JOURNAL OF VIROLOGY, Feb. 2004, p. 1352–1366 Vol. 78, No. 3
0022-538X/04/$08.00+0 DOI: 10.1128/JVI.78.3.1352–1366.2004

A *cis*-Acting Replication Element in the Sequence Encoding the NS5B RNA-Dependent RNA Polymerase Is Required for Hepatitis C Virus RNA Replication

Shihyun You,[1] Decherd D. Stump,[2] Andrea D. Branch,[2] and Charles M. Rice[1]*

Center for the Study of Hepatitis C, Laboratory of Virology and Infectious Disease, The Rockefeller University, New York, New York 10021,[1] and Division of Liver Disease, Department of Medicine, Mount Sinai School of Medicine, New York, New York 10029[2]

Received 23 July 2003/Accepted 10 October 2003

RNA structures play key roles in the replication of RNA viruses. Sequence alignment software, thermodynamic RNA folding programs, and classical comparative phylogenetic analysis were used to build models of six RNA elements in the coding region of the hepatitis C virus (HCV) RNA-dependent RNA polymerase, NS5B. The importance of five of these elements was evaluated by site-directed mutagenesis of a subgenomic HCV replicon. Mutations disrupting one of the predicted stem-loop structures, designated 5BSL3.2, blocked RNA replication, implicating it as an essential *cis*-acting replication element (CRE). 5BSL3.2 is about 50 bases in length and is part of a larger predicted cruciform structure (5BSL3). As confirmed by RNA structure probing, 5BSL3.2 consists of an 8-bp lower helix, a 6-bp upper helix, a 12-base terminal loop, and an 8-base internal loop. Mutational analysis and structure probing were used to explore the importance of these features. Primary sequences in the loops were shown to be important for HCV RNA replication, and the upper helix appears to serve as an essential scaffold that helps maintain the overall RNA structure. Unlike certain picornavirus CREs, whose function is position independent, 5BSL3.2 function appears to be context dependent. Understanding the role of 5BSL3.2 and determining how this new CRE functions in the context of previously identified elements at the 5′ and 3′ ends of the RNA genome should provide new insights into HCV RNA replication.

The first molecular clones of hepatitis C virus (HCV) were reported in 1989 (11). Comparative sequence analysis revealed that HCV is related to flavi- and pestiviruses (12), and HCV was subsequently placed in the family *Flaviviridae*. The genomic RNA of viruses in this family has a long open reading frame (ORF) flanked by nontranslated regions (NTRs) at the 5′ and 3′ termini (48). HCV initiates translation of the ORF via an internal ribosome entry site (IRES) (25, 51). Translation of the ORF yields a polyprotein that is cleaved co- and post-translationally by host and viral proteases. The HCV polyprotein gives rise to 10 viral proteins: core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B (3). NS5B, the viral RNA-dependent RNA polymerase, comprises the C-terminal portion of the polyprotein. The 5′ and 3′ NTRs of HCV have *cis*-acting replication elements (CREs) which are essential for the viral life cycle (14, 32, 33, 36, 70) and are known to bind cellular proteins (15, 21, 26, 46, 57, 67).

Coding regions of viral RNAs often have embedded nucleic acid signals and overlapping reading frames that encode proteins, as in ϕX174 (52), hepatitis B virus (16), and Rous sarcoma virus (27). Embedded RNA signals include promoters (29), nucleation sites for encapsidation (2, 30, 35, 49), and other types of CREs (5, 18, 40, 43, 44, 53). Among the best characterized is the picornavirus CRE (17, 18, 40, 43, 45, 71). The picornavirus CRE acts as a template for the uridylylation of VPg, the protein primer for genome replication, and thus plays a direct role in initiation of RNA replication. The exact sequence, structure, and location of these CREs vary. In poliovirus, it is in the coding region of the nonstructural protein 2C (19, 45, 71).

Considerable sequence diversity exists among HCV genomes, reflected by the many genotypes and subgenotypes (54, 55). The large number and diversity of the HCV sequences make it possible to use bioinformatics to identify regions of the main ORF that are likely to have a second function. These regions have been explored most extensively in the core-encoding region. Bioinformatics and confirmatory antibody testing indicate that the core gene region is overlapped by an alternate ORF that is expressed and stimulates specific immune responses (62, 65, 68). Conserved RNA elements are also present (54, 64), and the initial portion of the core-encoding region is reported to influence IRES function (24, 25, 47, 50, 66). Frame shift signals (8a, 10) also occur in this signal-intense region of the HCV genome.

In addition to the core region, the NS5B coding region may also harbor functional RNA elements. Several stem-loop (SL) structures have been proposed in this region (23, 56, 61). However, of 20 highly conserved synonymous codons present in the NS5B coding region, several cannot be accounted for by the stem-loops that have been proposed thus far (65). In this report, we used a combination of custom-designed sequence alignment software (65), thermodynamic folding programs, and classical comparative phylogenetic analysis to generate secondary structure models of six potential stem-loop elements in the NS5B coding region. We tested the functional significance of these structures by site-directed mutagenesis and RNA structure probing. We provide evidence for a CRE in the

* Corresponding author. Mailing address: Laboratory of Virology and Infectious Disease, Center for the Study of Hepatitis C, The Rockefeller University, 1230 York Ave., New York, NY 10021. Phone: (212) 327-7046. Fax: (212) 327-7048. E-mail: ricec@rockefeller.edu.

NS5B coding region that is essential for replication of subgenomic replicons in cell culture. This novel CRE is a component of a larger cruciform structure located in the 3′ end of the long ORF.

MATERIALS AND METHODS

Cell culture. Cell monolayers of the human hepatoma cell line Huh-7.5 (7) were grown in Dulbecco's modified minimal essential medium (DMEM; Invitrogen) supplemented with nonessential amino acids, 10% fetal bovine serum (FBS), 100 U of penicillin per ml, and 100 μg of streptomycin per ml. Cells were passaged after treatment with 0.05% trypsin–0.02% EDTA and seeded at a dilution of 1:4 to 1:5.

Comparative sequence analysis and building RNA structural models. A total of 208 sequences of HCV isolates were retrieved from GenBank and aligned with MegAlign (DNAstar, Madison, Wis.). A subset of 23 highly diverse sequences was selected and used to generate a consensus. The Windows version of Mfold was applied to a sliding window of 150 to 300 nucleotides of the consensus sequence, generating a series of potential RNA secondary structures. Structures within 10% of the energy minimum were examined to identify regions in which the consensus sequence produced a structure also present in the majority of the individual sequences. Priority was given to secondary structures located in regions known to have excessively conserved codons, defined as before (65). Comparative phylogenetic analysis (28) was applied to the 208 sequences to confirm or eliminate structural models. For this analysis, the 208 sequences were divided into six groups based on genotypes containing 112 (genotype 1), 29 (genotype 2), 13 (genotype 4), 3 (genotype 5), 25 (genotypes 6, 7, 8, and 11), and 26 (genotypes 3 and 10) sequences.

Construction of plasmids for RNA synthesis: TA cloning. PCR-based site-directed mutagenesis was used to introduce silent mutations into the NS5B coding region. For PCR, KlenTaqLA DNA polymerase (kindly provided by Wayne Barnes, Washington University, St. Louis, Mo.) was used.

Mutant 5BSL1. Two separate PCRs were carried out with pCon1/SG-Neo (I) plasmid as a template. Primers 2513 and 2516 were used to generate a 5′ PCR fragment. Primers 2515 and 2514 were used to generate an overlapping 3′ PCR fragment. For assembly, the 5′ and 3′ PCR fragments were mixed in an equal molar ratio and amplified using outer primers 2513 and 2514.

Mutant 5BSL2. Primers 2752 and 2755 were used for the 5′ PCR product, and primers 2753 and 2754 were used for 3′ PCR product. The 5′ and 3′ PCR products were mixed and amplified with outer primers 2752 and 2753.

PCR mutagenesis and assembly of mutants 5BSL3.1, 5BSL3.2, and 5BSL3.3 were performed using the same strategy. Flanking primers 2253 and 2254 were used for the second assembly PCR. Primer sequences used to produce the 5′ and 3′ PCR products are listed in Table 1. Assembled mutant PCR products were then subcloned into pCR2.1 by TA cloning (Invitrogen). Sequences were confirmed by automated nucleotide sequencing prior to cloning into the parental plasmid pCon1/SG-Neo (I).

The *pCon1/SG-Neo (I)-5BSL1* mutant was assembled in two steps. A *Bcl*I-*Bsu*36I-digested DNA fragment from the TA clone harboring the mutant 5BSL1 was subcloned into *Bsu*36I-*Bcl*I-digested pCon1/SG-Neo (I) without the internal *Bsu*36I-*Bsu*36I fragment, which was then recloned into the unique *Bsu*36I site in order to restore the Con1/SG-Neo (I) backbone. The orientation of the fragment was confirmed by digestion with appropriate restriction enzymes.

The *pCon1/SG-Neo (I)-5BSL2* mutant was constructed by a four-fragment ligation. The TA-5BSL2 clone was digested with *Sfi*I and *Nco*I, followed by ligation with the *Nco*I-*Xba*I, *Xba*I-*Xho*I, and *Xho*I-*Sfi*I fragments of pCon1/SG-Neo (I).

The rest of the mutant plasmids were constructed by three-fragment ligations. The *Kpn*I-*Spe*I fragment of each mutant TA clone was ligated with the *Kpn*I-*Xho*I and *Xho*I-*Spe*I fragments of pCon1/SG-Neo (I). Plasmids were purified by CsCl banding prior to linearization and in vitro transcription.

Insertion of CRE sequences between the stop codon of the Neo^r gene and the encephalomyocarditis virus (EMCV) IRES. To insert a 200-base sequence containing the SL3 cruciform structure after the stop codon of the Neo^r gene in pCon1/SG-Neo (I), a *Bsi*W1 restriction site was engineered adjacent to a *Pme*I site. Primers 3131 and 3132 were used for PCR, and the product was digested with *Pme*I and *Hin*dIII. The resulting fragment was ligated with *Pme*I-*Hin*dIII-digested pCon1/SG-Neo (I) or mutant p5BSL3.2 mutA. Primers 3133 and 3134 were used for PCR with the parental or the mutant templates. These PCR products contained the 200-base sequence of the 5BSL3 cruciform region flanked by *Pme*I and *Bsi*WI sites. The wild-type or mutant *Pme*I-*Bsi*WI fragment was then cloned into *Pme*I-*Bsi*WI-digested pCon1/SG-Neo (I) or p5BSL3.2 mutA. A schematic of these constructions is depicted below in Fig. 9A.

Insertion of CRE sequences following the ORF stop codon. pCon1/SG-Neo (I) was digested with *Bgl*II, and the staggered ends were filled in with the Klenow fragment of DNA polymerase I. After digestion with *Spe*I, the 688-base fragment was cloned into pCon1/SG-Neo (I) 3′ *Afl*II that had been digested with *Afl*II (filled in) and *Spe*I. pCon1/SG-Neo (I) 3′ *Afl*II (D. Jacobson, G. Randall, and C. M. Rice, unpublished data) harbors two nucleotide substitutions in the 3′ NTR variable region that create an *Afl*II site 17 nucleotides downstream of the polyprotein stop codon. The final construct therefore consisted of a duplication of 440 bases, including the C-terminal NS5B coding region with the wild-type 5BSL3 cruciform structure (see Fig. 9B, construct c). To introduce the same 440-base sequence in the context of 5BSL3.2 mutB, PCR was first performed with primers 2752 and 2252 using p5BSL3.2 mutB as a template. The PCR fragment was TA cloned, sequenced, and digested with *Bgl*II and *Afl*II. The resulting fragment was ligated with a 939-base *Bgl*II-*Mfe*I fragment of pCon1/SG-Neo (I) and the 10,374-base *Afl*II-*Mfe*I fragment of construct c to give rise to construct e. To construct d (see Fig. 9B), p5BSL3.2 mutB was digested with *Bgl*II, filled in, and then digested with *Spe*I. Plasmid e was digested with *Afl*II, filled in, and then digested with *Spe*I. The 688-base *Bgl*II-*Spe*I and 11,080-base *Afl*II-*Spe*I fragments were ligated to produce construct d.

In vitro RNA transcription. As described previously (7), plasmid DNAs were linearized with *Sca*I (New England Biolabs), followed by phenol-chloroform extraction and ethanol precipitation. RNA transcripts were synthesized at 37°C for 1.5 h in a 50-μl reaction volume containing 40 mM Tris-HCl (pH 7.9), 10 mM NaCl, 12 mM $MgCl_2$, 2 mM spermidine, 10 mM dithiothreitol, 3 mM NTP, 100 U of RNasin (Roche), 100 U of T7 RNA polymerase (Epicentre Technologies), and 3 μg of linearized DNA. At the end of the reaction, 10 U of DNase I (Roche) was added and incubated at 37°C for 20 min to remove the DNA template. DNase-treated RNAs were extracted with acid phenol-chloroform, followed by precipitation with isopropanol. After washing the RNA pellet with 75% ethanol, the RNAs were dissolved in RNase-free distilled water. The integrity of RNA was confirmed by agarose gel electrophoresis.

Transfection of Huh-7.5 cells and selection with G418. Subconfluent Huh-7.5 cells were trypsinized, collected, and washed twice with ice-cold RNase-free phosphate-buffered saline. The cells were resuspended in phosphate-buffered saline at 2.5×10^7 cells/ml. One microgram of the in vitro-transcribed RNA was mixed with 0.4 ml of the Huh-7.5 cell suspension, transferred to a 2-mm gap cuvette (BTX), and pulsed with a BTX ElectroSquarePorator (0.82 kV; five pulses; 99-μs pulse length, 1.1-s intervals). After 10 min at room temperature, the cells were transferred to 9.6 ml of DMEM–10% FBS. To measure the efficiency of G418-resistant colony formation, the transfected cells were plated at a series of densities (10^6, 10^5, and 10^4 cells). To maintain the total number of plated cells at 10^6 cells per 10-mm dish, cells transfected with polymerase-defective Con1/SG-Neo pol$^-$ RNA transcripts were used as feeders. After 48 h, the medium was changed to DMEM–10% FBS supplemented with G418 (Geneticin; Gibco Life Technologies) at 1 mg/ml. The medium was replaced every 4 days. After about 3 weeks, G418-resistant colonies were stained with 1% crystal violet in 50% ethanol.

Replicon RNA analysis and quantitation. After RNA transfection as described above, 7.5×10^5 cells were plated in one well of a six-well plate. On day 2 postelectroporation, cells were passaged 1:3 for later time points. The cells were maintained without drug selection and harvested at the indicated time points. Total RNA was isolated using TRIzol reagent (Gibco-BRL) according to the manufacturer's instructions. A 100-ng aliquot of total RNA was used to quantify HCV-specific RNA levels using an ABI Prism 7700 sequence detector (Applied Biosystems). Real-time reverse transcription (RT)-PCR amplifications were performed using the Platinum Quantitative RT-PCR ThermoScript One-Step system (Invitrogen, Life Technologies) and primers specific for the HCV 5′ NTR: 5′-CCTCTAGAGCCATAGTGGTCT-3′ (sense, 10 μM); 5′-CCAAATCTCCA GGCATTGAGC-3′ (antisense, 10 μM); and 6-carboxyfluorescein-CACCGGA ATTGCCAGGACGACCGG (probe, 10 μM; Applied Biosystems). RT reactions were incubated for 30 min at 50°C, followed by inactivation of the reverse transcriptase coupled with activation of *Taq* polymerase at 95°C for 5 min. After cooling to 25°C for 2 min, 40 cycles of PCR were performed with cycling conditions of 15 s at 95°C, 40 s at 50°C, and 30 s at 72°C. The reaction mix contained the glyceraldehyde-3-phosphate dehydrogenase detection mix from Applied Biosystems (VIC-MGBNFQ) for normalization. Synthetic HCV RNA standards of known concentration were included with each set of reactions and used to calculate a standard curve. The real-time PCR signals were analyzed in a multiplex format using SDS software (version 1.6.3; Applied Biosystems).

RNA secondary structure probing: enzymatic probing. One microgram of transcript RNA was treated with cobra venom RNase V_1 (Ambion) or RNase T_1

TABLE 1. Primers used in this study[a]

Primer	Sense	Enzyme or probe	Sequence
2252	−	*Afl*II	GCCTATTGGCCTTAAGTGTTTAGGTCCCCG
2253	+	*Kpn*I	TGGGGTACCGCCCTTGCGAG
2254	−	*Spe*I	CCCACTAGTGGCGCGCCTG
2513	+	*Bcl*I (5999)	GGACAGGCGCCCTGATCACGCCATG
2514	−	*Bsu*36I (6551)	GAAGAGCCCATCACGGCCTGAGG
2515	+	5B SL1 (6314)	TCCAGCAA**A**GC**G**GTTAACCACATCCGCTC**G**GT**T**TGGAAGGAC**C**T**C**CT**C**GAAGACACTGAGACACC
2516	−	5B SL1	**G**AG**G**AGGTCCTTCCA**A**AC**C**GAGCGGATGTGGTTAAC**C**GC**T**TTGCTGGATAGGTTCCGG
2517	+	5B SL3.2 B	GTTGCTGGTTAC**TCG**GG**C**GG**T**GACATATATCACAGCCT**CAGCA**G**G**GCCAGGCC**GA**G**G**TGGTTCATGTGG
2518	−	5B SL3.2 B	CCACATGAACCA**CC**T**C**GGCCTGGC**CC**T**GCTG**AGGCTGTGATATATGTC**A**CC**G**CC**CGA**GTAACCAGCAAC
2519	+	5B SL3.2 A	GGTTAC**TCG**GG**C**GG**T**GACAT**C**TACCA**TTCT**CT**CAGCA**G**G**GCC**A**G**G**CC**GA**G**G**TGG
2520	−	5B SL3.2 A	CCA**CC**T**C**GG**C**T**G**GC**CC**T**GCTG**AG**AGAA**TG**G**TA**G**ATGTC**A**CC**G**CC**CGA**GTAACC
2521	+	5B SL3.3	CCCCGCTGGTTCATGTGGTG**T**CT**C**CT**G**C**T**C**T**T**G**TCTGTCGG**T**GT**T**GG**G**ATCTATCTACTCCCC
2522	−	5B SL3.3	GGGGAGTAGATAGAT**C**CC**A**AC**A**CC**G**ACAGA**CA**A**G**AG**C**AG**G**AG**A**CACCACATGAACCAGCGGGG
2621	+	5B SL3.2 D	AGCGGGGGAGACAT**C**TA**C**CA**TTCT**CTGTCTCGTGCC
2622	−	5B SL3.2 D	GGCACGAGACAG**AGAA**TG**G**TA**G**ATGTCTCCCCCGCT
2623	+	5B SL3.2 E	CACAGCCTGTCTCG**G**GCC**A**GG**C**CCCGCTGGTTC
2624	−	5B SL3.2 E	GAACCAGCGGGG**CC**T**G**GC**C**CGAGACAGGCTGTG
2625	+	5B SL3.2 C	AGCGGGGGAGACAT**C**TA**C**CA**TTCT**CTGTCTCG**G**GCC**A**GG**C**CCCGCTGGTTC
2626	−	5B SL3.2 C	GAACCAGCGGGG**CC**T**G**GC**C**CGAGACAG**AGAAT**G**G**TA**G**AGGTCTCCCCCGCT
2752	+	6505	CTTTACGATGTGGTCTCCACC
2753	−	7394 (*Nco*I)	GCGCTAAGGCCATGGAGTCG
2754	+	SL2	GACGAGGCGAGCCT**T**CG**C**GCCTTCAC**C**GA**A**GC**A**ATGACT**C**G**C**TACTCTGCCCCC
2755	−	SL2	GGGGGCAGAGTA**GC**G**A**GTCAT**T**GC**T**TC**G**GTGAAGGC**G**CG**AA**GGCTCGCCTCGTC
2756	+	SL3.1	GCTGCGTCCCAGTTGGATT**CCAG**C**TCG**TGGTTCGTTGCTG
2757	−	SL3.1	CAGCAACGAACCA**CGA**G**CTGG**AATCCAACTGGGACGCAGC
2937	+	A68C	CGGGGGAGACAT**CT**ATCACAGCCTG
2938	−	A68C	CAGGCTGTGATA**G**ATGTCTCCCCCG
2939	+	U71C	GGGAGACATATA**C**CACAGCCTGTCTC
2940	−	U71C	GAGACAGGCTGTG**GT**ATATGTCTCCC
2941	+	C74U	GAGACATATATCA**T**AGCCTGTCTCGTG
2942	−	C74U	CACGAGACAGGCT**A**TGATATATGTCTC
2943	+	A75U/G76C/C77U	GACATATATCAC**TCT**CTGTCTCGTGCC
2944	−	A75U/G76C/C77U	GGCACGACACAG**AGA**GTGATATATGTC
2945	+	C77U	CATATATCACAG**T**CTGTCTCGTGCC
2946	−	C77U	GGCACGAGACAG**A**CTGTGATATATG
2947	+	C76A/U86G	CACAGCCTGTCT**AG**G**C**CCCGACCCCGC
2948	−	C76A/U86G	GCGGGGTCGGGC**CC**TAGACAGGCTGTG
2949	+	C76A/U86A	CACAGCCTGTCT**AGA**GCCCGACCCCGC
2950	−	C76A/U86A	GCGGGGTCGGGC**TC**TAGACAGGCTGTG
2951	+	U86G	CAGCCTGTCTCG**G**GCCCGACCCCGC
2952	−	U86G	GCGGGGTCGGGC**C**CGAGACAGGCTG
2953	+	C90A/A92G	CTGTCTCGTGCC**AG**G**C**CCCGCTGGTTC
2954	−	C90A/A92G	GAACCAGCGGGG**CC**T**G**GCACGAGACAG
2955	+	A92G	GTCTCGTGCCCG**G**CCCCGCTGGTTC
2956	−	A92G	GAACCAGCGGGG**C**CGGGCACGAGAC
3057	−	Primer 1	GAAAAAAACAGGATGGCCTATTG
3059	−	Primer 3	CTATTGGCCTGGAGTGTTTAG
3060	−	Primer 4	GAGTGTTTAGGTCCCCGTTC
3131	+	PmeI-*Bsi*W1	CTTCTGAGTTTAAACACGTACGACCACAACGGTTTCCCTCTAG
3132	−	*Hin*dIII	GTTGTTTGTCTTCAAGAAGCTTCC
3133	+	*Pme*I 5B SL3	AGCTTTGTTTAAACTCACTCCAATCCCGGC
3134	−	*Bsi*WI 5B SL3	CCGTTGTGGTCGTACGAGTGTTTAGGTCCCCGTTCATC
3298	+	A92C	GTCTCGTGCCCG**C**CCCCGCTGGTTC
3299	−	A92G	GAACCAGCGGGG**G**CGGGCACGAGAC
3300	+	A92U	GTCTCGTGCCCG**T**CCCCGCTGGTTC
3301	−	A92U	GAACCAGCGGGG**A**CGGGCACGAGAC
3302	+	C90A/A92A	GTCTCGTGCC**AG**A**C**CCCGCTGGTTC
3303	−	C90A/A92A	GAACCAGCGGGG**TC**TGGCACGAGAC

[a] Mutations are indicated in boldface. Generated restriction enzyme sites are underlined.

(U.S. Biochemicals) at 0°C for 30 min as described elsewhere (8). The reaction was carried out in a 10-μl reaction volume with 10 mM Tris-HCl (pH 7.0), 0.1 M KCl, 10 mM $MgCl_2$, 3 μg of yeast tRNA, and RNases. One or 0.1 U of RNase T_1 and 0.1 or 0.01 U of RNase V_1 were used. The reaction was extracted with acidic phenol-chloroform followed by ethanol precipitation.

Lead cleavage. Using a protocol adapted from reference 38, 1 μg of transcript RNA was incubated in a 10-μl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 10 mM Mg-acetate, 100 mM K-acetate, 3 μg of yeast tRNA, and a range of concentrations of lead acetate (0, 5, 10, 15, and 30 mM) for 5 min at 37°C. After the reaction, the volume was adjusted to 50 μl with RNase-free water and extracted with acid phenol-chloroform followed by ethanol precipitation.

Primer extension analysis. Oligonucleotides (MWG Biotechnologies) used for primer extension analyses were primer 1 (5′-GAAAAAAACAGGATGGCCTATTG-3′), primer 3 (5′-CTATTGGCCTGGAGTGTTTAG-3′), and primer 4 (5′-

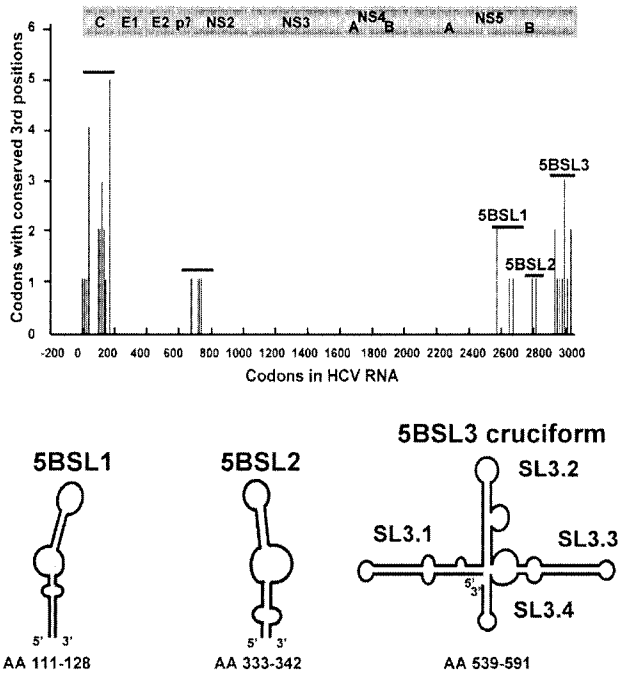


FIG. 1. Locations of codons with highly conserved third positions and diagrams of predicted RNA secondary structures in the NS5B coding region. Large clusters of third-base conservation were identified in the core and NS5B by phylogenetic analysis (65). Shown below are models of potential secondary structures identified in the NS5B coding region with methods including the RNA folding program Mfold. The locations of these predicted RNA structures (NS5B codon number) are indicated.

GAGTGTTTAGGTCCCCGTTC-3′). Each oligonucleotide was gel purified and 5′-end labeled with 10 U of T4 polynucleotide kinase (NEB) and 100 μCi of [γ-^{32}P]ATP (8,000 Ci/mmol; Amersham) in a reaction mixture containing 70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 100 mM KCl, and 1 mM β-mercaptoethanol at 37°C for 30 min. The unincorporated radioisotope was removed by passing the reaction mixture through QIAquick silica gel (Qiagen). Cleaved RNAs were mixed with the 5′-end-labeled oligonucleotide (5×10^4 cpm as determined by liquid scintillation counting) in a 20-μl reaction volume containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, and 10 mM dithiothreitol, prewarmed to 40°C for 2 min, and incubated for 50 min following addition of 100 U of Superscript II RNase H$^-$ reverse transcriptase (Gibco-Invitrogen). The cleavage sites on RNA were determined by analyzing the reverse-transcribed cDNA products separated by denaturing polyacrylamide gel (12% polyacrylamide–7 M urea) electrophoresis. Dideoxy sequencing reactions were carried out on unmodified RNA and run in parallel.

RESULTS

Models of RNA secondary structural elements in the HCV NS5B coding region. Previously, synonymous codons with highly conserved third-position nucleotides were identified throughout the main HCV ORF by analyzing eight diverse HCV sequences (65). Five clusters of such conserved codons were identified: two in the structural gene region and three in the NS5B coding region (Fig. 1). A cluster was defined as a region in which at least three highly conserved codons occurred within a 100-codon segment. Analysis revealed that the NS5B coding region did not contain a conserved alternative ORF similar to the one found in the core-encoding region (65). This suggested that the exceptionally conserved codons might be due to RNA structural elements and RNA signals embedded in the polymerase-encoding region.

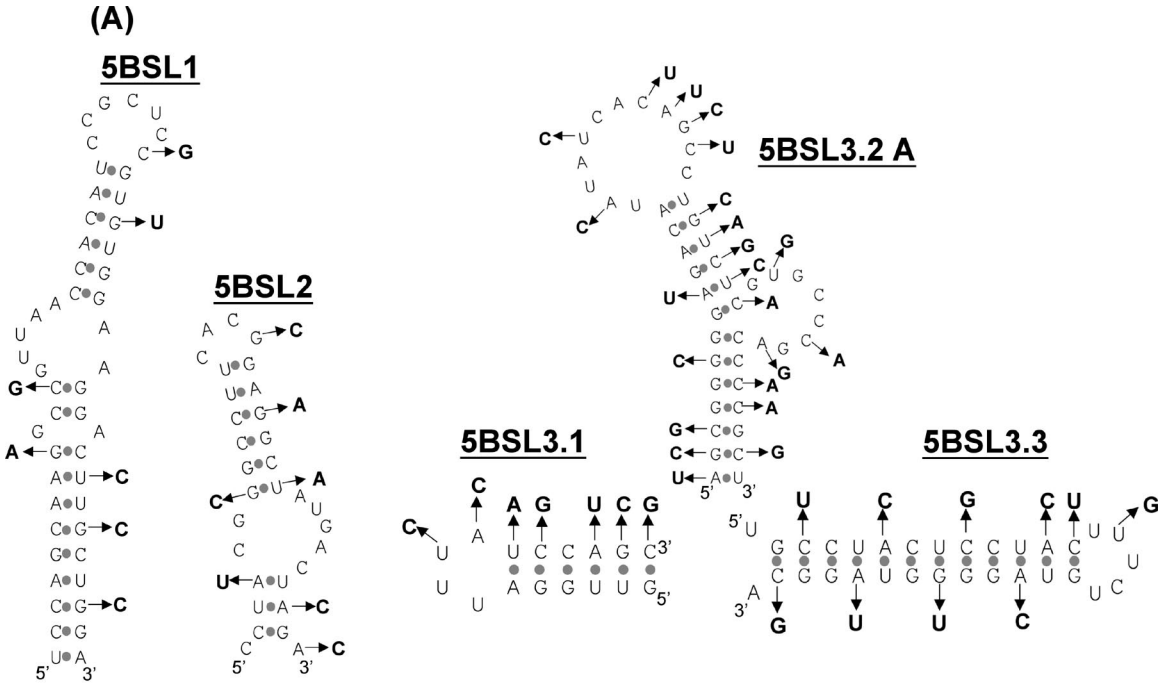
A combination of thermodynamic RNA folding programs and classical comparative phylogenetic analysis was used to build secondary structural models of RNA elements in the NS5B gene. In a reiterative process, a consensus sequence of 23 diverse NS5B sequences was generated. The Windows version of Mfold was then applied to a sliding window of 150 to 300 nucleotides, and a series of potential secondary structures was obtained. Structures that were also predicted to be present in the majority of the individual sequences were selected for further investigation. Comparative phylogenetic analysis of 208 sequences was used to identify compensatory base pair changes and to refine the structural models. This process yielded six predicted stem-loop elements (Fig. 1). Five have been proposed previously, either in whole or in part. 5BSL1 (codons 111 to 128) and 5BSL2 (codons 333 to 342) were originally noted by Hofacker and colleagues (23) and later modified by Tuplin et al. (61). 5BSL3.1, 5BSL3.3, and 5BSL3.4 were first noted by Smith and Simmonds (56).

The stem-loop element 5BSL3.2 has not been described previously. The 48 nucleotides that fold into 5BSL3.2 encode amino acids 555 to 571 of the NS5B. 5BSL3.2 is one of four stem-loop elements that may comprise a larger structure, which is depicted as a cruciform. The nucleotides comprising the cruciform encode the NS5B amino acids 539 to 591. These amino acids make up the C-terminal portion of the polymerase structure, a linker region, and a membrane anchor (52a), which are downstream of the thumb domain (37) of the protein.

Directed mutagenesis identified a potential CRE in the NS5B coding region. To examine the possible function of these RNA structures in the NS5B coding region, the subgenomic replicon system (6) was used. Con1/SG-Neo (I), the parental replicon in this study, contains an adaptive S2204I mutation in NS5A that greatly enhances replication (6). Multiple mutations were introduced into each element to disrupt its predicted RNA structure. In all cases, mutations were chosen to preserve the amino acid sequence of NS5B and to conserve the codon usage frequencies found in the HCV ORF. Codon choice was considered to be important, since silent mutations that introduce rare codons can affect the folding of the resulting protein (13), possibly by reducing the rate of translation and favoring the production of β-strands (1).

The multiple nucleotide changes designed to disrupt the predicted RNA structures are indicated in Fig. 2A. The Mfold algorithm was used to generate the lowest energy structure predicted for each mutant sequence and to determine its free energy. The predicted mutant RNA structures (data not shown) and their calculated free energies (Fig. 2B) indicated that the mutations should successfully disrupt the target structures.

To assess the impact of these substitutions on HCV RNA replication, we measured the G418 transduction efficiencies of replicons bearing these mutations. In vitro-transcribed RNAs were transfected into Huh-7.5 human hepatoma cells by electroporation. A polymerase-defective (pol$^-$) replicon was tested in parallel as a negative control. The initial set of mutations in the 5BSL3.2 structure were lethal. The 5BSL3.2 mutA replicon produced no G418-resistant colonies (Fig. 3A). In contrast, the G418 transduction efficiencies of the 5BSL1,



(B)

ΔG (kcal/mol)	5BSL1	5BSL2	5BSL3.1	5BSL3.2 A	5BSL3.3
wt	- 23.5	- 8.8	- 3.1	- 17.4	- 22.9
mut	- 14.0	- 2.5	1.5	- 9.4	- 3.5

FIG. 2. Targeting of each stem-loop structure by site-directed mutagenesis. (A) Predicted RNA secondary structures and the mutations that were introduced. (B) Lowest thermodynamic free energies of wild-type and mutant structures as determined by Mfold.

5BSL2, 5BSL3.1, and 5BSL3.3 mutants were similar to that of the parental replicon (Fig. 3A). These results indicate that 5BSL3.2 contains one or more essential RNA elements that play an important role in HCV RNA replication.

HCV RNA levels of each mutant were also quantified by real-time RT-PCR. The replication-defective subgenomic replicon (pol$^-$) was used to follow input RNA and to distinguish it from replicating RNA. As shown in Fig. 3B, at day 1 similar levels of RNA were present for all constructs tested, suggesting that mutations in the predicted structures did not affect RNA stability. Over time, the level of parental replicon RNA was maintained and even increased slightly, whereas that of the pol$^-$ mutant decreased. By day 3 postelectroporation, the level of parental RNA was significantly higher than that of the pol$^-$ mutant. By day 5, the RNA level of the parental replicon was 200-fold higher than that of the pol$^-$ mutant. Mutants 5BSL1, 5BSL2, 5BSL3.1, and 5BSL3.3 generated progeny RNA at about the same rate as the parent, paralleling the results of the G418 transduction assay. The RNA levels for these mutants were about ~150- to 300- fold higher than the RNA of the pol$^-$ mutant on day 5. In contrast, RNA levels of 5BSL3.2 mutA decreased over time, falling at about the same rate as the pol$^-$ mutant. This indicates that 5BSL3.2 is essential for viral RNA replication.

We then examined the proposed stem-loop structure of 5BSL3.2 in greater detail. Figure 4 summarizes a phylogenetic analysis of 208 GenBank sequences relative to Con1 (a consensus sequence of genotype 1b [41]). 5BSL3.2 has a lower helix, an upper helix, a terminal loop, and an internal loop (bulge). The phylogenetic data are presented for each predicted base pair and each unpaired nucleotide. All of the base pairs that are indicated in the diagram can be formed by at least 75% of the 208 sequences, except for the initial base pair, which is present in about 66% of the sequences. Fifteen of the 48 nucleotides in 5BSL3.2 are perfectly conserved among the 208 sequences. Four of the proposed base pairs are comprised of perfectly conserved bases and, thus, by the definitions used in comparative phylogenetic analysis it is not possible for them

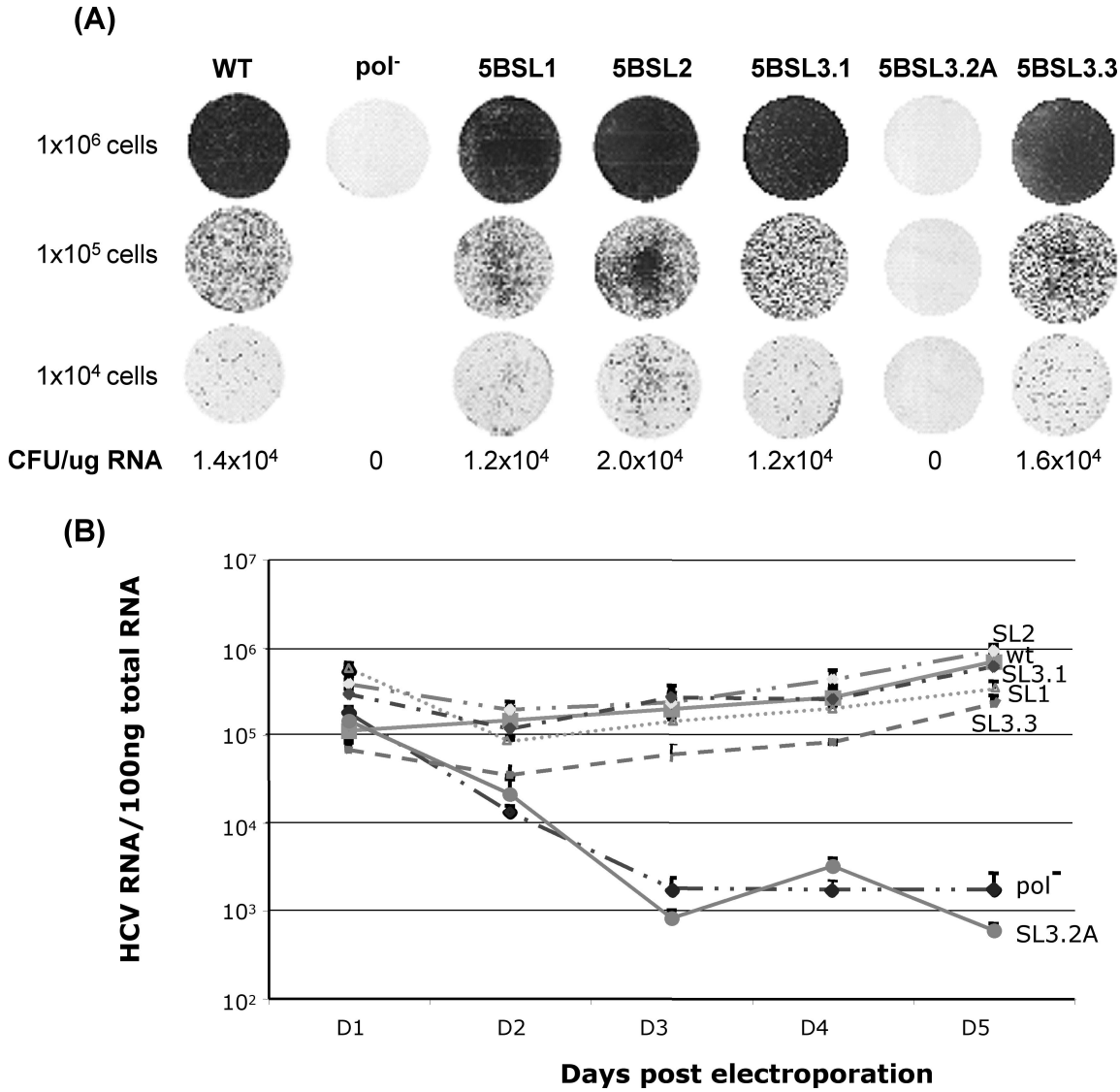


FIG. 3. Effects of silent mutations in stem-loop structures on replication of the subgenomic replicons. (A) Transduction efficiencies of the mutant constructs in a G418 drug selection assay. After electroporation, 10^6, 10^5, or 10^4 transfected cells were plated with feeders as described in Materials and Methods. After selection for 3 weeks, G418-resistant colonies were stained with crystal violet. Approximate transduction efficiencies (in CFU per microgram of RNA) are indicated. (B) Effects of mutations on HCV RNA levels at various times after transfection. Cells were harvested on day 1, 2, 3, 4, and 5 postelectroporation, and total RNA was harvested. One hundred nanograms of total RNA was analyzed by real-time RT-PCR (see Materials and Methods). HCV RNA was normalized to glyceraldehyde-3-phosphate dehydrogenase RNA. Similar results were obtained in two independent repetitions of the experiment.

to have supporting mutational data. The lower helix of 5BSL3.2 contains 8 bp. Three of the eight are supported by compensatory base pair changes, providing strong evidence that this helix exists in vivo. The upper helix contains six Watson-Crick base pairs. The nucleotides in three of these base pairs are perfectly conserved and thus are not amenable to comparative phylogenetic analysis. Two of the other 3 bp are supported by compensatory base pair changes. The third base pair in this group, which is located at the base of the helix, cannot be confirmed by phylogenetic analysis; however, its existence in the HCV RNA of the replicon is supported by structure probing data (see below).

RNA structure probing supports the proposed cruciform. The 5BSL3.2 element is predicted to be a part of a larger cruciform structure that is formed by 171 nucleotides comprising the C-terminal portion of the NS5B coding region and the first part of the 3′ NTR. Cleavage by various RNases and lead(II) was used to identify nucleotides within this 171-base region that are part of single-stranded or double-stranded structures. To allow possible long-range interactions between nucleotides in the cruciform structure and other parts of functional replicon RNA to be detected, we used the same 7,987-base subgenomic RNA transcripts for the structure probing experiments that we used in the transduction assays. Optimal

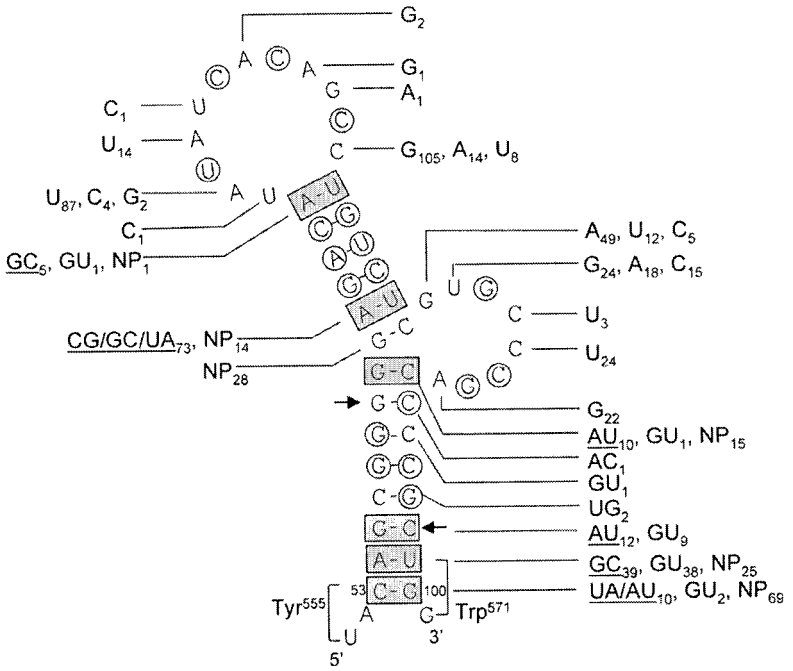


FIG. 4. Analysis of conserved nucleotides and covariant base pairs in 5BSL3.2. The sequence and predicted secondary structure of the parental replicon [Con1/SG-Neo (I)] were compared to 208 sequences from GenBank. The number and identity of nucleotides in the GenBank sequences that differ from those in the wild-type replicon are indicated. Base pairs in the 5BSL3.2 structure that were confirmed by compensatory base pair changes are shaded. Nucleotides that are perfectly conserved among all 208 sequences are circled. The number of sequences that have covariant base pairs, wobble base pairs, and those that are unable to form conventional base pairs are indicated. NP (nonpaired) represents pairs of bases that cannot form conventional base pairs. Arrows identify codons whose third-position nucleotides are highly conserved, as defined before (65). On the 5′ side, UAC is the codon for the NS5B amino acid Tyr^{555}. C53 is the 53rd base according to the 5BSL3 cruciform numbering scheme used below in Fig. 5, 7, and 8. On the 3′ side, UGG is the codon for Trp^{571}. In this analysis, the GenBank sequences were also divided into six groups: genotype 1; genotype 2; genotype 4; genotype 5; genotypes 6 to 9 and 11; and genotypes 3 and 10 (55, 60). Each group was analyzed individually (data not shown). There was general agreement, with one exception. The base pair located at the base of the upper helix does not appear to be formed by sequences of genotype 3 and 10.

Mg^{2+} and salt concentrations, and also oligonucleotides for primer extension analyses, were determined using the parental replicon RNA (Table 1; see also Materials and Methods).

The RNA structure probing confirmed many features of the predicted cruciform structure, particularly those in and around 5BSL3.2 and 5BSL3.3 (Fig. 5). RNase T_1, which cleaves after guanosine residues in single-stranded regions, detected G76 in the terminal loop and G85 and G87 in the internal loop of 5BSL3.2. Lead(II), a reagent that primarily cleaves bases in single-stranded loops or in bulges, reacted with bases in both loops of 5BSL3.2 and the terminal loop of 5BSL3.3. In addition, lead(II) cleavage confirmed the single-stranded character of the regions between 5BSL3.1 and 5BSL3.2 as well as between 5BSL3.2 and 5BSL3. Cleavages by RNase V_1, which prefers base-paired nucleotides or stacked helical single-stranded regions, were found in two distinct areas. One was in the upper helix of 5BSL3.2. Cleavage in this region supports the predicted interaction between bases A62 and U81. The other area with a high density of RNase V_1 cleavages was in the upper portion of the long helix of 5BSL3.3. Cleavage in this region supports the predicted interaction between C119 and G133. With longer exposure, additional RNase V_1 cleavages were detected within the lower helix of 5BSL3.2, indicating that C56 pairs with G97 and that G57 pairs with C96. Finally, RNase V_1 cleavage indicated that there might be an interaction between C46 and G9 in 5BSL3.1.

In general, enzymatic and chemical mapping data strongly support the main features of the predicted cruciform structure and the model of 5BSL3.2 in particular. Most of the unexpected cleavages were minor and mapped to bases that are near the ends of helices. RNase T_1 cleaved guanosine nucleosides near the base of 5BSL3.1 (G48 and G49) and near the bottom of the 5BSL3.3 helix (G109 and G129). These cleavages may reflect breathing or the presence of conformational isomers in the RNA preparation.

Interesting minor RNase T_1 cleavage sites, G80 and G63, were found in the upper helix of 5BSL3.2. G80 and its partner, C65, as well as G63 and its partner, C82, were perfectly conserved among the 208 sequences analyzed (Fig. 4). This degree of conservation cannot be explained by the need to preserve the coding sequence of the NS5B. The genetic code offers alternatives that would preserve both the amino acid sequence and allow a compensatory base pair change at this position. It is possible that the greater stability of the C-G pair, compared to A-U and G-U pairs, constrains the sequence and helps to explain the conservation. In addition, the susceptibility of the upper helix to both single- and double-strand-specific RNases and the unusual sequence conservation of C65-G80 and G63-

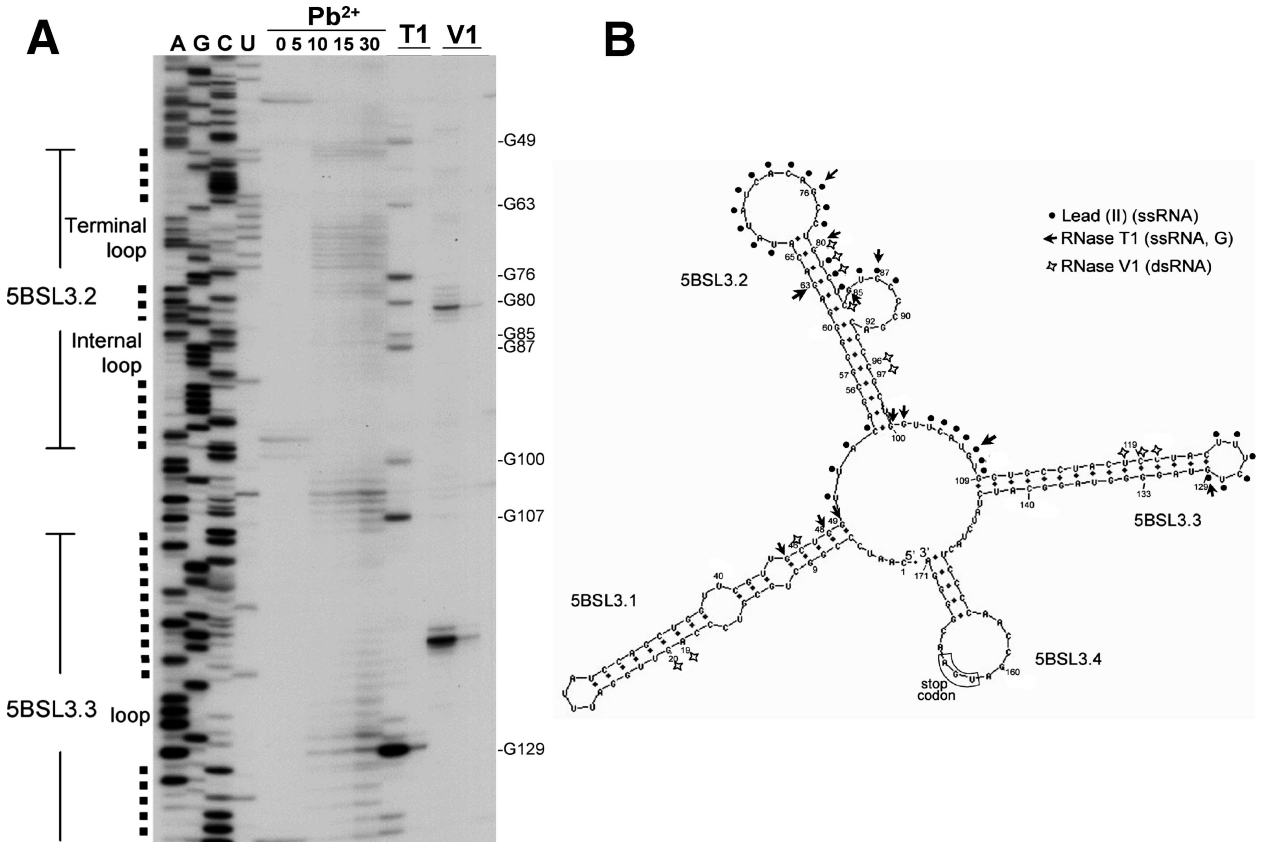


FIG. 5. RNA structure probing of the 5BSL3 cruciform. (A) Primer extension analysis of RNA treated with RNase T_1 (1 or 0.1 U) and RNase V_1 (0.1 or 0.01 U), or lead(II) acetate (0, 5, 10, 15, and 30 mM; see Materials and Methods). cDNA products were resolved in a 12% PAGE–7 M urea denaturing gel. (B) Comparison of the predicted RNA secondary structure with the results of enzymatic and chemical probing. As indicated, the symbols mark the position of cleavage or modification by the different probing reagents.

C82 could indicate that the upper helix undergoes conformational change in solution and that these base pairs participate in as-yet-unidentified interactions that may be important for the function of 5BSL3.2.

Both the structure and primary sequence of 5BSL3.2 are required for HCV RNA replication. A series of progressively more subtle mutations were introduced into Con1/SG-Neo (I) to define key features of the 5BSL3.2 element. In all cases, the amino acid sequence of NS5B and HCV codon usage were preserved. The original mutant analyzed, 5BSL3.2 mutA, contained 22 substitutions in both the stem and loop sequences. 5BSL3.2 mutB contained multiple substitutions within the base-paired regions and was also not viable (Fig. 6). RNA structure probing revealed that the silent changes in both mutA and mutB did indeed disrupt Watson-Crick interactions (data not shown).

To determine the impact of substitutions in loop nucleotides, 5BSL3.2 mutC, mutD, and mutE were produced. These three constructs separate the clusters of loop mutations present in mutA. mutC has changes in both loops and mutD has six changes within the terminal loop, whereas mutE has three changes in the internal loop. Remarkably, the G418 transduction assay revealed that all three sets of loop mutations were lethal. This implies that both the terminal and the internal loops contain primary sequence elements that are crucial for HCV RNA replication.

To gain further insight into the sequence requirements of the two 5BSL3.2 loops, more subtle substitutions were made and tested in the G418 transduction assay (Fig. 7). Single point mutations made at positions 71 (U71C) and 74 (C74U) blocked HCV RNA replication. In addition, the triple mutation of AGC to UCU at positions 75, 76, and 77 also blocked replication. In contrast, position 68 (A68C) could be changed without affecting HCV replication significantly. The mutant RNAs were predicted to fold into structures of similar shape and thermostability as wild-type RNA, and structural probing did not reveal any irregularities (data not shown). These results emphasize the importance of the primary sequence of the terminal loop.

The effects of substitutions in and around the internal loop suggested that the structure of this region, as well as its primary sequence, were important. Three mutant sequences contained substitutions in and around the 5′ portion of the internal loop. These mutations involve the codon that specifies Arg^{566} of NS5B. Mutant C84A/U86G was not viable; mutant C84A/U86A replicated at reduced levels, and mutant U86G replicated efficiently. The change of C84 to A would be expected to disrupt a G-C base pair at the bottom of the upper helix,

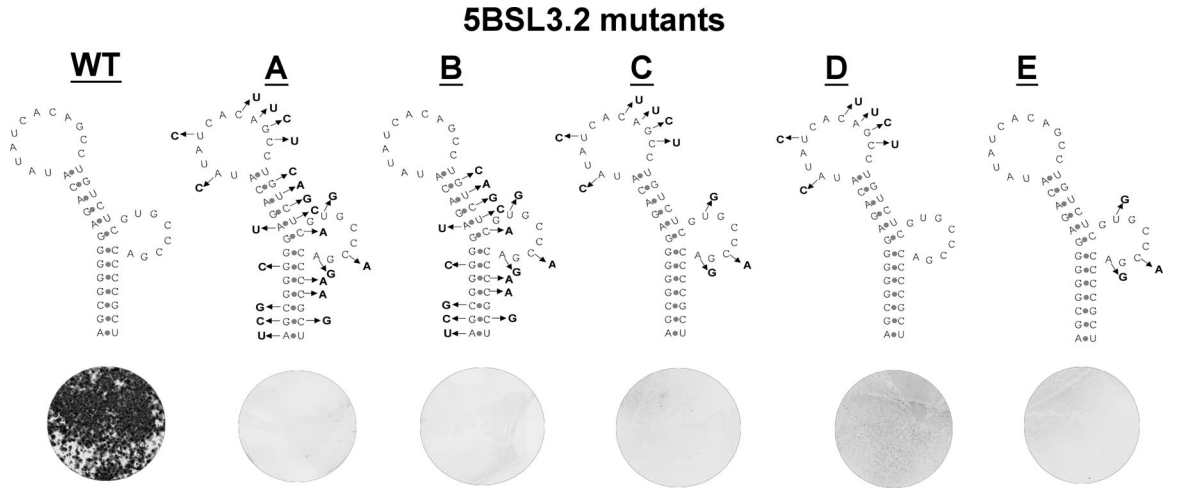


FIG. 6. Genetic analysis of mutant subgenomic replicons targeting the 5BSL3.2 structure. The silent mutations introduced into each construct are shown. Mutant 5BSL3.2 mutA contains mutations throughout the structure. 5BSL3.2 mutB contains the same changes as mutA except for the mutations in the terminal loop. Mutant 5BSL3.2 C has changes in both the terminal and internal loops, which are separated in mutD and mutE, respectively. The effects of these changes were examined by G418 drug selection. Plates with 10^6 transfected cells are shown.

potentially altering the overall structure of the 5BSL3.2 element. Such a disruption could explain the phenotypes of mutant C84A/U86G (lethal) versus mutant U86G (viable). The RNA structure probing data presented below help to account for the phenotype of mutant C84A/U86A (attenuated, but viable) by demonstrating that this replicon retained a structure similar to wild-type 5BSL3.2. Additional studies demonstrated that the sequence of the 3′ portion of the internal loop is critical. Mutant C90A was lethal (Fig. 7B), as might be expected because C90 is a perfectly conserved nucleotide (Fig. 4). Mutants A92C and A92U were viable, although the latter produced 10-fold fewer colonies in the G418 transduction assay. Mutant A92G was lethal, which was surprising given that G occurs at this position in many HCV sequences in GenBank.

C84A/U86G disrupts the upper helix of 5BSL3.2. RNA structure probing was carried out to examine possible alterations in 5BSL3.2 that might account for the nonviability of the C84A/U86G mutant compared to viable mutants C84A/U86A and U86G. The RNase T_1, RNase V_1, and lead(II) cleavage patterns of mutants C84A/U86A and U86G were similar to those of the parental RNA (Fig. 8). In contrast, the RNase T_1 digestion pattern of mutant C84A/U86G revealed two additional intense bands (Fig. 8). These bands resulted from enhanced cleavage after G63 and G80, two nucleotides that had not been mutated. This indicates that the C84A/U86G substitutions modified the structure of the 5BSL3.2 upper helix and that this alteration correlates with the failure of mutant C84A/U86G to replicate.

The differences noted between the digested products of C84A/U86A and those of C84A/U86G suggest that the substitution of A for C at position 84 can be tolerated, but not if G is also present at position 86. In fact, G85 and G87 of the mutant C84A/U86A were targeted by RNase T_1 strongly, but not in C84A/U86G. Having a mutation at C84A would disrupt the base pair at the bottom of the upper stem, which would enlarge the internal loop. This could increase the accessibility of G85 and G87 to RNase T_1 digestion. However, the C84A/U86G mutant did not have the same cleavage pattern as the C84A/U86A mutant, despite the fact that they both contain the C84A change. In the absence of a stable base pair at the bottom of the upper helix, a G at position 86 may interact with nucleotides that would otherwise contribute to this helix, thereby disrupting the overall structure. As noted above, G80 is a minor RNase T_1 cleavage site in wild-type RNA, indicating that the base pair C65-G80 most likely exists as a meta-stable interaction. Taken together, these observations suggest that base pairs in the upper helix establish an essential scaffold of 5BSL3.2. One likely possibility is that the helices hold the bases in the terminal and internal loops in an optimal orientation for interaction with other components of the viral replication complex.

Cruciform duplications and attempts at rescue. Some picornavirus CREs found in the polyprotein coding region can function when relocated to other genome positions. Such constructs, liberated from the constraint of maintaining the protein-coding sequence, can be very useful for fine mutational analysis of CREs. To examine this possibility for 5BSL3.2, we created a convenient insertion site in the replicon between the Neo^r gene and the EMCV IRES (Fig. 9A, construct b). Insertion of an additional wild-type or mutant (5BSL3.2 mutA) copy of 5BSL3 at this position was tolerated without any reduction in replication efficiency (Fig. 9A, constructs c and d). However, when these insertions were tested on the otherwise nonviable 5BSL3.2 mutA background, neither the mutant nor the wild-type 5BSL3 was able to restore replication (Fig. 9A, constructs e and f).

We also tested the 3′ NTR as a possible CRE insertion site. A 440-base sequence that included wild-type 5BSL3 (NS5B codons 447 to 591) was inserted downstream of the polyprotein stop codon (Fig. 9B; see also Materials and Methods). The presence of this duplication dramatically reduced replication efficiency even in the absence of 5BSL3.2 mutations (Fig. 9B, construct c). A construct containing the 5BSL3.2 mutB sequence in both the polyprotein and 3′ NTR positions was

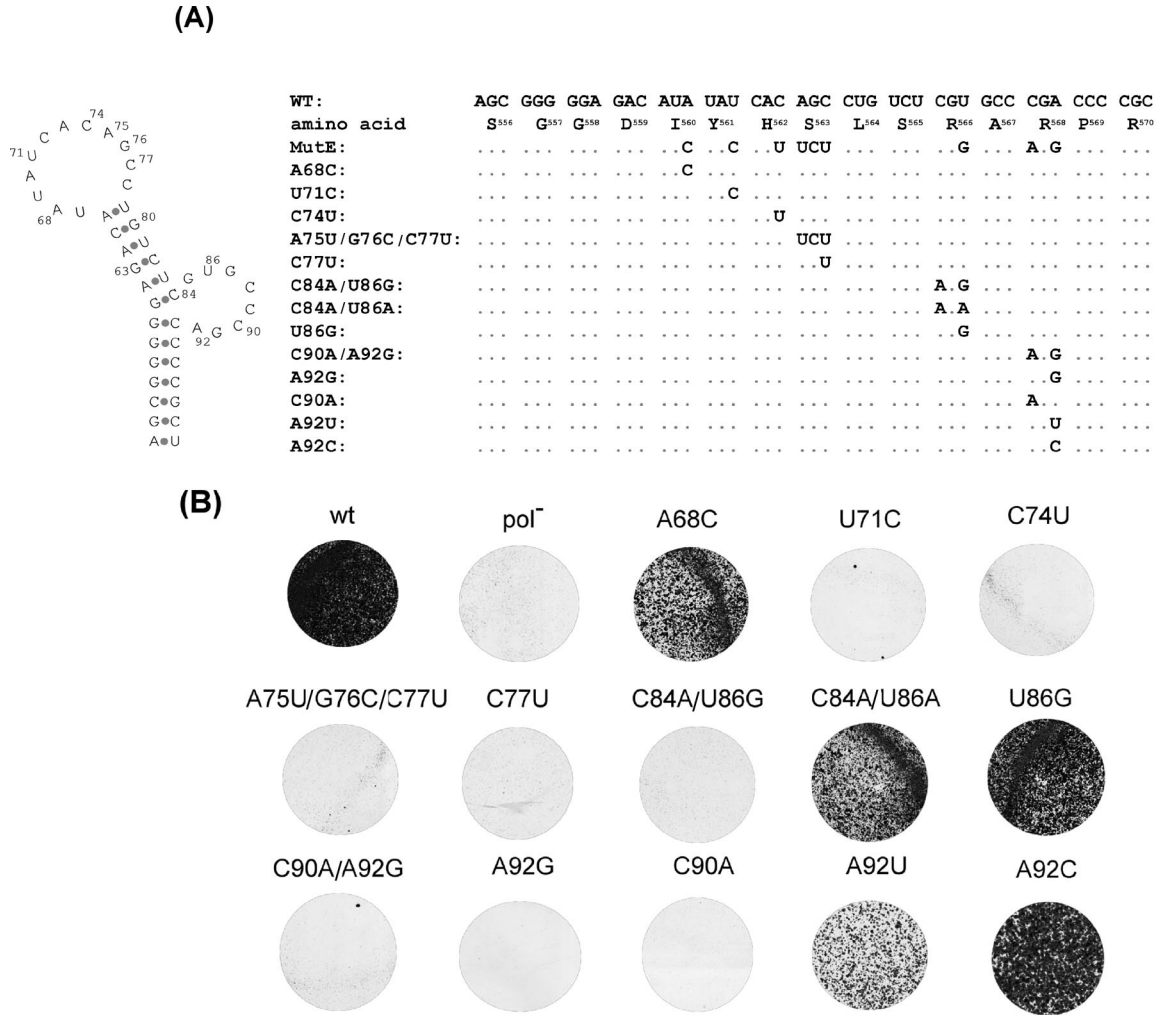


FIG. 7. Mutational analysis of the terminal and internal loops of 5BSL3.2. (A) One-, two-, or three-nucleotide substitutions were introduced at specific positions in the 5BSL3.2 structure, as indicated. (B) Effects of these mutations on G418 transduction efficiency. Plates with 10^6 transfected cells are shown.

nonviable, as expected (Fig. 9B, construct d). However, when the 3′ NTR copy was the wild-type sequence (Fig. 9B, construct e), a few G418-resistant colonies were obtained but in reduced numbers compared to the parental duplication construct (Fig. 9B, construct c). Although we have only examined a limited number of constructs, it appears that the context of the HCV NS5B CRE is important for optimal function.

DISCUSSION

We used custom-designed sequence alignment software, thermodynamic RNA folding programs, and comparative phylogenetic analysis to build models of six RNA stem-loop structures in the HCV NS5B gene. Five out of the six elements identified in this study were reported previously. 5BSL1 was first identified by Hofacker et al. (23) and later by Tuplin et al., who designated it SL7730 (61). 5BSL2 is the tip of a larger structure designated SL8376 by Tuplin et al. Both Hofacker et al. and Tuplin et al. mapped a structure designated SL8828 (codons 474 to 501), which did not appear in our search, and Tuplin et al. identified an additional structure that did not meet our selection criteria, SL8926 (codons 509 to 530). 5BSL3.1 (codons 539 to 554) and 5SB3.3 (codons 575 to 585) correspond to SL9011 and SL9118, respectively (61). 5BSLSL3.1 was first identified by Smith and Simmonds, who also described 5BSL3.3 and 5BSL3.4 (56). As a group, these studies and our investigation indicate that an unusual percentage of the nucleotides in the HCV NS5B coding region may form distinctive RNA structures. The cruciform, which includes the final 171 bases of the main ORF, may be part of an even-larger structure. The NS5B codons immediately upstream, 536, 537, and 538, have exceptionally conserved third-position nucleotides, as defined previously (65). Two stem-loop elements identified by Tuplin et al., SL8828 and SL8926, lie only slightly further upstream.

We used site-directed mutagenesis to analyze five of the elements identified in our search. In an initial screen, multiple mutations designed to disrupt the predicted RNA structures

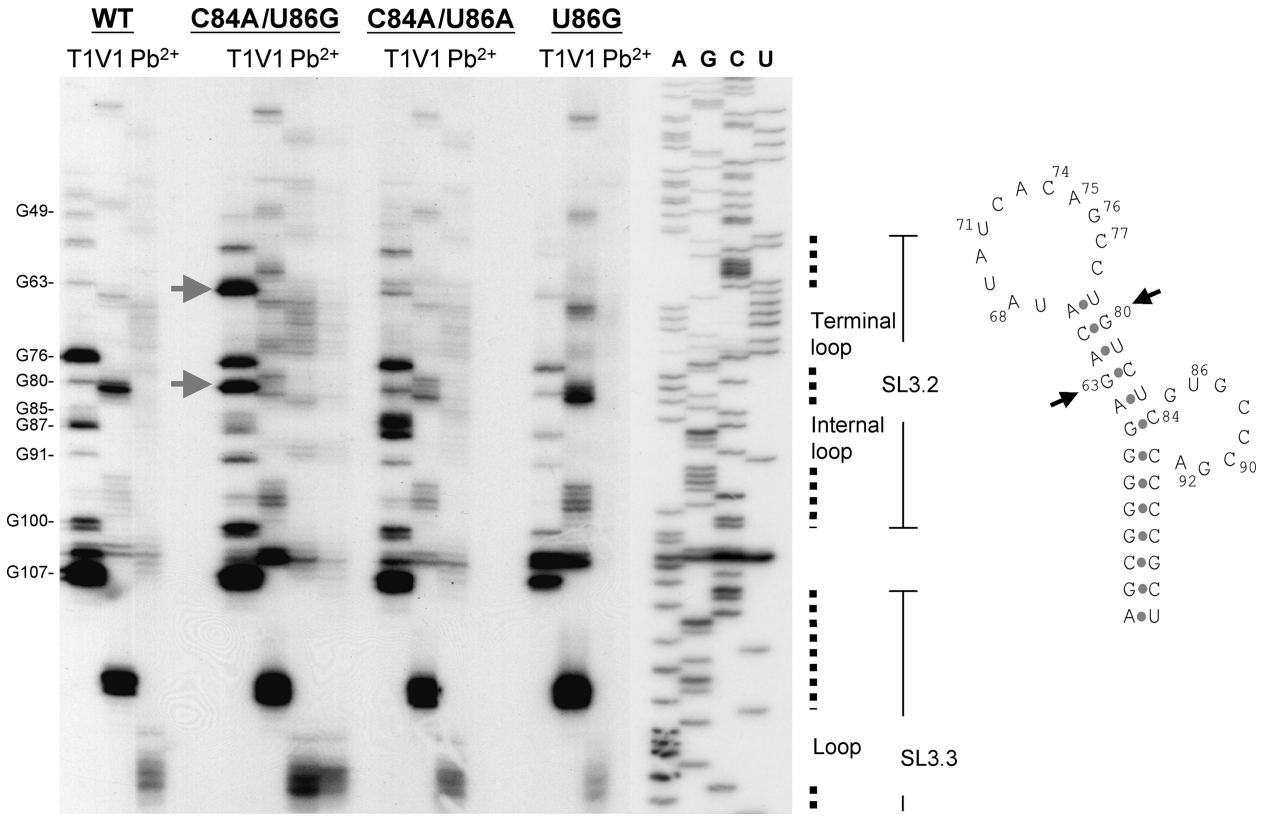


FIG. 8. Destabilization of the 5BSL3.2 upper helix by C84A/U86G. RNA structure probing was used to examine the 5BSL3.2 structure for mutants C84A/U86G, C84A/U86A, and U86G. RNAs were treated with RNase T_1 (1 U), RNase V_1 (0.1 U), or lead(II) (15 and 30 mM) followed by primer extension analysis (see Materials and Methods). cDNA products were resolved on a 12% PAGE–7 M urea denaturing gel. Unique T_1 cleavages observed for mutant C84A/U86G RNA are indicated by arrows. At the right, the locations of these cleavages are also shown on the secondary structure model of 5BSL3.2.

were engineered while preserving the amino acid sequence of the NS5B protein. Uniquely, mutations in 5BSL3.2 blocked replicon replication, indicating that 5BSL3.2 is an essential CRE. Mutations in 5BSL1, 5BSL2, 5BSL3.1, and 5BSL3.3 did not markedly diminish replication in cell culture. We also tested three different mutant subgenomic replicons containing different combinations of silent mutations of the SL8926 element, proposed by Tuplin et al. None could be distinguished from the wild-type replicon in the G418 selection assay (data not shown). However, these negative results should not be taken as definitive evidence that these RNA elements lack functional importance. They may exert subtle effects, function in different cellular environments, or perform other functions in the HCV life cycle (such as packaging) that were not assayed using the current replicon systems. We did not perform directed mutagenesis on 5BSL3.4; however, other studies suggest that this stem-loop may not be essential. 5BSL3.4 includes 8 nucleotides of the variable region of the 3′ NTR, and HCV RNA transcripts lacking these 8 bases and the subsequent 16 bases were still infectious in a chimpanzee (69). Deletion of the entire variable region significantly reduced replication of a subgenomic replicon but was not lethal (14).

Essential features of the CRE 5BSL3.2. 5BSL3.2 is about 50 bases in length and has an 8-bp lower helix, a 6-bp upper helix, a 12-base terminal loop, and an 8-base internal loop (bulge). Disruption of 5BSL3.2 had a striking impact on HCV replication. Many of the mutant replicons with changes in this element did not yield any G418-resistant colonies during drug selection. Directed mutagenesis and structure probing demonstrated that sequences in both loops are important for viability and indicated that the upper helix is an essential scaffold that is easily disrupted.

Primary sequence requirements in the loops of 5BSL3.2. Analysis of a series of mutants with one, two, and three base changes revealed that many of the nucleotides in the loops of 5BSL3.2 are highly constrained and may be invariable. Point mutations at five positions (71, 74, 77, 90, and 92) were lethal (Fig. 7). Structure probing did not reveal any irregularities in the cleavage products of RNAs with mutations in these locations (data not shown). These results indicate that the primary sequences of the loop regions, rather than their contributions to the global structure, are important.

In general, we found a high correlation between the results of directed mutagenesis and the results of phylogenetic sequence analysis. Typically, mutations that were lethal in the replicon were perfectly conserved among the 208 sequences analyzed, while substitutions that retained viability were represented in the set of natural sequences. For example, mutant A68C had a wild-type phenotype, and this position is variable in nature (Fig. 4). Conversely, mutant C74U is nonviable, and

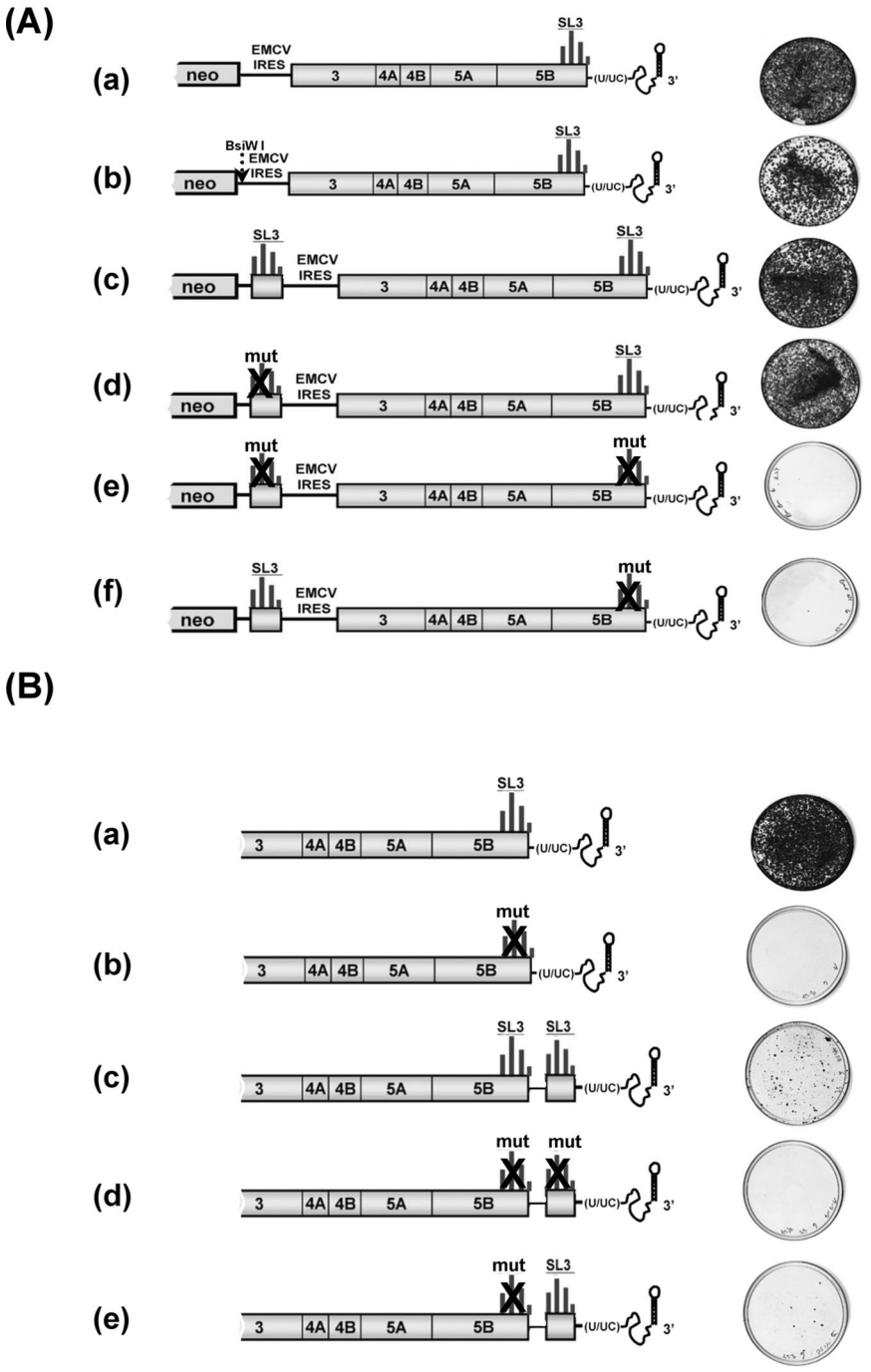


FIG. 9. Attempts to rescue the 5BSL3.2 mutant subgenomic replicon by adding a sequence harboring the cruciform structure. (A) Effects of an additional copy of the cruciform between the Neo^r gene and the EMCV IRES. Schematic diagrams of the constructs tested are shown. The phenotype of these mutant constructs was compared based on G418 transduction efficiency. (a) Parental subgenomic replicon, Con1/SG-Neo (I); (b) Con1/SG-Neo (I) with a *Bsi*WI restriction site inserted after the Neo stop codon; (c) wild-type copy of the cruciform inserted after the Neo^r gene; (d) mutant (mutA) copy of the cruciform on the parental background; (e) mutant (mutA) copy of the cruciform inserted in the context of 5BSL3.2 mutA; (f) wild-type copy of the cruciform inserted in the context of 5BSL3.2 mutA. (B) Effects of inserting an additional copy of the cruciform after the polyprotein stop codon. A schematic diagram and a crystal violet-stained plate indicating G418 transduction efficiency are shown for each construct. (a) Parental replicon; (b) 5BSL3.2 mutB; (c) parental replicon with an additional wild-type cruciform in the 3′ NTR; (d) 5BSL3.2 mutB with an additional mutB cruciform in the 3′ NTR; (e) 5BSL3.2 mutB with an additional wild-type cruciform in the 3′ NTR.

C74 is perfectly conserved in all 208 sequences. The discrepant results noted above for position 92 may indicate that a G at this site must be accompanied by a second sequence change elsewhere in the genome. By modifying our selection conditions, we are attempting to obtain revertants with second-site mutations that restore viability. Such revertants may reveal new RNA-RNA interactions involving bases in the 5BSL3.2 loops.

The upper helix is a scaffold. The importance of the upper helix is strongly supported by three mutants. Mutant C84A/U86G was not viable. This mutant is unable to form the C-G base pair that lies at the bottom of the upper helix in wild-type sequences (Fig. 8). Elimination of this C-G base pair and the change of the loop nucleotide 86 from U to G caused misfolding and disrupted the upper helix, as revealed by RNA structure probing data. Interestingly, mutant C84A/U86A was viable and maintained a similar pattern of cleavage products as wild type, providing a clear indication that preservation of the 5BSL3.2 structure correlates with viability. The contrast between C84A/U86G and C84A/U86A suggested that a G at position 86 is able to alter the upper helix, while an A at position 86 finds no suitable partners within the nucleotides that make up the stem. When the upper helix is intact, as in mutant U86G, the wild-type cleavage pattern and level of replication are preserved. This implies that one essential function of the upper helix may be to position the loop bases for interaction with specific binding partners in the HCV replication complex.

Location dependence of 5BSL3.2. Certain CREs of other viruses, such as human rhinovirus (43), foot-and-mouth disease virus (42), and poliovirus (18), can be translocated and remain functional. If the function of the 5BSL3.2 CRE were similarly independent of its location, it might be possible to rescue mutant replicons by adding a wild-type version of the 5BSL3 cruciform to the mutant replicon. To test this possibility, the wild-type 5BSL3 cruciform domain was inserted (in *cis*) into two different locations in mutant subgenomic replicons. When a wild-type 200-nucleotide-long segment was inserted after the stop codon of the Neo^r gene in the mutant subgenomic replicon, there was no rescue. In addition, there was no significant change in the level of replication in the wild-type replicon when either a wild-type or a mutant copy of the 5BSL3 cruciform domain was inserted into the wild-type replicon at this site (Fig. 9A, constructs c and d). These results suggest this structural element has no impact on the replicon when it is placed out of context near the 5′ terminus of the RNA. However, when the wild-type cruciform domain was inserted into the variable region of the 3′ NTR of the wild-type replicon, a significant reduction in replication was observed (Fig. 9B, construct c). It is not clear whether this reduction was due to the presence of two wild-type CREs competing for interacting partners when placed in close proximity or to the disruption of a higher-order RNA structure important for HCV replication. It is interesting that insertion of an IRES-Neo^r cassette into the 3′ NTR variable region also compromised replicon viability (D. Jacobson, G. Randall, and C. M. Rice, unpublished data). These observations, with two different sequences, suggest that insertions in this region may be poorly tolerated in general. Interestingly, when a wild-type copy of the cruciform was inserted into the 3′ NTR of a replicon with a mutated 5BSL3.2 element, a small number of G418-resistant colonies were produced. While these results are promising and suggest that the CRE may be able to function outside its normal polyprotein context, further work is needed to define the positional constraints and sequence context important for CRE 5BSL3.2 function.

Unlike CREs found in the HCV NTRs, 5BSL3.2 is in the coding region. From the standpoint of genome design, it is interesting to consider the potential advantages and disadvantages of placing CREs within coding sequences. The overlap of a CRE and a coding sequence places a constraint on the evolution of escape mutants if the protein stimulates a significant immune response, and it impedes the evolution of the CRE. However, the overlap offers the advantage of allowing a genome to be more compact, and it may give the CRE a unique ability to regulate translation of the coding gene. Direct communication with the translation machinery is known to be essential for CREs involved in the frameshift required for expression of the Rous sarcoma virus polymerase (27) and for CREs that influence protein conformation, e.g., by slowing the rate of translation (13, 59). A conserved pseudoknot in mRNAs of prion proteins may nucleate protein folding (1). However, the CRE in NS5B of HCV does not seem to act at the level of translation. Cell-free extracts from Huh-7.5 cells revealed no significant differences in the production of viral proteins between wild type and the 5BSL3.2 mutants (data not shown); however, if the CRE influences the conformation of the polymerase or translation termination, these effects would not have been detected in our assay.

Potential role of 5BSL3.2 in HCV replication. Currently, it is not clear how the 5BSL3.2 structure is involved in HCV replication. 5BSL3.2 is just upstream of the polyprotein stop codon, at a site that would appear to provide the CRE with ready access to newly synthesized polymerase. The placement of the CRE in the C-terminal portion of the coding region may simultaneously promote binding to the nascent polymerase and assembly of the minus-strand initiation complex. Such CREs are present in the intergenic region of RNA3 of brome mosaic virus (58) and at the S and M sites of the bacteriophage Qβ genome (44). Cheng et al. reported that the NS5B protein has a strong affinity for the 200 nucleotides upstream of the stop codon that terminates the main ORF, a region that includes 5BSL3.2. The interaction between the NS5B protein and the RNA of the NS5B coding region was strong enough to withstand competition with other HCV RNA regions, such as the 5′ and 3′ NTRs (9).

The life cycles of many RNA viruses require long-range interactions between RNA elements. These interactions can be mediated by base pairing or by proteins. In porcine reproductive and respiratory syndrome virus, a kissing interaction between stem-loop structures in the ORF and the 3′ NTR is required for RNA replication (63). Other examples of long-distance RNA-RNA interactions important for RNA virus replication have been found for flaviviruses (31, 72), flock house virus (39), RNA coliphage Qβ (34), and barley yellow dwarf luteovirus (20). In the case of the HCV 5BSL3.2 CRE, we did not detect any obvious interactions between essential bases in the loops and sequences elsewhere; rather, these regions were susceptible to cleavage by lead(II) and RNase T_1, indicating that they were single stranded under the experimental conditions. It is possible that bases in the 5BSL3.2 structure interact with CREs in other parts of the RNA genome but do so only in the presence of a viral protein(s) or host factor(s) that was not available. A protein-protein bridge between the 5′ cloverleaf and the 3′ poly(A) tail is required for poliovirus minus-strand RNA initiation (4, 22).

In summary, we have used phylogenetic modeling, site-directed mutagenesis, and RNA structure probing to identify a

new CRE, 5BSL3.2, that is important for HCV RNA replication. This CRE lies in the polyprotein coding sequence corresponding to NS5B codons 555 to 571. We are currently attempting to define the step(s) in HCV RNA replication that requires 5BSL3.2 and using genetic and biochemical approaches to identify protein factors and other RNA elements important for 5BSL3.2 function.

ACKNOWLEDGMENTS

We thank James Fan, Jack Heitpas, Hernan Jaramillo, and Merna Torres for expert technical assistance. We are also grateful to many colleagues for helpful discussions during the course of this work and to Brett Lindenbach, Margaret MacDonald, Joseph Marcotrigiano, and Timothy Tellinghuisen for critical reading of the manuscript.

This work was supported by grants from the Public Health Service (CA57973, AI40034, DK052071, and DA16156) and the Greenberg Medical Research Institute.

REFERENCES

1. **Barrette, I., G. Poisson, P. Gendron, and F. Major.** 2001. Pseudoknots in prion protein mRNAs confirmed by comparative sequence analysis and pattern searching. Nucleic Acids Res. **29:**753–758.
2. **Bartenschlager, R., M. Junker-Niepmann, and H. Schaller.** 1990. The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. J. Virol. **64:**5324–5332.
3. **Bartenschlager, R., and V. Lohmann.** 2000. Replication of hepatitis C virus. J. Gen. Virol. **81:**1631–1648.
4. **Barton, D. J., B. J. O'Donnell, and J. B. Flanegan.** 2001. 5′ cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. EMBO J. **20:**1439–1448.
5. **Baumstark, T., and P. Ahlquist.** 2001. The brome mosaic virus RNA3 intergenic replication enhancer folds to mimic a tRNA TpsiC-stem loop and is modified in vivo. RNA **7:**1652–1670.
6. **Blight, K. J., A. A. Kolykhalov, and C. M. Rice.** 2000. Efficient initiation of HCV RNA replication in cell culture. Science **290:**1972–1974.
7. **Blight, K. J., J. A. McKeating, and C. M. Rice.** 2002. Highly permissive cell lines for hepatitis C virus genomic and subgenomic RNA replication. J. Virol. **76:**13001–13014.
8. **Blight, K. J., and C. M. Rice.** 1997. Secondary structure determination of the conserved 98-base sequence at the 3′ terminus of hepatitis C virus genome RNA. J. Virol. **71:**7345–7352.

8a. **Boulant, S., M. Becchi, F. Penin, and J. P. Lavergne.** 2003. Unusual multiple recoding events leading to alternative forms of hepatitis C virus core protein from genotype 1b. J. Biol. Chem. **278:**45785–45792.

9. **Cheng, J. C., M. F. Chang, and S. C. Chang.** 1999. Specific interaction between the hepatitis C virus NS5B RNA polymerase and the 3′ end of the viral RNA. J. Virol. **73:**7044–7049.
10. **Choi, J., Z. Xu, and J. H. Ou.** 2003. Triple decoding of hepatitis C virus RNA by programmed translational frameshifting. Mol. Cell. Biol. **23:**1489–1497.
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. **Choo, Q.-L., K. H. Richman, J. H. Han, K. Berger, C. Lee, C. Dong, C. Gallegos, D. Coit, A. Medina-Selby, P. J. Barr, A. J. Weiner, D. W. Bradley, G. Kuo, and M. Houghton.** 1991. Genetic organization and diversity of the hepatitis C virus. Proc. Natl. Acad. Sci. USA **88:**2451–2455.
13. **Cortazzo, P., C. Cervenansky, M. Marin, C. Reiss, R. Ehrlich, and A. Deana.** 2002. Silent mutations affect in vivo protein folding in Escherichia coli. Biochem. Biophys. Res. Commun. **293:**537–541.
14. **Friebe, P., and R. Bartenschlager.** 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.
15. **Fukushi, S., M. Okada, T. Kageyama, F. B. Hoshino, K. Nagai, and K. Katayama.** 2001. Interaction of poly(rC)-binding protein 2 with the 5′-terminal stem loop of the hepatitis C-virus genome. Virus Res. **73:**67–79.
16. **Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay.** 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. Nature **281:**646–650.
17. **Gerber, K., E. Wimmer, and A. V. Paul.** 2001. Biochemical and genetic studies of the initiation of human rhinovirus 2 RNA replication: purification and enzymatic analysis of the RNA-dependent RNA polymerase 3D(pol). J. Virol. **75:**10969–10978.
18. **Goodfellow, I., Y. Chaudhry, A. Richardson, J. Meredith, J. W. Almond, W. Barclay, and D. J. Evans.** 2000. Identification of a *cis*-acting replication element within the poliovirus coding region. J. Virol. **74:**4590–4600.
19. **Goodfellow, I. G., D. Kerrigan, and D. J. Evans.** 2003. Structure and function analysis of the poliovirus cis-acting replication element (CRE). RNA **9:**124–137.
20. **Guo, L., E. M. Allen, and W. A. Miller.** 2001. Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. Mol. Cell **7:**1103–1109.
21. **Hahm, B., Y. K. Kim, J. H. Kim, T. Y. Kim, and S. K. Jang.** 1998. Heterogeneous nuclear ribonucleoprotein L interacts with the 3′ border of the internal ribosomal entry site of hepatitis C virus. J. Virol. **72:**8782–8788.
22. **Herold, J., and R. Andino.** 2001. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. Mol. Cell **7:**581–591.
23. **Hofacker, I. L., M. Fekete, C. Flamm, M. A. Huynen, S. Rauscher, P. E. Stolorz, and P. F. Stadler.** 1998. Automatic detection of conserved RNA structure elements in complete RNA virus genomes. Nucleic Acids Res. **26:**3825–3836.
24. **Honda, M., E. A. Brown, and S. M. Lemon.** 1996. Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. RNA **2:**955–968.
25. **Honda, M., L. H. Ping, R. C. Rijnbrand, E. Amphlett, B. Clarke, D. Rowlands, and S. M. Lemon.** 1996. Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA. Virology **222:**31–42.
26. **Isoyama, T., N. Kamoshita, K. Yasui, A. Iwai, K. Shiroki, H. Toyoda, A. Yamada, Y. Takasaki, and A. Nomoto.** 1999. Lower concentration of La protein required for internal ribosome entry on hepatitis C virus RNA than on poliovirus RNA. J. Gen. Virol. **80:**2319–2327.
27. **Jacks, T., and H. E. Varmus.** 1985. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science **230:**1237–1242.
28. **James, B. D., G. J. Olsen, and N. R. Pace.** 1989. Phylogenetic comparative analysis of RNA secondary structure. Methods Enzymol. **180:**227–239.
29. **Joost Haasnoot, P. C., R. C. Olsthoorn, and J. F. Bol.** 2002. The Brome mosaic virus subgenomic promoter hairpin is structurally similar to the iron-responsive element and functionally equivalent to the minus-strand core promoter stem-loop C. RNA **8:**110–122.
30. **Junker-Niepmann, M., R. Bartenschlager, and H. Schaller.** 1990. A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. EMBO J. **9:**3389–3396.
31. **Khromykh, A. A., P. L. Sedlak, and E. G. Westaway.** 2000. *cis*- and *trans*-acting elements in flavivirus RNA replication. J. Virol. **74:**3253–3263.
32. **Kieft, J. S., K. Zhou, A. Grech, R. Jubin, and J. A. Doudna.** 2002. Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. Nat. Struct. Biol. **9:**370–374.
33. **Kieft, J. S., K. Zhou, R. Jubin, and J. A. Doudna.** 2001. Mechanism of ribosome recruitment by hepatitis C IRES RNA. RNA **7:**194–206.
34. **Klovins, J., V. Berzins, and J. van Duin.** 1998. A long-range interaction in Qβ RNA that bridges the thousand nucleotides between the M-site and the 3′ end is required for replication. RNA **4:**948–957.
35. **Knaus, T., and M. Nassal.** 1993. The encapsidation signal on the hepatitis B virus RNA pregenome forms a stem-loop structure that is critical for its function. Nucleic Acids Res. **21:**3967–3975.
36. **Kolykhalov, A. A., S. M. Feinstone, and C. M. Rice.** 1996. Identification of a highly conserved sequence element at the 3′ terminus of hepatitis C virus genome RNA. J. Virol. **70:**3363–3371.
37. **Lesburg, C. A., M. B. Cable, E. Ferrari, Z. Hong, A. F. Mannarino, and P. C. Weber.** 1999. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat. Struct. Biol. **6:**937–943.
38. **Lindell, M., P. Romby, and E. G. Wagner.** 2002. Lead(II) as a probe for investigating RNA structure in vivo. RNA **8:**534–541.
39. **Lindenbach, B. D., J. Y. Sgro, and P. Ahlquist.** 2002. Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. J. Virol. **76:**3905–3919.
40. **Lobert, P. E., N. Escriou, J. Ruelle, and T. Michiels.** 1999. A coding RNA sequence acts as a replication signal in cardioviruses. Proc. Natl. Acad. Sci. USA **96:**11560–11565.
41. **Lohmann, V., F. Korner, J. O. Koch, U. Herian, L. Theilmann, and R. Bartenschlager.** 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science **285:**110–113.
42. **Mason, P. W., S. V. Bezborodova, and T. M. Henry.** 2002. Identification and characterization of a *cis*-acting replication element (CRE) adjacent to the internal ribosome entry site of foot-and-mouth disease virus. J. Virol. **76:**9686–9694.
43. **McKnight, K. L., and S. M. Lemon.** 1998. The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. RNA **4:**1569–1584.
44. **Miranda, G., D. Schuppli, I. Barrera, C. Hausherr, J. M. Sogo, and H. Weber.** 1997. Recognition of bacteriophage Qβ plus strand RNA as a template by Qβ replicase: role of RNA interactions mediated by ribosomal proteins S1 and host factor. J. Mol. Biol. **267:**1089–1103.
45. **Morasco, B. J., N. Sharma, J. Parilla, and J. B. Flanegan.** 2003. Poliovirus cre(2C)-dependent synthesis of VPgpUpU is required for positive- but not negative-strand RNA synthesis. J. Virol. **77:**5136–5144.
46. **Pudi, R., S. Abhiman, N. Srinivasan, and S. Das.** 2003. Hepatitis C virus internal ribosome entry site-mediated translation is stimulated by specific

interaction of independent regions of human La autoantigen. J. Biol. Chem. **278:**12231–12240.
47. **Reynolds, J. E., A. Kaminski, H. J. Kettinen, K. Grace, B. E. Clarke, A. R. Carroll, D. J. Rowlands, and R. J. Jackson.** 1995. Unique features of internal initiation of hepatitis C virus RNA translation. EMBO J. **14:**6010–6020.
48. **Rice, C. M., E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, and J. H. Strauss.** 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science **229:**726–733.
49. **Rieger, A., and M. Nassal.** 1995. Distinct requirements for primary sequence in the 5′- and 3′-part of a bulge in the hepatitis B virus RNA encapsidation signal revealed by a combined in vivo selection/in vitro amplification system. Nucleic Acids Res. **23:**3909–3915.
50. **Rijnbrand, R., P. J. Bredenbeek, P. C. Haasnoot, J. S. Kieft, W. J. M. Spaan, and S. M. Lemon.** 2001. The influence of downstream protein-coding sequence on internal ribosome entry on hepatitis C virus and other flavivirus RNAs. RNA **7:**585–597.
51. **Rijnbrand, R. C. A., and S. M. Lemon.** 2000. Internal ribosome entry site-mediated translation in hepatitis C virus replication, p. 85–116. *In* C. Hagedorn and C. M. Rice (ed.), Hepatitis C virus, vol. 242. Springer-Verlag, Berlin, Germany.
52. **Sanger, F., G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, C. A. Fiddes, C. A. Hutchison, P. M. Slocombe, and M. Smith.** 1977. Nucleotide sequence of bacteriophage phi X174 DNA. Nature **265:**687–695.
52a. **Schmidt-Mende, J., E. Bieck, T. Hugle, F. Penin, C. M. Rice, H. E. Blum, and D. Moradpour.** 2001. Determinants for membrane association of the hepatitis C virus RNA-dependent RNA polymerase. J. Biol. Chem. **276:**44052–44063.
53. **Schuppli, D., G. Miranda, S. Qiu, and H. Weber.** 1998. A branched stem-loop structure in the M-site of bacteriophage Qβ RNA is important for template recognition by Qβ replicase holoenzyme. J. Mol. Biol. **283:**585–593.
54. **Simmonds, P.** 1997. Clinical relevance of hepatitis C virus genotypes. Gut **40:**291–293.
55. **Simmonds, P., J. Mellor, T. Sakuldamrongpanich, C. Nuchaprayoon, S. Tanprasert, E. C. Holmes, and D. B. Smith.** 1996. Evolutionary analysis of variants of hepatitis C virus found in South-East Asia: comparison with classifications based upon sequence similarity. J. Gen. Virol. **77:**3013–3024.
56. **Smith, D. B., and P. Simmonds.** 1997. Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both ends of the genome. J. Mol. Evol. **45:**238–246.
57. **Spangberg, K., L. Wiklund, and S. Schwartz.** 2001. Binding of the La autoantigen to the hepatitis C virus 3′ untranslated region protects the RNA from rapid degradation in vitro. J. Gen. Virol. **82:**113–120.
58. **Sullivan, M. L., and P. Ahlquist.** 1999. A brome mosaic virus intergenic RNA3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA in vivo. J. Virol. **73:**2622–2632.
59. **Thanaraj, T. A., and P. Argos.** 1996. Protein secondary structural types are differentially coded on messenger RNA. Protein Sci. **5:**1973–1983.
60. **Tokita, H., H. Okamoto, H. Iizuka, J. Kishimoto, F. Tsuda, L. A. Lesmana, Y. Miyakawa, and M. Mayumi.** 1996. Hepatitis C virus variants from Jakarta, Indonesia classifiable into novel genotypes in the second (2e and 2f), tenth (10a) and eleventh (11a) genetic groups. J. Gen. Virol. **77:**293–301.
61. **Tuplin, A., J. Wood, D. J. Evans, A. H. Patel, and P. Simmonds.** 2002. Thermodynamic and phylogenetic prediction of RNA secondary structures in the coding region of hepatitis C virus. RNA **8:**824–841.
62. **Varaklioti, A., N. Vassilaki, U. Georgopoulou, and P. Mavromara.** 2002. Alternate translation occurs within the core coding region of the hepatitis C viral genome. J. Biol. Chem. **277:**17713–17721.
63. **Verheije, M. H., R. C. Olsthoorn, M. V. Kroese, P. J. Rottier, and J. J. Meulenberg.** 2002. Kissing interaction between 3′ noncoding and coding sequences is essential for porcine arterivirus RNA replication. J. Virol. **76:**1521–1526.
64. **Walewski, J. L., J. A. Gutierrez, W. Branch-Elliman, D. D. Stump, T. R. Keller, A. Rodriguez, G. Benson, and A. D. Branch.** 2002. Mutation Master: profiles of substitutions in hepatitis C virus RNA of the core, alternate reading frame, and NS2 coding regions. RNA **8:**557–571.
65. **Walewski, J. L., T. R. Keller, D. D. Stump, and A. D. Branch.** 2001. Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. RNA **7:**710–721.
66. **Wang, T. H., R. C. Rijnbrand, and S. M. Lemon.** 2000. Core protein-coding sequence, but not core protein, modulates the efficiency of cap-independent translation directed by the internal ribosome entry site of hepatitis C virus. J. Virol. **74:**11347–11358.
67. **Wood, J., R. M. Frederickson, S. Fields, and A. H. Patel.** 2001. Hepatitis C virus 3′X region interacts with human ribosomal proteins. J. Virol. **75:**1348–1358.
68. **Xu, Z., J. Choi, T. S. Yen, W. Lu, A. Strohecker, S. Govindarajan, D. Chien, M. J. Selby, and J. Ou.** 2001. Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. EMBO J. **20:**3840–3848.
69. **Yanagi, M., M. St. Claire, S. U. Emerson, R. H. Purcell, and J. Bukh.** 1999. In vivo analysis of the 3′ untranslated region of the hepatitis C virus after in vitro mutagenesis of an infectious cDNA clone. Proc. Natl. Acad. Sci. USA **96:**2291–2295.
70. **Yi, M., and S. M. Lemon.** 2003. 3′ nontranslated RNA signals required for replication of hepatitis C virus RNA. J. Virol. **77:**3557–3568.
71. **Yin, J., A. V. Paul, E. Wimmer, and E. Rieder.** 2003. Functional dissection of a poliovirus *cis*-acting replication element [PV-*cre*(2C)]: analysis of single- and dual-*cre* viral genomes and proteins that bind specifically to PV-*cre* RNA. J. Virol. **77:**5152–5166.
72. **You, S., and R. Padmanabhan.** 1999. A novel in vitro replication system for Dengue virus. Initiation of RNA synthesis at the 3′-end of exogenous viral RNA templates requires 5′- and 3′-terminal complementary sequence motifs of the viral RNA. J. Biol. Chem. **274:**33714–33722.