MHV LEADER RNA SECONDARY STRUCTURE AFFECTS BINDING TO THE NUCLEOCAPSID PROTEIN

Lisa M. Welter[1], Stephen A. Stohlman [1,2], and Robert J. Deans*[1]

Department of Microbiology[1] and Neurology [2]
University of Southern California School of Medicine
Los Angeles, California 90033

INTRODUCTION

Leader RNA is found at the 5' end of the mouse hepatitis virus (MHV) genomic RNA, at the 5' end of the seven viral mRNAs, and free in the cell. Leader RNA is synthesized by a transcriptional activity separate from the activities that synthesize both (-) sense genomic length RNA and the virus mRNAs[1]. It is believed to function as a primer, binding to complementary intergenic sites, on (-) sense template, situated 5' of each of the initiation sites for viral mRNAs[2]. Using monoclonal antibodies specific for the nucleocapsid (N) protein, immunoprecipitations of RNA/protein complexes from infected cells indicate that the N protein is complexed to: 1) genomic RNA; 2) viral mRNAs; and 3) even free leader containing RNA fragments as small as 60 nucleotides in length[3]. Northwestern blot analysis showed that the viral N protein exhibits RNA binding activity that is specific for viral leader containing RNA when expressed in the (+) sense[4,5]. However, this system has several limitations. First, in denaturing conditions, RNA/protein interactions which require the interaction of multiple protein subunits cannot be studied, and second, it is not possible to quantify relative affinities and binding characteristics.

The gel retardation assay (GRA) was used to define the N protein interaction with leader RNA. The advantages of this technique are: 1) the detection of weak nucleic acid/protein interactions; 2) the use of crude protein extracts thereby eliminating the requirement of purification of a single binding protein; 3) the visualization of multiple protein interactions with a single nucleic acid species; and 4) the nucleic acid protein interactions are able to occur freely in solution.

To study the RNA-protein interactions of MHV and to understand how N interactions with free leader, mRNA, or genomic RNA are related to viral assembly, we have used the GRA to define a high affinity interaction between leader RNA and the N protein. In this study we show the resolution of a high affinity RNA-protein complex formed in the presence of in vitro radiolabelled leader RNA transcripts, using MHV infected cell lysates as a source of N protein, and non-radiolabelled uninfected cytoplasmic RNA as a non-specific competitor. The specific RNA-protein complex is observed only with infected but not uninfected cell lysates. Through the use of the GRA we have been able to map the location of this binding activity to a proximal stem loop structure of leader RNA.

Coronaviruses and Their Diseases
Edited by D. Cavanagh and T.D.K. Brown
Plenum Press, New York, 1990

MATERIALS AND METHODS

In Vitro Transcript Preparation

The in vitro RNA transcripts used in this study were derived from Bluescribe (+), (BS+) plasmids (Stratagene, La Jolla, CA) containing a BamHI/DdeI fragment from the 5' end of mRNA from gene F of the A59 strain of MHV, inserted at the BamHI site of the BS+ plasmid[6]. A schematic map of these plasmids, and the in vitro transcripts produced are depicted in Figure 1. This construct, pBSL contains the 72 base leader plus 15 nucleotides extending into the E1 coding region[4]. Transcripts containing the 3' nucleotide sequences of leader RNA, starting at nucleotide 24, were derived by deleting the sequences between SmaI and SnaBI sites of pBSL. This plasmid is designated pBSL-SSΔ. Plasmid F82D was obtained from Dr. Michael Lai. This cDNA clone of genomic RNA initiates at nucleotide 56 of leader. Transcripts containing viral genomic sequences from 56 to 180, are obtained by digesting with the restriction endonuclease NarI which cleaves interval genomic sequences. Recombinant pT7 plasmids were first linearized with one of the restriction enzymes shown in Figure 1 to allow synthesis of transcripts of different lengths. Transcription was carried out as previously described[4].

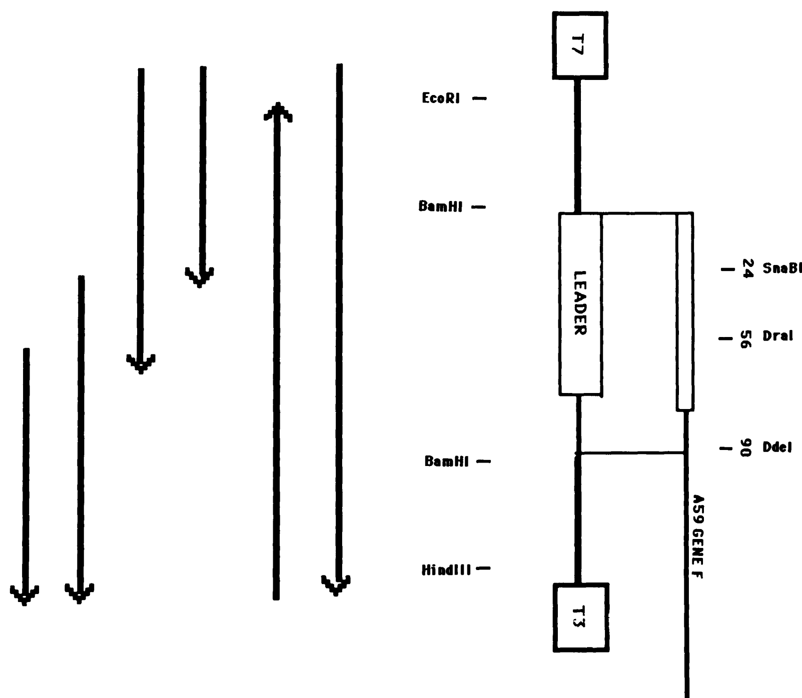

Figure 1. Schematic map of leader plasmids. The RNA transcripts containing leader sequences I-24, I-56 or all 72 nucleotides were derived from pBSL digested with SnaBI, DraI or HindIII respectively. RNA transcripts containing leader sequences beginning at nucleotide 24 of leader were derived from pBSL-SSΔ digested with HindIII. RNA transcripts beginning at nucleotide 56 of leader was derived from plasmid F82D digested with NarI.

Preparation of Protein Extracts

Cellular lysates containing the N protein were prepared from DBT cells infected with the A59 strain of MHV, as previously described[4,7]. N protein was also obtained by electroelution following resolution by SDS-polyacrylamide gel electrophoresis. Protein lysates, prepared as described above, were electrophoresed on 10% polyacrylamide gels[4]. Exterior lanes were loaded with ^{32}P labelled protein lysates of infected cells. Following electrophoresis, the ^{32}P-labelled N protein was identified by autography using the wet gel. A slice of gel containing the unlabelled N was removed and N was isolated by electroelution in a Schleicher and Schuell electroeluter using 25 mM Tris, 0.165 M glycine, pH 8.0 buffer as described[8]. Protein was quantitated by the Branford colorimetric microassay[9]. The identity of purified N was confirmed by Western blot analysis using the anti-N monoclonal antibody designated A1.10[2].

RNA Gel Retardation Assay

Binding reactions and gel conditions are a modification of those described by Schneider et al. (10). A typical 20 ul binding reaction contained 25 mM HEPES pH 7.5, 1 mM EDTA, 150 mM NaCl, 5 mM DTT, 0.5 mM PMSF, 10% glycerol, and 5 ug of total cytoplasmic mouse liver RNA as a nonspecific competitor. Prior to their addition, RNA transcripts were boiled for 1 minute and quenched on ice. The final reaction mixture contained 20-30 ng of ^{32}P labeled RNA transcript and 10-20 ug of the appropriate protein extract in the presence or absence of cytoplasmic liver RNA as a non-specific competitor. After incubation at room temperature for 10 minutes, 5 ul of loading dye containing 50% glycerol, 0.05% bromophenol blue and 10 mM DTT was added. 7% acrylamide gels (made from a 44% acrylamide: 0.8% w/v bisacrylamide stock) containing 5% glycerol in Tris-glycine buffer (50 mM Tris base, 0.33 M glycine, pH 8.5) were equilibrated overnight by storing them in Tris-glycine buffer. Prior to loading, gels were pre-run at 25 mAmps until constant current was obtained. Fresh running buffer was added prior to addition of the samples. Samples were then electrophoresed at room temperature for 2-4 hours at 25 mAmps (until the bromophenol blue reached the bottom of the gel).

RESULTS

Identification of Specific RNA/Protein Interactions in Infected Cell Lysates

The GRA was used to resolve specific RNA-protein interactions between MHV leader RNA sequences, and the proteins from viral infected cells or purified N protein. Prior to addition, ^{32}P-labeled in vitro transcripts were heated to 100°C for 1 minute, quenched on ice, and then incubated at 25°C with the appropriate protein. Figure 2 shows that the RNA transcript containing 72 nucleotides of (+) sense viral leader RNA migrates as a retarded heterogeneous population when incubated with infected cell lysate. Mouse liver RNA had previously been demonstrated by Northwestern blot analysis to be an effective competitor of non-specific RNA/protein interaction[4]. Figure 2 shows that the retarded population is resolved into a discrete band with the addition of nonspecific competitor.

Two lines of evidence indicated that the N protein is required for formation of the leader RNA/protein complex. First, the persistence of the discrete complex formed in the presence of competitor RNA is only seen with extracts prepared from infected cells and not with extracts from uninfected cells. In addition, we have confirmed that N protein, which has been isolated by elution from 10% SDS-polyacrylamide gels, generates the leader specific protein complex (Figure 3).

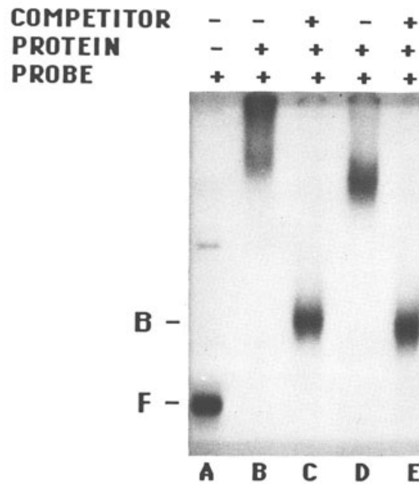

Figure 2. GRA of infected cell lysates binding to leader-containing transcripts, derived from pBSL digested with HindIII. The migration of transcript alone is shown in lane A. Radiolabeled in vitro transcripts were incubated at room temperature with 40 ug (lanes B and C) or 20 ug (lanes D and E) of infected cell lysate either alone (lanes B and D), or in the presence of 5 ug of liver RNA as a nonspecific competitor (lanes C and E).

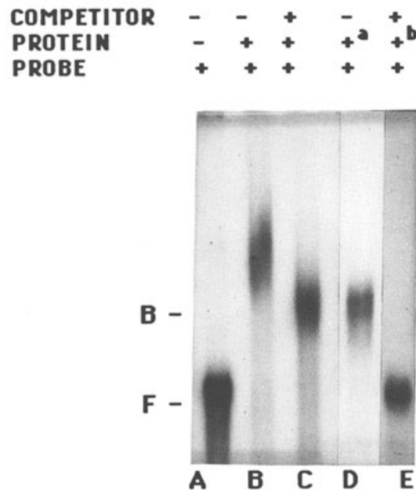

Figure 3. GRA of uninfected cell lysates or purified N protein. In vitro transcripts derived from pBSL, digested with HindIII, were incubated incubated with 20 ug of infected cell lysates (lanes B and C) in the absence (lane B) or presence (lane C) or 5 ug of liver RNA as a nonspecific competitor. Lane D shows the transcript incubated with 7.8 μg of purified N. Lane E shows the transcript incubated with 40 μg of unifected cell lysate and 5 μg of liver RNA as nonspecific competitor.

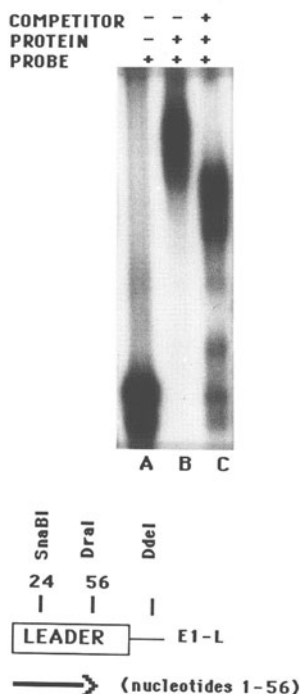

Figure 4. GRA of infected cell lysates binding to transcripts containing nucleotides I-56 of leader derived from pBSL digested with DraI. The migration of transcript alone is shown in lane A. Radiolabeled in vitro transcripts were incubated at room temperature with 20 ug of infected cell lysate (lanes B and C), in the absence (lane B), or in the presence of 5 ug of liver RNA as a nonspecific competitor (lane C).

Sequence Specificity Involved in Complex Formation

Three lines of evidence confirm the sequence specificity of the leader RNA/N-protein interaction. First, an in vitro transcript consisting solely of plasmid sequences is unable to form a specific complex. Secondly, there is no specific complex formation when 5S RNA is used as a probe. Finally, the MHV leader transcript in the (-) sense in the presence of nonspecific competitor, does not form a specific RNA-protein complex (data not shown). These results confirm that the RNA/protein interaction is specific for leader sequences in the (+) sense[11].

Mapping Binding Regions of Leader RNA

We have used the GRA to map the N protein binding region of MHV leader RNA. This was accomplished by testing RNA transcripts containing partial sequences of MHV leader RNA. RNA transcripts containing nucleotides 1 to 24 (data not shown), nucleotides 24 to 90 (Figure 6) or nucleotides 56 to 180 (Figure 5) did not exhibit specific binding. However, RNA transcripts containing nucleotide sequences from 1 to 56 (Figure 4) and from 1 to 90 (Figure 2) do exhibit specific binding activity. This data is summarized in Figure 7.

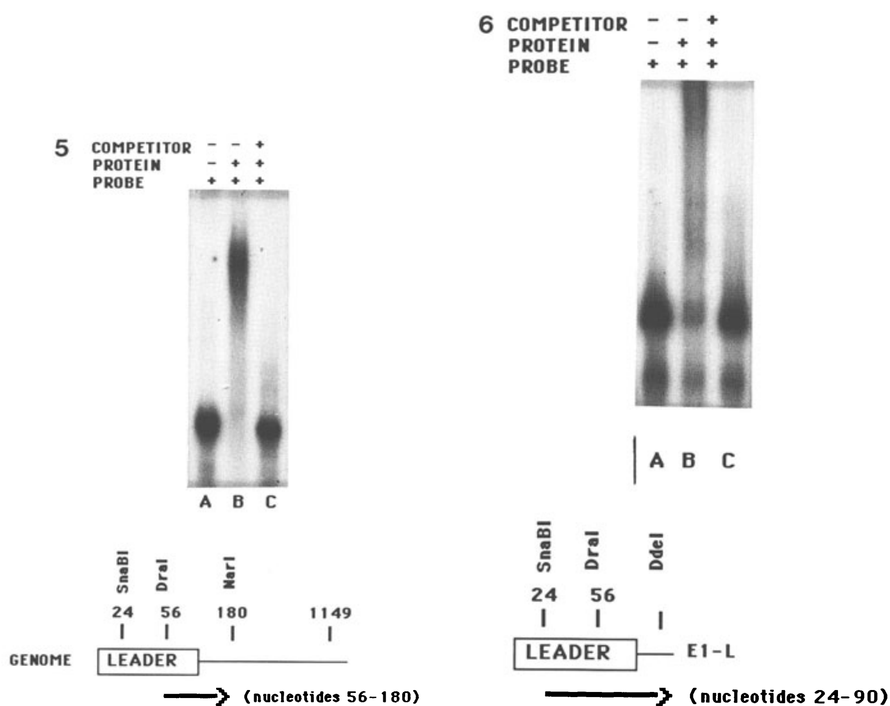

Figure 5. GRA of infected cell lysates binding to transcripts beginning at leader nucleotide 56, derived from pF82D digested with NarI. The migration of transcript alone is shown in lane A. Radiolabeled in vitro transcripts were incubated at room temperature with 20 ug of infected cell lysate (lanes B and C), in the absence (lane B), or in the presence of 5 ug of liver RNA as a nonspecific competitor (lane C).

Figure 6. GRA of infected cell lysates binding to leader-containing transcripts, derived from pBSL-SSΔ digested with HindIII. The migration of transcript beginning at nucleotide 24 of leader alone is shown in lane A. Radiolabeled in vitro transcripts were incubated at room temperature with 20 ug of infected cell lysate (lanes B and C), in the absence (lane B), or in the presence of 5 ug of liver RNA as a nonspecific competitor (lane C).

Secondary Structure Predictions of Leader RNA

We performed a computer analysis of potential secondary structure of free leader RNA (Zuker[12]) and this predicts a proximal stem loop structure (Figure 7). This structure is conserved between transcripts containing leader nucleotides 1 to 56 and 1 to 90. These RNAs exhibit specific binding (Figures 2 and 4). In contrast, the RNAs in which the stem loop structure is absent did not exhibit specific binding (Figures 5 and 6).

DISCUSSION

These experiments were initiated from our interest in studying the RNA/protein interactions involved in viral morphogenesis. The RNA binding property of the viral N protein has been previously demonstrated by the use of Northwestern blotting techniques and these data show specificity of binding for the leader RNA sequences

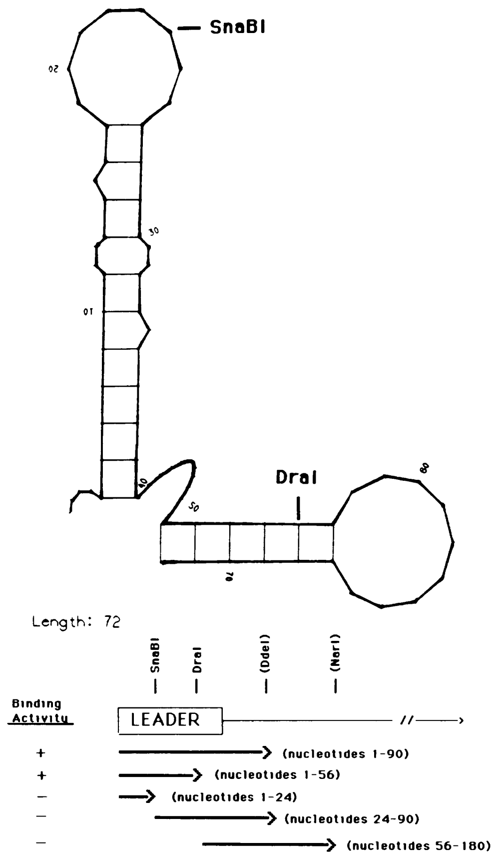

Figure 7. Summary of binding activity and predicted secondary structure of A59 leader sequences (I-72). The predicted secondary structure was generated by computer analysis using the program by Zuker[12]. Below is a summary of transcript binding activity observed in the gel retardation assay.

expressed in a (+) sense configuration[4,5]. We have also previously shown that monoclonal antibodies specific for the N protein could immunoprecipitate RNA/protein complexes[3]. These studies demonstrated that not only free viral leader RNA, but also viral genomic and mRNAs, which have leader sequences at their 5'-ends, are complexed with the N protein[3,4], consistent with the findings that all (+) sense encoded virus RNA species contain leader sequences at their 5' ends[13].

Virus subgenomic mRNAs are not packaged into mature virions at a detectable level, and hence other signals may be involved in directing genomic RNA into the assemble pathway. Examination of specific MHV leader RNA/N protein interactions will allow definition of the role of the N protein during the lytic cycle and persistence phases of MHV infection. For example, RNA sequences found on genomic RNA but absent from mRNA may control the cooperativity of N protein binding, such that RNA is appropriately presented or packaged into virions. Protein/protein interactions between the N protein complexed to free leader RNA and the viral polymerase may also be required for viral transcription. In addition, N may be involved in binding (+) sensed genomic RNA and mRNAs during replication and/or transcription in order to prevent anti-sense inhibition of gene expression.

The data presented in this report confirm our previous description of the specificity of the binding of MHV leader RNA in the (+) sense configuration to the MHV encoded N protein[4]. Our mapping data using RNA transcripts containing various leader sequences demonstrates that the leader RNA/N interactions may depend on secondary structure as well as specific nucleotide sequences, since only those transcripts exhibiting the secondary structure comparable to free leader exhibit specific binding.

We are presently constructing plasmids that contain mutations in the stem or loop site of the predicted conserved secondary structure. These mutations will determine if primary sequences are the sole determinant, or the extent to which these sequences in the correct context of secondary structure are required for recognition. Recently we have conducted competition experiments using purified N protein and the previously described RNA transcripts as homologous competitors. Those results confirm our mapping data in which only those transcripts with the designated secondary structure are able to compete for binding activity (data not shown). In addition, testing of the predicted secondary structure is being determined through the use of specific chemical modifications and cleavage, as well as using various ribonucleases whose activity are sensitive to secondary structure. The ability to isolate RNA complexes by GRA is also well suited for the application of in situ chemical cleavage and footprinting analysis, such as has been described by Murakawa et al.[14] and Nielsen et al.[15]. We are currently using these approaches to map the specific nucleotides of leader RNA which interact with the MHV N protein.

ACKNOWLEDGEMENTS

This work was supported by grant DMB-17148 from the National Science Foundation (SAS); GM 40146 from the N.I.H. (RJD) and a grant from the Medical and Life Insurance Foundation (RJD).

REFERENCES

1. P. R. Brayton, M. M. C. Lai, C. D. Patton and S. A. Stohlman, J. Virol. 42:847-853 (1982).
2. R. S. Baric, C. -K. Shieh, S. S. Stohlman and M. M. C. Lai, Virol. 156:342-354 (1987).
3. R. S. Baric, G. W. Nelson, J. Q. Fleming, R. J. Deans, J. G. Keck, N. Casteel and S. A. Stohlman, J. Virol. 62:4280-4287 (1988).
4. S. A. Stohlman, R. S. Baric, G. W. Nelson, L. H. Soe, L. M. Welter and R. J. Deans, J. Virol. 62:4288-4295 (1988).
5. S. G. Robbins, M. F. Frana, J. J. McGowan, J. F. Boyle and K. W. Holmes, Virol. 150:402-410 (1986).
6. J. Armstrong, S. Smeeken and P. Rottier, Nucl. Acids Res. 11:883-891.
7. M. M. C. Lai and S. A. Stohlman, J. Virol. 26:236-242 (1978).
8. E. Jacobs and A. Clad, Anal. Biochem. 154:583-589 (1986).
9. M. Bradford, Anal. Biochem. 72:248 (1976).
10. R. Schneider, I. Gander, U. Miller, R. Mertz and E. L. Wennacher, Nucl. Acids Res. 14, 1303-1317.
11. L. M. Welter, G. Nelsen, S. A. Stohlman and R. J. Deans, Submitted to Nucl. Acids Res. (1986).
12. M. Zuker and P. Stiegler, NAR 9:133-148 (1981).
13. M. M. C. Lai, Baric, R. S., Brayton, P. R. and Stohlman, S. A. Proc. Natl. Acad. Sci. USA 81:3626-3630 (1984).
14. G. J. Murakawa, B. C. Chen, M. D. Kuwabara, D. P. Nierlich and D. S. Sigman, Nucl. Acids Res. In Press.
15. P. E. Nielsen, C. Jeppesen and O. Buchardt, FEBS Letters 235:122-124 (1988).