STUDIES INTO THE MECHANISM FOR MHV TRANSCRIPTION

Mary C. Schaad, Wan Chen, Sheila A. Peel and
Ralph S. Baric

Program in Infectious Diseases, Department of Epidemiology, University of North Carolina at Chapel Hill Chapel Hill, North Carolina 27599-7400

ABSTRACT

Previous studies have demonstrated that the MHV genome is divided into seven transcriptional units which are transcribed from highly conserved intergenic start sites (UCU/CAAAC) into mRNA containing a common leader RNA at the 5' end and a coterminal 3' end. In this manuscript, we provide evidence that an additional transcriptional unit is encoded at the 3' end of the MHV genome and is transcribed from a perfect intergenic region into a leader-containing ~800 nt mRNA . This mRNA could potentially encode a small 17-18 kDa protein which is identical to the C-terminal third of the nucleocapsid gene.

INTRODUCTION

Mouse hepatitis virus, a member of the coronaviridae, contains a single-stranded plus polarity RNA of about 31 kb in length. The genomic RNA is enclosed within a helical nucleocapsid constructed from multiple copies of a 50 kDa nucleocapsid protein designated N. The nucleocapsid is surrounded by a bi-lipid envelope derived from internal host cell membranes and contains two or three virus specific glycoproteins. The S glycoprotein is 180/90 kDa in molecular weight and forms the distinct surface projections of the virus particle. The M glycoprotein gene is about 23 kDa molecular weight and contains O-linked glycosidic moieties. Some strains of MHV also contain a gene encoding a 65 kDa hemagglutinin esterase protein[1].

The mechanism of MHV transcription is controversial and remains under study. The majority of data suggest that the incoming genomic RNA is transcribed into a full-length negative-stranded RNA which acts as template for the synthesis of 6 or 7 subgenomic mRNAs. It is postulated that a free leader RNA of about 65-72 nucleotides in length is transcribed from the full-length minus-stranded RNA and acts in trans to prime transcription from highly conserved intergenic start sites (UCU/CAAAC) just 5' to each mRNA. The subgenomic mRNA can then act as template for the synthesis of subgenomic negative-stranded RNAs which function as templates for additional rounds of subgenomic mRNA synthesis[2,3,4].

In this study, we provide evidence that a previously unrecognized mRNA is transcribed from a highly conserved intergenic start site that is located within the N gene sequence at the 3' end of the genome. The new mRNA could potentially encode a 154-155 amino acid protein which is in frame with the C-terminal third of the nucleocapsid gene of MHV.

Coronaviruses, Edited by H. Laude and J.F. Vautherot
Plenum Press, New York, 1994

METHODS

Virus and Cells

Mouse hepatitis virus strains MHV-A59, MHV-S and MHV-1 were used throughout the course of this study. The virus was propagated on DBT cells maintained at 37°C in dMEM containing 10% Nu-serum, and 1% kanamycin/gentamicin.

Isolation of Viral mRNA and Replicative Form RNAs

Cultures of DBT cells were infected with MHV at a MOI of 10 and maintained at 37°C for 8-12 hrs postinfection. Medium was removed and the cells washed in ice cold PBS. To isolate RNA for replicative form (RF) analysis, intracellular RNA was isolated by lysing cells in 0.5 ml of Hypo TKM (10 mM Tris-Hcl, pH 7.5, 150 mM KCl, 1.5 mM $MgCl_2$) containing 100 ug/ml Proteinase K and 0.05% NP40. The nuclei were removed by slow speed centrifugation, supernatents extracted with multiple phenol, phenol/chloroform/isoamyl alcohol, and chloroform treatments then ethanol precipitated. To isolate intracellular RNA for mRNA analysis, the cells were lysed as previously described[2].

For RF analysis, the intracellular RNA was resuspended in 10 ul of water, adjusted to 1 X DNAse buffer (100 mM NaCl, 10 mM Tris, pH 7.8, 2 mM EDTA, 2 mM $MgCl_2$, 2 mM $CaCl_2$) and treated with 4 U of DNase 1 for 15 min at 30° C. Following DNase treatment, the sample was adjusted to 1 X RNAse buffer (700 mM NaCl, 10 mM Tris, pH 7.4, 30 mM EDTA) and treated with 50 ng/ml RNase A for 15 min at 30°C. The samples were loaded onto 1% agarose gels containing TBE (89 mM Tris, 89 mM Boric acid, 2 mM EDTA)[2].

PCR Cloning and Sequencing

Primers were designed from published sequences from the MHV-A59 N gene and leader RNA. The forward primer was 5'-TAAGAGTGATTGGCGTCCGTACG-3' which corresponded to nucleotides (nt) 3-25 in the leader RNA sequence. The reverse primer was derived from nt 920-944 in the N gene sequence and was 5'-GCAAGAATGGGGAACTGTGGATCAC-3'. Intracellular RNA was isolated from MHV-A59, MHV-S, and MHV-1-infected DBT cells. Two ug of RNA was reverse transcribed into cDNA in 1 X RT buffer (10 mM Tris, pH 8.8, 50 mM KCl, 0.1% Triton X-100), 5 uM oligo dT, 500 uM dNTPs, 0.5 U/ul RNASIN, 10 mM $MgCl_2$, 10 mM DTT and 8 U AMV reverse trans-criptase at 42°C for 1 hr. The reaction was stopped by the addition of 2 ul of 0.5 M EDTA, extracted with phenol/chloroform and then precipitated with ethanol. Approximately 30 ng of cDNA was amplified in a 50 ul reaction containing 1 x TAQ DNA polymerase buffer, 2 mM $MgCl_2$, 1.25 mM each dNTPs and 5 U of TAQ DNA polymerase (Promega). After 25 cycles, the PCR products were separated on 4.5% NuSieve (FMC) agarose gels and bands of the appropriate size were excised and electroeluted into 5 M potassium acetate. The eluted fragments were phenol/chloroform extracted and precipitated in ethanol. Pellets were washed 3X in 70% ethanol, 1X with 95% ethanol and vacuum dried.

PCR products were directly sequenced by the chain termination method using a sequenase kit (USB). Approximately 300 ng of purified PCR product was incubated with 20 pM primer and 10% DMSO and denatured at 97°C for 8.5 min. The reaction was cooled to room temperature for 1 min, and the extension reactions were performed on ice with a 1:10 dilution of the nucleotide mix following standard Sequenase version 2.0 protocols (USB).

RESULTS

Demonstration of a Small Eighth mRNA and RF RNA in MHV-Infected Cells

We have noted on occasion the presence of an additional small RNA less than 1,000 nt in length in MHV-infected cells. To determine if this small mRNA species was in fact a product of the MHV genome, cultures of DBT cells were infected with MHV-A59, treated with actinomycin D (10 ug/ml) at 2 hr postinfection, and radiolabeled with 100 uCi/ml ^3H-

Uridine from 3-7 hr postinfection. The RNA was isolated and separated on 1% agarose gels. As expected, the seven characteristic viral mRNAs previously reported in MHV-A59 infected cells were readily detected. In addition, a small <1000 nt RNA was also present in infected, but not uninfected cells (Figure 1).

If this small RNA is of viral origin, analysis of RF RNAs in infected cell lysates should demonstrate the presence of a corresponding double-stranded RF RNA. To test this hypothesis, cultures of DBT or 17CL1 cells were infected with MHV-A59, treated with actinomycin D, and radiolabeled with ^{3}H-Uridine. The intracellular RNA was digested with DNAse I and RNAse A and separated on 1% agarose gels. In agreement with

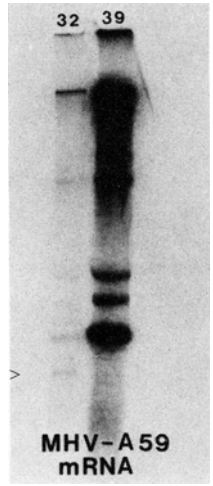

Figure 1. ^{3}H-Uridine-labeled MHV-A59 intracellular RNAs isolated at 10 hr postinfection from DBT cells labeled at 32°C and 39.5°C.

previous findings, seven RNase A resistant RF RNAs were present in MHV-infected cells[2]. In addition, an eighth subgenomic RF RNA was occasionally detected (data not shown). The size of this RF RNA was about 1600-1800 bp in length suggesting that it was involved in the transcription of an approximate 800-900 nt mRNA. The ability to detect the small mRNA and RF RNA by radiolabeling experiments was difficult, and at times their presence was not detectable suggesting that the mRNA is usually made in small quantities.

Table 1. Conservation of intergenic 7-8 in MHV

VIRUS	CONSENSUS INTERGENIC (Y=C or T)	PUTATIVE 7-8 INTERGENIC	PUTATIVE OPEN READING FRAME (AMINO ACIDS)
MHV-A59	TCYAAAC	TCCAAAC	154
MHV-JHM	TCYAAAC	TCCAAAC	155
MHV-1		TCCAAAC	155
MHV-3		TCCAAAC	154
MHV-S		TCCAAAC	154

14

Identification of a Potential Intergenic Start Site for mRNA 8 Transcription.

A highly conserved intergenic sequence UCU/CAAAC is present in the MHV genome just 5' to the start site for the synthesis of each subgenomic mRNA[1,3]. If an additional ~800-900 nt mRNA is encoded in the MHV genome, it seems likely that such a conserved start site would be present within the 3' end of N gene coding sequences. Nucleotide sequence analysis of such a "start" site among the different MHV strains is shown in Table 1. A perfect intergenic start site was present at nt 828 in the MHV-A59 N gene[5]. In addition, sequence analysis also revealed "start sites" at similar locations at the 3' end of the genome among other group II coronavirus strains (BCV, OC43, data not shown).

To determine whether the small mRNA found in MHV-infected cells initiated from this putative start site, oligodeoxynucleotide primers were chosen which were located at the 5' end of the leader RNA sequence and downstream from the potential intergenic region in the MHV N gene coding sequences. Intracellular RNA was isolated from MHV-A59, MHV-S, and MHV-1 infected cultures and reverse transcribed into cDNA. Following 25 cycles of PCR amplification, the expected 160 base pair product was detected in all MHV strains tested, but not in uninfected controls. Surprisingly, two additional PCR products were also identified in MHV-1 (120 bp) and MHV-S (100 bp) infected cells (Figure 2).

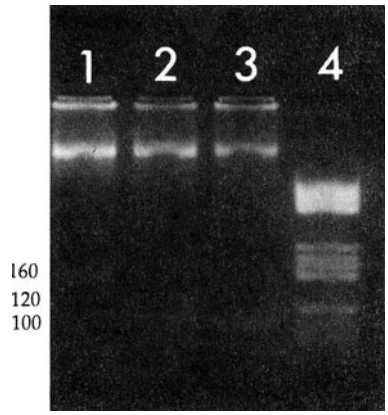

Figure 2. RNA extracted from MHV-1, MHV-S and MHV-A59-infected cells was reverse transcribed and amplified by 25 cycles of PCR using primers originating from the 5' end of the leader RNA sequence (nt 3-25) and within the N Gene (nt 920-944). Lanes 1,2 and 3: MHV-A59, MHV-1 and MHV-S. Lane 4: DNA Molecular Weight Marker V (Boehringer Mannheim).

Partial Sequence Analysis of the MHV mRNA 8 Gene

The partial sequence of the MHV-A59, MHV-S and MHV-1 "mRNA 8" and its start sites are shown in Figure 3. The sequence in this region was identical to those previously published[5]. As expected, all of the PCR-derived products contained the MHV leader RNA sequence at the 5' end of the mRNA consistent with the replication strategy of the virus. The perfect intergenic present at nt 828 was used as the start site for mRNA 8 synthesis from all MHV strains tested (Figure 3). In addition to the transcripts initiating from the perfect intergenic, MHV-1 and MHV-S had transcripts initiating from two "imperfect" intergenics. mRNAs initiating from either the perfect intergenic (UCCAAAC) in the MHV strains examined, or from the second start site in MHV-1 (CCAAUC), would encode a 154-155 amino acid protein, which was in frame with the carboxy terminus of the N protein[5] (Figure 4). mRNA initiated from the second start site in the MHV-S N gene (UAAAC) was, however, unique to MHV-S infected cells and started downstream from the ATG initiation codon for the 154-155 amino acid protein. This ORF would encode a putative 28 amino acid protein out of frame with respect to N. Unless additional start sites are located within the MHV-A59 and MHV-1 genomes to encode a mRNA that expresses the putative

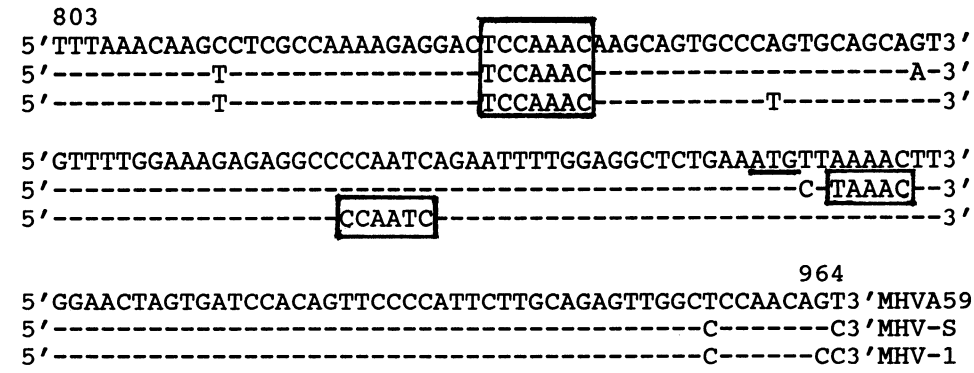

Figure 3. PCR amplified products spanning the putative mRNA 8 start site for MHV-A59, MHV-S and MHV-1 were sequenced. Boxed nucleotide sequences represent the start sites for mRNA 8 synthesis in each MHV strain. The initiation codon ATG is underlined and is in frame with the N gene sequence.

28 amino acid protein, it seems unlikely that this protein plays any significant role in virus replication.

DISCUSSION

The MHV genome is divided into seven transcriptional units which are synthesized from highly conserved intergenic start sites (UCU/CAAAC) into mRNA containing a common leader RNA sequence derived from the 5' end of the genome and a 3' co-terminal end[1]. In this manuscript, we provide evidence suggesting that a new transcriptional unit is encoded within the N gene at the 3' end of the MHV genome. The mRNA 8 transcript is initiated from a perfectly conserved intergenic sequence at nt 828 in the MHV-A59 sequence into a mRNA of about 900 nt in length. The ability to detect this mRNA and its corresponding RF RNA using radiolabeling experiments was highly variable suggesting that it was made in relatively small amounts or was difficult to detect because of its size and location in the gel.

Although we have not demonstrated the putative protein product from this mRNA, it should encode a 154-155 amino acid protein of about 17-18 kDa which is identical to the C-terminal third of the N gene of MHV. Interestingly, a variety of N "degradation" products have also been described in MHV infected cells[6]. It is unclear whether any of these products may represent the product of mRNA 8. While a variety of co-terminal truncated genes, which function very differently during infection, have previously been described in the SV40 and hepatitis delta virus genomes[7,8], this represents the first such demonstration of this type of genetic organization in the MHV genome. This transcriptional unit may also be expressed during other group II coronavirus infections. Highly conserved intergenic sequences are also present within the BCV and HCV-OC43 genomes which could be transcribed into mRNA encoding a truncated protein which is in frame with the C-terminus of the nucleocapsid protein. In addition, a small RNA of <1,000 nt and its corresponding RF RNA have been demonstrated in BCV-infected cells[4].

The N gene functions in nucleocapsid formation, transcription and pathogenicity. Sequence analysis among the different MHV strains indicate that the N protein can be divided into three highly conserved domains[5]. The putative mRNA 8 product would encode amino acids 300-454 containing a portion of the basic domain II and all of the acid domain III[5]. A second internal open reading frame (nt 65-688) is also present within the N gene which may encode a 207 amino acid protein[5] (Figure 4). A ts mutant, Alb4, has been isolated which contains a 28 amino acid in frame deletion in the N gene sequence between amino acids 380 and 408[9]. The finding of a new mRNA which potentially encodes a protein which is identical to amino acids 300-454 in the N gene raises questions about the actual role of the nucleocapsid gene in virus transcription and pathogenicity. It is also unclear

whether the deletion in Alb4 preferentially alters the function of the N protein, the function of the truncated C-terminal product of mRNA 8, or both.

The finding of two "illegitimate" start sites in the MHV-S and MHV-1 genomes are puzzling, but support previous studies that imperfect start sites can be used to initiate mRNA transcription[10]. In the case of MHV-1, the CCAAUC "start" site is also conserved in the MHV-A59 and MHV-S genomes, but leader-containing transcripts are present only in MHV-1-infected cells. Interestingly, in MHV-1 and B1-infected cells, new mRNA species (2-2,3-1) have been identified which start at a very similar intergenic region (UC/UUAAUC) suggesting that CCAAUC may also represent a functional intergenic in the MHV-1 genome[10]. It was suggested that the presence of three UCUAA repeats encoded within the leader sequence at the 5' end of the MHV-1 genome may permit the use

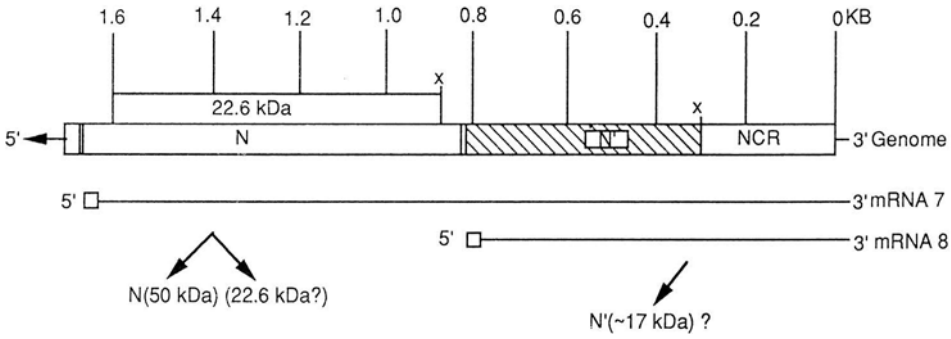

Figure 4. Tentative genetic map of the 3' end of the MHV genome. The boxed regions represent likely ORFs that are encoded in the 3' end of the genome in the N gene sequence. Solid double bars represent the location of perfect intergenic "start" sequences and the x denotes the termination codons for these ORFs. Hatched lines represent the putative 17-18 kDa protein that is identical to the C-terminus of the N gene and may be encoded by mRNA 8.

of these particular intergenic regions[10]. In the case of the "imperfect" start site in the MHV-S genome, a specific mutation is present within this intergenic region (UAAAC) that is not present in MHV-1 or MHV-A59 genomes (AAAAC), and this may account for initiation of mRNA synthesis. Since this mRNA is MHV-S specific and contains a 28 amino acid open reading frame that is probably not expressed in other strains of MHV, it seems likely that this gene is not critical for virus growth. Additional studies are needed to determine the role of any of these gene products in virus replication.

REFERENCES

1. M.M.C. Lai. *Ann. Rev. Microbiol.* 44:303 (1990).
2. S.G. Sawicki and D.L. Sawicki. *J. Virol.* 64:1050 (1990).
3. S. Makino, M. Joo, and J.K. Makino. *J. Virol.* 65: 6031-6041 (1991).
4. M.A. Hofmann, P.B. Sethna, and D.A. Brian. *J.Virol.* 64: 4108-4114 (1990).
5. M.M.Parker and P. S. Masters. *Virology* 179:463 (1990).
6. S.G. Robbins, M.F. Frana, J.J. McGowan, J.F. Boyle, and K.V. Holmes. *Virology* 150:402 (1986).
7. J. Tooze DNA Tumor Viruses, Molecular Biology of Tumor Viruses, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1981).
8. F-L. Chang, P-J. Chen, S-J. Tu, C-J. Wang, and D-S. Chen. *PNAS USA* 88:8490 (1991).
9. C.A. Koetzner, M.M. Parker, C.S. Ricard, L.S. Sturman, and P.S. Masters. *J. Virol.* 66:1841 (1992).
10. N. La Monica, K. Yokomori, and M.M.C. Lai. *Virology* 8:402 (1992).