CELLULAR PROTEIN hnRNP-A1 INTERACTS WITH THE 3'-END AND THE INTERGENIC SEQUENCE OF MOUSE HEPATITIS VIRUS NEGATIVE-STRAND RNA TO FORM A RIBONUCLEOPROTEIN COMPLEX

Xuming Zhang,[1,4] Hsin-Pai Li,[2] Wenmei Xue,[2] and Michael M. C. Lai[1,2,3]

[1]Department of Neurology
[2]Department of Molecular Microbiology and Immunology
[3]Howard Hughes Medical Institute
 University of Southern California School of Medicine
 Los Angeles, California 90033
[4]Department of Microbiology and Immunology
 University of Arkansas for Medical Sciences
 Little Rock, Arkansas 72205

1. ABSTRACT

We previously showed that several cellular proteins specifically bind to the 3'-end and the intergenic sequences of the negative-strand RNA of mouse hepatitis virus (MHV), and proposed that these distant RNA sequences can be brought together by cellular and viral proteins (Furuya and Lai, 1993; Zhang et al., 1994; Zhang and Lai, 1995). The cellular protein p35 has been identified as a heterogeneous nuclear ribonucleoprotein (hnRNP) A1. We have now expressed hnRNP-A1 as a glutathione-S-transferase (GST) fusion protein and demonstrated that the amino terminal two-thirds of hnRNP-A1 interacted with the two MHV regulatory RNA sequences (3'-end and intergenic sequences) through protein-RNA interaction while its carboxy-terminal glycine-rich domain mediated homomeric (protein-protein) interactions. In a partially reconstituted reaction, in which the two MHV RNA fragments and the purified GST-hnRNP-A1 fusion protein were mixed, an RNP complex was formed. Depletion of either hnRNP-A1 or one of the RNA components abolished the complex formation. These results indicate that hnRNP-A1 can mediate the formation of an MHV RNP complex, which includes both the negative-strand leader and intergenic sequences. Site-directed mutagenesis revealed that mutations in the MHV intergenic se-

quences, which inhibited MHV RNA transcription, also inhibited the RNP complex formation. Deletion analysis showed that the amino terminal RNA-binding domains of hnRNP-A1 is essential for the RNP complex formation while the carboxy-terminal protein-binding domain enhanced the complex formation by 90-fold. These findings provide direct evidence demonstrating that the negative-strand leader RNA and intergenic sequences can form an RNP complex mediated by cellular protein hnRNP-A1.

2. INTRODUCTION

Heterogeneous nuclear ribonucleoprotein (hnRNP) is a collective term representing a number of proteins (termed A to U) that are associated with pre-mRNAs or hnRNAs to form RNP complexes in mammalian cells (Dreyfuss et al. 1993). hnRNP A1 is the first and best characterized hnRNP protein. It contains 320 amino acids deduced from the cDNA sequence (Buvoli et al., 1988) with an estimated molecular weight of 34 kDa. Its N-terminal two-thirds (residues 1 to 196) forms two tandemly arranged RNA-binding domains (RBD) which bind to the 5'- and 3'-splicing consensus sequences of pre-mRNAs (Dreyfuss et al., 1993). The C-terminal domain (residues 197 to 320) is very rich in glycine (40%), and is designated as a glycine-rich domain. This domain contains a 36 residue sequence analogous to an "RGG box", an RNA-binding motif first identified in hnRNP U (Kiledjian and Dreyfuss, 1992). Recently, it has been shown that the glycine-rich domain mediates the protein-protein interaction of A1 with itself and with other hnRNP basic "core" proteins and some SR-proteins in vitro (Cartegni et al., 1996). Downstream of the RGG box is the M9 domain, which was shown to mediate both nuclear import and export of the A1 protein (Siomi and Dreyfuss, 1995; Weighardt et al., 1995, Pollard et al., 1996). A1 has also been shown to be involved in RNA-reannealing and alternative RNA-splicing. Its diverse and specific protein-RNA and protein-protein interactions have been implicated to function in mRNA maturation in mammalian cells (Dreyfuss et al., 1993).

We previously showed that two cellular proteins with molecular weights of 35 and 38 kDa (p35/38) specifically bound to the 3'-end and the intergenic (IG) sequences of the negative-strand RNA of mouse hepatitis virus (MHV) (Furuya and Lai, 1993; Zhang and Lai, 1995). By performing site-directed mutagenesis of the IG sequence, we further demonstrated that p35/38 is possibly involved in the regulation of MHV RNA transcription, since the efficiency of p35/38 binding to the IG sequence correlated with the amount of mRNAs transcribed (Zhang and Lai, 1995). We hypothesized that these regulatory RNA sequences of MHV can be brought together by cellular and viral proteins to form a transcription initiation complex (Zhang et al, 1994; Zhang and Lai, 1995). Recently, we have isolated the cellular protein p35 and identified it as hnRNP-A1 (Li et al., 1997). Protein expression and UV cross-linking experiments further confirmed that hnRNP-A1 had the same RNA-binding properties to MHV RNAs as p35/38.

In the present study, we investigated the role of the cellular protein hnRNP-A1 (previously designated p35) in the formation of an MHV RNP complex by analyzing the interactions between hnRNP-A1 and MHV RNAs. We present direct evidence demonstrating that both protein-RNA and protein-protein interactions are involved in the formation of an MHV RNP complex. Our results demonstrated that the RNA-binding domain of hnRNP-A1 mediate the binding of hnRNP-A1 to both the 3'-end and the IG sequence of the negative-strand RNA of MHV, whereas the carboxy-terminal glycine-rich domain exclusively involved in protein-protein interaction. Furthermore, mutations of the intergenic sequence of MHV RNA, which inhibited the protein-RNA interaction, also inhibited the RNP com-

plex formation. Deletion analyses of hnRNP-A1 protein indicated that protein-RNA interaction is required for an RNP complex formation while the protein-protein interaction significantly enhanced the complex formation. This study establishes the important components in the proposed model involving protein-RNA interaction, and thus contributes to the understanding of the complex mechanism of coronavirus mRNA transcription.

3. MATERIALS AND METHODS

3.1. Construction of cDNA Clones

The cDNA encoding the murine hnRNP A1 gene was synthesized from RNAs isolated from a murine astrocytoma cell line (DBT) by reverse transcription (RT) and polymerase chain reaction (PCR) according to the procedures described previously (Zhang et al., 1994). It was then cloned into the *Bam*HI and *Eco*RI sites of pBluescript and pGEX 4–1, resulting in pBS-mA1 and pGST-mA1, respectively. To generate deletion constructs of A1, pGST-mA1 DNA was used as a template for PCR amplification using various sets of primers. These PCR fragments were then digested with *Bam*HI and *Eco*RI and directionally cloned into the pGEX4–1, resulting in GST-fusion plasmids containing various domains of A1.

3.2. Expression of GST Fusion Proteins

The procedure for expressing GST fusion proteins was as described previously (Smith and Johnson, 1988; Ausubel et al., 1989).

3.3. In Vitro Transcription and in Vitro Translation

In vitro transcription and translation were carried out according to the methods recommended by the manufacturer (Promega).

3.4. UV Cross-Linking of RNA-Protein Complex

The UV crosslinking study was carried out to determine the interaction of hnRNP-A1 protein with MHV RNA sequences. The purified GST-fusion proteins were used as a protein source, and the ^{32}P-labeled, *in vitro*-transcribed MHV RNA was used as a probe. The procedure for UV crosslinking experiments was essentially as described previously (Zhang and Lai, 1995).

3.5. Protein Binding Assay

A protein binding assay was performed to determine the protein-protein interaction as described previously (Ausubel et al., 1989).

3.6. In Vitro Formation and Detection of the RNP Complex

The *in vitro* formation and detection of the RNP complex was based on the procedures described by Wills (1996) and Lamond and Sproat (1996) with slight modifications. Briefly, three basic components [the *in vitro*-transcribed cold RNA fragment of the 3'-end

182 nt sequence, XN182(-); the *in vitro*-transcribed ^{32}P-labeled intergenic sequence of 118 nt, ^{32}P-IG7-SM118(-); and the purified GST-A1 fusion protein] were mixed in a reaction buffer of 100 µl containing 300 mM potassium chloride, 1.5 mM ATP, 5 mM creatine phosphate, 10 mM creatine phosphokinase, 2 mM magnesium chloride, and 40 U of RNasin. The mixture was incubated at 30 oC for 1 h. To identify RNP complex formation, the biotin-labeled riboprobe [biotin-56(+)], which is complementary to the 3'-end of XN182(-) RNA, was added to the reaction, and the incubation continued for an additional hour. Preblocked streptavidin-agarose beads (Sigma) were added to the reaction and incubated at 4 oC on a rocking platform for 1 h, followed by 4 washes with wash buffer WB300 (20 mM HEPES-potassium hydroxide, pH7.9, 300 mM potassium chloride, and 0.1% nonidet P-40). After adding 200 µl of Bonner's RNA extraction buffer (7 M urea, 350 mM sodium chloride, 10 mM Tris, pH8.0, 10 mM ethylenediamine tetraacetic acid, 1% sodium dodecyl sulfate) and 20 µl of 0.2 mM glycogen, RNA was extracted with phenol/chloroform, and analyzed by denaturing PAGE containing 6% polyacrylamide as described (Lamond and Sproat, 1996). The dried gel was autoradiographed and/or scanned in a PhosphorImager and the amount of radioactivity of specific bands was determined.

4. RESULTS AND DISCUSSION

4.1. Mapping the RNA- and Protein-Binding Domains of hnRNP-A1

Although several functional domains of hnRNP-A1 have been identified, it is not known which domain(s) specifically binds MHV RNA. To understand how an MHV RNP complex is formed, it will be necessary to determine which domains of hnRNP-A1 bind MHV regulatory RNA sequences. For this purpose, we employed PCR to generate 9 deletion fragments by deleting one or more domains of hnRNP-A1, and expressed them in *E. coli* as GST-fusion proteins. We then used purified GST fusion proteins and one of the two ^{32}P-labeled riboprobes [NX182(-) RNA or IG7-SM118(-) RNA] for UV crosslinking studies. As shown in Fig. 1, the full-length GST-A1 and five carboxy-terminal deletion GST-fusion proteins (ΔIX, ΔCT, 2xRBD, RBD I, RBD II) bound to both XN182(-) and IG7-SM118 (-) RNAs. In contrast, all GST fusion proteins with amino-terminal deletions (Gly, RGG, CT, and VIII) did not bind to the two RNAs. These results are consistent with the previous findings on the general RNA-binding properties of hnRNP-A1 (Dreyfuss et al., 1993). Surprisingly, the RGG box, which has been shown to contribute to RNA-binding and to be an RBD in hnRNP-U (Kiledjian and Dreyfuss, 1992), did not bind to MHV RNA. We conclude from these results that the amino-terminal two RBDs are the only domains capable of binding the complementary strand of MHV 5'-untranslated region and the IG sequence.

We further examined the protein-binding sites of hnRNP-A1 using the full-length and truncated GST-fusion proteins and the *in vitro*-translated, ^{35}S-labeled full-length hnRNP-A1 in a protein-binding assay. As shown in Fig. 1, GST-A1 fusion proteins containing the complete or portions of the glycine-rich domain retained the protein-binding activity, whereas GST-A1 fusion proteins lacking this domain did not. These results demonstrate that the carboxy terminal glycine-rich domain of hnRNP-A1 possesses the protein-binding activity with itself, consistent with the published data on human hnRNP-A1 (Cartegni et al., 1996). The binding specificity was further confirmed by using GST protein as a negative control. No background binding activity was observed between GST and hnRNP-A1 under this experimental condition (data not shown).

A

2xRBD   Gly

| RBD I | RBD II | RGG | CT | *Clone* |

B

Binding property of A1

	NX182 (-) RNA	IG(-) RNA	A1 protein
A1	+	+	+
ΔIX	+	+	+
ΔCT	+	+	+
2xRBD	+	+	−
RBD I	+	+	−
RBD II	+	+	−
Gly	−	−	+
RGG	−	−	+
CT	−	−	+
VIII	−	−	+

Figure 1. Summary of deletion analyses on protein-RNA and protein-protein interactions. (A) Schematic diagram of the structure of hnRNP-A1 and its deletion derivatives, which are expressed as glutathione-*S*-transferase (GST)-fusion proteins. RBD, RNA-binding domain; Gly, glycine-rich domain; RGG, arginine-glycine-glycine box; CT, C-terminus. The name of each construct is shown under *Clone* which reflects its functional domain(s). (B) Interactions between GST-A1 fusion proteins and NX182(-) RNA or IG(-) RNA, and between GST-A1 proteins and in vitro-translated A1 protein were determined by UV cross-linking and protein-binding assays, respectively. Bound proteins were analyzed by SDS-PAGE containing 10% polyacrylamide. (+), positive reaction; (-), negative reaction.

4.2. In Vitro Formation of an RNP Complex

We previously hypothesized that cellular and viral proteins and coronavirus regulatory RNA sequences form a transcription initiation complex by protein-RNA and protein-protein interactions, thus regulating viral RNA transcription (Zhang et al., 1994). To test this hypothesis, we developed an in vitro assay to determine whether the two distant *cis*-acting regulatory sequences of MHV RNA can form a complex with cellular protein hnRNP-A1. When the two RNA fragments [IG7-SM118(-) RNA was labeled with ^{32}P and XN182(-) RNA was unlabeled] were mixed with purified GST-A1 protein, an RNP complex containing these three basic components was formed (Fig. 2, lane 1). This complex was then isolated by affinity selection with a biotin-labeled antisense riboprobe, biotin-56(+), which is complementary to the 3'-end 56 nt of the unlabeled XN182(-) RNA and subsequently with streptavidin agarose beads. When GST-A1 was replaced with GST or omitted, no specific RNP complexes were formed (Fig. 2, lane GST; data not shown). Similarly, when either XN182(-) RNA or biotin-56(+) was omitted in the reaction, no specific RNP complexes were formed (data not shown). We conclude that the two MHV RNA sequences can form an RNP complex by cellular protein hnRNP-A1. This is the first demonstration of such an RNP complex in vitro.

4.3. Mutations of the IG Sequence Affect RNP Complex Formation

To determine whether protein-RNA interaction is critical in RNP complex formation, we chose two mutant MHV IG sequences, MG(-) and MGG(-), which have been shown previously to abrogate the protein-RNA interaction and transcription efficiency (Zhang and Lai, 1995). MG(-) has one G mutation in the consensus sequence (UCUAAA*C* to UCUAAA*G*) while MGG(-) has two G mutations (UC*U*AAAC to UC*G*AAA*G*). The amounts of the RNP formed in MG(-) and MGG(-) were 16 and 9% of the wild-type, re-

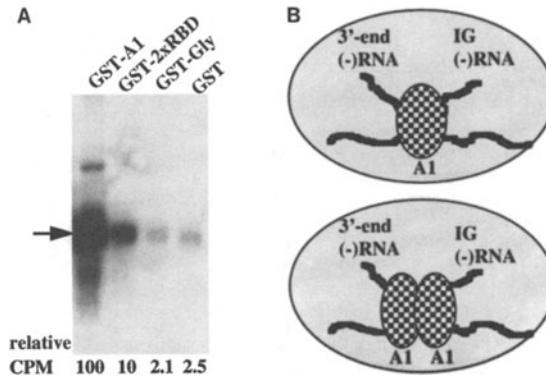**Figure 2.** Protein-RNA and protein-protein interactions are involved in the formation of an MHV RNP complex in vitro. (A) In vitro RNP complex formation. The full-length and two deletion constructs of hnRNP-A1 were expressed in *E. coli* as GST-fusion proteins, which were then purified with glutathione-Sepharose beads and used for studies of RNP complex formation. After PAGE separation, gels were autoradiographed. The ^{32}P-labeled IG7(-) RNA recovered from the complex is indicated by an arrow. The radioactivity of this band in each lane was determined by a PhosphorImager (Life Technology), and the absolute CPMs were converted to relative CPMs to the value of the full-length hnRNP-A1, which was set to 100 (shown at the bottom). The proteins used in each reaction are indicated at the top. (B) Two possible models of the RNP complex formation in vitro.

spectively (data not shown), indicating that protein-RNA interaction is important for the RNP complex formation.

4.4. Effects of Deletion of hnRNP-A1 on RNP Complex Formation

We further determined whether both protein- and RNA-binding activities of hnRNP-A1 are required for the formation of an RNP complex. To do this, we performed the RNP reconstitution reaction by mixing the truncated GST-fusion proteins with the two RNA fragments [NX182(-) RNA and ^{32}P-IG7(-) RNA] in a reconstitution reaction buffer. As shown in Fig. 2A, the RNP complex was significantly reduced (10 % of the full-length) when the complete carboxy-terminal glycine-rich domain was deleted (lane GST-2xRBD), while almost no RNP complex was formed (2.1% of the full-length), when the amino-terminal RBD was deleted (lane GST-Gly). The radioactivity of the complex formed by GST-Gly was comparable to that by GST (lane GST, 2.5% of the full-length). These results indicate that the RBDs are essential for the formation of the RNP complex, again suggesting the importance of the protein-RNA interaction in RNP complex formation. They also indicate that the glycine-rich domain is not absolutely required for, but enhances the RNP complex formation, suggesting that protein-protein interaction is an auxiliary function in RNP complex formation in vitro.

These results suggest two possible senarios in the formation of an RNP complex (Fig. 2B). One is that a single hnRNP-A1 protein is able to bind to the two separate RNA sequences (3'-end and intergenic sequences of the negative-strand RNA) to form a complex without the involvement of protein-protein interaction. This is most likely to be mediated by the two RNA-binding domains of hnRNP-A1 through direct protein-RNA interaction, as it was shown that each RNA-binding domain independently binds the 3'-

end and intergenic sequences (Fig. 1). On the other hand, the formation of such a complex may require two hnRNP-A1 proteins through protein-protein interaction, as evidenced by the deletion analysis (Fig. 2A). Both possibilities may operate in vitro and in vivo. It should be noted that in our *in vitro*-RNP reaction, the amounts of RNAs and hnRNP-A1 protein may be significantly large as compared to the intracellular milieu, in which viral RNAs are restricted to certain compartments and are surrounded by diverse macromolecules in excess; the cis-acting sequences of MHV RNA might not be easily accessible to a single hnRNP-A1 protein molecule as shown here *in vitro*. Thus, we can envisage that if such a complex were formed *in vivo*, protein-protein interactions of hnRNP-A1 would have a greater role. This assumption is consistent with the findings that protein interactions of hnRNP-A1 are essential in the assembly of a splicing complex in mammalian cells by associating with other hnRNP and SR proteins (Dreyfuss et al., 1993). Whether other MHV RNA sequences and viral and cellular proteins are also involved in the formation of an MHV RNP complex remains to be determined.

ACKNOWLEDGMENTS

REFERENCES

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., 1994, *Current protocols in molecular biology*. John Wiley & Sons Inc., New York.

Buvoli, M., Biamonti, G., Tsoulfas, P., Bassi, M. T., Ghetti, A., Riva, S., and Morandi, C., 1988, cDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms, *Nucl. Acids Res.* 16:3751–3770.

Cartegni, L., Maconi, M., Morandi, E., Cobianchi, F., Riva, S., and Biamonti, G., 1996, hnRNP A1 selectively interacts through its Gly-rich domain with different RNA-binding proteins, *J. Mol. Biol.* 259:337–348.

Dreyfuss, G., Matunis, M.J., Pinol-Roma, S., and Burd, C.G, 1993, hnRNP proteins and the biogenesis of mRNA, *Annu. Rev. Biochem.* 62:289–321.

Furuya, T. and Lai, M.M.C., 1993, Three different cellular proteins bind to the complementary sites on the 5'-end positive- and 3'-end negative-strands of mouse hepatitis virus RNA, *J. Virol.* 67:7215–7222.

Kiledjian, M. and Dreyfuss, G., 1992, Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box, *EMBO J.* 11: 2655–2664.

Lamond, A., and Sproat, B.S., 1996, Isolation and characterization of ribonucleoprotein complexes. p103–140. In: *RNA processing: A practical approach* (Higgins, S.J., and Hames, B.D., eds.). New York:Oxford University Press.

Li, H.-P., Zhang, X.M., Comai, L., Duncan, R., and Lai, M.M.C., 1997, Identification of the hetergeneous nuclear ribonucleoprotein (hnRNP) A1 as the cellular protein binding to the regulatory RNA regions of mouse hepatitis virus, in: *Coronaviruses and Arteriviruses* (eds. L. Enjuanes, S. Siddell, and W. Spaan), Plenum, New York.

Pollard, V.W., Michael, W.M., Nakielny, S., Siomi, M.C., Wang, F., and Dreyfuss, G., 1996, A novel receptor-mediated nuclear protein import pathway, *Cell* 86, 985–994.

Simoi, H., and Dreyfuss, G., 1995, A nuclear localization domain in the hnRNP A1 protein, *J. Cell Biol.* 129:551–560.

Smith, D.B. and Johnson, K.S., 1988, Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase, *Gene* 67:31–40.

Weighardt, F., Biamonti, G., and Riva, S, 1995, Nucleo-cytoplasmic distribution of human hnRNP proteins: a search for the targeting domains in hnRNP A1, *J. Cell Sci.* 108: 545–555.

Will, C.L., Kastner, B., Luehrmann, R., 1996, Analysis of ribonucleoprotein interactions. p141–177. In: *RNA processing: A practical approach* (Higgins, S.J., and Hames, B.D., eds.). New York:Oxford University Press.

Zhang, X.M, Liao, C.-L. and Lai, M.M.C., 1994, Coronavirus leader RNA regulates and initiates subgenomic mRNA transcription, both in trans and in cis, *J. Virol.* **68**:4738–4746.

Zhang, X.M., and Lai, M. M. C., 1995, Interactions between the cytoplasmic proteins and the intergenic (promoter) sequence of murine hepatitis virus RNAs: Correlation with the amounts of subgenomic mRNA transcribed, *J. Virol.* **69**: 1637–1644.