Determination of the Secondary Structure of and Cellular Protein Binding to the 3'-Untranslated Region of the Hepatitis C Virus RNA Genome

TAKAYOSHI ITO¹ AND MICHAEL M. C. LAI^{1,2*}

Howard Hughes Medical Institute¹ and Department of Molecular Microbiology and Immunology,²
University of Southern California School of Medicine, Los Angeles, California 90033-1054

Received 5 May 1997/Accepted 14 August 1997

Hepatitis C virus (HCV) contains a positive-stranded RNA genome of approximately 9.5 kb. Despite the overall sequence diversity among individual HCV isolates, the 3'-end 98 nucleotides (nt) of the HCV RNA, which constitute part of the 3'-untranslated region (3'-UTR), are highly conserved. This conserved region may contain the *cis*-acting signals for RNA replication involving possibly both viral and cellular proteins. We carried out RNase digestion studies, which revealed that this 98-nt region contains three stem-loops but may also assume alternative structures. We further performed UV cross-linking experiments to detect cellular proteins that bound to this region. A 58-kDa cellular protein (p58) was detected. Its binding site was mapped to the stem-loops 2 and 3, which are the most conserved region of the 3'-UTR. Site-directed mutagenesis studies revealed that both stem structures and specific nucleotide sequence within the two loops are important for p58 binding. Mutations that disrupted stem structures abolished protein binding, while the compensatory mutations restored its binding. This region also contains partial sequence similarity to the reported consensus binding sequence for polypyrimidine tract-binding protein (PTB) (a 57-kDa protein). The UV-cross-linked protein could be immunoprecipitated with the anti-PTB antibody, and the recombinant PTB bound to the HCV 3'-UTR with the same binding specificity as p58, establishing that this protein is PTB. However, the reported PTB-binding sequence was not sufficient, but rather the entire stem-loops 2 and 3 were required, for PTB binding; thus, its binding specificity is significantly different from the reported PTB-binding sequence requirement. This protein was detected in both the nuclei and cytoplasm of most mammalian cell lines tested and human primary hepatocytes. PTB may participate in the regulation of HCV RNA synthesis or translation.

Hepatitis C virus (HCV) is now recognized as the principal agent of parenterally transmitted non-A, non-B hepatitis. The viral genome has been cloned and sequenced (5) and shown to be a 9.4-kb single-stranded, positive-sense RNA, encoding a large polyprotein of about 3,010 to 3,033 amino acids (5, 17). The polyprotein is cleaved at multiple sites to produce at least 10 structural and nonstructural proteins (21), with structural protein genes being located at the 5' side, preceding those encoding the nonstructural proteins. The genomic organization and putative amino acid sequence of the polyprotein of HCV have led to its categorization as a new genus of the *Flaviviridae* family (17, 24). Although the overall sequence of HCV RNA shows significant diversity within the coding region among various isolates (see reference 3 for a review), the 5'-untranslated region (5'-UTR) and the 3'-end 98 nucleotides (nt) of the HCV RNA are highly conserved (2, 11, 18, 32, 38). Conceivably, both the 5'- and 3'-end sequences of HCV RNA are important for viral RNA replication and/or translation.

Higher-order RNA structures in the 3'-UTR of positive-strand RNA viruses are generally thought to play a role in the initiation of negative-strand RNA synthesis. For example, a pseudoknot structure at the 3' terminus of the poliovirus RNA is an important element for viral RNA replication (15), and a

kissing interaction between the two predominant hairpin loops in the picornavirus 3'-UTR may also be important for RNA replication (23, 26). The 3'-UTRs of flaviviruses, such as yellow fever virus, also form stem-loop (SL) structures, which likely function as *cis*-acting signals for the initiation of viral RNA synthesis (28). By computer analysis, the HCV 3'-UTR has been suggested to assume several SL structures (18, 32, 38); however, their existence has yet to be confirmed. The HCV 3'-UTR consists of three domains: a highly variable sequence of 21 to 39 nt, followed by a UC-rich sequence of variable length (73 to 98 nt) (38) and, at the very 3' end, a 98-nt highly conserved sequence (18, 32). The sequence and probable structural conservation of the very 3'-end 98 nt suggests their functional importance, most likely in viral RNA synthesis or translation.

The regulation of viral RNA synthesis or translation by either 5'- or 3'-end sequences likely involves not only RNA structure but also viral and cellular proteins. In this study, we set out to examine whether the 3'-end 98 nt of HCV RNA interact with cellular proteins. A 58-kDa cellular protein was shown to bind to this region specifically and was identified to be polypyrimidine tract-binding protein (PTB). This protein was detected in both the nuclei and cytoplasm of most mammalian cell lines and human primary hepatocytes. The protein binding requires two stem structures and specific sequences in the loop region of the 98 nt RNA, thus confirming the existence of the higher-order structure within the 3'-UTR of HCV RNA. This binding specificity is significantly different from that for the reported PTB-binding consensus sequence. The PTB

* Corresponding author. Mailing address: Department of Molecular Microbiology and Immunology, University of Southern California School of Medicine, 2011 Zonal Ave., HMR-503, Los Angeles, CA 90033-1054. Phone: (213) 342-1748. Fax: (213) 342-9555. E-mail: michlai@hsc.usc.edu.

FIG. 1. Schematic diagrams of 3'-UTR transcripts used in this study. The cDNA of 98 nt at the 3' end or the entire 3'-UTR of HCV RNA containing the T7 RNA polymerase promoter (large open arrow) was cloned into the *EcoRI* site of pCR2.1 vector (Invitrogen) with the TA cloning system. The plasmids were linearized with *EcoRI* and transcribed with T7 RNA polymerase to generate HCV-X(+), HCV-X(-), or HCV-3'(+)-Full. HCV-3'(+)-Full contains the entire 3'-UTR including the X region (open box) and its upstream variable sequence (hatched box) in the 3'-UTR of HCV RNA plus 13 nt of poly(U) (black box). All of the other RNAs were transcribed directly from the PCR products amplified from these two plasmids, using site-specific primers containing the T7 promoter. Boxed region, HCV sequence; line, vector sequence.

binding to the 3'-UTR may play a role in the regulation of HCV RNA synthesis.

MATERIALS AND METHODS

Construction of cDNA clones. The scheme for construction of HCV-X(+), the plasmid used for in vitro synthesis of the 3'-end 98 nt of the HCV RNA, is shown in Fig. 1. The 3'-end 98 nt (X region) of HCV sequence reported previously (18, 32) were amplified by PCR using 5' 60-nt and 3' 60-nt complementary sequences of this region as primers and cloned into pCR2.1 vector (Invitrogen) by the TA cloning method (39). The resulting pCR-HCV-X(+) clone was confirmed by sequencing (29). HCV-X(+), which contains the full-length cDNA sequence (except the 3'-end 98 nt) of a genotype 1b HCV RNA plus an additional 13 nt of poly(T) sequence at the 3' end. HCV-X(-), which generates the antisense HCV RNA in the X region, was constructed by PCR amplification of pCR-HCV-X(+), using a reverse primer (5'-TAATACGACTACTATAACATGATCTGCAGAGAGG), which contains the T7 promoter sequence (boldfaced) and the sequence complementary to the 3' end of the X region. This PCR product was gel purified and directly used as a template for in vitro transcription. For constructs transcribing various truncated HCV RNAs in the X region (Fig. 1), PCR amplification was performed on pCR-HCV-X(+) plasmid DNA, using site-specific primers containing a T7 promoter. The PCR products were gel purified and directly used as templates for in vitro transcription. For RNAs containing site-directed mutations in the X region, the mutant cDNAs were amplified from HCV-X(+), by PCR with site-specific primers containing the desired substitutions and subcloned into

pCR2.1 vector. Various mutant RNAs were generated from the cDNA clones by in vitro transcription as described below.

In vitro RNA transcription. To synthesize RNA in vitro, HCV-X(+) and HCV-3'(+)-Full plasmid DNAs were linearized with *EcoRI* and subjected to in vitro transcription using T7 RNA polymerase (Promega). To generate truncated RNAs, gel-purified PCR products [HCV-X(-); HCV-X(+) 5'-78, -53, -42, -28, and -20, 3'-46, -72, and -77, and Δ31-40; and HCV-3'(+)-Full-5'-28] as described above were used directly for in vitro transcription without cloning.

In vitro transcription was carried out as recommended by the manufacturer (Promega). ³²P-labeled RNAs were obtained by transcription in the presence of 50 μCi of [α-³²P]UTP (3,000 Ci/mmol [ICN Biomedicals]), 12.5 μM UTP, and 500 μM each ATP, GTP, and CTP. Unlabeled RNAs were transcribed in the presence of 500 μM each ATP, GTP, CTP, and UTP. After transcription at 37°C for 2 h, 1 U of RQ DNase I (Promega) was added to the reaction mixture to digest DNA templates, and the transcribed RNAs were purified with a Sephadex G-25 column equilibrated in STE buffer (10 mM Tris-HCl, 1 mM EDTA, 0.1 mM NaCl [pH 7.5]) to remove unincorporated ribonucleotides and deoxyribonucleotides from DNA digestion. The radioactivity of the transcription product was determined with a liquid scintillation counter (LC6000 IC; Beckman).

Preparation of cell extracts. Cells from different cultured cell lines (HeLa, DBT, COS7, HepG2, and Molt4) and from mouse and human primary hepatocytes were harvested and washed twice in cold phosphate-buffered saline. Human primary hepatocytes were obtained from a surgically resected liver of an HCV-free patient at the USC Medical Center. The liver tissues used in this study had been approved by our Institutional Review Board. Both nuclear and cytoplasmic extracts were prepared as previously described (7, 9), with slight modification. Briefly, 5 × 10⁸ to 1 × 10⁹ cells were resuspended in 5 packed-cell volumes of buffer A (10 mM HEPES [pH 7.9], 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM dithiothreitol [DTT]) and incubated for 10 min on ice. After centrifugation for 10 min at 2,000 rpm in a 224 swing bucket rotor (International Equipment Company), cells were resuspended in 2 packed-cell volumes of buffer A and lysed by 15 to 20 strokes in a glass Dounce homogenizer (type b pestle). The homogenate was centrifuged for 10 min at 2,000 rpm in a 224 swing bucket rotor; the pellet was used for nuclear extract preparation, and the supernatant was used for cytoplasmic extract preparation.

For nuclear extract preparation, the pellet was resuspended in 0.67 volume of buffer C (20 mM HEPES [pH 7.9], 25% [vol/vol] glycerol, 0.42 M NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride, 0.5 mM DTT) and incubated on ice for 30 min before centrifugation at 13,000 rpm for 30 min. The supernatant was collected and dialyzed against UV cross-linking binding buffer (5 mM HEPES [pH 7.8], 25 mM KCl, 2 mM MgCl₂, 3.8% glycerol, 0.1 mM EDTA, 2 mM DTT). Finally, the sample was centrifuged at 13,000 rpm for 20 min, the supernatant was collected, and aliquots were frozen at -80°C for subsequent experiments.

For the preparation of cytoplasmic extract, the supernatant was mixed with 0.11 volume of buffer B (0.3 M HEPES [pH 7.9], 1.4 M KCl, 0.03 M MgCl₂) and centrifuged at 40,000 rpm (100,000 × g) in a Beckman 50 Ti ultracentrifuge rotor for 60 min at 4°C. The supernatant was collected and dialyzed against UV cross-linking binding buffer, and aliquots were frozen as described above. Protein concentrations were determined with a Bio-Rad protein assay kit.

UV cross-linking of RNA-protein complexes. UV cross-linking of RNA to protein was performed as previously described (9), with slight modification. Briefly, different amounts of proteins extracted from either cytoplasm or nuclei of various cell lines were first mixed with 20 μg of yeast tRNA in a final volume of 20 μl of binding buffer (5 mM HEPES [pH 7.8], 25 mM KCl, 2 mM MgCl₂, 3.8% glycerol, 0.1 mM EDTA, 2 mM DTT). In competition assays, different amounts of unlabeled RNA were mixed with a fixed amount of cellular extracts in the binding buffer. The amounts of the unlabeled competitor RNAs were predetermined with a spectrophotometer. Following an incubation at 30°C for 10 min, ³²P-labeled RNA (10⁶ cpm; 30 ng of RNA) was added, and the reaction mixture was incubated for an additional 10 min. Subsequently, the reaction mixture was irradiated with UV light (254 nm) (UV Stratalation 2400; Stratagene) on ice at a distance of 13 cm for 10 min. After irradiation, RNase A was added to a final concentration of 2 mg/ml, and the mixture was incubated for 30 min at 37°C, boiled for 3 min in the sample buffer (10% glycerol, 200 mM DTT, 3% sodium dodecyl sulfate [SDS], 62.5 mM Tris-HCl [pH 6.8]), and electrophoresed on a 10% polyacrylamide gel containing 0.1% SDS. After drying, the gel was exposed to X-ray film.

Thermodynamic calculation of RNA secondary structure. The optimal and suboptimal secondary structures for wild-type and truncated RNA molecules were predicted by the method of Zuker and Stiegler (40) with the Mfold program in the Genetics Computer Group sequence analysis software package.

Enzymatic probing of RNA secondary structure. The RNA products transcribed from HCV-X(+) were treated with RNase-free DNase I (RQ DNase I; Promega), extracted with phenol-chloroform and recovered by precipitation with ethanol. The purified RNA transcripts were resuspended in water, heated for 2 min at 75°C, slowly cooled to 42°C, and placed on ice. The transcripts were then partially digested with nuclease S1 (Promega), RNase T₁ (Pharmacia), or cobra venom nuclease V1 (V1; Pharmacia) as described elsewhere (15, 37). For S1 digestion, various amounts of nuclease S1 were mixed with RNA in 4 μl of reaction mixture containing 280 mM NaCl, 50 mM sodium acetate (pH 4.5), and 4.5 mM zinc sulfate and incubated at 37°C for 5 min. For RNase T₁ or V1

FIG. 2. UV cross-linking analysis of HeLa cell extracts. (A) UV cross-linking of proteins from cytoplasmic and nuclear extracts of HeLa cells to positive-strand RNA of the 3'-end 98 nt of HCV RNA. Various amounts (2, 5, or 20 μ g) of proteins were mixed with 32 P-labeled HCV-X(+) RNA, UV irradiated, digested with RNase A, and separated by SDS-PAGE on 10% polyacrylamide gels. (-), no cellular extract. Arrows indicate the 115-, 58-, 49-, and 39-kDa proteins. Molecular mass markers are indicated on the left. (B) UV cross-linking using HCV-3'(+)/Full (left) or HCV-X(-) (right) RNA. Computer images were generated by Adobe Photoshop, version 3.0.

digestion, various amounts of the enzymes were mixed with RNA in 4 μ l of TMK buffer (30 mM Tris-HCl [pH 7.4], 10 mM $MgCl_2$, 270 mM KCl) and incubated at 37°C for 10 min. After incubation, the RNA was purified by extraction with phenol-chloroform and recovered by precipitation with ethanol. The cleavage sites were detected by primer extension using a primer complementary to the 3' end of the RNA. For this purpose, the primer oligonucleotide was end labeled with [γ - 32 P]ATP, using phage T4 polynucleotide kinase. The partially digested RNA was dissolved in 4 μ l of annealing buffer (40 mM Tris-HCl [pH 7.5], 25 mM $MgCl_2$, 50 mM NaCl), mixed with 1 μ l of 32 P-labeled primer, and heated at 75°C for 3 min. After the reaction mixture was cooled to 42°C, 1 U of avian myeloblastosis virus reverse transcriptase (RT) in 10 μ l of RT buffer (50 mM Tris-HCl [pH 7.5], 75 mM KCl, 10 mM DTT, 100 μ g of bovine serum albumin/ml, 0.5 mM deoxynucleoside triphosphates) was added, and the mixture was incubated at 37°C for 15 min. Reaction products were mixed with 15 μ l of formamide sample buffer (95% formamide, 20 mM EDTA), boiled for 2 min, and separated by electrophoresis in 6% polyacrylamide gels containing 8 M urea.

Immunoprecipitation of p58-RNA complex with a monoclonal anti-PTB antibody. After UV cross-linking of HeLa cytoplasmic extracts to HCV-X(+) RNA and RNase treatment as described above, the sample was diluted to 500 μ l with NETS buffer (50 mM Tris-HCl [pH 7.4], 5 mM EDTA, 1 mM DTT, 100 mM NaCl, 0.05% Nonidet P-40) and mixed with monoclonal anti-PTB antibody (kindly provided by E. Wimmer, State University of New York, Stony Brook). The immunocomplexes were immobilized on protein A-Sepharose 4B beads saturated with anti-mouse immunoglobulin G antibodies (Sigma). The unbound materials were washed five times with the same buffer. The bound proteins were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) followed by autoradiography. Parallel reaction mixtures were made with unrelated antibodies, a monoclonal anti-TATA-binding protein (TBP) antibody (TFIID; Calbiochem) and a monoclonal anti-FLAG M2 antibody (Kodak), and served as controls.

Cloning of PTB cDNA and expression of recombinant PTB proteins. cDNA clones encoding PTB were obtained by RT-PCR amplification of mRNA from HeLa cells, using two primers corresponding to the 5' and 3' ends of PTB-coding sequence (25) plus *Bam*HI sites. The cDNA was subcloned into the pCR2.1 vector (Invitrogen) by the TA cloning method (39). The resulting pCR-PTB clone was confirmed by sequencing (29). To obtain a cDNA coding for a glutathione *S*-transferase (GST)-PTB fusion protein, pCR-PTB was digested by *Bam*HI and subcloned into the *Bam*HI site of pGEX-4T-1 vector (Pharmacia Biotech). The GST-PTB fusion protein was induced by isopropylthiogalactopyranoside in *Escherichia coli* according to a published method (8), purified by binding to glutathione-Sepharose 4B beads (Pharmacia), and recovered by elution with 5 mM reduced glutathione.

RESULTS

A cellular protein binds to the 3'-end 98 nt of HCV RNA. To analyze the possible cellular factors modulating the functions of the 3'-UTR of HCV RNA, we performed UV cross-linking

between the 32 P-labeled 3'-end 98 nt of HCV RNA [HCV-X(+)] (Fig. 1) and both cytoplasmic and nuclear extracts of HeLa cells. After RNase digestion, the cross-linked proteins were resolved by SDS-PAGE. The results showed that a major protein of 58 kDa (p58) and several minor proteins (p115, p49, and p39) were detected. These proteins were detected in both the nuclear and cytoplasmic extracts (Fig. 2A). Since HCV RNA replication probably takes place in the cytoplasm of infected cells, as inferred from studies of flavivirus replication (4), we first concentrated on the cytoplasmic extracts. When negative-strand RNA of this region [HCV-X(-)] (Fig. 1) was used for UV cross-linking, no protein band was detected even after prolonged exposure (Fig. 2B), suggesting that the binding was specific for the viral genomic-sense RNA. The same set of proteins was detected when the entire 3'-UTR of HCV RNA [HCV-3'(+)/Full] (Fig. 1), including the variable sequence, the UC-rich sequence, and X region, was used (Fig. 2B). Since these RNAs contained six bases of vector sequence at the 3' end, we also generated PCR products amplified from HCV-X(+) and HCV-3'(+)/Full plasmids, using site-specific primers for in vitro transcription to synthesize the respective HCV RNAs without any vector sequences. These vector-free RNAs also bound identical proteins, indicating that these proteins bound to HCV sequence (data not shown).

To further determine the binding specificity of these proteins, we used an unlabeled homologous RNA [HCV-X(+)] and various unrelated RNAs, including the 3'-end 90 nt of positive-sense mouse hepatitis virus (MHV) RNA and the 3'-end 182 nt of negative-sense MHV RNA (9), to perform competition analysis. The unlabeled competitor RNA was added to the binding reactions at 30°C for 10 min prior to the addition of the radiolabeled probe. When increasing amounts of the unlabeled homologous RNA (1- to 50-fold molar excess) were preincubated with a fixed amount of cytoplasmic extracts, only p58 binding was inhibited (Fig. 3, lanes 2 to 4). None of the other minor proteins (p115, p49, and p39) were affected at all. In contrast to the homologous RNA, neither of the MHV RNAs affected the amount of p58 binding (Fig. 3, lanes 5 to 10). We conclude that the p58 specifically bound to the posi-

FIG. 3. Competition analysis of UV cross-linked proteins. The unlabeled competitor RNAs used are indicated above each lane. HCV-X(+), MHV 3'-end 90 nt of (+) strand, and MHV 3'-end 182 nt of negative-strand RNA were transcribed from HCV-X(+), pNH2, and pNX1 (9), respectively. A fixed amount of the radiolabeled HCV-X(+) RNA (2×10^6 cpm; 60 ng of RNA) was UV cross-linked with a fixed amount of cytoplasmic extracts (60 μ g). Lanes: (-), no competitor; 2 to 4, homologous competition assay with various amounts of unlabeled RNAs (1-, 5-, and 50-fold molar excesses over the radiolabeled probe, respectively); 5 to 7 and 8 to 10, heterologous competition assays with various amounts of either the 3'(+)- or 3'(-) end of unlabeled MHV RNA (1-, 5-, and 50-fold molar excesses over the radiolabeled probe, respectively). The arrow indicates the 58-kDa protein. Computer images were generated by Adobe Photoshop, version 3.0.

tive-sense 3'-end 98 nt of HCV RNA. Most of the other minor protein bands detected were probably nonspecific RNA-binding proteins.

Mapping of the protein-binding site on the 3' end of HCV RNA. To dissect the sequence requirement for p58 binding, we used a series of truncations of the X region (Fig. 1) for UV cross-linking studies. The results showed that deletions from the 3' end did not affect p58 binding until more than 56 nt from the 3' end had been deleted (Fig. 4A); the 5'-42 RNA still

FIG. 4. Mapping of the protein-binding site on the 3'-UTR of HCV RNA by using truncation mutant RNAs. Increasing amounts of HeLa cytoplasmic extracts (0, 2, and 20 μ g) were incubated with various 32 P-labeled HCV-X(+) RNAs of different sizes (A) or HCV-3'-(+)Full RNAs of different sizes (B) and UV irradiated as for Fig. 2. Only the 58-kDa protein band of the SDS-polyacrylamide gels is shown. Computer images were generated by Adobe Photoshop, version 3.0. WT, wild type.

retained a significant protein-binding activity. Any additional deletions from the 5' end [X(+) 5'-28 and 5'-20] almost completely abolished p58 binding, although 5'-28 retained a very low level of the binding activity. Therefore, the 5'-end 42 nt of the X region contains the minimum binding domain for p58 (Fig. 4A). In contrast, none of the 5'-end-truncated RNAs bound p58. Particularly, deletions of as few as 21 nt from the 5' end [X(+) 3'-77] completely abolished the binding, indicating that the first 21 nt of the X region contain the critical sequences required for p58 binding. However, the binding of the 5'-end 42 nt was weaker than that of the 5'-end 53 nt (Fig. 4A). Furthermore, deletion of nt 31 to 40 almost completely abolished p58 binding. Therefore, we conclude that the 5'-end 21 nt of the X region contain the primary binding site for p58, but internal sequences (nt 30 to 50) also are critical for the protein binding.

To assess whether the UC-rich sequence and the upstream variable sequence contribute to p58 binding, we examined the complete 3'-UTR RNA and a corresponding RNA with a deletion of 70 nt at the 3' end [3'(+)-Full and 3'(+)-Full-5'-28] (Fig. 1). Figure 4B shows that the binding patterns of the full-length 3'-UTR and its deletion mutant were similar to those of the X region alone, with only a marginal increase of binding activity, indicating that the binding sequence resides almost completely in the X region, particularly the 5'-end 42 nt.

Determination of the secondary structure of the 3' end of HCV RNA. Computer analysis of the X region based on the Zuker program (40) predicted the presence of three SL structures, including SL3 within the first 20 nt from the 5' end and SL2 within nt 23 to 49 (Fig. 5A) (18). The inclusion of the variable 3'-UTR sequence and UC-rich region in the complete 3'-UTR did not alter the structure of these predicted SLs in the X region (Fig. 5B). The deletion analysis described above suggested that both SL3 and SL2 are required for p58 binding. It should be noted that although the three-SL structure as shown in Fig. 5 (18) represents the predicted optimal structure, several alternative suboptimal models have ΔG values very close to that of the optimal structure (data not shown). In all of these models, SL1 is invariable, whereas SL2 and SL3, which are included in the binding region of p58, can assume several alternative structures.

To determine whether p58 binding involves these putative secondary structures, we first probed the actual structure of the X region of HCV RNA by partial RNase digestion. Unlabeled 98-nt RNA of the X region was incubated with various concentrations of nuclease S1 and RNase T₁, both of which digest single-stranded RNA, or V1, which digests double-stranded RNA, and analyzed by primer extension using a 32 P-end-labeled primer complementary to the 3' end of the RNA. Figure 6 shows that the RNase digestion pattern was consistent with the presence of SL2, although SL3 could not be established unequivocally due to the poor digestion in this region. There were also a few overlaps between the single-strand-specific sites and double-strand-specific sites, suggesting that the conformation of this region may be flexible, probably consisting of both the optimal and suboptimal structures. Another possibility is that the SL structures in this X region may form higher-order tertiary RNA structures; e.g., GCUG in the loop 2 (Fig. 6A) may have a kissing interaction with UAGC in the SL3.

Site-directed mutagenesis of SL2 and SL3. To analyze the sequence and structural requirements for p58-RNA binding, we examined the protein-binding patterns of several site-specific mutants of SL2 and SL3. The predicted secondary structures of the mutant RNAs are shown in Fig. 7A.

We first studied the role of the potential kissing interaction

FIG. 5. Predicted models of the secondary structures of the 3'-end 98 nt (X region) (A) and the entire 3'UTR (B) of HCV RNAs. Modeling was performed with the computer program MulFold (Genetics Computer Group). The 3'-end 98 nt consist of three SL structures indicated as SL1, SL2, and SL3 (18). The ΔG values for the structure of the X region (A) and the entire 3'UTR (B) are -39.4 and -34.9 kcal/mol, respectively. The 5' end of the X region is designated nt 1, and the remaining sequences are numbered in the 5'→3' direction.

between SL2 and SL3 in p58 binding. We constructed HCV-X(+):CUG₃₃₋₃₅→AGU, in which the CUG in the loop 2 was replaced by its complementary sequence, and HCV-X(+):CU₆₋₇→AG/UAG₁₄₋₁₆→ACU, in which the CU and UAG nucleotides in SL3 were replaced by their complementary sequences (Fig. 7A-a and -b). These mutations disturb the possible kissing interaction but maintain the overall structure of SL2 and SL3. Fig. 7B-a and -b show that the 3-nt substitutions in the loop of SL2 abolished almost all of the p58 binding, but the mutation of the 5 nt in SL3, which preserve the SL3 stem structure, did not affect binding. The compensatory mutation that restored the possible kissing interaction between SL2 and SL3 [HCV-X(+):CU₆₋₇→AG/UAG₁₄₋₁₆→ACU/CUG₃₃₋₃₅→AGU] did not restore p58 binding (Fig. 7B-c). These results suggest that p58 binding requires the specific nucleotide sequence in the loop of SL2 and probably the stem structure, but not sequence, of SL3 as well. The possible kissing interaction between the loops of SL2 and SL3 is not required.

We further constructed several mutations which disrupted the stem structures or altered the loop sequences of either SL2 or SL3: HCV-X(+):AGC₁₅₋₁₇→CUG, in which the stem structure of the SL3 was destroyed, and its compensatory mutation, HCV-X(+):GCU₅₋₇→CAG/AGC₁₅₋₁₇→CUG (Fig. 7A-d and -e); HCV-X(+):UCUU₁₁₋₁₄→AAAA, in which the UCUU in the loop of SL3 was replaced by the AAAA sequence (Fig. 7A-f); and HCV-X(+):UCA₂₃₋₂₅→GAC, in which the stem structure of SL2 was destroyed, and its compensatory mutation, HCV-X(+):UCA₂₃₋₂₅→GAC/UGA₄₇₋₄₉→GUC (Fig. 7A-g and -h). Figure 7B-d to -h shows that disruption of the stem structure of either SL2 or SL3 abolished p58 binding and that their compensatory mutations restored binding; in addition, substitution of the loop sequences of SL3 abolished p58 binding. These results strongly suggested that p58 binding requires the stem structures of the SL2 and SL3, and p58 also recognizes specific sequences in the loops of both SL2 and SL3. Furthermore, since HCV-X(+):AGC₁₅₋₁₇→CUG, which is predicted to form an alternative suboptimal structure (Fig. 7A-d), failed to bind p58, we conclude that p58 binding involves both the loops and stems of the SL2 and SL3 of the optimum structure of the X region.

Sequence homology of the p58-binding site in the HCV 3'-UTR with the PTB-binding sequence. It was noted that the

5'-end sequence of SL3, which contains part of the p58-binding site, partially matched the 3' half of the reported consensus RNA sequence for PTB binding (Fig. 8A) (31). PTB, which is a 57-kDa protein, has been reported to bind to the genomic RNAs of several RNA viruses, including the 5'-UTR of HCV RNA (1, 13). To investigate whether the PTB consensus sequence was responsible for p58 binding, we linked HCV-X(+) RNA to the remaining 5' half of the reported PTB-binding consensus sequence (5'-GCUGCCU-3') [PTB-HCV-X(+)] (Fig. 8B), generating an RNA that contains the complete PTB-binding consensus sequence. Figure 8C shows that binding of p58 to this RNA did not differ significantly from binding to the HCV-X(+) RNA. We also investigated the SL2 loop deletion mutant of HCV-X(+) linked to the complete PTB-binding consensus [PTB-HCV-X(+) Δ31-40]. The results showed that the addition of the complete PTB-binding sequence did not restore the p58-binding activity of this mutant RNA. These results suggest that the reported PTB-binding consensus sequence is not sufficient to account for p58 binding. Thus, p58 binds to both SL2 and SL3.

Immunoprecipitation of the p58-HCV 3'-UTR complex. Although the site-specific mutagenesis studies indicated that the sequence requirement for p58 binding was significantly different from the reported binding specificity of PTB, the facts that these two proteins have similar electrophoretic mobilities and that the PTB-binding sequence is present in the p58-binding domain prompted us to examine whether the p58 bound to the 3'-UTR of HCV was PTB. The UV-cross-linked proteins from the cytoplasmic extract of HeLa cells were immunoprecipitated with a monoclonal antibody against PTB. Figure 9 shows that the p58 protein cross-linked to the HCV X region was precipitated with the anti-PTB antibody but not by unrelated antibodies. Furthermore, the p58 protein cross-linked to the mutant RNA HCV-X(+):GCU₅₋₇→CAG/AGC₁₅₋₁₇→CUG (Fig. 7A-e), which has a mutated PTB-binding sequence, was also precipitated by the anti-PTB antibody. These observations demonstrated that the p58 is indeed the PTB. In addition, the p58 binding to HCV-X(+) RNA was reduced by competition with poly(U), but not by poly(A) or poly(C) (data not shown), consistent with the reported properties of PTB (25). Furthermore, the p58 protein cross-linked to both wild-type and mutant HCV RNAs could be completely immunoprecipitated by the anti-PTB antibody; very little residual p58 protein was

FIG. 6. Determination of the secondary structure of the 3' end of HCV RNA. (A) Schematic drawing of the secondary structure model as shown in Fig. 5. The sites cut by either nuclease S1 or RNase T₁ (T1) are indicated by triangles; those cut by both nuclease S1 (S1) and RNase T₁ are indicated by filled triangles. The sites cut by V1 are indicated with circles (weak signal) and double circles (strong signal). (B) Primer extension with HCV-X(+) RNA after treatment with S1, RNase T₁ or V1. After each treatment with various amounts of enzyme (0.5, 0.05, and 0.01 U of S1; 0.1, 0.05, and 0.01 U of RNase of T₁; 0.5, 0.1, and 0.05 U of V1), ³²P-labeled primer oligonucleotide (complementary to nt 79 to 98 in the X region) was annealed to the RNA and extended with reverse transcriptase. Lane C, without treatment. Sequencing ladders were run in parallel to serve as size markers. Computer imagings were generated by Adobe Photoshop, version 3.0.

detected in the supernatant of the immunoprecipitation reaction (data not shown), indicating that PTB is the only cellular protein bound to the 3'-UTR of HCV RNA.

Recombinant PTB binds to the 3' end of HCV-RNA. To further establish that PTB binds to the 3' end of HCV-RNA, we performed UV cross-linking experiments to examine a purified GST-PTB fusion protein for its binding to HCV-X(+). Figure 10 shows that the recombinant PTB could bind to the

wild-type HCV-X(+). Furthermore, a mutation that destroyed the SL2 stem structure [HCV-X(+):UCA₂₃₋₂₅→GAC] (Fig. 7A-g) abolished PTB binding, and its compensatory mutation [HCV-X(+):UCA₂₃₋₂₅→GAC/UGA₄₇₋₄₉→GUC] (Fig. 7A-h) restored binding; also, HCV-X(+):UCUU₁₁₋₁₄→AAAA, which had mutations in the loop region of SL3 (Fig. 7A-f), bound the recombinant PTB only weakly. These binding patterns are similar to the binding specificity of p58 derived from the HeLa cell extracts (Fig. 7). Therefore, we conclude that the p58 cross-linked to the HCV 3'-UTR is PTB.

Detection of p58 in different cell lines. To determine the cellular distribution of p58, we examined the UV cross-linking patterns of the cellular extracts from four different cell lines (DBT, COS7, HepG2, and Molt4) and from mouse and human primary hepatocytes. Molt4 cells and human hepatocytes have been reported to support HCV replication (14, 30). Figure 11 shows that p58 was detected in all cell lines and primary hepatocytes. No significant differences in the amount of protein bound to the X region were observed. These results indicated that this protein is expressed in hepatocytes and various cell lines of different origins.

DISCUSSION

In this study, we have demonstrated by UV cross-linking analysis that a 58-kDa cellular protein binds specifically to the 3'-end 98 nt of the HCV RNA. This protein is present in both the cytoplasm and nuclei of most mammalian cell lines and human hepatocytes (Fig. 11). Immunoprecipitation experiments and UV cross-linking experiments using the recombinant PTB established that this protein is PTB. Partial deletion and site-directed mutagenesis of the 98-nt sequence mapped the PTB-binding domain to a stretch of 42 nt, which assumes two SLs. Significantly, the 3'-end 98 nt constitute the most conserved region in the viral genome, and among these 98 nt, the PTB-binding region is particularly conserved, being almost invariable among many HCV isolates sequenced so far (38). Such a conservation of sequence suggests the functional importance of this sequence. We have further shown that minor nucleotide substitutions within certain domains of the 98-nt sequence abolished PTB binding, indicating the rigid sequence requirement for PTB binding. Thus, PTB may play a significant role in HCV biology. Although the function of the HCV 3'-UTR is not yet known, it is conceivably involved in RNA replication or translation.

The data from the enzymatic probing analysis of the p58-binding RNA structure and inference from the RNA sequence requirement for p58 binding established that the PTB-binding region consists of two SLs. This structure is the most favorable structure for the HCV 3'-UTR, as predicted from computer analysis. However, the possible existence of multiple suboptimal structures of the HCV 3'-UTR is suggested from the partial overlaps between the single-strand- and double-strand-RNA-cutting sites in the enzymatic probing studies (Fig. 6). Particularly significant is that the second-most-stable RNA structure predicted from computer analysis is the predicted structure for HCV-X(+):AGC₁₅₋₁₇→CUG RNA (Fig. 7A-d). This one, i.e., the two-SL model, has a ΔG value very close to that of the optimal model, i.e., the three-SL model, indicating the strong likelihood of the existence of this alternative RNA structure. Interestingly, this RNA did not bind p58 at all. Therefore, it appears that among the multiple possible structures of the HCV 3'-UTR, only the predicted optimal structure will bind p58. Site-directed mutagenesis established that SL2 and SL3 are both involved in p58 binding.

Parts of the SL3 region partially matched the reported con-

FIG. 7. Site-directed mutagenesis of the SL2 and SL3 regions of HCV-X(+). (A) Computer-predicted secondary structure of HCV-X(+) RNAs with site-directed mutations. Each RNA was generated from a plasmid containing HCV-X(+) cDNA with indicated nucleotide substitutions (boldfaced). The details of these RNAs are described in the text. (B) UV cross-linking analysis. Increasing amounts of HeLa cytoplasmic proteins (0, 2, and 20 μg) were incubated with various ³²P-labeled probes described above. Only the p58 band of the SDS-polyacrylamide gels is shown. Computer imagings were generated by Adobe Photoshop, version 3.0.

sensus binding sequence for PTB (31) (Fig. 8A), also known as heterogeneous nuclear ribonucleoprotein I, and p57 (10), which is involved in spliceosome assembly (25). Immunoprecipitation of the PTB-HCV-X(+) RNA complex by the PTB-specific antibody established that PTB is the only protein binding to the HCV 3'-UTR. Also, p58-HCV-X(+) RNA binding was inhibited by poly(U) but not by poly(A) or poly(C), similar to the properties of PTB. However, the site-specific mutagenesis studies indicated that the sequence and structure requirement for PTB binding is significantly different from the re-

ported PTB-binding consensus RNA sequence. For example, site-specific mutations in SL3, which altered the reported PTB-binding primary sequence but not the SL3 stem structure (e.g., Fig. 7A-b and -e), did not affect protein binding; in contrast, mutations in SL2, which disrupted the loop sequence or stem structure, abolished p58 binding. These studies conclusively showed that p58 recognizes both the primary loop sequences and stem structures of SL2 and SL3. These structural and sequence requirements for protein binding were also observed with the recombinant PTB. Furthermore, the inclusion of the

FIG. 8. (A) Comparison of the sequences of the reported PTB-binding consensus and X region of the HCV 3'-UTR. (B) Schematic diagrams of the modified HCV-X(+) RNAs. HCV-X(+) and HCV-X(+) Δ31-40 are described in Fig. 1. PTB-HCV-X(+) contains an additional GCUCCU sequence upstream of the HCV-X(+) (in boxes), and PTB-HCV-X(+)Δ31-40 is its mutant with a loop 2 deletion. (C) UV cross-linking analysis. Increasing amounts of HeLa cytoplasmic proteins (0, 2, and 20 μg) were incubated with various ³²P-labeled probes described above. Only the 58-kDa protein is shown. Computer imagings were generated by Adobe Photoshop, version 3.0.

complete PTB-binding RNA sequence as previously reported (31) did not significantly increase protein binding. Thus, the significance of the reported PTB-binding consensus sequence (31) for PTB binding to the HCV 3'-UTR is not clear. Multiple binding sites for PTB have been identified in the poliovirus 5'-UTR (12), and the complexity of PTB-binding sequences in

FIG. 9. Immunoprecipitation of the UV-cross-linked proteins with an anti-PTB antibody. A monoclonal anti-PTB antibody (lane 2) and two unrelated antibodies, a monoclonal anti-TBP (lane 3) and a monoclonal anti-FLAG M2 (lane 4) antibody, were used to immunoprecipitate the proteins cross-linked to HCV-X(+) RNA (A) or HCV-X(+):GCU_{5,7}→CAG/AGC_{15,17}→CUG RNA (B). Lane 1, UV cross-linked p58 without immunoprecipitation. Only the p58 band of the SDS-polyacrylamide gels is shown. Computer imagings were generated by Adobe Photoshop, version 3.0.

FIG. 10. UV cross-linking of the recombinant PTB to HCV RNA. GST-PTB protein (30 ng) was UV cross-linked to the radiolabeled HCV-X(+) RNA (wild type [WT]) and several mutant RNAs (f, g, and h) depicted in Fig. 7. Arrow indicates the GST-PTB protein (86 kDa). Computer imagings were generated by Adobe Photoshop, version 3.0.

the internal ribosomal entry site of encephalomyocarditis virus RNA also has been reported (34–36). Furthermore, the reported PTB-binding consensus sequence does not exist in poliovirus and encephalomyocarditis virus internal ribosomal entry sites. It is possible that PTB has a broader RNA-binding sequence requirement than was previously reported.

It has been reported that PTB exists together with a 100-kDa cellular protein as part of a large complex, which is required for RNA splicing in HeLa cells (25). We speculate that PTB, together with the viral or other cellular proteins, may form a ribonucleoprotein complex in the HCV 3'-UTR and play important roles in the replication or translation of viral RNA. PTB has also been reported to bind to the 5'-UTR of HCV RNA and function in internal initiation of translation (1), although this claim has been disputed (16, 27). Recent studies suggest that direct or indirect RNA-RNA interactions between the 5' and 3' ends of various cellular or viral RNAs can regulate translation, which usually starts from the 5' end of the RNA (6, 19, 20). Replication and transcription of several RNA viruses have also been shown to require direct or indirect interactions between the 5' and 3' ends of the viral RNA (20, 22, 33). The finding that PTB binds to both the 5' and 3' ends of HCV RNA suggests that PTB may play a role in the cross-talk between the two ends and thus participate in viral replication or translation.

FIG. 11. UV cross-linking of cytoplasmic extracts from five different cell lines and from mouse and human primary hepatocytes. Twenty micrograms of each cytoplasmic extract was UV cross-linked to the radiolabeled HCV-X(+) RNA. (-), no cellular extract. The arrow indicates the 58-kDa protein. Computer imagings were generated by Adobe Photoshop, version 3.0.

ACKNOWLEDGMENTS

We thank F. Lahser and E. Wimmer for the gift of anti-PTB antibody and Daphne Shimoda for assistance in preparing the manuscript. We also thank Kunitada Shimotohno, Institute for Virus Research, Kyoto University, for sharing his similar results prior to publication.

This work was partially supported by research grant AI 40038 from the National Institutes of Health. T.I. is a Research Associate and M.M.C.L. is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

- Ali, N., and A. Siddiqui. 1995. Interaction of polypyrimidine tract-binding protein with the 5' noncoding region of hepatitis C virus RNA genome and its functional requirement in internal initiation of translation. *J. Virol.* **69**: 6367–6375.
- Bukh, J., R. H. Purcell, and R. H. Miller. 1992. Sequence analysis of the 5' noncoding region of hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **89**:4942–4946.
- 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.
- Chambers, K. H., C. S. Hahn, R. Galler, and C. M. Rice. 1990. Flavivirus genome organization, expression and replication. *Annu. Rev. Microbiol.* **44**:649–688.
- 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.
- Danthinne, X., J. Seurinck, F. Meulewaeter, M. van Montagu, and M. Cornelissen. 1993. The 3' untranslated region of satellite tobacco necrosis virus RNA stimulates translation in vitro. *Mol. Cell. Biol.* **13**:3340–3349.
- Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a solution extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475–1489.
- Frangioni, J. V., and B. G. Neel. 1993. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal. Biochem.* **210**:179–187.
- Furuya, T., and M. M. C. Lai. 1993. Three different cellular proteins bind to complementary sites on the 5'-end-positive and 3'-end-negative strands of mouse hepatitis virus RNA. *J. Virol.* **67**:7215–7222.
- Ghetti, A., S. Pinol-Roma, W. M. Michael, C. Morandi, and G. Dreyfuss. 1992. hnRNP I, polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res.* **20**:3671–3678.
- 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.
- Hellen, C. U. T., T. V. Pestova, M. Litterst, and E. Wimmer. 1994. The cellular polypeptide p57 (pyrimidine tract-binding protein) binds to multiple sites in the poliovirus 5' nontranslated region. *J. Virol.* **68**:941–950.
- Hellen, C. U. T., G. W. Witherell, M. Schmid, S. H. Shin, T. V. Pestova, A. Gil, and E. Wimmer. 1993. A cytoplasmic 57 kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. *Proc. Natl. Acad. Sci. USA* **90**:7642–7646.
- Ito, T., J. Mukaigawa, J. Zuo, Y. Hirabayashi, K. Mitamura, and K. Yasui. 1996. Cultivation of hepatitis C virus in primary hepatocyte culture from patients with chronic hepatitis C results in release of high titre infectious virus. *J. Gen. Virol.* **77**:1043–1054.
- Jacobson, S. J., D. A. M. Konings, and P. Sarnow. 1993. Biochemical and genetic evidence for a pseudoknot structure at the 3' terminus of the poliovirus RNA genome and its role in viral RNA amplification. *J. Virol.* **67**:2961–2971.
- Kaminski, A., S. L. Hunt, J. G. Patton, and R. J. Jackson. 1995. Direct evidence that polypyrimidine tract binding protein (PTB) is essential for internal initiation of translation of encephalomyocarditis virus RNA. *RNA* **1**:924–938.
- Kato, M., M. Hijikata, Y. Ootsuyama, M. Nakagawa, S. Ohkoshi, T. Sugimura, and K. Shimotohno. 1990. Molecular cloning of human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc. Natl. Acad. Sci. USA* **87**:9524–9528.
- 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.
- Kwon, Y. K., and N. B. Hecht. 1993. Binding of a phosphoprotein to the 3' untranslated region of the mouse protamine 2 mRNA temporally represses its translation. *Mol. Cell. Biol.* **13**:6547–6557.
- Lahser, F. C., L. E. Marsh, and T. C. Hall. 1993. Contributions of the brome mosaic virus RNA-3 3'-nontranslated region to replication and translation. *J. Virol.* **67**:3295–3303.
- Lin, C., B. D. Lindenbach, B. M. Pragai, D. W. McCourt, and C. M. Rice. 1994. Processing in the hepatitis C virus E2-NS2 region: identification of p7 and two distinct E2-specific products with different C termini. *J. Virol.* **68**:5063–5073.
- Luo, G., W. Luytjes, M. Enami, and P. Palese. 1991. The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure. *J. Virol.* **65**:2861–2867.
- Melchers, W. J. G., J. G. J. Hoenderop, H. J. B. Slot, C. W. A. Pleij, E. V. Pilipenko, V. I. Agol, and J. M. D. Galama. 1997. Kissing of the two pre-dominant hairpin loops in the coxsackie B virus 3' untranslated region is the essential structural feature of the origin of replication required for negative-strand RNA synthesis. *J. Virol.* **71**:686–696.
- Okamoto, H., K. Kurai, S. I. Okada, K. Yamamoto, H. Lizuka, T. Tanaka, S. Fukuda, T. Tsuda, and S. Mishiro. 1992. Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology* **188**:331–341.
- Patton, J. G., S. A. Mayer, P. Tempst, and B. Nadal-Ginard. 1991. Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing. *Genes Dev.* **5**:1237–1251.
- Philipenko, E. V., K. V. Poperechny, S. V. Maslova, W. J. G. Melchers, H. J. B. Slot, and V. I. Agol. 1996. Cis-element, oriR, involved in the initiation of (–) strand poliovirus RNA: a quasi-globular multi-domain RNA structure maintained by tertiary ('kissing') interactions. *EMBO J.* **15**:5428–5436.
- 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.
- 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: implication for flavivirus gene expression and evolution. *Science* **229**:726–735.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
- Shimizu, Y. K., A. Iwamoto, M. Hijikata, R. H. Purcell, and H. Yoshikura. 1992. Evidence for in vitro replication of hepatitis C virus genome in a human T-cell line. *Proc. Natl. Acad. Sci. USA* **89**:5477–5481.
- 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.
- Tanaka, T., N. Kato, M.-J. Cho, K. Sugiyama, and K. Shimotohno. 1996. Structure of the 3' terminus of the hepatitis C virus genome. *J. Virol.* **70**:3307–3312.
- Wertz, G. W., S. Whelan, A. LeGrone, and L. A. Ball. 1994. Extent of terminal complementarity modulates the balance between transcription and replication of vesicular stomatitis virus RNA. *Proc. Natl. Acad. Sci. USA* **91**:8587–8591.
- Witherell, G. W., A. Gil, and E. Wimmer. 1993. Interaction of polypyrimidine tract binding protein with the encephalomyocarditis virus mRNA internal ribosomal entry site. *Biochemistry* **32**:8268–8275.
- Witherell, G. W., C. S. Schultz-Witherell, and E. Wimmer. 1995. Cis-acting elements of the encephalomyocarditis virus internal ribosomal entry site. *Virology* **214**:660–663.
- Witherell, G. W., and E. Wimmer. 1994. Encephalomyocarditis virus internal ribosomal entry site RNA-protein interactions. *J. Virol.* **68**:3183–3192.
- Wu, H.-N., Y. J. Wang, C.-F. Hung, H.-J. Lee, and M. M. C. Lai. 1992. Sequence and structure of the catalytic RNA of hepatitis delta virus genomic RNA. *J. Mol. Biol.* **223**:233–245.
- Yamada, N., K. Tanihara, A. Takada, T. Yorihozi, M. Tsutsumi, H. Shimomura, T. Tsuji, and T. Date. 1996. Genetic organization and diversity of the 3' noncoding region of the hepatitis C genome. *Virology* **223**:255–261.
- Zhou, M. Y., S. E. Clark, and C. E. Gomez-Sanchez. 1995. Universal cloning method by TA strategy. *BioTechniques* **19**:34–35.
- Zuker, M., and P. Stiegler. 1981. Optimal computer folding of large RNA sequences using thermodynamic and auxiliary information. *Nucleic Acids Res.* **9**:133–148.