Hepatitis C Virus Internal Ribosome Entry Site-mediated Translation Is Stimulated by Specific Interaction of Independent Regions of Human La Autoantigen*

Renuka Pudi‡§, Saraswathi Abhiman¶‖, Narayanaswamy Srinivasan¶, and Saumitra Das‡ ‡‡**

From the ‡Department of Microbiology and Cell Biology and the ¶Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

The human La autoantigen has been shown to interact with the internal ribosome entry site (IRES) of hepatitis C virus (HCV) *in vitro*. Using a yeast three-hybrid system, we demonstrated that, in addition to full-length La protein, both N- and C-terminal halves were able to interact with HCV IRES *in vivo*. The exogenous addition of purified full-length and truncated La proteins in rabbit reticulocyte lysate showed dose-dependent stimulation of HCV IRES-mediated translation. However, an additive effect was achieved adding the terminal halves together in the reaction, suggesting that both might play critical roles in achieving full stimulatory activity of the full-length La protein. Using computational analysis, three-dimensional structures of the RNA recognition motifs (RRM) of the La protein were independently modeled. Of the three putative RRMs, RRM2 was predicted to have a good binding pocket for the interaction with the HCV IRES around the GCAC motif near the initiator AUG and RRM3 binds perhaps in a different location. This observation was further investigated by the filter-binding and toe-printing assays. The results presented here strongly suggest that both the N- and C-terminal halves can interact independently with the HCV IRES and are involved in stimulating internal initiation of translation.

Hepatitis C virus (HCV)[1] is a single-stranded, positive-sense enveloped RNA virus classified in a separate genus of the family Flaviviridae (1). HCV has been shown to be the primary causative agent of non-A, non-B viral hepatitis, which often leads to development of chronic hepatitis, cirrhosis, or hepatocellular carcinoma (2, 3). The HCV genome RNA is ~9.5 kb in length and consists of a 5′-untranslated region (UTR), a long open reading frame encoding the viral polyprotein, and a 3′-UTR (4). The translation initiation of HCV occurs by cap-independent mechanism that is directed by the highly structured 5′-UTR (341 nt long), consisting of four stem-loops and a pseudoknot structure. The 5′-UTR contains unique *cis*-acting element called "internal ribosome entry site" (IRES) that mediates 5′-cap-independent internal initiation of translation (5–9). The HCV IRES element requires most of the 5′-UTR, except the first 40 nt, and extends to a short stretch (30–40 nt) of sequence downstream of the initiator AUG codon (6, 10). Hepatitis C virus IRES is significantly different from that of picornaviruses (*e.g.* poliovirus) with respect to length, secondary and tertiary structure of the 5′-UTR RNA (11). Also in contrast to picornavirus IRES, HCV IRES does not have a strict requirement for the canonical initiation factors (eIFs) other than eIF2 and eIF3 (12). A number of cellular polypeptides have been shown to specifically interact with the HCV 5′-UTR RNA, which include polypyrimidine-tract binding protein (p57/60) (13), human La antigen (La, p50/52) (14), poly(rC)-binding protein 2 (15), heterogeneous nuclear ribonucleoprotein L (p68) (16), and ribosomal protein factors S9 and S5 (12). Interaction of these proteins with the 5′-UTR may be important for translation and/or replication of the HCV genome, although the precise roles are not known yet.

The La protein (also called SS-B) was originally identified as an autoantigen that was recognized by sera from patients with systemic lupus erythematosus and Sjogren's syndrome (17). Homologues of La protein have been identified in fruit fly, (*Drosophila melanogaster*) and in yeast (*Saccharomyces cerevisiae*) (18). Human La antigen is an RNA-binding protein, belonging to the RNA recognition motif (RRM) super family. It is a 50/52-kDa protein, which is predominantly localized within the nucleus and which functions in the maturation of RNA polymerase III transcripts and the unwinding of double-stranded RNA (19–22). La protein has been shown to be associated with U1RNA (23), telomerase RNA (24), adenovirus VA RNA (25), vesicular stomatitis virus leader RNA (26), influenza virus (27), Sindbis virus (28), and the HIV TAR element (29). The other targets of the La protein binding include the 5′-UTRs of poliovirus (30), hepatitis C virus (14), encephalomyocarditis virus (31), and Bip mRNA (32). Interestingly, it has been shown that La protein specifically interacts with both the 5′- and 3′-UTR of hepatitis C virus RNA (33). La protein plays a functional role in internal initiation of translation where addition of purified La to RRL in the *in vitro* translation assays using either HCV or poliovirus IRES resulted in stimulation of the translation activity (14, 30). Sequestration of La in RRL inhibits HCV IRES-mediated translation, which can be rescued by exogenous addition of purified La protein. Different approaches of La depletion were employed to demonstrate the

[1] The abbreviations used are: HCV, hepatitis C virus; UTR, untranslated region; IRES, internal ribosome entry site; La, human La antigen; RRM, RNA recognition motif; nt, nucleotide(s); eIF, eukaryotic initiation factor; HIV, human immunodeficiency virus; TAR, *trans*-activation response element; RRL, rabbit reticulocyte lysate; CAT, chloramphenicol acetyltransferase; SCR, structurally conserved region; β-gal, β-galactosidase, 3-AT, 3-amino-1,2,4-triazole; Ni-NTA, nickel-nitrilotriacetic acid; TE, Tris-EDTA.

Role of La N- and C-terminal Halves on HCV IRES Function

functional requirement of La protein on HCV IRES-mediated translation. A small RNA called I-RNA (60 nt) originally isolated from *S. cerevisiae* has been shown to bind to La protein and inhibit HCV IRES-mediated translation (34). Similarly, a SELEX RNA generated against La protein has been used to compete with HCV 5′-UTR for the La protein binding to block the IRES-mediated translation (35). However, it has been noted that HCV IRES has a lower requirement of La protein, compared with poliovirus IRES (36). Similarly, in absence of La protein, X-linked inhibitor of apoptosis protein IRES-mediated translation is shown to be severely affected (37) and Bip IRES activity is enhanced severalfold in presence of La protein (32). Interestingly, binding of La protein to the HIV TAR sequence alleviates the translational repression exerted by the TAR sequence on a downstream reporter gene (38) and, in the case of encephalomyocarditis virus, La protein alleviates inhibitory activity of surplus polypyrimidine tract-binding protein (31).

La protein has been shown to have three putative RRMs and a basic region followed by a stretch of acidic region at the C terminus (39). A truncated La protein (1–194), which still contained RRM and bound the poliovirus 5′-UTR, failed to stimulate poliovirus translation or correct the aberrant translation in RRL (40), suggesting that binding of the La protein with the RNA is insufficient to explain the activity. The C terminus also contains a homodimerization domain; deletion of this domain has been shown to abrogate the ability of La protein to enhance IRES-mediated translation of poliovirus (41).

In this report we have demonstrated that both N- and C-terminal halves of La protein are able to interact with HCV IRES *in vivo* by the yeast three-hybrid assay system. The relative binding affinities have been determined using filter-binding assays. Furthermore, the biological relevance of these interactions was studied by exogenous addition of purified full-length and truncated La proteins in the *in vitro* translation system using bicistronic constructs in rabbit reticulocyte lysate. Interestingly, both the N- and C-terminal halves of La proteins were able to stimulate the HCV IRES-mediated translation unlike the case in poliovirus IRES, suggesting that the deletion of the C terminus does not abrogate the ability of La protein to enhance the translation of hepatitis C virus. To date, no crystallographic or NMR structure information is available for either the full-length or individual domains of La protein. Using computational analysis, we have searched for traditional RRMs in the La protein. We have generated the three-dimensional model of the RRM regions of the La protein by comparative modeling on the basis of the closest homologues of known structures determined using NMR or crystallography. The predicted secondary structure of the region 105–208 residues of La (referred to as RRM2) showed similarity with that of RRM-containing proteins of known structures. Additionally, using comparative sequence and structural analysis, it is suggested that the RRM2 within the N-terminal domain of La protein binds HCV RNA at the GCAC motif near the initiator AUG. The prediction was further investigated by the filter-binding and toe-printing assays, which clearly demonstrated much higher RNA binding affinity of the RRM2 at this region. The results constitute the first report to demonstrate physical interaction of La protein with the viral RNA *in vivo* and provide direct evidence of the *in vitro* interactions of RRM2 of La protein with the HCV IRES around the region encompassing the initiator AUG and the GCAC motif.

EXPERIMENTAL PROCEDURES

Plasmids—HCV 1b encoding plasmid, pCV, was generously given by Dr. Akio Nomoto and Dr. Tsukiyama-Kohara (University of Tokyo, Tokyo, Japan). HCV 5′-UTR along with 42-nt (18–383 nt) coding sequence was PCR-amplified from plasmid pCV and cloned between *Hin*dIII and *Eco*RI sites of the mammalian expression vector pcDNA3

(Invitrogen) to generate pcDHCV-383. Similarly nt 1–341 of the HCV 5′-UTR were PCR amplified from the clone pT7DC1–341 (a generous gift from Dr. Aleem Siddiqui, University of Colorado, Denver, CO) and cloned into *Hin*dIII sites of pcDNA3 vector to generate pcDHCV-341. The bicistronic reporter vector, pRL HCV 1b, containing two reporter genes (*Renilla* luciferase and firefly luciferase) separated by the HCV IRES was generously provided by Dr. Richard M. Elliot (University of Glasgow, Glasgow, Scotland, UK). Poliovirus 5′-UTR containing bicistronic construct, p2–5′, was a generous gift from the laboratory of Dr. Nahum Sonenberg (McGill University, Montreal, Canada). The cDNA clone encoding human La autoantigen, pETLa, was obtained from Dr. Jack Keene (Duke University, Durham, NC). La coding sequence was amplified by PCR using the following primers: 5′-GACCG-GATCCATGGCTGAAAATGG-3′ and 5′-CGTAGAATTCCTACTGGTC-TCCAG-3′. The amplified product was subcloned into *Bam*HI and *Eco*RI sites of pRSET-A vector (Invitrogen). The two deletion constructs pRSET-A La1–208 (La-N) and pRSET-A La209–408 (La-C) were constructed using the primers 5′-GACCGGATCCATGGCTGAAAATGG-3′, 5′-GGCCGAATTCTTTAGCTCTTAATT-3′, 5-ATATGGATCCCAGGA-GCAAGAAGC-3′, and 5′-CGTAGAATTCCTACTGGTCTCCAG-3′. The deletion constructs pRSET-A La1–100 (RRM1) and pRSET-A La101–208 (RRM2) were also constructed using the primers 5′-GACCGGAT-CCATGGCTGAAAATGG-3′, 5′-GATCGAATTCACTTCA GGTAGGGG-3′, 5′-GGCCGGATCCGAAGTGACTTGGGA-3′, and 5′-CGTAGAATTC-CTACTGGTCTCCAG-3′. Similarly, pRSET-A La209–300 (RRM3) was amplified by PCR and cloned in pRSET-A vector in *Bam*HI and *Eco*RI sites using the following primers: 5′-ATATGGATCCCAGGAGCAAGA-AGC-3′ and 5′-TCTCGAATTCTCCCAAGTCACCTT-3′. All the PCRs were carried out with 35 cycles, each cycle consisting of denaturation (95 °C for 40 s), annealing (55 °C for 1 min), and extension (68 °C for 1 min) using *Pfx* DNA polymerase (Invitrogen).

Purification of La Full-length and Truncated Proteins Using Ni-NTA-Agarose Column—*Escherichia coli* BL21 (DE3) cells were transformed with pRSET-A vectors containing either the full-length or the deletion mutants of La. Single colonies were inoculated into 5 ml of LB broth containing 75 μg/ml ampicillin and grown at 37 °C incubator shaker at 200 rpm speed until A_{600} reached 0.6. The cultures were induced with 0.6 mM isopropyl-1-thio-β-D-galactopyranoside and grown for another 4 h. The cells were pelleted and resuspended in 200 μl of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole). The extract was made by sonication. The above crude extracts were mixed with 0.25 volume of Ni-NTA-agarose slurry (Qiagen) and kept for rocking at 4 °C for 2 h. The lysate was loaded onto a column, and the flow-through was discarded. The column was washed with 5 ml of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 40 mM imidazole). The bound protein was eluted with 500 μl of elution buffer containing 500 mM imidazole. The eluted proteins were dialyzed at 4 °C for 4–6 h in 500 ml of dialysis buffer (50 mM Tris, pH 7.4, 100 mM KCl, 7 mM β-mercaptoethanol, 20% glycerol), aliquoted, and stored in a −70 °C freezer.

In Vitro Transcription—mRNAs were transcribed *in vitro* from different linearized plasmid constructs under T7 promoters in run-off transcription reactions. The HCV bicistronic construct, pRL HCV 1b, was linearized with *Hin*dIII, and poliovirus bicistronic construct, p2–5′, was linearized with *Xho*I downstream of firefly luciferase to be used as templates for run-off RNA synthesis. The linear DNA were electrophoresed on agarose gels and extracted using silica beads (Bangalore Genie) and then transcribed using a Ribomax large scale RNA production system-T7 (Promega).

Radiolabeled mRNAs were transcribed *in vitro* using T7 RNA polymerase (Promega) and [α-^{32}P]uridine triphosphate (PerkinElmer Life Sciences). The pcDHCV-383 and pcDHCV-341 were linearized with *Eco*RI, gel eluted, and transcribed *in vitro* to generate the ^{32}P-labeled RNA. The transcription reaction was carried out under standard conditions (Promega protocol) using 2.5 μg of linear template DNA at 37 °C for 1 h 30 min. After alcohol precipitation, the RNA was resuspended in 25 μl of nuclease-free water. 1 μl of the radiolabeled RNA sample was spotted onto DE81 filter paper, washed with phosphate buffer, and dried, and the incorporated radioactivity was measured using liquid scintillation counter.

Filter-binding Assay—The [α-^{32}P]HCV 18–383 RNA or [α-^{32}P]HCV 1–341 RNA was incubated with the proteins at 30 °C for 15 min in RNA binding buffer (containing 5 mM HEPES, pH 7.6, 25 mM KCl, 2 mM $MgCl_2$, 3.8% glycerol, 2 mM dithiothreitol, and 0.1 mM EDTA), and loaded onto nitrocellulose filters equilibrated with 2 ml of RNA binding buffer. The filters were then washed twice with 2 ml of binding buffer and dried, and the counts retained were measured in liquid scintillation counter. The graph was plotted with protein concentration (nM) on *x* axis and the percentage of bound RNA as the percentage of counts

retained, on the *y* axis. The relative affinity constants were calculated as the protein concentrations at which 50% RNA was bound.

Yeast Transformation—The RNA-hybrid and the protein-hybrid vectors were co-transformed into yeast host strain, L40uraMS2, by lithium acetate method. The overnight culture grown in 10 ml of YPD medium (1% yeast extract, 2% peptone, and 2% dextrose) was diluted into 50 ml of YPD such that A_{600} is 0.4 and grown for an additional 2–4 h. The cells were pelleted, washed with 40 ml of $1\times$ TE (10 mM Tris, pH 7.5, 1 mM EDTA), resuspended in 2 ml of $1\times$ LiAc/0.5$\times$ TE buffer (100 mM lithium acetate, pH 7.5, and 0.5$\times$ TE) and incubated at 30 °C for 30 min. 1 μg of each plasmid and 100 μg of herring sperm DNA were mixed with 100 μl of the yeast suspension and 700 μl of $1\times$ LiAc, 40% polyethylene glycol 3350, $1\times$ TE, and incubated at 30 °C for 30 min. 88 μl of Me$_2$SO was added, and heat shock was given for 15 min at 42 °C. The cells were washed with $1\times$ TE and plated onto His-selective plates and incubated at 30 °C. The transformants growing on His$^-$ plates were later spotted as different dilutions onto plates containing 10 mM 3-aminotriazole.

Liquid β-Galactosidase Assay—The yeast colonies were inoculated into 5 ml of selective medium and grown overnight. The cells were harvested and suspended in 100 μl of Z buffer (60 mM Na$_2$HPO$_4$, 40 mM NaH$_2$PO$_4$, 10 mM KCl, 1 mM MgSO$_4$, pH 7.0), 1 mM phenylmethylsulfonyl fluoride, and 1% SDS. Cells were disrupted by repeated cycles of vortexing and cooling in presence of glass beads. 0.8 ml of Z buffer, 1 mM phenylmethylsulfonyl fluoride, and 1% SDS was added, and 100 μl was set aside for protein estimation. To the remaining, 0.2 ml of 4 mg/ml ONPG in 200 mM phosphate buffer, pH 7.0, was added and incubated at room temperature until yellow color developed. The reaction was stopped by adding 0.5 ml of 1 M Na$_2$CO$_3$, and A_{420} was measured.

In Vitro Translation—*In vitro* translation of the capped bicistronic mRNAs were carried out in micrococcal nuclease-treated rabbit reticulocyte lysates (RRL, Promega Corp.). Briefly, 12.5-μl reaction mixtures contained 8.75 μl of RRL containing 0.25 μl each of minus methionine and minus leucine amino acid mixtures and 10 units of RNasin (Promega Corp.). The reaction mixtures were incubated at 30 °C for 1 h and 30 min. 2 μl of the reaction mixtures were assayed for both the *Renilla* and firefly luciferase activity according to Promega protocol using the Dual Luciferase reporter assay system. To assay the CAT reporter gene activity, 2 μl of 1:10 dilution of the reactions mixtures were added to 38 μl of 0.25 M Tris-HCl, pH 7.4, 2 μl of [^{14}C]chloramphenicol in total volume of 90 μl. After incubation at 37 °C for 10 min, 8 μl of 10 mM acetyl-CoA was added and incubated further for 1 h at 37 °C followed by ethyl acetate (500 μl) extraction. The upper organic phase was separated and evaporated to dryness. The residues were resuspended in 25 μl of ethyl acetate and spotted onto silica gel TLC plates, followed by chromatography using a mixture of chloroform/methanol (95:5). The TLC plate was dried and exposed to PhosphorImager, and bands corresponding to acetylated and unacetylated forms were quantified. The relative CAT activity was expressed as a ratio of acetylated form to the total.

Primer Extension Inhibition Analysis—Primer extension inhibition (toe-printing) assay was performed as described previously (12). Briefly, increasing concentration of purified His-tagged La full-length or deletion proteins (200, 400, and 600 ng) were incubated with 2.5 pmol of *in vitro* transcribed RNA corresponding to the HCV IRES (18–383), and binding reaction was performed in a final volume of 20 μl at 30 °C for 20 min. To this reaction, ^{32}P-end-labeled primer complimentary to 25 nucleotides of the 3′ end of the HCV-383 was added and allowed to extend using 3 units of avian myeloblastosis virus-reverse transcriptase (Promega) at 30 °C for 1 h. The cDNAs were alcohol-precipitated, resuspended, and compared with the dideoxynucleotide sequence ladders by electrophoresis on a 6% polyacrylamide, 7 M urea denaturing gel.

Comparative Modeling—Traditional RRMs in the La protein were searched using the motif search tools such as PROSITE. The known related crystal and NMR structures of RNA-bound complexes were used as the basis in modeling the three-dimensional structure of RRM2-RNA complex. At least six different proteins with complex crystal or NMR structures of RRMs bound to RNA or DNA are available. The highest sequence identity between any of the putative RRMs of La proteins and the RRMs of known structure is on the order of 14%, which is very low. Such sequence identities are possible even between two entirely unrelated proteins of completely different fold. Hence we identified the best homologues for comparative analysis and modeling on the basis of highest compatibility of the sequences of putative RRMs in the La protein with each of the known structures that are bound to RNA. The structural features considered for compatibility analysis are solvent accessibility, hydrogen bonding pattern, and secondary structures. It is important to choose a minimal subset of known structures as the basis under low sequence similarity situation, as it has been shown earlier

that choosing all the possible structures as basis can result in high errors in the model (42).

Thus, two constraints were used in choosing the template structures: (*a*) the RRM of known structure should be bound to RNA, and (*b*) the compatibility of the sequences of RRMs in the La protein with the chosen structures should be highest. Two different known structures of RRMs of RNA-bound complexes used in the modeling were: "splicing factor U2B" from human and U1A spliceosomal protein.

The suite of programs encoded in *COMPOSER* (42) and incorporated in *SYBYL* (Tripos Inc., St. Louis, MO) was used to generate the three-dimensional models of RRMs in La protein. The sequences of the known structures of RRM proteins were aligned on the basis of their three-dimensional structures and refined using the structural features such as solvent accessibility and secondary structures and relationships such as hydrogen bonding pattern (43, 44). The optimal superposition of Cα atoms in the template structures resulted in the identification of the structurally conserved regions (SCRs) and structurally variable regions. The mean of the topologically equivalent Cα atoms weighted by the square of the sequence identity (42) between the RRM in the La protein and each one of the template structures define the framework for the family of RRM proteins of known structure. The sequence of an RRM of La was then aligned with the known structures and the framework. The regions of RRM of the La protein equivalent to the framework region were modeled from the SCRs of the template structures. To model a given SCR of RRM-La, the equivalent SCR in a homologue of known structure with the best local sequence identity with RRM-La is superimposed on the framework. The variable regions are modeled by identifying a suitable segment from a known structure in the data bank. A search is made for a segment having the desired number of residues and the proper end-to-end distances across the three "anchor" Cα at the either side of the putative loop such that the loop can be fitted joining the contiguous conserved regions. A template matching approach (45) to rank the candidate loops was also used. The best ranking loop with no short contacts with the rest of the protein is fitted using the ring-closure procedure of F. Eisenmenger.[2] Side chains are modeled either by extrapolating from the equivalent positions in the basis structure where appropriate or by using rules derived from the analysis of known protein structures (46).

Energy Minimization—The *COMPOSER*-generated model was energy minimized in *SYBYL* using the *AMBER* force-field (47). During the initial cycles of energy minimization, the backbone was kept rigid and side chains alone were moved. Subsequently, all atoms in the structure were allowed to move during minimization. This approach kept disturbance of the backbone structure to the minimum. Energy minimization was performed until all short contacts and inconsistencies in geometry were rectified. During the initial stages of minimization, the electrostatic term was not included, as the main objective was to relieve steric clashes and to rectify bad geometry.

Modeling of the RRM-RNA Complex—The NMR structure of human U1A protein bound to RNA (48) has been used as the template to model the binding of RNA in the RRMs of La protein. The choice of 1aud is made on the basis of the highest sequence similarity of 1aud with RRM2 and RRM3 of La. We have identified the regions of U1A protein that interacts with RNA and placed the RNA in an identical orientation in the RRM models of La protein. This has been achieved by optimally superimposing the RNA-interacting regions of U1A on the equivalent regions of the RRM model of La and carrying RNA from the U1A protein complex in the process.

RESULTS

La Protein Interacts with HCV IRES in Vivo in Yeast Three-hybrid System—Previously, it has been demonstrated that La autoantigen interacts with the 5′-untranslated region of hepatitis C virus RNA *in vitro* and the C terminus effector domain modulates the binding efficiency (35). In fact several *trans*-acting factors have been shown to interact with the HCV 5′-UTR, but so far none of the interactions have been shown *in vivo*. To investigate whether La protein interacts with the HCV IRES RNA *in vivo*, we took advantage of the genetic screening method by using yeast three-hybrid system. The system allows detection of the target RNA (bait) and potential protein (prey) interaction in a *S. cerevisiae* strain containing integrated reporters (histidine and β-galactosidase). Interaction between

---

[2] F. Eisenmenger, unpublished results.

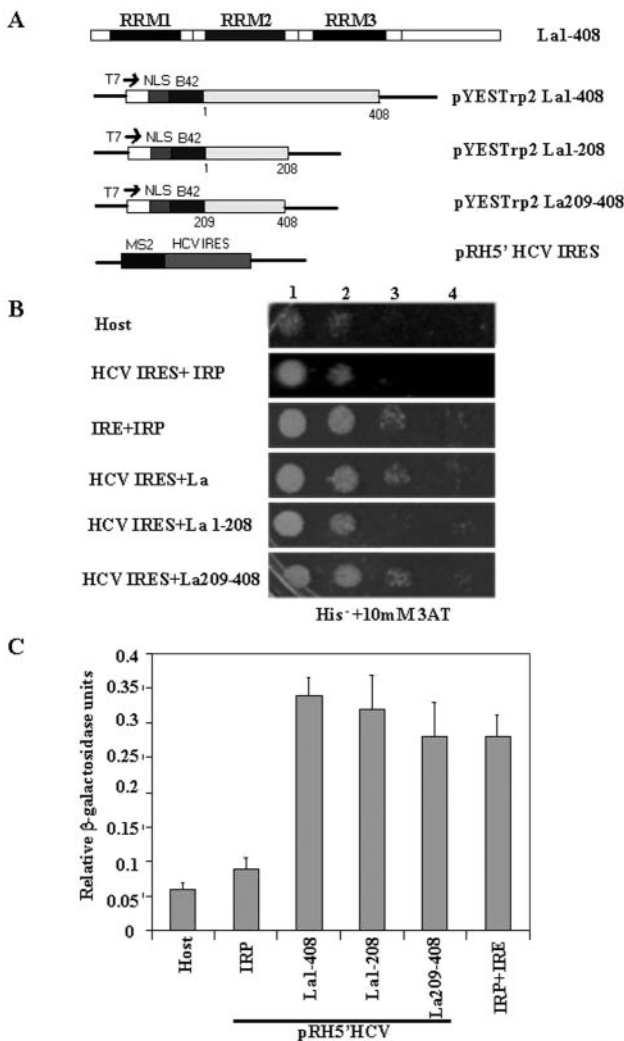*Role of La N- and C-terminal Halves on HCV IRES Function*

FIG. 1. **Yeast three-hybrid system analysis.** *A*, schematic representation of the full-length La protein showing the relative positions of the three putative RRMs. The protein hybrids and RNA hybrid generated in pYESTrp2 and pRH5′ vectors, respectively, are shown. *B*, three-hybrid yeast strain, *S. cerevisiae* L40uraMS2, was transformed with plasmids carrying different B42 activation domain fusions and MS2RNA fusions. The transformants were selected on −Trp/Ura/His medium and subsequently tested for the expression of reporter gene activity. The positive three-hybrid interactions were further confirmed by growth on the above medium supplemented with 10 mM 3-aminotriazole. *C*, the relative β-galactosidase activity of each transformant is plotted as *vertical bars*. The mean and the standard deviations from the three independent experiments are shown. The cotransformants in each case are indicated *below* the *bars*. IRE-IRP interaction represents the positive control.

the *bait* RNA and the *prey* protein results in the reconstitution of transcriptional activation of the reporter genes (*His* and *LacZ*), changing the phenotype of the yeast cells. Positive interactions are detected by selection on His⁻ plates and assaying the β-galactosidase reporter gene activity. For this purpose, RNA hybrid vector was generated by cloning the HCV IRES (18–383 nt) into the pRH5′ vector (Invitrogen), which produced HCV IRES as a fusion to MS2 RNA. The protein hybrid vectors were generated by cloning the full-length La1–408 as well as the deletions, La1–208 (La-N) and La209–408 (La-C) into pYESTrp2 vector, which produces the above proteins as a fusion to B42 activation domain (Fig. 1*A*). The pYESTrp2 protein hybrid vectors were co-transformed with the pRH5′-HCV RNA hybrid vector into the host *S. cerevisiae* (L40uraMS2). The colonies showing histidine (His)-positive phenotype were selected on His-negative plates. Different dilutions of the trans-

formants were then assayed for the ability to grow on His-negative plates containing 10 mM 3-amino-1,2,4-triazole (3-AT) for further screening of specific interaction (Fig. 1*B*). The transformants were also assayed for β-galactosidase activity (β-gal), the relative β-galactosidase activities resulting from the interactions are shown in Fig. 1*C*. The host alone was taken as background. When pYESTrp2 La1–408 was co-transformed with pRH5′-HCV construct, the growth of the transformants on His⁻-3-AT plates and almost 6-fold increase in β-galactosidase activity over the background indicated that La protein does interact with the HCV IRES in the *in vivo* conditions also. To confirm that this interaction was because of the bridge formed by the MS2-HCV hybrid RNA and the B42-La protein, several controls were used. MS2-HCV with IRP did not show considerable growth at higher dilutions on His⁻-3-AT plates and there was no significant increase in β-gal activity. As expected, the positive controls, MS2-IRE and the B42-IRP hybrid showed strong interaction (Fig. 1*C*). MS2 alone with B42-La did not show significant β-gal activity (data not shown).

Similarly, to investigate whether the La-N or La-C can interact independently with the HCV IRES RNA *in vivo*, either the pYESTrp2La1–208 or pYESTrp2La209–408 vector was co-transformed with the pRH5′-HCV RNA hybrid vector. Interestingly, both La-N and La-C deletions demonstrated growth on selective plates (His⁻, 10 mM 3-AT) and modest increase (4–5-fold) in β-galactosidase activity over the negative controls, indicating significant interaction with the HCV IRES RNA (Fig. 1). Additionally, we have also investigated the La protein binding with the HCV-RNA *in vitro* by UV-cross-linking assays using the full-length and the deletions La1–208 (La-N) and La209–408 (La-C). Results showed that, in addition to the full-length protein, both the N- and C-terminal halves of La protein are able to interact with the HCV IRES *in vitro* (data not shown).

Both N- and C-terminal Halves of La Protein Are Necessary to Enhance HCV IRES-mediated Translation—Human La autoantigen has been shown to enhance HCV IRES-mediated translation. Previously, it has been demonstrated that a mutant La protein comprising the C-terminal half (La229–408) is capable of binding HCV 5′-UTR and stimulating the HCV IRES-mediated translation in RRL, albeit to a lesser extent compared with the full-length La protein (35). To further characterize the La protein requirement for the HCV IRES function, the effect of addition of purified N- and C-terminal halves of La proteins was determined by *in vitro* translation reactions. The bicistronic construct containing two reporter genes, *Renilla* luciferase and firefly luciferase, flanked by the HCV IRES was used in this experiment. The initiation of cap-independent translation, occurring internally from HCV 5′-UTR, resulted in the synthesis of firefly luciferase, whereas cap-dependent translation produced the *Renilla* luciferase. The capped bicistronic RNA was translated in the presence of increasing concentration of either full-length or the truncated La proteins; the luciferase activities were measured using the Stop and Glow dual luciferase assay system (Promega) and plotted against La concentration (Fig. 2*A*). The results showed gradual stimulation of luciferase activity up to ~4-fold with the increase in concentration (25, 50, and 75 ng; Fig. 2*A*, *lanes 2–4*) of full-length La protein. Furthermore, addition of either the N- or the C-terminal half of La protein (12.5, 25, and 37.5 ng, Fig. 2*A*, *lanes 5–7* and *8–10*) showed modest increase (up to 2.5–3-fold) in HCV IRES-mediated translation. The result suggests that the HCV IRES-mediated translation could be stimulated by specific interaction of independent regions of the La protein. However, the translation stimulation by either La-N or La-C was relatively less, compared with the full-length. Thus, we
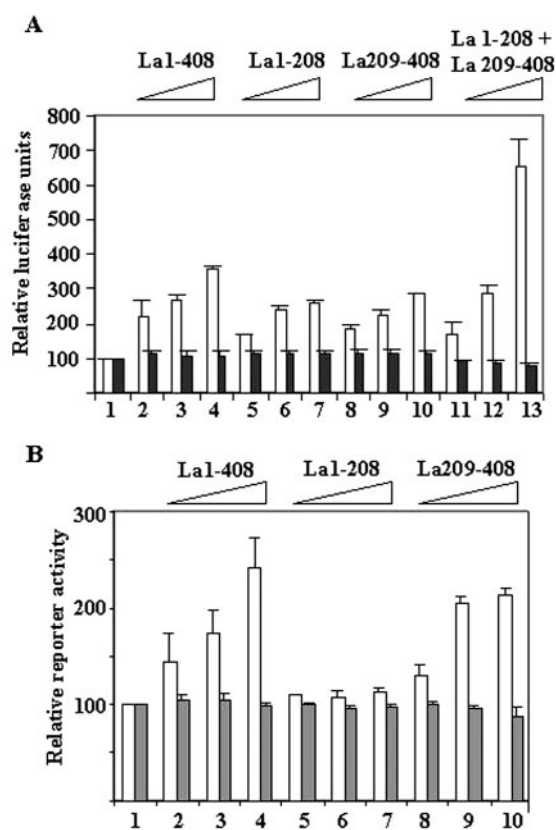
FIG. 2. **The effect of exogenous addition of purified full-length and truncated La proteins on the HCV and poliovirus IRES-mediated translation.** *A*, 1 μg of capped HCV bicistronic RNA was translated in RRL and supplemented with increasing concentrations of either purified His-tagged full-length La1–408 (25, 50, and 75 ng) or La1–208 (12.5, 25, and 37.5 ng) or La209–408 (12.5, 25, and 37.5 ng) or equimolar amounts of La1–208 and La209–408 proteins together (6.25, 12.5, or 18.75 ng each) as indicated. Both the reporter genes', *Renilla* luciferase and firefly luciferase, activities were measured, and the relative luciferase units were plotted. The *open bars* represent the firefly luciferase activity, and the *shaded bars* represent the *Renilla* luciferase activity. *B*, a similar experiment was performed with 1 μg of capped CAT poliovirus Luc bicistronic RNA supplemented with full-length La1–408 (250, 500, and 750 ng), La1–208 (125, 250, and 375 ng), or La209–408 (125, 250, and 375 ng) proteins. *Open bars* represent firefly luciferase activity, and *shaded bars* represent the CAT activity.

investigated whether an additive effect on HCV IRES-mediated translation could be achieved by adding both La-N and La-C together in the same reaction. For this, increasing concentrations of La-N and La-C proteins were added together in equimolar amounts (6.25, 12.5, and 18.75 ng each) in the translation reactions, and the luciferase activities were measured. The result showed gradual increase of the HCV IRES-mediated translation up to ~6.5-fold (Fig. 2*A*, *lanes 11–13*), which is equivalent to the level expected with the addition of full-length La protein. However, no significant increase was observed in the *Renilla* luciferase activity (Fig. 2*A*, *shaded bars*). The results suggest that both N- and C-terminal halves might play a critical role in achieving full stimulatory activity of the full-length protein.

To further explore the role of N- and C-terminal halves of La protein in poliovirus IRES-mediated translation, similar experiments were carried out using poliovirus bicistronic construct (Fig. 2*B*). Addition of increasing concentrations of full-length La protein (250, 500 and 750 ng; Fig. 2*B*, *lanes 2–4*) gradually enhanced poliovirus IRES-mediated translation of firefly luciferase up to 2.5-fold. Interestingly, addition of C-terminal half alone resulted in significant stimulation, up to 2.2-fold (Fig. 2*B*, *lanes 8–10*). However, the N-terminal half lacking the C-ter-

minal effector domain failed to stimulate the poliovirus IRES activity (Fig. 2*B*, *lanes 5–7*). Cap-dependent translations of CAT gene from the poliovirus bicistronic RNA were not altered with the addition of either the full-length or the deleted La proteins (Fig. 2*B*, *shaded bars*). Taken together, the results strongly suggest that, indeed, La protein enhances HCV IRES-mediated translation, and it appears that deletion of C-terminal effector domain does not abrogate the ability of La protein to enhance HCV IRES-mediated translation.

Computational Analysis Revealed a Consensus RNA Recognition Pattern in the RRMs of Known Structures—To understand the structural basis of the interaction of La protein with HCV RNA, we have searched for traditional RRMs (49–52) in the La protein using the motif search tools. The only region identified as RRM in the La protein ranges from residue position 105 to 200, which is referred to as RRM2. However, because the overall sequence similarity with proteins of known structure is low, we confirmed the similarity by comparing predicted secondary structure of RRM2 with the observed secondary structures in the RRM-containing proteins of known structure. Correspondence between observed and predicted secondary structures is excellent (data not shown), supporting the view that the region 105–200 could be a RRM domain. We have assessed the capability of this predicted RRM2 to bind RNA using three-dimensional structural models. We have used the known related crystal and NMR structures of RNA bound complexes as the basis in modeling the three-dimensional structure of RRM2-RNA complex. We have also analyzed the other putative RRMs of La protein, RRM1 and RRM3, in a similar way.

To start with, we analyzed by comparing the known three-dimensional structural complexes of the RRM-containing proteins. As mentioned under "Experimental Procedures," not all the known structures of RRMs are available bound to RNA and the best sequence similarity between the putative RRMs in La protein and the known structures is as low as 14%, which is not significant. We have used the constraints that the structures useful for modeling and analysis should be bound to RNA and they should show best compatibility between sequence of RRMs in La with the observed structural features such as solvent accessibility and secondary structure. The known structures that passed these conditions are: spliceosomal U2B″-U2A′ protein (Ref. 53; code 1a9n), human U1A protein (Ref. 48; code 1aud), hairpin ribozyme inhibitor (Ref. 54; code 1hp6), and U1A spliceosomal protein (Ref. 55; code 1urn). However, it should be noted that 1aud, 1hp6, and 1urn practically refer to the same protein, although the lengths and sequences of bases in the RNA molecules bound to the known structural complexes vary substantially. Despite being practically identical sequences, 1aud, 1hp6, and 1urn represent independent complex structure determination. We had ensured in the comparative modeling that we take only one of these three structures apart from 1a9n in generating the framework of the model. Thus, although these three structures were considered for comparative analysis, we have ensured that our modeling and other results are not biased.

A detailed investigation of each of these complex structures revealed that, apart from a short N-terminal region, two loops are involved in a series of interactions with the RNA. These interactions are extremely well conserved within the homologous protein-RNA complex structures whose structure-based sequence alignment is shown in Fig. 3. Fig. 4*A* shows the optimal superposition of the known structures with a consensus stretch of RNA. One of the RNA-interacting loops is centered around residue position 50 (1aud numbering, Fig. 3), which links two contiguous β-strands, and the other is located

FIG. 3. **Structure-based sequence alignment.** Sequences of RRM1, -2, and -3 of La protein were aligned along with the sequences of known structures of RRM proteins. The structural environments of the residues in the known structures are encoded in the representation: *uppercase*, solvent-inaccessible; *lowercase*, solvent-accessible; *italics*, positive ϕ, one of the Ramachandran angles; *bold*, side-chain hydrogen bonded to the main-chain amide; *underline*, side-chain hydrogen bonded to the main-chain carbonyl. The *numbers* within *parentheses* represent the first residue of the given protein in a block. Conserved α-helical and β-strand regions are also indicated. Figure was produced using JOY (60).
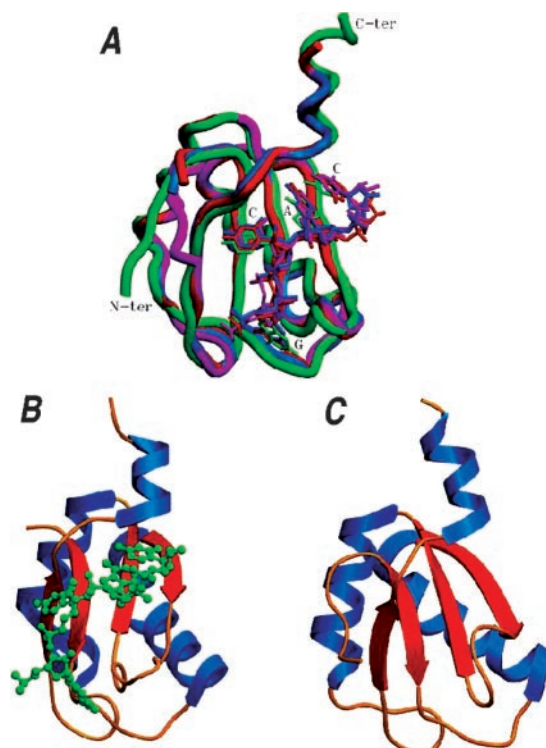

FIG. 4. **Computational analysis of RNA recognition motifs.** *A*, superposition of four known structures of RRM-containing proteins bound to RNA. The Cα trace of the proteins are shown in different *colors* after the optimal match of the Cα positions. The consensus GCAC base sequence regions of the bound RNA in these structures are also shown. High similarity in the region of RNA bound to the protein can be seen. *B* and *C*, three-dimensional models of RRM2 bound to the RNA with base sequence GCAC (*panel B*) and RRM3 (*panel C*). Helical and β-strand regions are highlighted in different *colors*. Slightly different conformations in the region of RNA binding of RRM2 and RRM3 can be noticed. *Panels A–C* were produced using SETOR (61).

around position 85, which links a β-strand with the C-terminal helix. A large majority of the interactions between the protein and RNA involve the polar atoms in the main chain of the polypeptide chain. Thus, although the RNA interacting regions are predominantly located in the loops, these are well conserved within the RRM proteins of known structure (Fig. 4A). It is well known that the length and conformation of the "equivalent" loops in homologous proteins vary enormously (42). However, it can be seen in Fig. 3 that, although the RNA binding regions are located in the loops, no insertion or deletion is present in the RNA-interacting loops. Although there are many residues conserved in the β-β loop, the sequence at the β-α loop is conserved as QYAKTDSD. These conservative features are probably the result of the need-based requirement of RNA binding.

Despite the fact that the bound RNA sequences and sizes vary enormously, the region of RNA interacting with the two loops of the protein is a stretch of four bases with a very well conserved conformation with an overall shape resembling an "L" shape (Fig. 4A). In three of the four structures, the sequence of this region of four bases is GCAC, and in 1a9n the sequence is GCAG. It can be concluded that the known structures of RRM proteins with RNA bound analyzed here represent a recognition pattern involving a highly conserved sets of loops in the protein and a four-base stretch of GCAC/G sequence in the RNA. Interestingly, a similar GCAC motif is observed in the HCV IRES RNA as well.

Structure-based Analysis of RNA Binding Capability of RRMs in the La Protein—The overall sequence similarity as seen in Fig. 3 is poor between RRM1 and the known structures with RRM. This feature could result in overall high structural variations in RRM1, even in case RRM1 adopts the overall fold as in the known structures. Significantly, in the β-β loop (around position 50 of 1aud), there is an insertion involving a Phe residue in RRM1. The loop conformation here is markedly

different from the conserved loop conformation seen in the known structures. Most importantly, in the C-terminal β-α region of RRM1, there are three prolyl residues in a stretch of six residues. Clearly, this stretch is likely to adopt a polyproline conformation, which is not appropriate in the place of a β-strand as lack of amide group in three places will preclude proximity of another β-strand to form a β-sheet structure. Obviously, a Pro-rich region cannot be accommodated in the helix as well that follows the β-strand in question. If the Pro-rich region forms a loop linking the β-strand and the following helix, then the polyproline conformation of the loop is very different from the bent loop conformation seen in the "equivalent" loop of the homologous proteins of known structure. Further, the main chain amide is unavailable, in three positions, for interacting with RNA. Thus, first of all it is not clear whether RRM1 would form the canonical RRM fold. Even if the fold of RRM1 is similar to that of known RRM folds, the putative RNA-interacting regions are likely to adopt conformations very different from those in the RRM proteins of known structure. Thus, the RRM1 of La, if it is capable of binding to RNA, is likely to bind in a novel mode. A three-dimensional structure-based explanation provided here is consistent with the proposals of Sobel and Wolin (56) and Ohndorf *et al.* (57).

From Fig. 3 it can be seen that the sequence of RRM2 of La protein involves no insertion or deletion at the two critical loop regions compared with RRM proteins of known structure. There are a few residue changes in the loop regions compared with the known structures. However, a detailed investigation of the model suggests that all the residues in the two loops of

RRM2 are comfortably placed in three dimensions. Fig. 4*B* shows the modeled structure of RRM2 bound to RNA. Almost all the residues in the two loops are generally well exposed to the surroundings in the known structures. The nature of substitution of residues seen in RRM2 is conservative, involving good accommodation of the loop sequences in the conformation seen in the equivalent regions of the known structures. Thus, the backbone conformations at the two RNA binding loops of RRM2 are likely to be similar to those seen in the homologous proteins of known structure. This will result in orientation of backbone amide and carbonyl groups at the loop regions similar to that seen in the known structures enabling RNA to bind at RRM2. Because of the predicted high similarity in the backbone structure of RRM2 with the closest homologues, it is suggested that RRM2 might bind to the GCAC motif in the HCV IRES RNA.

The overall sequence similarity of RRM3 of La protein with the RRMs of known structure is significant and is an indication of the retention of the fold. The sequence of the RRM3 could be comfortably modeled in the RRM fold (Fig. 4*C*); however, there is a deletion of a residue in the β-β loop, which should alter the conformation of the loop to some extent. Another putative RNA binding loop (β-α) has no deletion or insertion involved, and the nature of residues are such that the backbone conformation of this loop in RRM3 is likely to be similar to that of the equivalent region in the homologous structures. Thus, all of the backbone region involved in RNA binding is preserved except for a deletion in the β-β loop. This may not prevent the RNA binding, but may accommodate RNA with sequence motifs different from GCAC.

RRM2 of La Protein Has Greater Affinity for Binding with the HCV IRES RNA—RRM1 and RRM2 are located at the N-terminal half, and the RRM3 is located at the beginning of the C-terminal half of La protein (Fig. 1*A*). Our computational modeling data predict that the RRM2 is the most plausible RRM for the interactions with the HCV IRES. To test this hypothesis, we have performed filter-binding assay using the full-length and the La deletion proteins corresponding to each RRM: La1–100 (RRM1), La101–208 (RRM2), and La209–300 (RRM3). Full-length and truncated La proteins were over-expressed in *E. coli* (BL21) and purified by passing through an Ni-NTA column. ^{32}P-Labeled HCV IRES RNA (18–383 nt) probe (10 fmol) was separately incubated with increasing concentration of each of the deletions. The amount of bound RNA was determined by binding to nitrocellulose filters and plotted to obtain the saturation curve (Fig. 5*A*). The linear portions of each curve were subjected to regression analysis. The apparent dissociation constant (K_d) was calculated in each case as the protein concentration at which 50% of the RNA bound at saturation was retained on the filter. Results showed that La protein binds to HCV IRES RNA with relatively high affinity, with an apparent K_d of 0.12 μM. As predicted by the computer modeling data, the filter-binding assay results clearly demonstrated that RRM2 protein binds to HCV IRES RNA with relatively high affinity, with an apparent K_d of 0.14 μM, closely similar to that of full-length La protein. However, the protein corresponding to RRM3 showed relatively lower affinity with an apparent K_d of 0.23 μM. RRM1 protein alone failed to show significant binding with the HCV IRES RNA, as demonstrated before in an earlier report.

The computational analysis of the three-dimensional structures of the RRMs of the La protein predicted that RRM2 might interact with the HCV IRES around the GCAC motif near the initiator AUG. To confirm that, we have used HCV IRES mutant (HCV 5′-UTR 1–341) lacking the initiator AUG and GCAC motif as probe RNA in the filter-binding assays. The full-length
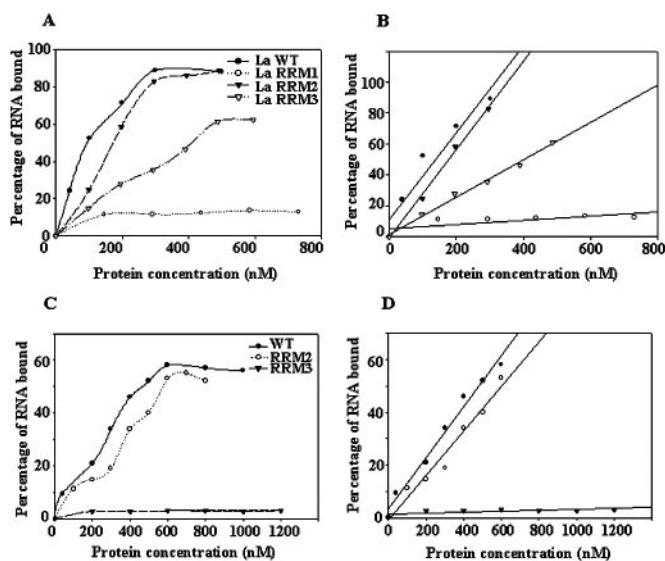

FIG. 5. **Filter-binding assay to determine the relative binding affinities of the RRMs.** *A*, [^{32}P]HCV IRES RNA (18–383 nt) was bound to increasing concentrations of either full-length or truncated La proteins corresponding to different RRMs (La1–100 (RRM1), La101–208 (RRM2), and La209–300 (RRM3)) as indicated on the *x* axis. The amount of bound RNA was determined by binding to nitrocellulose filter. The percentage of bound RNA was plotted against the protein concentration (nM). *B*, the linear region of the curves in *panel A* were plotted and the apparent K_d was calculated as the protein concentration at which 50% of RNA was retained. *C*, filter-binding assays were performed as described above using [^{32}P]HCV 5′-UTR RNA (1–341) and increasing concentrations of either full-length and deletion La proteins (La101–208 (RRM2) and La209–300 (RRM3)) as indicated. The percentage of bound RNA was plotted against the protein concentration (nM). *D*, the linear regions of the curves in *panel C* were plotted and the apparent K_d was calculated as the protein concentration at which 50% of RNA was retained.

La protein and RRM3 showed significant binding with the mutant HCV IRES with the apparent K_d of 0.24 and 0.37 μM, respectively. However, the RNA binding with RRM2 was drastically reduced possibly because of lack of contact points on the RNA (Fig. 5, *C* and *D*).

Toe-print Analysis Shows the Contact Points of RRM2 around GCAC Motif near Initiator AUG of the HCV IRES RNA—To precisely map the RRM2 contact points on the HCV IRES RNA, toe-printing assays were performed. Increasing concentrations of either purified full-length La protein or deletions were incubated with 5 pmol of *in vitro* transcribed RNA corresponding to the HCV IRES (18–383 nt). To this complex, ^{32}P-end-labeled primer complimentary to 24 nucleotides of the 3′ end of the HCV-383 was added and extended using avian myeloblastosis virus-reverse transcriptase. The resulting extended products were analyzed on a 6% polyacrylamide, 7 M urea denaturing gel (Fig. 6, *panels A* and *B*). For precise mapping of the contact points, a DNA sequencing reaction, corresponding to the HCV 383 RNA and using the same end-labeled primer for priming, was run alongside. The results demonstrated specific reverse transcriptase pauses (toe-prints) with the addition of increasing amount of proteins indicating possible protein binding sites. Four specific toe-prints corresponding to C-334, A-342, C-355, and A-364 were observed around the initiator AUG, which showed dose dependence with the addition of increasing concentrations of full-length La protein (Fig. 6*A*). RRM2 also showed similar toe-prints; however, prominent pauses were observed at A-342, C-355, and A-364, whereas, the contact point C-334 was relatively lesser (Fig. 6*B*). RRM3 did not show significant increase in the intensity of the toe-prints at these points. The results suggest that the primer extension was inhibited at several points around 342–364 nt, perhaps
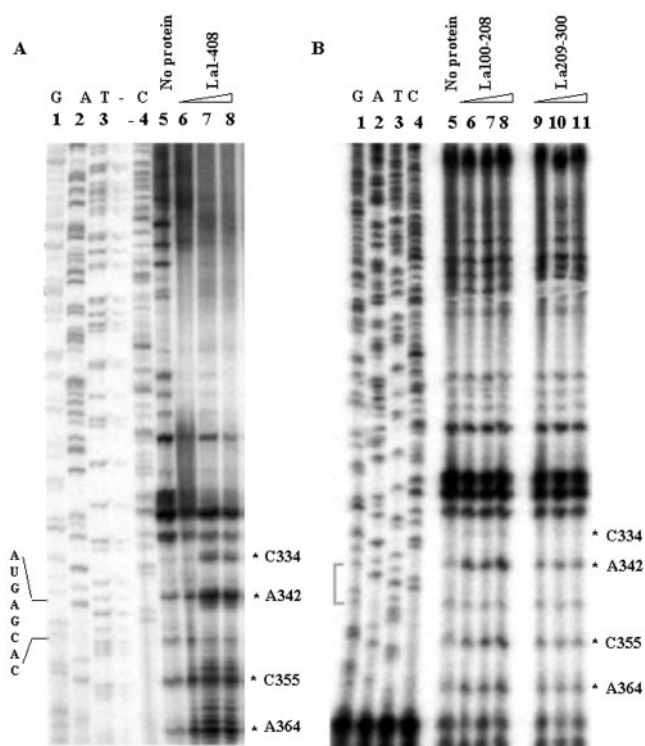
Role of La N- and C-terminal Halves on HCV IRES Function

FIG. 6. **Primer extension inhibition (toe-printing) analysis.** *A*, HCV IRES RNA (18–383 nt) was incubated with increasing amounts (200, 400, and 600 ng) of La full-length protein (*lanes 6–8*), as indicated *above* the *lanes*, and analyzed by primer extension. *Lane 5* shows the no protein control. The cDNA products terminated at the sites indicated on the *right*. The sequence corresponding to initiator AUG and GCAC is indicated on the *left* of the panel. *B*, similarly HCV IRES RNA (18–383 nt) was incubated in absence of any protein (*lane 5*) or presence of increasing concentrations of (200, 400, and 600 ng) of either La100–208 (RRM2) protein (*lanes 6–8*) or La209–308 (RRM3) protein (*lanes 9–11*). The major toe-prints are indicated on the *right* and the GCAC motif near AUG is marked on the *left* of the panel. *Lanes 1–4* in each panel show the DNA sequencing ladder corresponding to the HCV 18–383 RNA obtained by using the same end-labeled primer.

because of binding of RRM2 with this region encompassing the AUG (342–344 nt) and the GCAC motif (346–349 nt).

DISCUSSION

La protein interaction with the 5′-UTR has been shown to play a pivotal role in enhancing IRES-mediated translation of hepatitis C RNA. In this report we have extended the study and further characterized the interaction in relation to *trans*-activation of HCV IRES function. Previously it was shown that the C-terminal domain of La protein is responsible for binding with the HCV 5′-UTR, largely because of the presence of a stretch of basic amino acid residues in this region. In contrast, the N-terminal domain was indicated to play a critical role in interaction with poliovirus 5′-UTR.

In this report, we demonstrate for the first time the *in vivo* interaction of HCV IRES with La protein and also show that both N- and C-terminal halves are able to interact independently with the RNA (Fig. 1). HCV IRES-mediated translation could be stimulated by independent regions of La protein to some extent but an additive effect was obtained by supplementing La-N and La-C proteins, suggesting that both might play a critical role in achieving full stimulatory activity (Fig. 2). By computational analyses we predicted that RRM2 of La protein might bind HCV IRES at the GCAC motif near initiator AUG (Figs. 3 and 4), which was confirmed by filter-binding and toe-printing assays (Figs. 5 and 6).

The binding affinity of full-length La protein for the HCV

IRES (18–383) was found to be significantly high. The major contributor to the binding is apparently derived from RRM2, as suggested from the computational modeling data and the filter-binding assay results. RRM2 showed highest affinity (K_d of 0.14 μM) compared with RRM3 (K_d of 0.23 μM). RRM1 failed to show considerable binding affinity for the HCV IRES. This is consistent with our prediction made based on structural modeling, as well as the earlier observation that the La deletion 1–103 (comprising of RRM1 alone) did not interact with the HCV 5′-UTR. However, the deletion La1–160, where part of RRM2 was included, showed binding with the HCV 5′-UTR (35).

The region corresponding to amino acids 110–190 of human La protein represents a canonical RRM designated as RRM2, which is conserved in diverse organisms. The degree of conservation is consistent with the idea that RRM2 might contribute to overall affinity of the HCV IRES RNA. This is also supported by our observations in computational modeling, filter-binding, and toe-printing assays. Among the three putative RRMs in the La protein, the sequence pattern in RRM2 is observed to be the closest to the RRMs of known crystal or NMR structure. Highly conserved RNA binding regions of known structure coupled with analysis of RRM-RNA interactions prompts us to predict that RRM2 of La protein is likely to bind to GCAC motif, which is located in the HCV IRES. A partially conserved sequence motif, CACAA, has been observed to be present in several viral RNAs known to interact with the La protein (35). Our computational analysis revealed the presence of GCAC motif in all the known complex structures analyzed that showed structural homology with RRM2 of human La protein. Interestingly, a GCAC motif (346–349 nt) was detected immediate downstream of the initiator AUG and a GCAU sequence (337–340 nt) in the upstream. Primer extension inhibition assay (toe-printing) with the HCV IRES and La full-length and RRM2 demonstrated specific toe-prints at the AUG (342 nt) and few nucleotides downstream (355 and 364 nt). Because this region encompasses the GCAC (346–349 nt), it is conceivable that the RRM2 might contact the GCAC motif as predicted by computational modeling. Additionally, in the filter-binding assay when HCV IRES probe lacking the GCAC motif and AUG (1–341 nt) was used, the RRM2 binding was almost abrogated although it showed highest affinity for binding to HCV IRES RNA (18–383 nt). The results indicate that perhaps RRM2 does not bind GCAU in absence of AUG. However, the binding affinity of RRM3 with HCV IRES was only marginally affected in absence of GCAC and AUG. Additionally, RRM3 did not show any specific pause sites at the above region. It is possible that RRM3 binds to HCV 5′-UTR in a location different from the site that binds RRM2. Consistent with that fact, the full-length La protein was still able to bind to HCV 1–341 RNA probe (lacking AUG and GCAC motif) perhaps through RRM3. The fact that RRM3 binds to a different region other than RRM2 attests to the possible role of La protein as RNA chaperone to hold the RNA-protein complex at the IRES for proper positioning of the 40 S ribosome onto the initiator AUG.

Although RRM1 does not bind RNA on its own, it contributes significantly in recognition of UUU-OH-containing RNAs (52). Additionally, it has been demonstrated earlier that deletion of the N-terminal residues of RRM1 in mutants 22–408 and 26–408 decreased the affinity for binding to HIV leader RNA (29). Analysis of the known structures of RRMs shows highly similar structures for RRMs bound to RNA and unbound forms. Consistent with this observation, although RRM2 and RRM3 of La protein may not alter their domain structures significantly as a result of binding to RNA, it is possible that the relative orientation between the RRM domains of La could alter because of

gross structural changes to suit the orientation of bound RNA. Thus, RRM1 might influence the access of RNA to RRM2 and RRM3. Model of the three-dimensional structure of RRM3 reveals interesting differences in the RNA binding region. These changes are predicted not to alter the backbone conformation essential for RNA binding, but may alter the specificity to the binding base sequence.

In yeast three-hybrid analysis, longer RNA baits sometimes reduce the reporter gene signal. However, specific interactions have been detected with RNA baits as long as 1600 nt (58). Keeping in mind the multiple contact points of La protein within HCV 5′-UTR, to test the interaction with the truncated La proteins, we preferred to use a longer RNA (HCV 18–383) as bait rather than using smaller deletions. We did observe significant increase in β-galactosidase activity over the control, although the signals were not as high as expected for optimal length of RNA bait. Additionally, to avoid signals from nonspecific interaction, we have investigated the growth pattern of the transformed yeast in presence of 10 mM 3-AT and included various negative controls. In addition, MS2-HCV (18–383) RNA hybrid did not show any interaction with another RNA-binding protein, IRP, indicating high specificity of the assay.

Our *in vitro* translation results showed dose-dependent stimulation of both HCV and poliovirus IRES-mediated translation with the addition of increasing concentration of full-length purified La protein. However, we have deliberately used a lower range of protein concentrations to monitor the marginal differences (if any) on the HCV IRES-mediated translation upon addition of the truncated La proteins. Perhaps because of this reason, in our assay we could observe stimulation only up to ~4-fold (below saturation level) upon addition of full-length La protein as opposed to earlier reports, where maximum 6-fold increase was demonstrated (35). Interestingly, addition of La-N or La-C protein alone could not stimulate the HCV IRES-mediated translation to the extent observed with full-length La protein. However, addition of La-N and La-C together in the same reaction could achieve ~6.5-fold stimulation of HCV IRES activity, as expected with full-length La protein. This observation indicates that perhaps both the terminal halves are involved in some critical protein interactions. Deletion of either half loses the contribution of some of the factors in stimulating HCV IRES activity.

The C-terminal half (termed as the effector domain) of La protein has been shown to possess the dimerization domain, and La homodimerization is essential for the *trans*-activation of poliovirus IRES function (41). The fact that the N-terminal half alone was capable of stimulating HCV IRES-mediated translation implies that perhaps La protein dimerization may not be an essential prerequisite for the HCV-RNA binding activity and *trans*-activation of HCV IRES-mediated translation.

HCV IRES requires a limited number of *trans*-acting factors as compared with the picornavirus IRES. The dimerization domain in the C-terminal half might be involved in protein-protein interaction with other *trans*-acting factors that could be critical for poliovirus IRES-mediated translation but may not be required for HCV IRES-mediated translation. Alternatively, a long helix-rich region has been predicted near the center of the La protein (59), which is almost equally divided into the N- and C-terminal deletions. The portions of this predicted helix could also exist in a stable conformation and can mediate protein-protein interactions. Further experiments are in progress to distinguish between the possibilities, which would shed more light on the actual mechanism of the La protein-mediated stimulation of IRES-mediated translation.

REFERENCES

1. Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W., and Houghton, M. (1989) *Science* **244,** 359–362
2. Houghton, M., Weiner, A., Han, J., Kuo, G., and Choo, Q.-L. (1991) *Hepatology* **14,** 381–388
3. Saito, I., Miyamura, T., Ohbayashi, A., Harada, H., Katayama, T., Kikuchi, S., Watanabe, Y., Koi, S., Onji, M., Ohta, Y., Choo, Q.-L., Houghton, M., and Kuo, G. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 6547–6549
4. Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T., and Shimotohno, K. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 9524–9528
5. Tsukiyama-Kohara, Z., Iizuka, N., Kohara, M., and Nomoto, A. (1992) *J. Virol.* **66,** 1476–1483
6. Wang, C., Sarnow, P., and Siddiqui, A. (1993) *J. Virol.* **67,** 3338–3344
7. Reynolds, J. E., Kaminski, A., Carroll, A. R., Clarke, B. E., Rowlands, D. J., and Jackson, R. J. (1996) *RNA* **2,** 867–878
8. Lu, H.-H., and Wimmer, E. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 1412–1417
9. Kamoshita, N., Tsukiyama-Kohara, K., Kohara, M., and Nomoto, A. (1997) *Virology* **233,** 9–18
10. Rijnbrand, R., Bredenbeek, P., van der Straaten, T., Whetter, L., Inchauspe, G., Lemon, S., and Spaan, W. (1995) *FEBS Lett.* **365,** 115–119
11. Rijnbrand, R. C. A., and Lemon, S. M. (2000) *Current Topics in Microbiology and Immunology*, pp. 85–111, Springer-Verlag, Berlin
12. Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J., and Hellen, C. U. (1998) *Genes Dev.* **12,** 67–83
13. Ali, N., and Siddiqui, A. (1995) *J. Virol.* **69,** 6367–6375
14. Ali, N., and Siddiqui, A. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 2249–2254
15. Fukushi, S., Okada, M., Kageyama, T., Hoshino, F. B., Nagai, K., and Katayama K. (2001) *Virus Res.* **73,** 67–79
16. Hahm, B., Kim, Y. K., Kim, J. H., Kim, T. Y., and Jang, S. K. (1998) *J. Virol.* **72,** 8782–8788
17. Tan, E. M. (1989) *Adv. Immunol.* **44,** 93–151
18. Yoo, C. J., and Wolin, S. L. (1994) *Mol. Cell. Biol.* **14,** 5412–5424
19. Gottlieb, E., and Steitz, J. A. (1989) *EMBO J.* **8,** 841–850
20. Gottlieb, E., and Steitz, J. A. (1989) *EMBO J.* **8,** 851–861
21. Maraia, R. J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 3383–3387
22. Maraia, R. J., and Intine, R. V. (2001) *Mol. Cell. Biol.* **21,** 367–379
23. Madore, S. J., Wieben, E. D., and Pederson, T. (1984) *J. Biol. Chem.* **259,** 1929–1933
24. Ford, L. P., Sway, J. W., and Wright, W. E. (2001) *RNA* **7,** 1068–1075
25. Francoeur, A. M., and Mathews, M. B. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 6772–6776
26. Kurilla, M., and Keene, J. D. (1983) *Cell* **34,** 837–845
27. Park, Y.-W., and Katze, M. G. (1995) *J. Biol. Chem.* **270,** 28433–28439
28. Pardigon, N., and Strauss, J. H. (1996) *J. Virol.* **70,** 1173–1181
29. Chang, Y.-N., Kenan, D. J., Keene, J. D., Gatignol, A., and Jeang, K.-T. (1994) *J. Virol.* **68,** 7008–7020
30. Meerovitch, K., Svitkin, Y. V., Lee, H. S., Lejbkowicz, F., Kenan, D. J., Chan, E. K., Agol, V. I., Keene, J. D., and Sonenberg, N. (1993) *J. Virol.* **67,** 3798–3807
31. Kim, Y. K., and Jang, S. K. (1999) *J. Gen. Virol.* **80,** 3159–3166
32. Kim, Y. K., Back, S. H., Jungmin, R., Lee, S. H., and Jang, S. K. (2001) *Nucleic Acids Res.* **29,** 5009–5016
33. Spangberg, K., Wiklund, L., and Schwartz, S. (2001) *J. Gen. Virol.* **82,** 113–120
34. Das, S., Ott, M., Yamane, A., Tsai, W., Gromier, M., Lahser, F., Gupta, S., and Dasgupta, A. (1998) *J. Virol.* **72,** 5638–5647
35. Ali, N., Pruijn, G. J. M., Kenan, D. J., Keene, J. D., and Siddiqui, A. (2000) *J. Biol. Chem.* **275,** 27531–27540
36. Isoyama, T., Kamoshita, N., Yasui, K., Iwai, A., Shiroki, K., Toyoda, H., Yamada, A., Takasaki, Y., and Nomota, A. (1999) *J. Gen. Virol.* **80,** 2319–2327
37. Holcik, M., and Korneluk, R. G. (2000) *Mol. Cell. Biol.* **20,** 4648–4657
38. Svitkin, Y. V., Pause, A., and Sonenberg, N. (1994) *J. Virol.* **68,** 7001–7007
39. Goodier, J. L., Fan, H., and Maraia, R. J. (1997) *Mol. Cell. Biol.* **17,** 5823–5832
40. Svitkin, Y. V., Meerovitch, K., Lee, H. S., Dholakia, J. N., Kenan, D. J., Agol, V. I., and Sonenberg, N. (1994) *J. Virol.* **68,** 1544–1550
41. Craig, A. W. B., Svitkin, Y. V., Lee, H. S., Belsham, G. J., and Sonenberg, N. (1997) *Mol. Cell. Biol.* **17,** 163–169
42. Srinivasan, N., and Blundell, T. L. (1993) *Protein Eng.* **6,** 501–512
43. Sali, A., and Blundell, T. L. (1990) *J. Mol. Biol.* **212,** 403–428
44. Zhu, Z.-Y., Sali, A., and Blundell, T. L. (1992) *Protein Eng.* **5,** 43–51
45. Topham, C. M., McLeod, A., Eisenmenger, F., Overington, J. P., Johnson, M. S., and Blundell, T. L. (1993) *J. Mol. Biol.* **229,** 194–220
46. Sutcliffe, M. J., Hayes, F., and Blundell, T. L. (1987) *Protein Eng.* **1,** 385–392
47. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984) *J. Am. Chem. Soc.* **106,** 765–784
48. Allain, F. H.-T., Howe, P. W. A., Neuhaus, D., and Varani, G. (1997) *EMBO J.* **16,** 5764–5772
49. Query, C. C., Bentley, R. C., and Kenan, D. J. (1989) *Cell* **57,** 89–101
50. Kenan, D. J., Query, C. C., and Keene, J. D. (1991) *Trends Biochem. Sci.* **16,** 214–220
51. Birney, E., Kumar, S., and Krainer, A. R. (1993) *Nucleic Acids Res.* **21,** 5803–5816

52. Maraia, R. J., Kenan, D. J., and Keene, J. D (1994) *Mol. Cell. Biol.* **14,** 2147–2158
53. Price, S. R., Evans, P. R., and Nagai, K. (1998) *Nature* **394,** 645–650
54. Rupert, P. B., and Ferre-d'Amare, A. R. (2001) *Nature* **410,** 780–786
55. Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., and Nagai, K. (1994) *Nature* **372,** 432–438
56. Sobel, S. G., and Wolin, S. L. (1999) *Mol. Biol. Cell* **10,** 3849–3862
57. Ohndorf, U.-M., Steegborn, C., Knijff, R., and Sondermann, P. (2001) *J. Biol.*
58. Martin, F., Michel, F., Zenklusen, D., Muller, B., and Schumperli, D. (2000) *Nucleic Acids Res.* **28,** 1594–1603
59. Chambers, J. C., Kenan, D., Martin, B. J., and Keene, J. D. (1988) *J. Biol. Chem.* **263,** 18043–18051
60. Mizuguchi, K., Deane, C. M., Johnson, M. S., Blundell, T. L., and Overington, J. P. (1998) *Bioinformatics* **14,** 617–623
61. Evans, S. V. (1993) *J. Mol. Graph.* **11,** 134–138

Chem. **276,** 27188–27196