JOURNAL OF VIROLOGY, Mar. 1995, p. 1637–1644 Vol. 69, No. 3
0022-538X/95/$04.00+0

Interactions between the Cytoplasmic Proteins and the Intergenic (Promoter) Sequence of Mouse Hepatitis Virus RNA: Correlation with the Amounts of Subgenomic mRNA Transcribed

XUMING ZHANG AND MICHAEL M. C. LAI*

Howard Hughes Medical Institute and Department of Microbiology, University of Southern California School of Medicine, Los Angeles, California 90033-1054

Received 29 August 1994/Accepted 2 December 1994

Previous studies suggested that coronavirus RNA transcription involves interaction between leader RNA and the intergenic (IG) sequences, probably via protein-RNA interactions (X. M. Zhang, C.-L. Liao, and M. M. C. Lai, J. Virol., 68:4738–4746, 1994; X. M. Zhang and M. M. C. Lai, J. Virol., 68:6626–6633, 1994). To determine whether cellular proteins are involved in this process, we performed UV cross-linking experiments using cytoplasmic extracts of uninfected cells and the IG (promoter) sequence between genes 6 and 7 (IG7) and the 5′ untranslational region of mouse hepatitis virus genomic RNA. We demonstrated that three different cellular proteins (p70, p48, and p35/38) bound to the promoter sequence of the template RNA. Deletion analyses of the template RNA mapped the binding site of p35/38 at the consensus transcription initiation signal. In contrast, the binding of p70 and p48 was less specific. p35/38 is the same protein as the one previously identified to bind to the complementary strand of the leader RNA; its binding affinity to the leader was approximately 15 times stronger than that to IG7. Site-directed mutagenesis of the IG sequence revealed that mutations in the consensus sequence of IG7 (UCUAAUCUAAAC to UCGAAAC and GCUAAAG), which resulted in reduced subgenomic mRNA transcription, also caused correspondingly reduced levels of p35/38 binding. These results demonstrated that the extent of protein binding to the IG sequences correlated with the amounts of subgenomic mRNAs transcribed from the IG site. These studies suggest that these RNA-binding proteins are involved in coronavirus RNA transcription and may represent transcription factors.

Among RNA viruses, coronavirus, as exemplified by mouse hepatitis virus (MHV), employs unique strategies for RNA replication and transcription, probably as a consequence of the unusual structure and properties of its RNA genome and RNA polymerase. MHV contains a single-stranded, positive-sense RNA genome of approximately 31 kb (10, 15). In MHV-infected cells, six to seven species of virus-specific subgenomic mRNAs, in addition to the genomic RNA, which form a nested set at the 3′ end, are synthesized (12, 16). Each mRNA contains a leader sequence of 72 to 77 nucleotides (nt), which is derived from the 5′ end of the genomic RNA (11, 13, 25). At the 3′ end of the leader, there are varying numbers (two to four copies) of a pentanucleotide sequence (UCUAA) among different MHV strains. In most mRNAs, only the 5′-most open reading frame is translated. The putative RNA polymerase is encoded by gene 1, which is 22 kb in length and located at the 5′ end of the viral genome (15). An intergenic (IG; previously named IS) sequence containing a consensus UCUAAAC or similar sequence, which is homologous to the pentanucleotide (UCUAA) repeats at the 3′ end of the leader, is present between each gene (3, 23, 24). The IG sequence serves as a promoter and transcription initiation site for subgenomic mRNA synthesis (8, 19). It has been postulated that subgenomic mRNA synthesis is the result of interactions between the pentanucleotide sequence of the leader and the consensus IG sequence through complementary sequences (9, 10). In support of this, it has been shown that both the site-specific mutations of the IG sequence and the copy number of the UCUAA sequence in the leader could affect the efficiency of mRNA transcription (8, 14, 18, 27). The fusion of the leader sequence with the mRNA body sequence occurred between the UCUAA repeats of the leader and the consensus IG sequence (11, 13, 24, 25). However, how the leader and IG sequence interact remains unanswered.

Recent studies further suggest that the regulation of coronavirus RNA transcription involves not only interaction between the IG sequence and a *trans*-acting leader sequence but also a *cis*-acting leader sequence at the 5′ end of the genomic RNA (29). This model came from studies using a defective interfering (DI) RNA reporter system (17). Furthermore, several studies demonstrated that the *trans*-acting leader RNA was incorporated into subgenomic mRNAs (7, 17, 29). These results are consistent with a discontinuous transcription mechanism which involves a free *trans*-acting leader RNA. The involvement of three disparate RNA components, some of which do not have complementary sequences between them, suggests that these interactions involve more than direct RNA-RNA interactions.

More recently, we found that a particular strain of MHV, JHM2c, which has a deletion of a 9-nt sequence (UUUAU AAAC) immediately downstream of the leader RNA, transcribes subgenomic mRNA species containing a whole array of heterogeneous leader fusion sites (28). The leader RNA is fused with mRNAs at sites either upstream or downstream of the consensus IG sequence. These regions bear little or no sequence homology with the leader. Moreover, some of the leader-mRNA fusion sequences contain a duplication of portion of the leader sequence or an insertion of nontemplated sequences which are not present in either the leader or tem-

* Corresponding author. Phone: (213) 342-1748. Fax: (213) 342-9555.

plate RNA. A similar, though much less conspicuous, heterogeneity of leader-mRNA fusion has also been reported in another MHV DI system (27). If the leader RNA binds to the IG sequence solely through the complementary sequence, then how do these leaders tether to the regions upstream or downstream of the consensus IG sequence, where there are no homologous sequences? Also, how does the leader RNA in *cis* normally interact with a distant IG region (up to 30 kb apart)? It has been suggested that, as in DNA-dependent RNA transcription, some cellular proteins first bind to RNA regulatory sequences (leader and IG) through protein-RNA interactions; these distant RNA components are then brought together through protein-protein interactions to form a transcription initiation complex with the putative RNA-dependent RNA polymerase (27–29). This model predicts that these cellular proteins act as transcription factors.

Indeed, by performing UV light cross-linking with cytoplasmic extracts, we have identified three different cellular proteins binding to the complementary strands of the 5′ untranslated region (UTR) of MHV genomic RNA (5). A 55-kDa protein (p55) binds to the 5′ end of the positive-strand RNA, while p35/38 binds to the 3′ end of the negative-strand RNA. Their binding sites were mapped to the region encompassing the pentanucleotide repeats of the leader (5). However, the functional significance of these RNA-binding proteins was not clear. In the present study, we further identified several cellular proteins binding to the promoter sequence (IG7, IG sequence between genes 6 and 7). One of the proteins (p35/38) is the same as that binding to the leader RNA, but its binding affinities to the IG sequence and leader are significantly different. Furthermore, site-directed mutagenesis revealed that mutations in the consensus sequence of IG7 resulted in the concordant changes in the amounts of protein binding and mRNA transcribed. This study thus provides the first evidence that cellular RNA-binding proteins are involved in coronavirus RNA synthesis, suggesting that these RNA-binding proteins are transcription factors.

MATERIALS AND METHODS

Plasmid construction: pIG7-SS237. The cDNA clone pT7-4N(+), which represents the 3′ 2.2 kb of the MHV RNA genome, encompassing the 3′ end of gene 6, IG7, and the complete gene 7, was constructed in pT7-4 vector as described previously (1). This plasmid was linearized with *Acc*I and subjected to PCR amplification to generate a 0.23-kb DNA fragment containing IG7 and the surrounding sequences. PCR was performed using a pair of primers (5′IG7, 5′-TAGCCCGGGTAGACGGTGTTAGCG-3′; 3′IG7, 5′-TTCCCGGGTCCA CGCTCGGTTTGG-3′); the DNA product was digested with *Sma*I and cloned into the *Sma*I site of pBluescript, resulting in pIG7-SS237. pIG7-SS237 was used for construction of all the deletion mutants used in this study.

Deletion mutants. pIG7-SS237 was digested with *Sma*I, and the *Sma*I-*Sma*I fragment (insert) was purified by low-melting-point agarose gel electrophoresis. This fragment was then digested with *Hin*cI, *Fok*I, and *Mse*I, and the resulting fragments, *Sma*I-*Hin*cI (152 nt) and *Hin*cI-*Sma*I (85 nt), *Sma*I-*Fok*I (135 nt) and *Fok*I-*Sma*I (102 nt), and *Sma*I-*Mse*I (118 nt) and *Mse*I-*Sma*I (119 nt), respectively, were blunt ended by T4 DNA polymerase and cloned into the blunt-ended *Sal*I and *Xba*I sites of pBluescript. These constructs are termed pIG7-SH152, pIG7-HS85, pIG7-SF135, pIG7-FS102, pIG7-SM118, and pIG7-MS119, respectively (Fig. 1). Fragment SH152 was further digested with *Bst*UI, and the resulting fragments, *Sma*I-*Bst*UI (83 nt) and *Bst*UI-*Hin*cI (70 nt), were cloned as described above, generating pIG7-SB83 and pIG7-BH70, respectively. The structure and orientation of these clones are illustrated in Fig. 1. The sequence and orientation of these constructs were confirmed by dideoxyribonucleotide chain termination sequencing (20).

Mutants generated by site-directed mutagenesis. To introduce mutations into the consensus IG sequence, we incorporated mutated nucleotides into the 3′ primers for PCR amplification. pIG7-SS237 DNA was used as a template. The 5′ primer 5′IG7 and one of the two 3′ primers (3′MG, 5′-CCTTAAAGTTTCGAT TCTCAACAATGCGGT-3′; 3′MGG, 5′-CCTTAAACTTTAGCTTCTCAACA ATGCGGT-3′) were used to generate mutants IG7-MG and IG7-MGG, respectively. The sequences of the wild-type IG7 (SM118) and mutants (IG7-MG and IG7-MGG) are shown in Fig. 5A. The PCR fragments were digested with *Sma*I and cloned into the blunt-ended *Sac*I and *Xho*I sites of pBluescript.

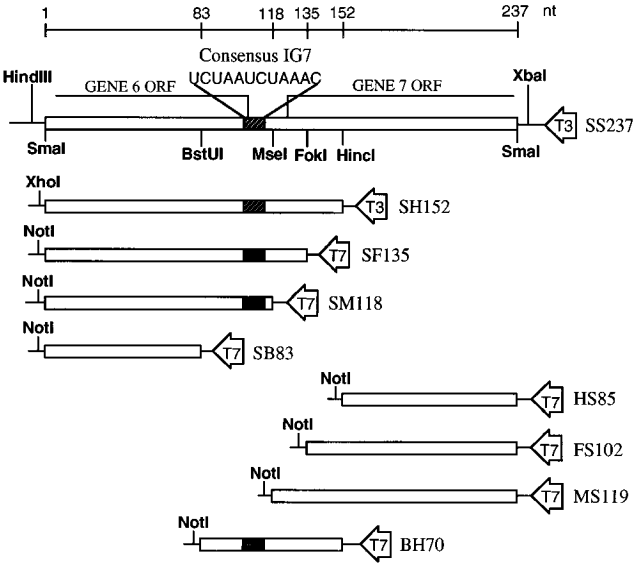


FIG. 1. Structures of IG7 and its deletion derivatives. Only the MHV sequence (open boxes) and the neighboring vector sequences (thin lines) are shown. The names of the constructs, which include the first letters of the enzyme sites at both ends and the length of the MHV-specific sequence, are indicated on the right of each construct. The direction of the constructs from left to right is 5′ to 3′; thus, the orientations of the transcripts from T7 or T3 promoter as indicated will be negative sense. The positions of the deletion clones are aligned with the parental sequence SS237. The MHV-specific sequence in SS237 includes the 3′ end (109 nt) of the gene 6 open reading frame (ORF), the 5′ end (114 nt) of the gene 7 ORF, and the sequence between genes 6 and 7. The enzyme sites shown inside SS237 were used for construction of deletion derivatives, and the enzyme sites shown within the vector sequence were used for digestion for runoff transcription.

To construct the DI RNA-chloramphenicol acetyltransferase (CAT) reporter vector (17, 29) containing the mutated IG sequence, IG7 of plasmid p25CAT (17) was replaced with mutant IG7 by PCR mutagenesis. Briefly, PCR was performed using one of two 5′ primers (5′MGCAT, 5′-TAACTAGTAGAATC GAAACTTTAAGGAATGGAGAAAAAAAT-3′; 5′MGGCAT, 5′-TAACTA GTAGAAGCTAAAGTTTAAGGAATGGAGAAAAAAAT-3′), paired with the 3′ primer 3CAT542 (29) and p25CAT DNA (17) as a template. The PCR products were digested with *Spe*I and *Bsp*EI, and the *Spe*I-*Bsp*EI fragments were cloned into the *Spe*I and *Bsp*EI sites of p25CAT to replace its original *Spe*I-*Bsp*EI fragment, resulting in pDECAT-MG and pDECAT-MGG, respectively (see Fig. 5A). The sequences of these mutants were confirmed by dideoxyribonucleotide chain termination sequencing (20).

In vitro RNA transcription. To generate negative-strand RNAs, pIG7-SS237 and pNX1-182 (5) were linearized with *Hin*dIII and pIG7-SH152 was linearized with *Xho*I. The linearized plasmid DNAs were subjected to in vitro transcription using T3 RNA polymerase. pIG7-SF135, pIG7-SM118, pIG7-SB83, pIG7-HS85, pIG7-FS102, pIG7-MS119, and pIG7-BH70 were linearized with *Not*I and transcribed with T7 RNA polymerase. The orientation of the inserts and the polarity of the transcripts are shown in Fig. 1. For synthesis of a nonspecific competitor, pBluescript was linearized with *Xba*I and transcribed with T7 RNA polymerase to generate a 60-nt RNA which represents cloning sites and internal sequences between T7 and T3 promoters of the vector.

In vitro transcription was carried out according to the manufacturer's recommended procedure (Promega). ^{32}P-labeled RNAs were obtained by transcription in the presence of 50 μCi of [α-^{32}P]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, CTP, GTP, and UTP. After transcription at 37°C for 1 h, 1 U of RQ DNase I (Promega) was added to the reaction to eliminate 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 activity of the radiolabeled probes was determined with a liquid scintillation counter (LC6000 IC; Beckman).

For transcriptional analysis, DI RNAs were prepared by in vitro transcription by T7 RNA polymerase from *Xba*I-linearized plasmid DNAs (p25CAT, pDE CAT-MG, and pDECAT-MGG) as described previously (29).

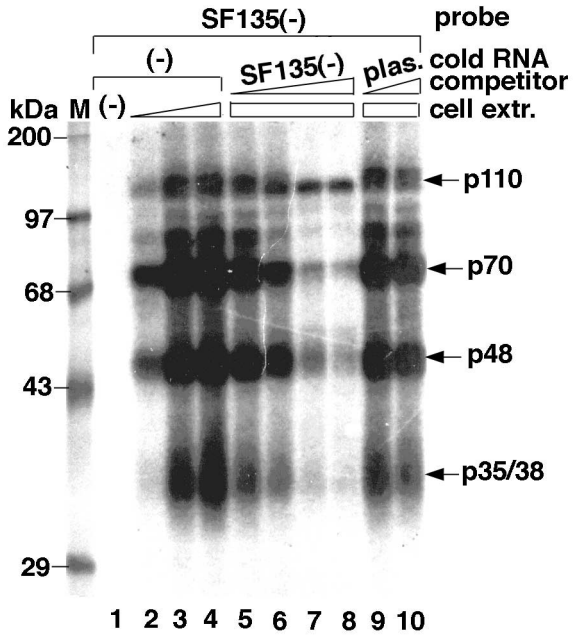


FIG. 2. UV cross-linking of cytoplasmic extracts with IG7-SF135(−) RNA. Lanes 1 to 4, UV cross-linking with a fixed amount of the radiolabeled probe (100 pg of RNA in each lane) and various amounts of cytoplasmic extracts, without the addition of unlabeled competitor (lane 1, no cytoplasmic extract; lanes 2 to 4, with 2, 10, and 20 μg of cytoplasmic protein, respectively); lanes 5 to 8, homologous competition assay with a fixed amount of cytoplasmic extracts (20 μg of protein in each lane) and various amounts of unlabeled RNAs (2-, 10-, 25-, and 50-fold molar excesses over the radiolabeled probe, respectively); lanes 9 and 10, nonspecific competition assay with a fixed amount of cytoplasmic extracts (20 μg of protein in each lane) and various amounts of unlabeled plasmid RNAs (25- and 50-fold molar excesses over the radiolabeled probe, respectively). The protein markers (lane M) and molecular sizes are indicated on the left. The arrows and names on the right side indicate the protein bands. Computer imagings were generated by Adobe Photoshop, version 2.5.1 LE.

Preparation of cytoplasmic extracts. DBT cells (6) were grown in 10-cm-diameter dishes to confluence, washed with cold phosphate-buffered saline (PBS), scraped off with a rubber policeman, and centrifuged at 2,000 × *g* for 5 min at 4°C. The cell pellets were resuspended in 1 ml of cold PBS (approximately 2.5×10^7 cells per ml) and stored at −80°C. Cytoplasmic extracts were prepared by the method described previously (5, 22) with slight modifications. Briefly, 1 ml of frozen cell preparation was thawed on ice and pelleted by centrifugation in an Eppendorf tube for 1 min. Cell pellets were resuspended in 400 μl of the extraction buffer (10 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.9], 10 mM KCl, 0.1 mM EDTA, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride) at 4°C by gentle pipetting. Cells were then allowed to swell on ice for 15 min. After addition of 25 μl of 10% Nonidet P-40, the mixture was vortexed vigorously for 20 s and centrifuged for 1 min in a microcentrifuge. The supernatant was collected, and the total protein concentration was determined with a protein assay kit (Bio-Rad).

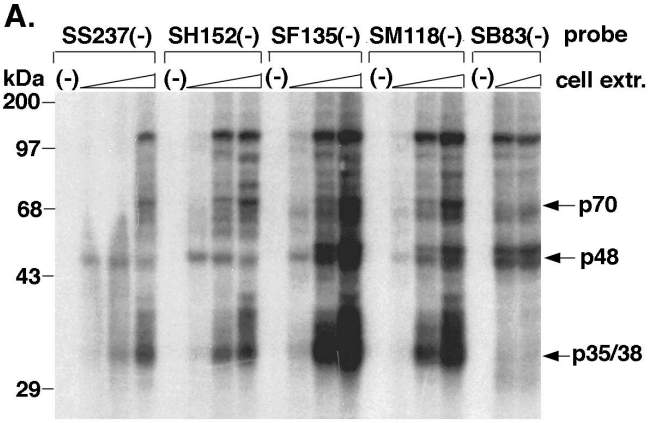
UV cross-linking of RNA-protein complex. UV cross-linking experiments were carried out essentially as described previously (5) with slight modification. Briefly, all reactions were carried out in a sequencing microplate (Stratagene). Different amounts of cytoplasmic extracts were mixed with 10 μg of yeast tRNA in 20 μl of the binding buffer (25 mM KCl, 5 mM HEPES [pH 7.8], 2 mM $MgCl_2$, 0.1 mM EDTA, 3.8% glycerol, 2 mM dithiothreitol). In competition assays, different amounts of unlabeled RNAs were mixed with a fixed amount of cellular extracts in the binding buffer. The amounts of the unlabeled competitor RNAs were determined with a spectrophotometer. Following an incubation at 30°C for 10 to 15 min, approximately 100 pg of ^{32}P-labeled RNA (2.5×10^5 cpm/μl) was added, and the reaction mixture was incubated for an additional 10 to 15 min. The reaction mixture was then irradiated with UV light (254 nm [UV Stratalinker 2400; Stratagene]) on ice at a distance of 13 cm for 10 min. After irradiation, RNase A was added to a final concentration of 1 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 dithiothreitol, 3% sodium dodecyl sulfate [SDS], 62.5 mM Tris-HCl [pH 6.8]), and electrophoresed on a 10% polyacrylamide gel containing 0.1% SDS at 50 V overnight. After drying, the gel was exposed to X-ray film, and in some instances, protein bands of interest were quantitated with a radioanalytic scanner (Ambis). The X-ray films were then scanned with a ScanJet IIc (Hewlett-Packard), and the computer imagings were saved in a tagged image file format with Adobe Photoshop 2.5.1 LE software.

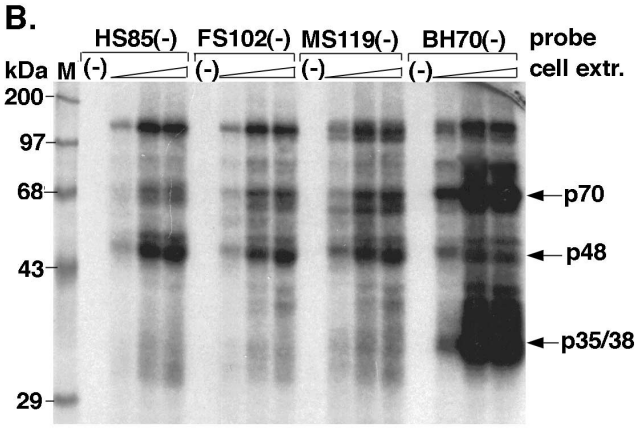
Virus infection, RNA transfection, preparation of cell lysates, and CAT assay. DBT cells (6) were infected with MHV strain JHM(2) and transfected with in vitro-transcribed DI RNAs by the DOTAP method (29). Cell lysates were prepared at 8 h posttransfection and subjected to CAT assay by the procedures described previously (29).

RESULTS

Cellular proteins bind to the negative-strand IG sequence of MHV RNA. Our previous study demonstrated that several cellular proteins bind to the positive and negative strands of the 5′ UTR of MHV RNA (5), which are the sequence involved in the regulation of RNA synthesis (10). It is thus of interest to determine whether other regulatory sequences of MHV RNA also bind cellular proteins. Since IG sequences are the initiation sites (promoter sequences) of subgenomic mRNAs and IG7 is the best-studied promoter of MHV RNA, we chose IG7 for this study. Since the template for mRNA synthesis should be negative-strand RNA, we first used SS237(−) and SF135(−) RNAs, both of which represent negative strands and include the IG sequence (Fig. 1), for UV cross-linking experiments with the cytoplasmic extracts of uninfected DBT cells. The results showed that four major proteins, p110, p70, p48, and p35/38, bound to IG7-SF135(−) RNA (Fig. 2, lanes 2 to 4). With increasing amounts of cytoplasmic extracts, the amounts of cross-linked proteins increased correspondingly (Fig. 2, lanes 2 to 4). Similar results were obtained with SS237(−) RNA (data not shown). To determine whether the binding of these proteins was specific, we used an unlabeled homologous RNA [IG7-SF135(−)] and a nonspecific RNA (vector plasmid) to perform competition assays. The unlabeled competitor RNA was added to the binding reactions at 30°C for 10 to 15 min prior to the addition of the radiolabeled probe. When increasing amounts of the unlabeled homologous RNA (2- to 50-fold molar excess) were preincubated with a fixed amount of cytoplasmic extracts (20 μg of cellular proteins), the binding of most of the proteins was inhibited, except that the binding of p110 was not efficiently inhibited (Fig. 2, lanes 5 to 8). In contrast, the nonspecific vector sequence (25- to 50-fold molar excess) did not affect the amount of protein binding (Fig. 2, lanes 9 and 10). We conclude that at least three cytoplasmic proteins (p70, p48, and p35/38) specifically bound to IG7-SF135(−) RNA, while p110 binding was probably not specific.

Mapping of the protein-binding sites on IG7. To dissect the sequence requirement for protein binding, a series of deletion constructs were transcribed in vitro with either T7 or T3 RNA polymerase to generate negative-strand RNA fragments of different lengths (Fig. 1). Figure 3 shows that p35/38 bound only to the fragments which contain the consensus IG sequence [constructs SS237(−), SH152(−), SF135(−), and SM118(−) in Fig. 3A and BH70(−) in Fig. 3B] and not RNAs without the IG sequence [e.g., constructs SB83(−) in Fig. 3A and HS85(−), FS102(−), and MS119(−) in Fig. 3B]. From these binding studies, we conclude that the 35-nt sequence (nt 83 to 118), including the consensus IG7 sequence, is essential for p35/38 binding. In contrast, the binding of p70 and p48 was only slightly decreased in some of the deletion clones, suggesting that the sequence requirement for p70 and p48 was not very specific or that these two proteins bound to a longer stretch of RNA sequence. The results are summarized in Fig. 3C and D. It is noted that some RNA fragments bound these proteins better than the others did. For example, SF135(−) RNA bound more proteins than even the larger RNAs [SS237(−)] did, suggesting that overall RNA conformation





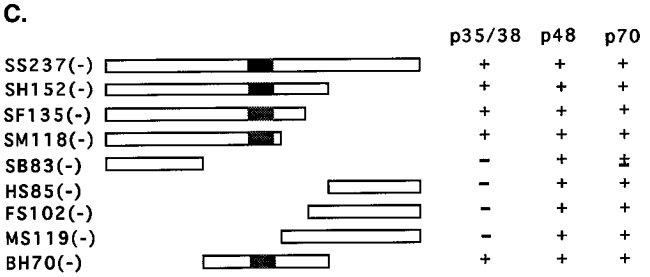
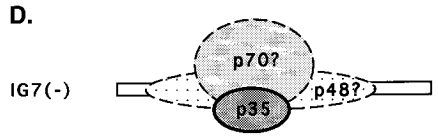


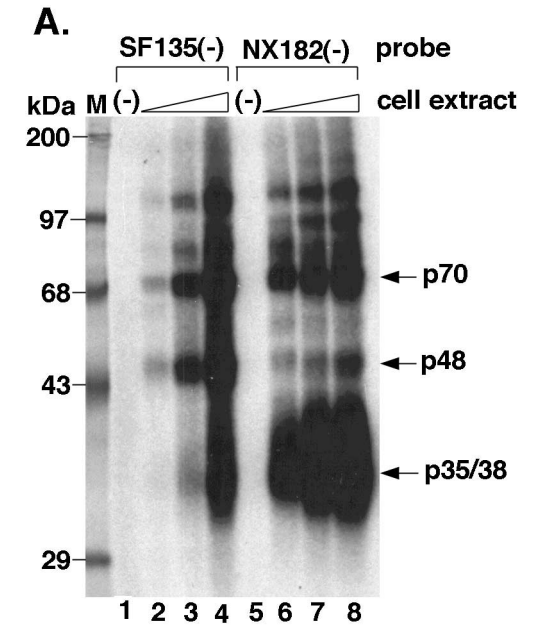
FIG. 3. Mapping of the protein binding sites on the negative-strand IG7 RNAs by using downstream deletions (A) and upstream deletions (B). Increasing amounts of cytoplasmic extracts [2, 10, and 20 μg of cellular protein in the left, middle and right lanes, respectively, in each panel, except for the panel SB83(−), which had only two lanes with 10 and 20 μg of protein, respectively] were mixed with ^{32}P-labeled RNAs and UV irradiated. Lanes (−), no cytoplasmic extracts added; lane M, molecular mass standards. Arrows on the right show the 70-kDa (p70), 48-kDa (p48), and 35- and 38-kDa (p35/38) proteins. (C) Summary of the protein-binding properties of IG7 and its deletion derivatives. +, binding positive; −, no binding; ±, weak binding. (D) Schematic diagram of cellular proteins bound to the IG7(−), showing the relative binding sites of each protein. The question marks indicate that the specificity and precise site of protein-binding on the negative-strand RNAs could not be firmly determined. Computer imagings were generated by Adobe Photoshop, version 2.5.1 LE.

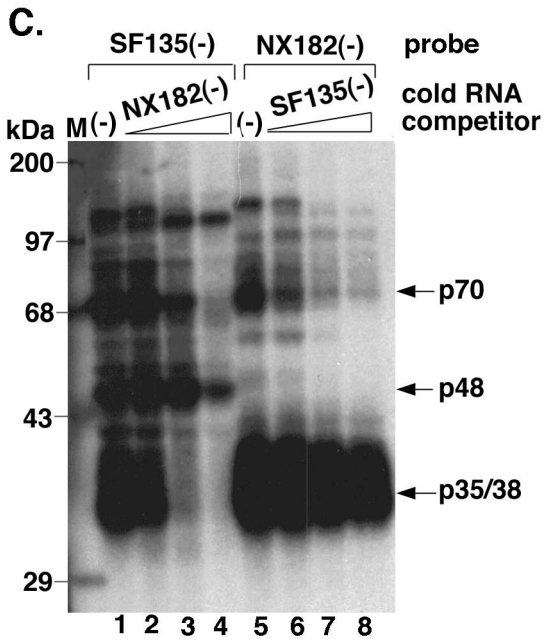


could affect the level of protein binding. Nevertheless, the binding of all of the proteins was similarly affected in each RNA fragment, suggesting that these proteins bind to specific RNA sequences rather than a conformational structure.

Comparison of the proteins bound to the IG sequence and to the 5′ UTR of MHV RNA. We previously demonstrated that cellular protein p35/38 bound to the negative-strand RNA of the 5′ UTR (5). In this study, a protein of similar molecular weight was found to bind to the negative-strand RNA of IG7. To determine whether these are the same proteins, the protein binding to IG7-SF135(−) RNA was directly compared with that binding to NX182(−) RNA, which represents the 5′-most 182 nt of DI RNA (5). In the previous study (5), only p35/38 was found to bind to NX182(−) RNA, but a slightly modified procedure for the preparation of cellular extracts as described in Materials and Methods and longer exposure of the gel done in this study showed that p70 also bound to NX182(−) RNA (Fig. 4A). A small amount of p48 also bound to NX182(−) RNA when a large amount (20 μg) of cellular lysates was used. Thus, the protein-binding patterns between the two RNAs were similar when a large amount of cytoplasmic extracts was used. However, the binding affinities of these proteins to these two RNA regions were significantly different (Fig. 4A). p35/38 bound much more strongly to NX182(−) RNA, while p48 bound almost exclusively to IG7(−) RNA. The differences in binding affinities of these proteins to IG7-SF135(−) and NX182(−) RNAs were determined quantitatively with a radioanalytic scanner. The results showed that p35/38 bound to NX182(−) RNA 15-fold more strongly than to IG7-135(−) RNA, whereas p48 bound to IG7-SF135(−) RNA 10-fold more strongly than to NX182(−) RNA (Fig. 4B). The large difference in the intensity of protein bands between the NX182(−) RNA and IG7-SF135(−) RNA most likely represented differences in protein-binding affinities of the two RNAs but was not due to the different specific activities of the probes, because the specific activity of radiolabeled NX182(−) RNA was only 20% higher than that of IG7-SF135(−) RNA [NX182(−) RNA contains 49 UTPs, while IG7-SF135(−) RNA contains 39 UTPs]. Approximately the same amounts of p70 bound to both RNAs (Fig. 4B). We conclude from these data that p35/38 predominantly bound to the negative-strand 5′ UTR while p48 predominantly bound to the negative-strand IG7.

To confirm that the proteins bound to both RNAs were the same, we performed cross competition analyses. The results showed that the binding of p70 and p35/38 to the radiolabeled IG7-SF135(−) RNAs was inhibited efficiently by unlabeled NX182(−) RNAs in a concentration-dependent manner (Fig. 4C, lanes 1 to 4). In contrast, the binding of p48 was poorly inhibited, consistent with the finding that p48 bound strongly to IG7-SF135(−) RNA but very poorly to NX182(−) RNA. The reciprocal competition experiments showed that the binding between the radiolabeled NX182(−) RNA and p35/38 could not be efficiently inhibited by unlabeled IG7-SF135(−) RNA (Fig. 4C, lanes 5 to 8), consistent with the finding that p35/38 binding to NX182(−) RNA was much stronger than that to IG7-SF135(−) RNA. In contrast, p70 binding was inhibited efficiently in both cross competition experiments, suggesting that p70 binds equally strongly to the 5′ UTR and IG7. Very little p48 bound to NX182(−) RNA. These results indicate that (i) the same cellular proteins (p70 and p35/38) bound to the negative strands of both the IG sequence and 5′ UTR,





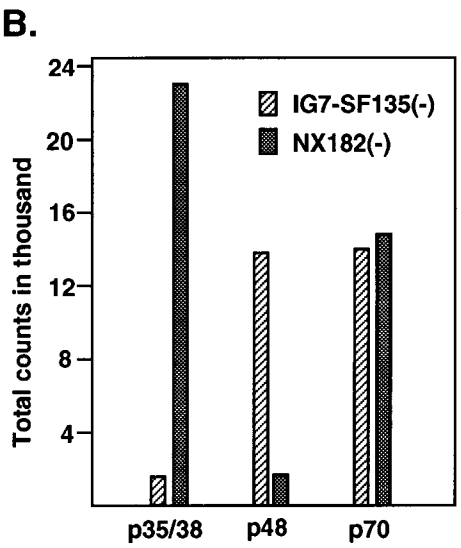


FIG. 4. Comparison of the protein-binding properties between the negative-strand RNAs of IG7 and the 5′ UTR. (A) UV cross-linking of cellular proteins with IG7-SF135(−) and NX182(−) RNAs. Increasing amounts of cytoplasmic extracts (2 μg of protein in lanes 2 and 6, 10 μg in lanes 3 and 7, and 20 μg in lanes 4 and 8) were mixed with the same amount of ^{32}P-labeled IG7-SF135(−) RNAs (lanes 1 to 4) and NX182(−) RNAs (lanes 5 to 8) and UV-irradiated. Lanes (−) (lanes 1 and 5), no cytoplasmic extracts added; lane M, molecular mass standards. Arrows on the right indicate the protein bands. (B) Diagrammatic presentation of the binding affinities of the two RNAs [IG7-SF135(−) and NX182(−)] with three different cellular proteins (p35, p48 and p70). The intensity of each protein band from lanes 3 and 7 in panel A was quantitated with the Ambis radioanalytic scanner and is presented as total counts. (C) Cross competition assay. A fixed amount of cytoplasmic extracts (20 μg of protein in lanes 1 to 4 and 2 μg in lanes 5 to 8) was preincubated with various amounts of unlabeled NX182(−) RNAs (lanes 1 to 4) and IG7-SF135(−) RNAs (lanes 5 to 8) (2-, 25-, and 50-fold molar excesses over radiolabeled RNAs in lanes 2 and 6, 3 and 7, and 4 and 8, respectively) prior to the addition of ^{32}P-labeled, heterologous RNAs. Lanes (−), no competitor RNAs added; lane M, molecular mass standards. Arrows on the right indicate the protein bands of interest. Computer imagings were generated by Adobe Photoshop, version 2.5.1 LE.

(ii) p35/38 predominantly bound to the negative-strand 5′ UTR, and (iii) p48 predominantly bound to IG7.

Site-specific mutagenesis of IG7 in RNA-protein interaction. Our modified model of transcriptional regulation of coronavirus RNA proposes that cellular or viral proteins interact with the regulatory sequences of viral RNAs and that these proteins, in turn, interact with each other to bring the regulatory RNA sequences together with viral RNA polymerase to form a transcriptional initiation complex (17, 29). Conceivably, these cellular proteins may act as transcription factors. To test this hypothesis, we constructed two different IG mutants (Fig. 5A), which have previously been shown to have significantly reduced transcriptional efficiencies (8, 18), for both transcriptional studies and UV cross-linking with cytoplasmic extracts. The results showed that the relative binding affinities of these IG mutants with p35/38 correlated approximately with the relative amounts of the subgenomic mRNAs transcribed from each IG sequence. Specifically, mutant MGG, which has a deletion of a UCUAA sequence and two additional single-nucleotide substitutions (U → G and C → G) in the remaining consensus sequence (UCUAAAC to GCUAAAG) (Fig. 5A), did not exhibit any transcriptional activity (CAT activity was virtually the same as the background level; Fig. 5A), consistent with the previous results obtained by Northern (RNA) blot analysis (18). Correspondingly, MGG(−) RNA did not bind p35/38 and bound slightly less p48 and p70 (Fig. 5B). Mutant MG, which has a deletion of a UCUAA sequence and a C → G substitution (UCUAAAC to UCGAAAC) (Fig. 5A), had CAT activity 50-fold lower than that of the wild-type IG7 (Fig. 5A). Correspondingly, this IG mutant bound significantly less p35/38 than the wild-type RNA (Fig. 5B). These results demonstrated that the binding efficiencies of p35/38 to the different IG7 sequences correlated roughly with the amounts of subgenomic mRNAs transcribed, suggesting that binding of p35/38 to IG7 is required for subgenomic mRNA transcription and that the cellular RNA-binding protein(s) is a transcription factor(s) for coronavirus RNA synthesis.

DISCUSSION

In this study, we have identified several cytoplasmic proteins that were cross-linked by UV irradiation to the promoter se-

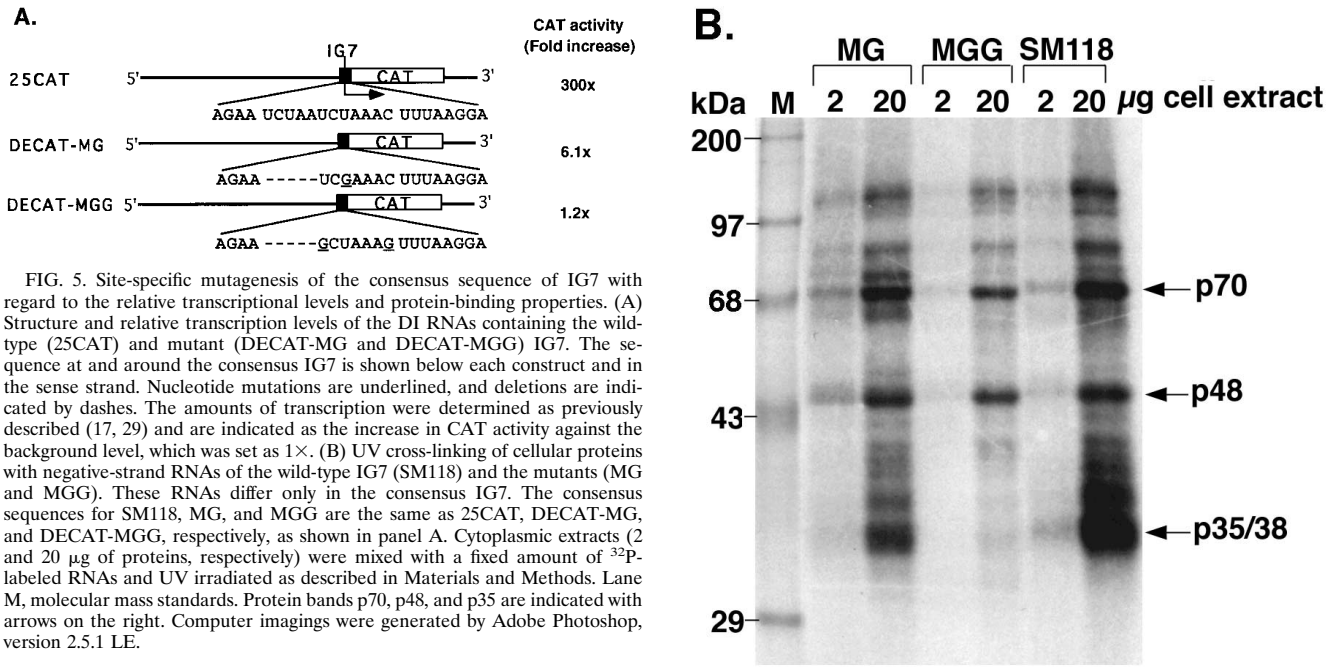


FIG. 5. Site-specific mutagenesis of the consensus sequence of IG7 with regard to the relative transcriptional levels and protein-binding properties. (A) Structure and relative transcription levels of the DI RNAs containing the wild-type (25CAT) and mutant (DECAT-MG and DECAT-MGG) IG7. The sequence at and around the consensus IG7 is shown below each construct and in the sense strand. Nucleotide mutations are underlined, and deletions are indicated by dashes. The amounts of transcription were determined as previously described (17, 29) and are indicated as the increase in CAT activity against the background level, which was set as 1×. (B) UV cross-linking of cellular proteins with negative-strand RNAs of the wild-type IG7 (SM118) and the mutants (MG and MGG). These RNAs differ only in the consensus IG7. The consensus sequences for SM118, MG, and MGG are the same as 25CAT, DECAT-MG, and DECAT-MGG, respectively, as shown in panel A. Cytoplasmic extracts (2 and 20 μg of proteins, respectively) were mixed with a fixed amount of ^{32}P-labeled RNAs and UV irradiated as described in Materials and Methods. Lane M, molecular mass standards. Protein bands p70, p48, and p35 are indicated with arrows on the right. Computer imagings were generated by Adobe Photoshop, version 2.5.1 LE.

quence of coronavirus template RNA. Some of the same proteins also bound to the template strand of the leader RNA at the 5′ end of the genome. Furthermore, we demonstrated by site-directed mutagenesis that the extents of the interactions between the promoter sequence (IG) and cellular proteins correlated with the amounts of subgenomic mRNA transcribed. These results provide direct evidence supporting the recently proposed model of coronavirus RNA transcription which proposes that the interactions among the regulatory RNA sequences for transcription are mediated by protein-RNA and protein-protein interactions rather than by direct RNA-RNA interactions (27–29). This is in contrast to the classical model of leader-primed transcription and other discontinuous transcription mechanisms (9, 21), which are based on direct RNA-RNA interactions between the leader RNA and the IG sequence. Recent findings that leader-mRNA fusions can occur at regions bearing no sequence homology between the leader and the IG sequence (27, 28) also suggest that direct RNA-RNA interaction does not necessarily lead to mRNA transcription. Finally, previous site-specific mutagenesis studies of IG sequences failed to explain why different mutations within the IG sequence affected the efficiency of transcription to different extents, although all these mutations affected the sequence complementarity (8, 18, 27). Our data here indicated that the extents of p35/38 binding to these mutant IG sites correlated with the amounts of mRNAs transcribed, suggesting that protein-RNA interactions are critical for the initiation of mRNA transcription from the IG site (28, 29). Thus, the cellular proteins, particularly p35/38, can be considered transcription factors for coronavirus RNA transcription and/or replication. In this modified transcription model, cellular factors first bind to the IG sites and the leader sequences on the template RNA and also the *trans*-acting leader RNA (5); these protein factors then interact with each other to bring the regulatory RNA elements together to form a transcription initiation complex (29). This proposed mechanism is very similar to DNA-dependent RNA transcription. At the present, there is no evidence for the existence of protein-protein interactions among the various cellular proteins identified. Undoubtedly, the characterization of these cellular proteins will shed further light on the mechanism of coronavirus RNA transcription. These investigations are currently in progress.

Our previous data suggested that both the *cis*-acting leader sequence and the IG sequence are the regulatory sequences for subgenomic mRNA transcription (17, 29). Our studies here showed that both elements bound to cellular proteins, some of which are common to both RNA regions (p35/38 and p70) but others of which (e.g., p48) are distinct. We have also shown in the previous UV cross-linking study that the inclusion of an additional 5′-end genomic sequence (1.2 kb) of negative strand with the leader sequence did not increase the amounts or species of proteins bound (5), suggesting that the 5′-end leader sequence is the only sequence within the 5′ 1.2 kb of viral genome capable of binding cellular proteins. Our studies here also showed that the deletion of the IG consensus sequence eliminated p35/38 binding. These results combined suggest that the cellular proteins indeed bind primarily the RNA regulatory elements for coronavirus transcription. Since some of the proteins (e.g., p35/38 and p70) bind to both regions, conceivably these proteins could interact with themselves to form a transcription initiation complex.

By comparing the IG sequence with the 5′ UTR, we noticed that the binding affinities of the different proteins with the IG sequence were significantly different from those with the 5′ UTR. For example, p35 bound to the negative-strand RNA of the 5′ UTR at least 15 times more strongly than to the IG sequence. Conceivably, the amounts of p35 bound to the 5′ UTR may also regulate the efficiency of subgenomic mRNA transcription. The cellular proteins which have been identified so far to bind to the MHV RNA regulatory elements are summarized in Fig. 6. It should be noted here that p55 binds exclusively to the positive-strand leader sequence and not to the negative strand (5); nor does it bind to the IG site of either strand (Fig. 2 and data not shown). It is tempting to suggest that p55 directly or indirectly interacts with the proteins on the IG sequence, thus bringing the leader RNA in *trans* to the IG

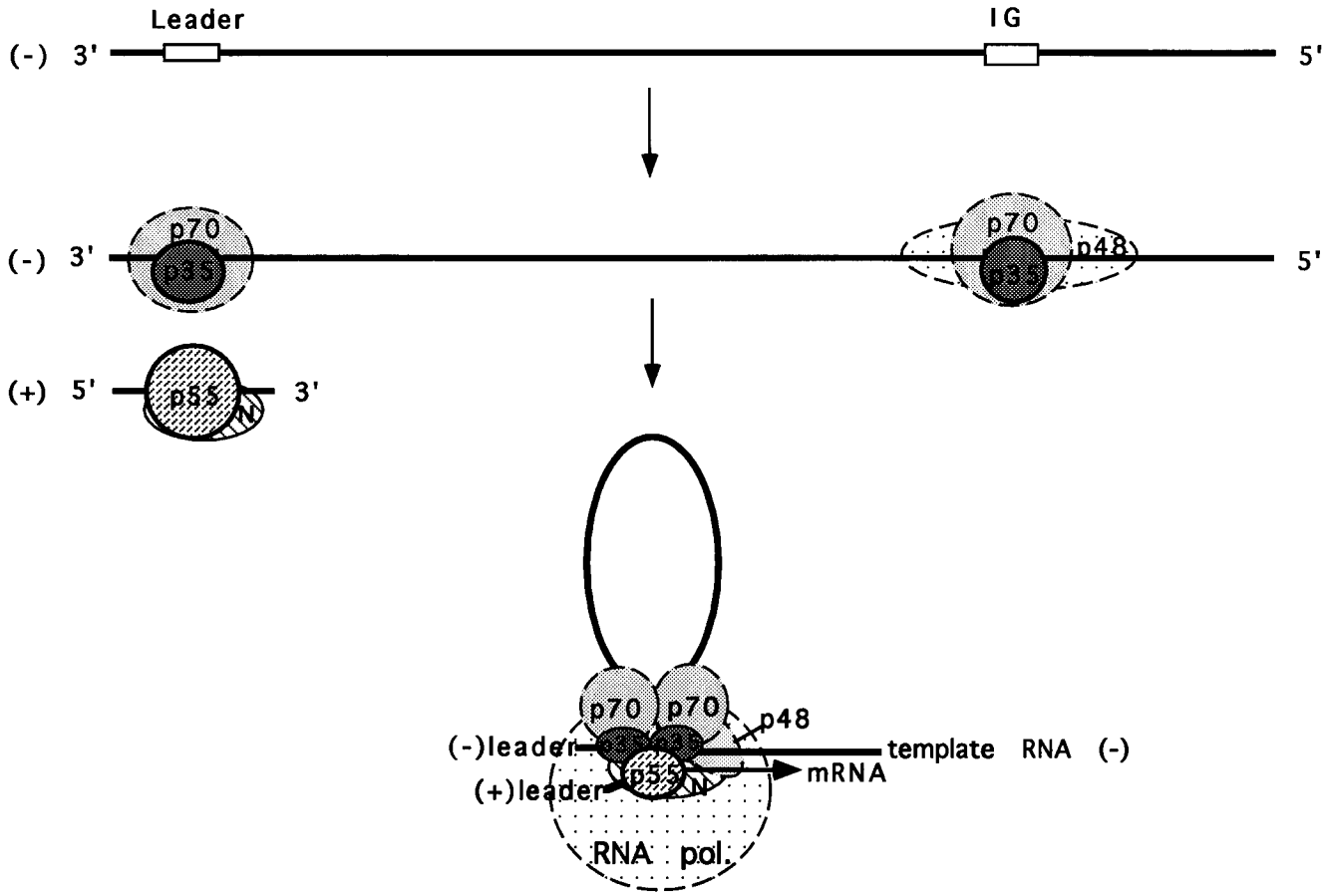


FIG. 6. Model of coronavirus RNA transcription through protein-RNA and protein-protein interactions. The open boxes indicate the UCUAA repeats at the 3′ end of the leader and the consensus IG sequence. The proteins bound to the different regions were identified in reference 5 and this study. N, viral nucleocapsid protein. Its binding to the 3′ end of leader RNA was demonstrated elsewhere (26). The presence of the putative viral RNA polymerase, N, p70, and p48 in the complex is hypothetical. In this putative complex, the negative-strand RNA serves as a template, and the leader RNA (positive sense) in *trans* binds to the negative-strand IG RNA and serves as a primer. The subgenomic mRNA transcription starts at the IG region inside the complex.

region. Very likely, viral RNA polymerase and other viral proteins, such as N protein (4, 26), also are required for complex formation. The involvement of the N protein in the formation of the transcription complex has been suggested from the findings that the N protein binds to the 3′ end of the leader RNA (26) and that the monoclonal antibody against the N protein inhibited MHV RNA synthesis (4). Conceivably, other possible *cis*-acting RNA regulatory elements, such as the 3′ end of genomic RNA, also participate in the formation of such a transcription complex. Once the leader RNA (positive strand) is associated with the negative-strand IG sequence of the template RNA in the complex, the leader RNA may then serve as a primer to initiate subgenomic mRNA transcription (Fig. 6). Results from a number of recent studies are compatible with this modified transcription model involving RNA-protein interactions (5, 17, 27–29). It should be noted that in this study, only the binding of p35/38 was found to correlate with the amounts of mRNA transcribed. The binding of p70 and p48 is less specific, and the role of these proteins in MHV RNA synthesis is less certain. It is likely that the transcription efficiency of MHV RNA is regulated by multiple factors.

In this study, we showed that an extra species of cellular protein (p70) which was not detected previously (5) bound to the negative strand of the 5′ UTR. Since the same NX182(−) RNA was used in both studies, the most likely explanation is that the procedure for preparation of cellular lysates used in this study allowed the more efficient release of p70 from a cellular compartment, such as cellular membranes. It has been shown previously that MHV RNA synthesis takes place in the membrane fraction (2). Another explanation is that in this study, we used more lysates for UV cross-linking studies; the more weakly binding proteins may have escaped detection previously. This is evident for p48, whose binding with NX182(−) RNA could not be detected when less than 10 μg of cellular proteins was used, as demonstrated here and previously (5). In contrast, when a large amount of cellular proteins (20 μg) was used, a weak band of p48 was detectable (Fig. 4A).

It should be noted that IG7 used in this study is a short RNA fragment (237 nt), containing only one IG sequence. The full-length viral genomic RNA, which contains multiple IG sequences, may have different protein-binding capabilities because of different RNA conformations. It has been shown here that the protein-binding affinities of different deletion clones were slightly different (Fig. 3), suggesting that overall RNA conformation could affect protein binding. However, the binding of these proteins is sequence specific rather than merely dependent on RNA conformation, as suggested from the studies of the IG7 mutants (Fig. 5). Thus, the full-length viral RNA likely binds the same proteins. Nevertheless, because the genome-length template RNA contains seven distinct IG sites, it

is possible that competition between the different IG sites affects the extent of protein binding. Regardless, the identification of these cellular proteins bound to the regulatory sequences of MHV RNA (the leader and IG sequence) and the correlation of the protein-binding affinity with the transcription efficiency at the IG site provide the first direct evidence that protein-RNA interactions are involved in the regulation of coronavirus mRNA transcription.

ACKNOWLEDGMENTS

This work was supported by a Public Health Service research grant AI19244 from the National Institutes of Health. X.Z. is supported by a postdoctoral training fellowship from NIH. M.M.C.L. is an investigator of the Howard Hughes Medical Institute.

REFERENCES

1. **Armstrong, J., S. Smeekens, and P. Rottier.** 1983. Sequencing of the nucleocapsid gene from murine coronavirus MHV-A59. Nucleic Acids Res. **11:** 883–891.
2. **Brayton, P. R., M. M. C. Lai, C. D. Patton, and S. A. Stohlman.** 1982. Characterization of two RNA polymerase activities induced by mouse hepatitis virus. J. Virol. **42:**847–953.
3. **Budzilowicz, C. J., S. P. Wilczynski, and S. R. Weiss.** 1985. Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. J. Virol. **53:**834–840.
4. **Compton, S. R., D. B. Rogers, K. V. Holmes, D. Fertsch, J. Remenick, and J. J. McGowan.** 1987. In vitro replication of mouse hepatitis virus strain A59. J. Virol. **61:**1814–1820.
5. **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.
6. **Hirano, N., K. Fujiwara, S. Hino, and M. Matsumoto.** 1974. Replication and plaque formation of mouse hepatitis virus (MHV-2) in mouse cell line DBT culture. Arch. Gesamte Virusforsch. **44:**298–302.
7. **Jeong, Y. S., and S. Makino.** 1994. Evidence for coronavirus discontinuous transcription. J. Virol. **68:**2615–2623.
8. **Joo, M., and S. Makino.** 1992. Mutagenic analysis of the coronavirus intergenic consensus sequence. J. Virol. **66:**6330–6337.
9. **Lai, M. M. C.** 1986. Coronavirus leader-primed transcription: an alternative mechanism to RNA splicing. Bioessays **5:**257–260.
10. **Lai, M. M. C.** 1990. Coronavirus: organization, replication and expression of genome. Annu. Rev. Microbiol. **44:**303–333.
11. **Lai, M. M. C., R. S. Baric, P. R. Brayton, and S. A. Stohlman.** 1984. Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. Proc. Natl. Acad. Sci. USA **81:**3626–3630.
12. **Lai, M. M. C., P. R. Brayton, R. C. Armen, C. D. Patton, C. Pugh, and S. A. Stohlman.** 1981. Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from the hepatotropic strain MHV-3. J. Virol. **39:**823–834.
13. **Lai, M. M. C., C. D. Patton, R. S. Baric, and S. A. Stohlman.** 1983. Presence of leader sequences in the mRNA of mouse hepatitis virus. J. Virol. **46:**1027–1033.
14. **La Monica, N., K. Yokomori, and M. M. C. Lai.** 1992. Coronavirus mRNA synthesis: identification of novel transcription initiation signals which are differentially regulated by different leader sequences. Virology **188:**402–407.
15. **Lee, H. J., C. K. Shieh, A. E. Gorbalenya, E. V. Koonin, N. La Monica, J. Tuler, A. Bagdzhadzhyan, and M. M. C. Lai.** 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology **180:**567–582.
16. **Leibowitz, J. L., K. C. Wilhemsen, and C. W. Bond.** 1981. The virus-specific intracellular RNA species of two murine coronaviruses: MHV-A59 and MHV-JHM. Virology **114:**39–51.
17. **Liao, C. L., and M. M. C. Lai.** 1994. Requirement of the 5′-end genomic sequence as an upstream *cis*-acting element for coronavirus subgenomic mRNA transcription. J. Virol. **68:**4727–4737.
18. **Makino, S., and M. Joo.** 1993. Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. J. Virol. **67:**3304–3311.
19. **Makino, S., M. Joo, and J. K. Makino.** 1991. A system for study of coronavirus mRNA synthesis: a regulated expressed subgenomic defective interfering RNA results from intergenic site insertion. J. Virol. **65:**6031–6041.
20. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
21. **Sawicki, S. G., and D. L. Sawicki.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J. Virol. **64:**1050–1056.
22. **Schreiber, E., P. Matthias, M. M. Müller, and W. Schaffner.** 1989. Rapid detection of octamer binding proteins with 'mini-extracts' prepared from a small number of cells. Nucleic Acids Res. **17:**6419.
23. **Shieh, C. K., H. J. Lee, K. Yokomori, N. La Monica, S. Makino, and M. M. C. Lai.** 1989. Identification of a new transcriptional initiation site and the corresponding functional gene 2b in the murine coronavirus RNA genome. J. Virol. **63:**3729–3736.
24. **Shieh, C. K., L. H. Soe, S. Makino, M. F. Chang, S. A. Stohlman, and M. M. C. Lai.** 1987. The 5′-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. Virology **156:**321–330.
25. **Spaan, W., H. Delius, M. Skinner, J. Armstrong, P. Rottier, S. Smeekens, B. A. M. van der Zeijst, and S. G. Siddell.** 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. EMBO J. **2:**1839–1844.
26. **Stohlman, J. A., R. S. Baric, G. N. Nelson, L. H. Soe, L. M. Welter, and R. J. Deans.** 1988. Specific interaction between coronavirus leader RNA and nucleocapsid protein. J. Virol. **62:**4288–4295.
27. **van der Most, R. G., R. J. de Groot, and W. J. M. Spaan.** 1994. Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: a study of coronavirus transcription initiation. J. Virol. **68:** 3656–3666.
28. **Zhang, X. M., and M. M. C. Lai.** 1994. Unusual heterogeneity of leader-mRNA fusion in a murine coronavirus: implications for the mechanism of RNA transcription and recombination. J. Virol. **68:**6626–6633.
29. **Zhang, X. M., C.-L. Liao, and M. M. C. Lai.** 1994. Coronavirus leader RNA regulates and initiates subgenomic mRNA transcription both in *trans* and in *cis*. J. Virol. **68:**4738–4746.