 3′-terminal sequence joined, at exactly the same junction, to the sequences of the synthetic oligonucleotides. Barring the possibility that this 3′ end was generated by the action of contaminating RNases or premature termination during HCV replication, these results suggest that the 3′ end has been reached. Given this structure for the HCV 3′ NTR, it is now clear why previous attempts have failed to define the correct 3′-terminal sequences. 3′-terminal oligo(U) was reported by several groups who used tailing with poly(A) polymerase followed by oligo(dT)-primed cDNA synthesis (16, 20, 22). Addition of 3′-terminal poly(A) would produce a 3′ terminus which could self-prime in the internal poly(U) tract and prevent copying of the 3′-terminal element during first-strand cDNA synthesis. Alternatively, the 3′-terminal SL structure may be an inefficient substrate for poly(A) polymerase, leading to selective tailing of partially degraded RNAs. Similar problems may have occurred in attempts to determine the 5′-terminal sequence of negative-strand RNA after dA tailing of first-strand cDNA (10). During preparation of this report, Tanaka et al. published a report of a study in which rapid amplification of 5′ cDNA ends, with dC tailing of first-strand cDNA and oligo(dG)-primed second-strand synthesis, was used to determine the 5′ sequence of minus-strand RNA isolated from HCV-infected liver (genotype 1b) (33). The sequence found by these investigators is in complete agreement with the one reported here for the HCV 1b subtype. Together with our results, these data show that the 5′ terminus of intracellular minus-strand RNA is complementary to the 3′-terminal sequence of genome RNA, as expected. In contrast to the report for HCV-1 (15), we found no evidence for 3′-terminal poly(A) for HCV-H, which is also genotype 1a. However, it should be noted that these investigators used tagged oligo(dT) primers for cDNA synthesis followed by PCR amplification, cloning, and sequence analysis. This method assumes a 3′ poly(A) or polypurine tract and would select for such RNAs even if they were present in low abundance.

Although some positive-strand RNA viruses contain 3′-terminal poly(A) (e.g., poliovirus and Sindbis virus), many others terminate with conserved RNA sequences which can often be folded into stable secondary structures (e.g., bromoviruses). Given this new information on HCV and recent results for pestiviruses and flaviviruses, common and distinct features of their 3′ NTRs are apparent. Within each genus, 3′ NTR sequences immediately following the long ORF tend to be poorly conserved and are often variable in length. This region is typically followed by highly conserved RNA sequences. Although initial reports suggested otherwise, none of the genera appears to terminate genome RNAs with long homopolymer tracts. Pestivirus 3′ NTRs are about 225 bases in length and terminate

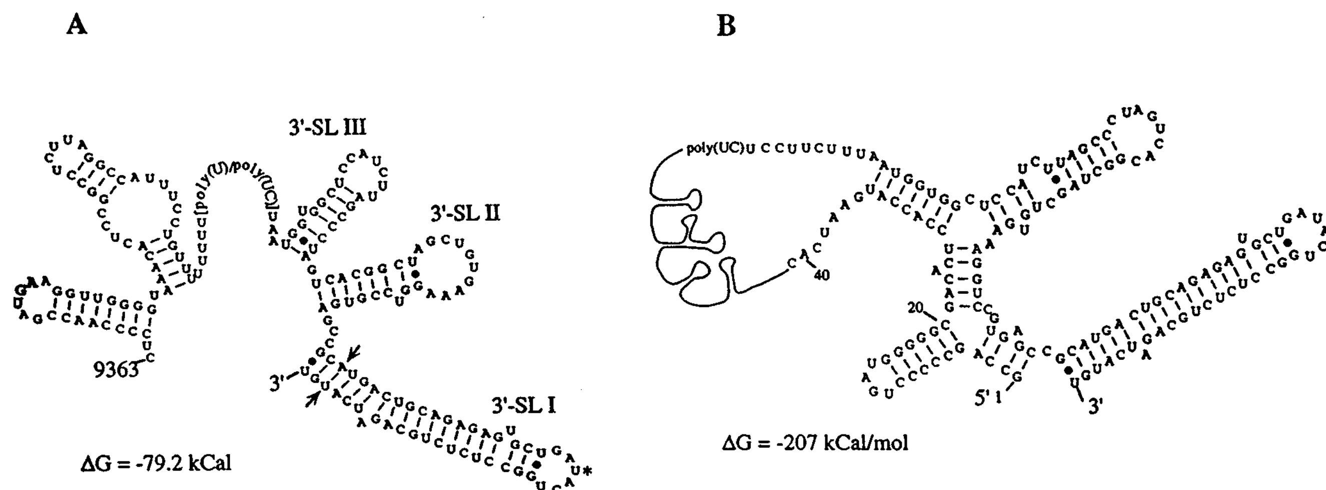


FIG. 7. Secondary structure predictions. (A) Computer-predicted secondary structure of the HCV-H 3′ NTR RNA and limited upstream sequence ($\Delta G = -79.2$ kcal/mol). SL structures in the 3′-terminal 98-base RNA element are labeled (3′-SL I to III); the ORF UGA termination codon is shown in boldface type. A variable base pair in the stem of 3′-SL I which is T-A in the WS isolate is indicated by an arrows. A variable position in the loop of 3′-SL I is indicated by an asterisk. (B) Predicted secondary structure for an HCV-H RNA sequence composed of the 358 bases from the 5′ terminus fused to 300 bases from the 3′ terminus. The ΔG value for the entire structure is −207 kcal/mol; only the 5′-terminal 41 nt and the 3′-terminal 110 nt are shown.

with a stretch of three to five C residues (12). Multiple secondary structures, including structures at or near the 3′ terminus (12), can be predicted in this region. For members of the *Flavivirus* genus, 3′ NTRs are more variable in length, but all contain a predicted 3′-terminal secondary structure of ~100 bases (9). Upstream from these 3′-terminal structures are short sequences which are highly conserved among mosquito or tick-borne flaviviruses but not shared by these two groups (35). Interestingly, certain tick-borne flaviviruses do contain a homopolymeric tract within the 3′ NTR, but it consists of poly(A) rather than poly(U) (35). Although not yet classified, there is one example of a virus with a 3′ NTR structure more reminiscent of that determined for HCV. This virus, called GBV-B, is one of the two newly cloned and sequenced GB hepatitis agents (28). These isolates (GBV-A and GBV-B) are most closely related to HCV; both appear to have positive-strand genome RNAs of ~9 to 10 kb and a single long ORF encoding proteins with significant homology to those of HCV. In the case of GBV-B, the 3′ NTR consists of 27 bases, a poly(U) tract, and an additional sequence of 49 bases. Other than the poly(U) tract, however, the GBV-B 3′-terminal sequence shows no significant homology with the HCV-H 3′ NTR and is not predicted to form a stable secondary structure (20a).

The function(s) of the HCV 3′ NTR, including the highly conserved 3′-terminal element, remains to be determined. For other RNA viruses, conserved terminal sequences or structures play critical roles in initiation of minus- and plus-strand RNA synthesis and in packaging of viral RNAs. Such processes are mediated via interactions with *trans*-acting proteins encoded by the virus or host and, in some cases, other *cis* RNA elements elsewhere in the genome. For instance, conserved tRNA-like structures at the 3′ termini of bromovirus RNAs are required for initiation of minus-strand synthesis. For negative-strand viruses, such as influenza virus and vesicular stomatitis virus, conserved sequences at the 5′ and 3′ termini can base pair and constitute the *cis* regulatory elements for transcription, replication, and packaging (23, 24). In the case of ilarviruses, positive-strand plant RNA viruses, a conserved RNA sequence near the 3′ terminus binds coat protein and is believed to be the signal for encapsidation; this interaction is also required for an early step in RNA replication (3). Terminal *cis* RNA elements important for translation and RNA replication have also been identified for positive-strand animal viruses, including alphaviruses (31), flaviviruses (7a), nodaviruses (4, 5), and picornaviruses (37). These examples focus attention on the newly identified 98-base RNA sequence at the 3′ terminus of HCV genome RNA as a possible *cis* element involved in RNA replication, modulation of translation or RNA stability, or encapsidation of genome RNA. In addition, the presence of an internal polypyrimidine tract in the HCV 3′ NTR is intriguing. This region may bind the 57- to 62-kDa polypyrimidine tract-binding proteins (PTB) (14, 29). PTB binding could be important for determining the higher-order structure of the 3′ NTR and might interact with other factors involved in RNA replication. In this regard, it is of interest that three distinct PTB binding sites have been defined in the 5′ NTR, and PTB-associated factors are required for efficient translation initiation by the HCV internal ribosome entry site (1). PTB is believed to be a homodimer which, in theory, might initiate or stabilize interactions between the HCV 5′ NTR and the 3′ NTR. Such interactions could be important for modulating translation versus replication of HCV genome RNA.

Besides the potential importance of the 3′ NTR for HCV replication and recovery of infectious HCV RNA from cDNA, the apparent conservation of the 3′ element may have important applications for HCV diagnosis and therapy. Determination of the levels of HCV RNA in patient plasma and tissue samples is important not only for diagnosis of HCV infection in the absence of an antibody response but also for monitoring the efficacy of therapeutic regimens. Current methods, such as quantitative RT-PCR or branched DNA, rely on conserved RNA targets in the HCV genome which can be either genus, type, or subtype specific. If further studies confirm that the 3′ element is indeed highly conserved and common to all members of the HCV genus, detection of this sequence may be a useful alternative for diagnosis of HCV infection. In terms of therapy, highly conserved elements in RNA virus genomes have, as mentioned above, been shown to be essential for efficient virus replication. Such elements, via interaction with viral and/or host factors, function in translation of the incoming viral RNA, as promoters for negative- and positive-strand

RNA synthesis, and as signals for selective packaging of viral RNAs. Hence, a conserved 3′ element in the HCV genome, which is likely to be important for one or more of these processes, presents an attractive therapeutic target. Identification of compounds which block interaction of this element with its cognate host or viral factors or gene therapy approaches using this element as an RNA decoy in transplanted hepatocytes may prove useful in eradicating or controlling chronic HCV infections.

ACKNOWLEDGMENTS

We thank Richard Mazzarella and David Schlesinger for help with the human genomic DNA analysis and David Gretch, Jay Hoofnagel, and Peter Simmonds for patient serum. We are also grateful to our colleagues for helpful discussions during the course of this work and to Sean Amberg, Elena and Ilya Frolov, Brett Lindenbach, and Karen Reed for critical reading of the manuscript.

This work was supported by Public Health Service grant CA57973.

REFERENCES

1. **Ali, N., and A. Siddiqui.** 1995. Interaction of polypyrimidine tract-binding protein with the 5′ noncoding region of the hepatitis C virus RNA genome and its functional requirement in internal initiation of translation. J. Virol. **69:**6367–6375.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1993. Current protocols in molecular biology. Greene Publishing Associates, New York.
3. **Baer, M. L., F. Houser, L. S. Loesch-Fries, and L. Gehrke.** 1994. Specific RNA binding by amino-terminal peptides of alfalfa mosaic virus coat protein. EMBO J. **13:**727–735.
4. **Ball, L. A.** 1995. Requirements for the self-directed replication of flock house virus RNA 1. J. Virol. **69:**720–727.
5. **Ball, L. A., and Y. Li.** 1992. *cis*-acting requirements for the replication of flock house virus RNA 2. J. Virol. **67:**3544–3551.
6. **Barnes, W. M.** 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. Proc. Natl. Acad. Sci. USA **91:**2216–2220.
7. **Boyer, J.-C., and A.-L. Haenni.** 1994. Infectious transcripts and cDNA clones of RNA viruses. J. Gen. Virol. **198:**415–426.

7a. **Breedenbeek, P. J., and C. M. Rice.** Unpublished data.

8. **Bukh, J., R. H. Miller, and R. H. Purcell.** 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. Semin. Liver Dis. **15:**41–63.
9. **Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice.** 1990. Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. **44:**649–688.
10. **Chen, P.-J., M.-H. Lin, K.-F. Tai, P.-C. Liu, C.-J. Lin, and D.-S. Chen.** 1992. The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5′ termini of viral genomic and antigenomic RNA. Virology **188:** 102–113.
11. **Choo, Q.-L., G. Kuo, A. J. Weiner, L. R. Overby, D. W. Bradley, and M. Houghton.** 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science **244:**359–362.
12. **Deng, R., and K. V. Brock.** 1993. 5′ and 3′ untranslated regions of pestivirus genome: primary and secondary structure analyses. Nucleic Acids Res. **21:** 1949–1957.
13. **Francki, R. I. B., C. M. Fauquet, D. L. Knudson, and F. Brown.** 1991. Classification and nomenclature of viruses: fifth report of the international committee on taxonomy of viruses. Arch. Virol. (Suppl.) **2:**223.
14. **García-Blanco, M. A., S. F. Jamison, and P. A. Sharp.** 1989. Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. Genes Dev. **3:**1874–1886.
15. **Han, J. H., V. Shyamala, K. H. Richman, M. J. Brauer, B. Irvine, M. S. Urdea, P. Tekamp-Olson, G. Kuo, Q.-L. Choo, and M. Houghton.** 1991. Characterization of the terminal regions of hepatitis C viral RNA: identification of conserved sequences in the 5′ untranslated region and poly(A) tails at the 3′ end. Proc. Natl. Acad. Sci. USA **88:**1711–1715.
16. **Hayashi, N., H. Higashi, K. Kaminaka, H. Sugimoto, M. Esumi, K. Komatsu, K. Hayashi, M. Sugitani, K. Suzuki, O. Tadao, C. Nozaki, K. Mizuno, and T. Shikata.** 1993. Molecular cloning and heterogeneity of the human hepatitis C virus (HCV) genome. J. Hepatol. **17**(Suppl. 3)**:**S94–S107.
17. **Houghton, M.** 1996. Hepatitis C viruses. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Fields virology, p. 1035–1058. Raven Press, New York.
18. **Inchauspe, G., S. Zebedee, D.-H. Lee, M. Sugitani, M. Nasoff, and A. M. Prince.** 1991. Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates. Proc. Natl. Acad. Sci. USA **88:**10292–10296.
19. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
20. **Kato, N., M. Hijikata, Y. Ootsuyama, M. Nakagawa, S. Ohkoshi, T. Sugimura, and K. Shimotohno.** 1990. Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. Proc. Natl. Acad. Sci. USA **87:**9524–9528.

20a. **Kolykhalov, A. A.** Unpublished data.

20b. **Lemm, J. A., M. Heise, and C. M. Rice.** Unpublished data.

21. **Mazzarella, R., and A. K. Srivastava.** 1994. Physical linkage of expressed sequence tags (ESTs) to polymorphic markers on the X chromosome. Hum. Mol. Genet. **3:**1095–1101.
22. **Okamoto, H., S. Okada, Y. Sugiyama, K. Kurai, H. Iizuka, A. Machida, Y. Miyakawa, and M. Mayumi.** 1991. Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions. J. Gen. Virol. **72:** 2697–2704.
23. **O'Neill, R. E., and P. Palese.** 1994. Cis-acting signals and trans-acting factors involved in influenza virus RNA synthesis. Infect. Agents Dis. **3:**77–84.
24. **Pattnaik, A. K., L. A. Ball, A. W. LeGrone, and G. W. Wertz.** 1995. The termini of VSV DI particle RNAs are sufficient to signal RNA encapsidation, replication, and budding to generate infectious particles. Virology **206:** 760–764.
25. **Rice, C. M.** 1996. *Flaviviridae*: the viruses and their replication. *In* B. N. Fields, D. M. Knipe and P. M. Howley (ed.), Fields virology, p. 931–960. Raven Press, New York.
26. **Shimizu, Y. K., R. H. Purcell, and H. Yoshikura.** 1993. Correlation between the infectivity of hepatitis C virus in vivo and its infectivity in vitro. Proc. Natl. Acad. Sci. USA **90:**6037–6041.
27. **Shimotohno, K., Y. Tanji, Y. Hirowatari, Y. Komoda, N. Kato, and M. Hijikata.** 1995. Processing of the hepatitis C virus precursor protein. J. Hepatol. **22:**87–92.
28. **Simons, J. N., T. J. Pilot-Matias, T. P. Leary, G. J. Dawson, S. M. Desai, G. G. Schlauder, A. S. Muerhoff, J. C. Erker, S. L. Buijk, M. L. Chalmers, C. L. Vansant, and I. K. Mushahwar.** 1995. Identification of two flavivirus-like genomes in the GB hepatitis agent. Proc. Natl. Acad. Sci. USA **92:**3401–3405.
29. **Singh, R., J. Valcárcel, and M. R. Green.** 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science **268:**1173–1176.
30. **Stoffa, R. J., H. M. White, A. A. Kolykhalov, C. M. Rice, E. Brunt, and M. Peters.** Interferon therapy for chronic hepatitis C: a long term follow up. Submitted for publication.
31. **Strauss, J. H., and E. G. Strauss.** 1994. The alphaviruses: gene expression, replication, evolution. Microbiol. Rev. **58:**491–562.
32. **Takamizawa, A., C. Mori, I. Fuke, S. Manabe, S. Murakami, J. Fujita, E. Onishi, T. Andoh, I. Yoshida, and H. Okayama.** 1991. Structure and organization of the hepatitis C virus genome isolated from human carriers. J. Virol. **65:**1105–1113.
33. **Tanaka, T., N. Kato, M.-J. Cho, and K. Shimotohno.** 1995. A novel sequence found at the 3′ terminus of hepatitis C virus genome. Biochem. Biophys. Res. Commun. **215:**744–749.
34. **Tokita, H., S. M. Shrestha, H. Okamoto, M. Sakamoto, M. Horikita, H. Iizuka, S. Shrestha, Y. Miyakawa, and M. Mayumi.** 1994. Hepatitis C virus variants from Nepal with novel genotypes and their classification into the third major group. J. Gen. Virol. **75:**931–936.
35. **Wallner, G., C. W. Mandl, C. Kunz, and F. X. Heinz.** 1995. The flavivirus 3′-noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne encephalitis virus. Virology **213:**169–178.
36. **Wang, C., and A. Siddiqui.** 1995. Structure and function of the hepatitis C virus internal ribosome entry site. Curr. Top. Microbiol. Immunol. **203:**99–115.
37. **Wimmer, E., C. U. Hellen, and X. Cao.** 1993. Genetics of poliovirus. Annu. Rev. Genet. **27:**353–436.