Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 3626–3630, June 1984
Biochemistry

Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus

(murine coronavirus/primer extension/long-terminal-repeat homology)

MICHAEL M. C. LAI, RALPH S. BARIC, PETER R. BRAYTON*, AND STEPHEN A. STOHLMAN

Departments of Microbiology and Neurology, University of Southern California School of Medicine, Los Angeles, CA 90033

Communicated by Peter K. Vogt, February 17, 1984

ABSTRACT **Mouse hepatitis virus, which replicates in cytoplasm, contains leader RNA sequences at the 5′ end of the virus-specific mRNAs. We have sequenced this leader RNA by synthesizing cDNA from a synthetic oligodeoxyribonucleotide primer (15-mer) that is complementary to the sequences at the junction site between the leader and body sequences of the mRNAs. The leader sequences on each mRNA have exactly the same size, which span ≈70 nucleotides. Leader cDNA fragments obtained from several mRNA species were sequenced and found to be identical. Computer analysis of the leader RNA sequences shows that they share extensive sequence homology with the long-terminal-repeat region of several mammalian sarcoma viruses, suggesting possible common functions. This is a novel case of spliced leader sequences in the mRNAs of a cytoplasmic virus. An identical leader sequence is also present at the 5′ end of the virion genomic RNA. The leader RNA is thus probably encoded by the virion genomic RNA template and is fused to the different body sequences of the various mRNAs. Since conventional RNA splicing is not involved, a novel mechanism for fusing two noncontiguous RNA segments in the cytoplasm must be utilized during viral transcription. Several minor cDNA bands longer than the leader were also synthesized, suggesting the possible presence of partially homologous sequences in other parts of the genome RNA.**

Mouse hepatitis virus (MHV), a member of the Coronaviridae, contains a single-stranded 60S RNA of M_r 5.4 × 10^6, which is of positive polarity (1). The virus replicates exclusively in the cytoplasm of several established mouse cell lines (2, 3) and matures by budding into endoplasmic reticulum (4). Upon entry into the cell, the genomic RNA is first transcribed by a virus-specific "early" RNA polymerase (5) into a genomic-sized negative-stranded RNA (6). This RNA species serves as the template for the synthesis of positive-stranded genomic RNA and six subgenomic mRNA species by a "late" RNA polymerase (5, 7). These mRNAs range from M_r 0.6 × 10^6 to M_r 5.4 × 10^6 and have a nested-set structure containing sequences starting from the 3′ end of the genomic RNA that extend for various distances toward the 5′ end (7). We have recently shown that these mRNAs contain a stretch of leader RNA sequences at their 5′ ends (8), which are not derived from a precursor RNA by the conventional RNA splicing mechanism. More likely, a novel mechanism, occurring in the cytoplasm and involving an independently synthesized free leader RNA as the primer for mRNA synthesis, might be employed in the generation of MHV mRNAs (9).

The presence of a "spliced" leader RNA at the 5′ ends of each MHV mRNA is puzzling since these viruses replicate exclusively in the cytoplasm. To understand such a unique mechanism for RNA "splicing," it is essential to determine the nature and sequence of the leader RNA. We have now used primer extension to determine the consensus leader RNA sequences in each mRNA species. The results showed that identical leader RNA sequences are present in at least the majority of mRNAs as well as virion genomic RNA of MHV. This is a novel case of spliced leader sequences in the mRNAs of a cytoplasmic virus.

MATERIALS AND METHODS

Preparation of MHV Intracellular RNA. The intracellular RNA of the A59 strain of MHV was isolated from the infected SAC (−) or DBT cells at 6 hr after infection in the presence of 2 μg of actinomycin D per ml (7). The RNA was extracted with $NaDodSO_4$/phenol as described (7) and used directly for reverse transcription without further purification.

For preparation of individual mRNAs, the total MHV intracellular RNA was mixed with ^{32}P-labeled MHV intracellular RNA (7) and separated by electrophoresis on 1% agarose gels at 90 V for 5 hr (7). Following autoradiography, the individual MHV mRNA species were extracted from the gels according to published procedures (7).

Primer Extension. The synthetic oligodeoxyribonucleotide (15-mer) 5′ A-G-T-T-T-A-G-A-T-T-A-G-A-T-T 3′ was prepared by Bruce Kaplan (Department of Molecular Genetics, The City of Hope Medical Center, Duarte, CA). An aliquot (0.5 μg) of this oligonucleotide was end-labeled at the 5′ end with [γ-^{32}P]ATP by polynucleotide kinase (10), mixed with MHV intracellular RNA, heat-denatured at 100°C for 30 sec, and then hybridized in a total volume of 20 μl containing 80% formamide, 0.4 M NaCl, 0.01 M Pipes buffer (pH 7.0), and 2 mM EDTA at 37°C for 5 hr (11). After ethanol precipitation and several washings, the hybrids were incubated in a buffer containing 1 mM (each) dATP, dCTP, dGTP, and TTP, 10 mM $MgCl_2$, 120 mM KCl, 30 mM mercaptoethanol, and 10 units of reverse transcriptase at 37°C for 1 hr. After alkaline digestion of the RNA template, the reaction products were analyzed by electrophoresis on 12% polyacrylamide gels either with or without 7 M urea according to published procedures (12). The extended primers of various sizes were eluted from the gel and used for further analysis (12).

Hybridization Analysis of the MHV Intracellular RNA. RNA from virus-infected cells was denatured by dimethyl sulfoxide/glyoxal and separated by electrophoresis according to McMaster and Carmichael (13). After electrophoresis, the RNA was transferred to nitrocellulose paper and hybridized to the ^{32}P-labeled leader-specific cDNA (see *Results*) according to the procedure of Thomas (14).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MHV, mouse hepatitis virus; LTR, long terminal repeat.

*Present address: Laboratory of Oral Medicine, National Institute of Dental Research, National Institutes of Health, Bethesda, MD 20205

DNA Sequencing. DNA sequence analysis was performed by a modification of Maxam–Gilbert procedure (12, 15).

RESULTS

Extension of a Synthetic Primer in the Leader Region. Our previous studies suggested that a T1 oligonucleotide, no. 19, is likely to contain sequences created by the fusion between a leader RNA and the body segments of the individual mRNAs of MHV (16). This oligonucleotide has been mapped at the 5′ end of the mRNAs (8). Recently, the map location of oligonucleotide 19 has been confirmed by comparison with the sequences of the cDNA clones of mRNA 7 (17). To characterize the complete sequence of the leader RNA, we synthesized an oligodeoxyribonuclotide complementary to part of oligonucleotide 19. This synthetic 15-mer (5′ A-G-T-T-T-A-G-A-T-T-A-G-A-T-T 3′) was labeled at the 5′ end with [γ-^{32}P]ATP by polynucleotide kinase and hybridized to the MHV-specific mRNAs. Complementary DNA was then synthesized from the hybrid by reverse transcriptase. As shown in Fig. 1*A*, the primer oligodeoxyribonucleotide bound to the MHV mRNAs was extended by reverse transcriptase to yield a single DNA fragment of roughly 70 nucleotides. The primer alone, without hybridizing to RNA, could not be extended. This result suggests that the synthesized cDNA probably corresponds to the consensus 5′-end sequences of the MHV mRNAs, since only one distinct major DNA band was observed. Several minor DNA bands were also occasionally observed, which were more evident when the DNA products were analyzed by electrophoresis in polyacrylamide gels containing 7 M urea (Fig. 1*B*). These bands are constant in size, although the relative amounts of each DNA fragment varied from experiment to experiment. The molar quantities of these fragments are always lower by at least a factor 10 than that of the primary product, the leader-specific fragment (Fig. 1*B*).

The sequences of the major DNA fragment and some of the minor DNA bands were determined by the Maxam–Gilbert procedure (12, 15). The DNA sequences of the major fragment and the corresponding complementary RNA sequences are shown in Fig. 2. These RNA sequences contain sequences corresponding to oligonucleotides 10 and 82 (cap-N-U-A-A-G), which we have shown to be present within the leader RNA region (8, 16). The sequence also shows that no AUG codon is present and that the 3′ half of this RNA segment is poor in G, whereas the 5′ half is rich in G. There is a 7-nucleotide (A-A-A-U-C-U-A) mirror-image direct repeat present within nucleotides 57–70. A comparison with known nucleic acid sequences by computer search revealed that this leader RNA does not represent any cellular DNA or RNA. However, this leader RNA shows an extensive sequence homology with the long-terminal-repeat (LTR) sequence of mammalian sarcoma viruses (Fig. 3). This homology is clustered at the 3′ end of the R region. Of the 72 nucleotides, 30 are shared between the MHV leader RNA and LTR sequences of mammalian sarcoma viruses, including Moloney murine, simian, and feline sarcoma viruses (18). The significance of such homology is unclear.

The minor DNA bands, which are smaller than the complete leader cDNA (Fig. 1*B*), were found to have partial sequences representing the 5′ side of the major cDNA product (Fig. 2). They are, therefore, likely to be premature termination products of *in vitro* reverse transcription. The DNA bands larger than the major leader-specific DNA fragment are of considerable interest. However, due to the difficulty in obtaining enough radioactivity, only partial sequences have been obtained. DNA 5 was found to contain the complete sequences of the major leader cDNA plus an additional 15 nucleotides at the 3′ end of the DNA, which correspond to additional sequences at the 5′ end of the mRNA. The cDNA band 6, on the other hand, contains sequences completely unrelated to the leader cDNA. The origin of these DNA bands is unclear and will be discussed below.

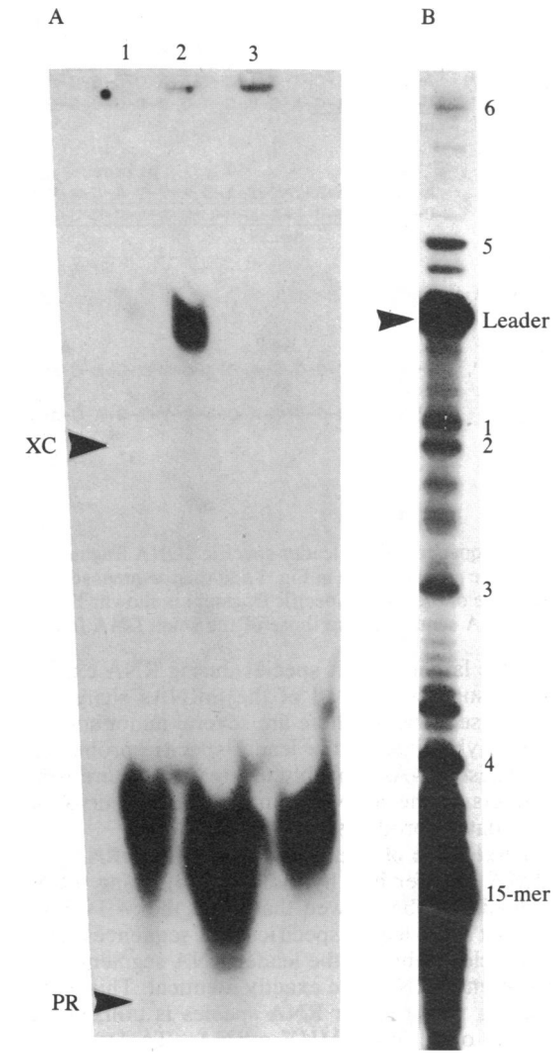

FIG. 1. Primer extension of the unfractionated total mRNAs of MHV. The synthetic oligomer (15-mer), which is complementary to T1 oligonucleotide 19, was labeled with ^{32}P at the 5′ end, hybridized to intracellular mRNAs, and used as a primer for reverse transcription. The cDNA synthesized was boiled at 100°C for 5 min and analyzed by electrophoresis on a 12% polyacrylamide gel (*A*) or a 12% polyacrylamide gel containing 7 M urea (*B*). (*A*) lane 1, free 15-mer; lane 2, 15-mer annealed to total intracellular mRNAs and "primer-extended"; lane 3, 15-mer annealed to uninfected cellular RNAs and primer extended. XC, xylene cyanol; PR, phenol red. The marked DNA bands in *B* were eluted and sequenced as shown in Fig. 2. The arrowhead in *B* marks the position of the smallest *Hae* III fragment (72 nucleotides) of ϕX174 replicative form DNA. The cDNA bands 1–6 were eluted and used for sequence analysis (Fig. 2).

Hybridization of the Leader RNA to All of the mRNAs. Previous studies showed that some of the MHV mRNAs contained identical T1 oligonucleotides within the leader RNA segments (8, 16). To extend this finding to all of the mRNAs, we synthesized a leader-specific cDNA from mRNA 7 using the synthetic 15-mer as a primer. This probe represents the entire leader RNA sequences of mRNA 7 and was used to hybridize to the intracellular RNA of MHV by using RNA transfer blot analysis (14). As shown in Fig. 4, this probe hybridized to all of the subgenomic mRNAs of MHV. Particularly interesting is that mRNA 4, which was usually undetectable with metabolically labeled virus-specific RNA (Fig. 4, lane B) (7), was clearly detected by hybridization. mRNAs 1 and 2 hybridized to a lower extent, probably due to degra-

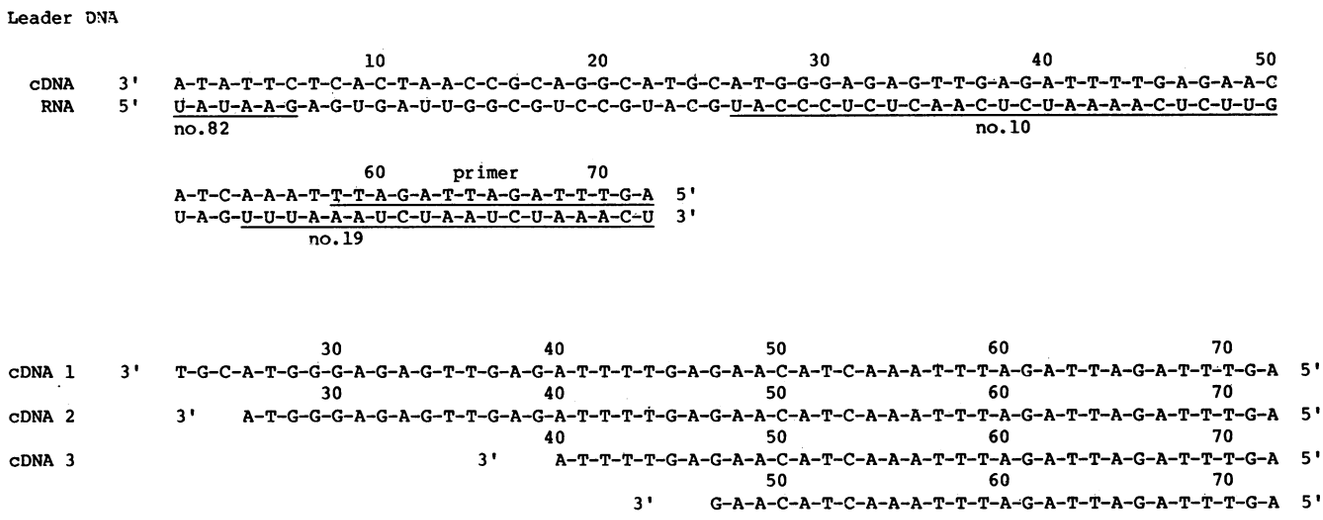

FIG. 2. Sequences of the leader-specific cDNA fragments and the corresponding leader RNA. The DNA fragments were eluted from the polyacrylamide gels as shown in Fig. 1 and then sequenced by a modification of the Maxam–Gilbert procedure (12, 15). Only the corresponding RNA sequence of the leader-specific fragment is shown. The positions of oligonucleotides 82, 10, and 19 (16) are denoted in the RNA sequence. The other DNA sequences are those of the small DNA fragments shown in Fig. 1*B*.

dation of the larger mRNA species during RNA extraction. This result suggests that all of the mRNAs share at least some leader sequences. There are several minor small RNA bands that hybridized to the leader-specific probe (Fig. 4, lane A). These RNAs probably are the premature termination products of the mRNAs (unpublished observation) or their degradation products.

Primer Extension of Each Individual MHV mRNA. Analysis by RNA transfer blot hybridization by using mRNA 7-specific leader probe showed that all of the MHV mRNAs share at least some leader-specific RNA sequences. However, it is not clear whether the leader RNA segments in each of the different mRNAs are exactly identical. This would be expected if a single leader RNA species is utilized for the transcription of all of the MHV mRNAs. To determine the exact size and sequence of the leader RNA segment in each mRNA, we performed primer extension of the synthetic 15-mer hybridized to the individual mRNA. Genomic and subgenomic mRNAs were eluted from the gel as shown in Fig. 4, lane B, and hybridized separately to the 5′-end-labeled synthetic 15-mer. Complementary DNA was then synthesized with reverse transcriptase, and the products were analyzed by electrophoresis on denaturing gels. As shown in Fig. 5, a major cDNA band of 72 nucleotides was synthesized from all of the MHV mRNA species. The sizes of the major products were identical for all of the mRNAs and they were identical to the major cDNA produced from the total unfractionated virus-specific mRNA (Fig. 5, lane T). The major DNA products derived from mRNAs 7, 6, and 5 were sequenced by the Maxam–Gilbert method and found to have sequences identical to that of the consensus sequence obtained from the leader DNA synthesized from the unfractionated mRNA, as shown in Fig. 2 (data not shown). Although the leader DNAs from mRNAs 1, 2, and 3 were not sequenced, they are likely to have identical sequences since all of them have identical sizes and their premature termination products are also similar (see below). This result suggests that all of the mRNAs contain identical leader sequences. Furthermore, the same predominant leader DNA band was also observed when the virion genomic RNA was used for primer extension (Fig. 5, lane G). This DNA band also has the same sequence as that of the leader segment in mRNA 7 (data not shown). Thus, the identical leader sequences also exist at the 5′ end of the virion RNA.

Several minor DNA bands were also synthesized from the individual mRNAs hybridized to the synthetic primer. These DNAs had the same size for all of the mRNA species and were identical to those found with the unfractionated virus-specific mRNAs, although the relative quantities vary with preparations. They are therefore likely to be premature termination products of reverse transcription.

DISCUSSION

This study presents evidence that most, if not all, of the mRNA species of MHV contain a stretch of identical leader

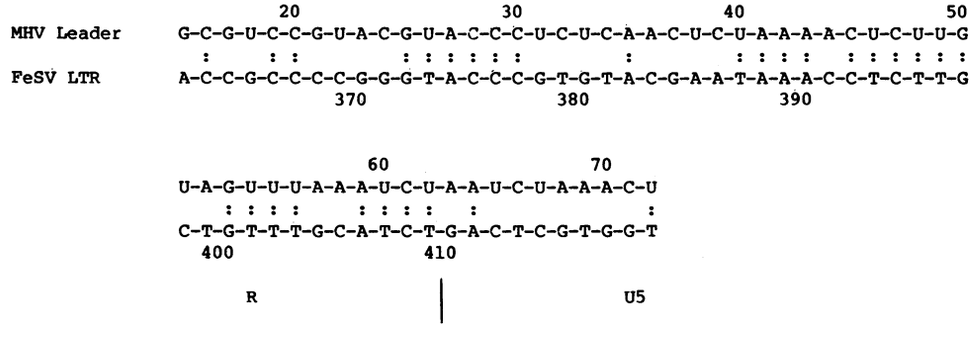

FIG. 3. Sequence comparison between the MHV leader RNA and the LTR region of Gardner–Arnstein feline sarcoma virus (FeSV). The FeSV sequences were obtained from Hampe *et al.* (18). The matched nucleotides are denoted by two dots. The R and U5 domains of LTR regions are marked.

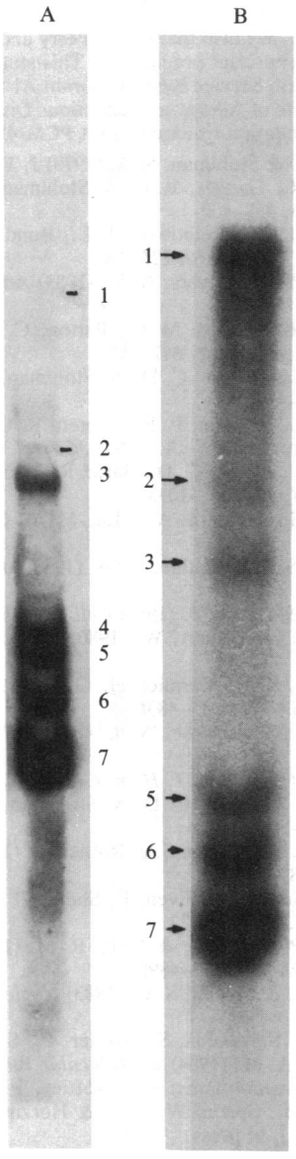

FIG. 4. Hybridization of the intracellular RNAs to the leader-specific cDNA. Lane A, the MHV-specific intracellular RNAs were denatured by dimethyl sulfoxide/glyoxal, separated by electrophoresis on 1% agarose, transferred to nitrocellulose paper, and hybridized to the leader-specific probe, made by primer extension from the synthetic 15-mer that had been hybridized to the purified mRNA 7. Lane B, ^{32}P-labeled RNA from the MHV-infected cells was separated by electrophoresis as shown in lane A. The bands 1–7 represent mRNAs 1–7, respectively.

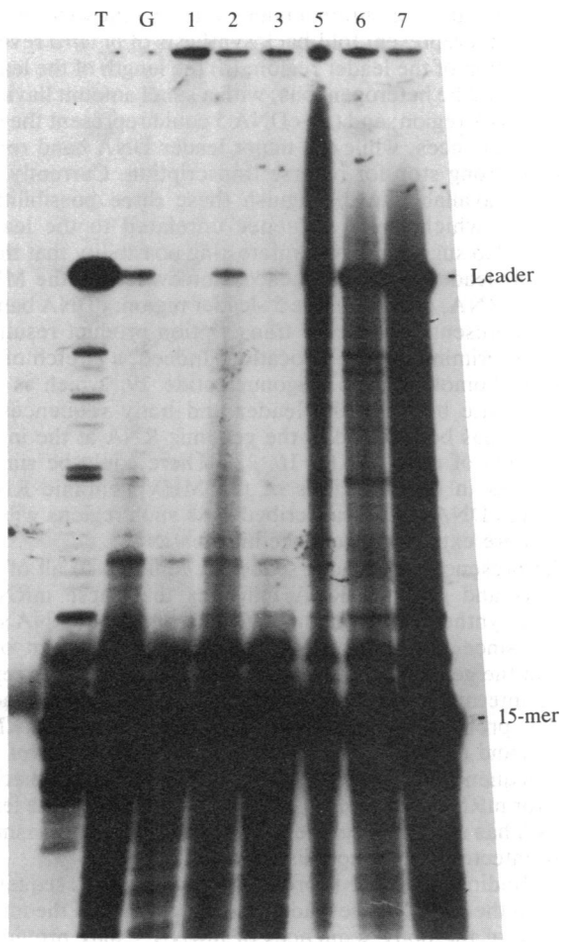

FIG. 5. Primer extension of the individual mRNA of MHV. The same experiments as that in Fig. 1 were performed, with the exception that individual MHV mRNAs were used in this experiment. The ^{32}P-labeled intracellular RNA was separated by electrophoresis as shown in Fig. 4, lane B, and each RNA species was eluted separately (7). The purified RNAs were then used for primer extension by using the synthetic 15-mer as the primer. The cDNAs synthesized were analyzed by electrophoresis on 12% polyacrylamide gels containing 7 M urea. Lane T, total intracellular RNA; lane G, genomic RNA; lanes 1–3 and 5–7, mRNAs 1–3 and 5–7, respectively.

sequences of roughly 70 nucleotides. This is a novel case of spliced leader sequences in the mRNAs of a cytoplasmic RNA virus. These data support conclusions previously reached by oligonucleotide fingerprinting studies that showed that the leader RNA sequences are present only once at the 5′ ends of MHV mRNAs (8). In this study, we have further shown that this leader RNA segment is also present at the 5′ end of the genomic RNA and that it is not a cellular RNA species, thus establishing that the leader RNA originated from the viral genomic RNA sequences. Therefore, the mechanism of MHV RNA synthesis is distinct from that of influenza virus, which utilizes 5′ ends of cellular mRNAs as primers for its own mRNA synthesis (19).

The consensus sequence of the leader RNA as determined in this study is similar, but not identical, to the 5′-end sequences of a cDNA clone of mRNA 7 of the JHM strain of MHV (20). It is, however, different from the sequences published on the 5′ ends of the A59 strain of MHV (17). Since the primer extension approach used in our study would detect the consensus leader sequence present in the majority of the mRNAs, rather than a single mRNA molecule, the sequences determined by us would more likely represent the true leader RNA of MHV. Indeed, the published 5′-end sequence of A59 MHV (17) has recently been shown to be the result of cloning artifacts (21). It is noteworthy that the 5′ end of the leader RNA, 5′ U-A-U-A-A-G-, agrees with that predicted from the oligonucleotide analysis (16).

The homology between the MHV leader RNA and the LTR of mammalian sarcoma viruses is very interesting. The R region contains the capping site and poly(A) addition signal (18) and also provides the jumping mechanism for reverse transcriptase. None of these functions appears to be conserved in MHV. There might be some common unknown functions or common evolutionary origin between the MHV leader and the LTR region. Further study on the mechanism of MHV RNA synthesis might shed light on this issue.

The origins of the cDNA products larger than the consensus leader cDNA are very curious. There are three possible interpretations for cDNA 5, which contains the entire leader

sequences plus 15 additional nucleotides: (*i*) the extra nucleotides could represent fold-back synthesis of *in vitro* reverse transcription of the leader region; (*ii*) the length of the leader region could be heterogeneous, with a small amount having a longer RNA region; and (*iii*) cDNA 5 could represent the real leader sequences, while the major leader DNA band represents a strong stop for reverse transcription. Currently, no data are available to distinguish these three possibilities. cDNA 6, which has a sequence unrelated to the leader cDNA, also suggests a very interesting possibility that there are short leader-related sequences somewhere in the MHV genomic RNA, other than the 5′-leader region. cDNA band 6 could represent the reverse transcription product resulting from the priming at such a location. Indeed, a stretch of sequences homologous to oligonucleotide 19, which is the junction site between the leader and body sequences of mRNAs, has been found in the genomic RNA at the initiation points of mRNA 7 (8, 16, 22). There could be similar sequences in other regions of the MHV genomic RNA. Whether cDNA 6 was transcribed from such regions will require more extensive sequence information.

The presence of a leader sequence common to all of the mRNAs and genomic RNA indicates that these mRNAs must be synthesized by joining two noncontiguous RNA segments, since the leader RNA sequences are present only once in the genomic RNA (8). Previous studies on the replicative intermediate RNA of MHV suggests that the leader RNA is probably synthesized independently as a free RNA species and then used as the primer for the synthesis of the body sequences of the mRNAs (9). This is a novel mechanism for mRNA synthesis. The presence of such a free leader RNA has recently been demonstrated in the cytoplasm of MHV-infected cells (unpublished observation).

The finding that there is probably a homologous sequence between the 3′ end of the leader RNA segment and the initiation site of the body sequences of mRNA 7 may provide a mechanism for the binding of the putative free leader RNA to the mRNA initiation sites on the negative-stranded RNA template. We do not know whether similar homology exists between the leader and the initiation sites of other MHV mRNAs. If such homology exists and the extent of homology differs for each individual mRNA, the binding efficiency of the free leader RNA to different mRNA initiation sites on the negative-stranded template would be different. Since it has been shown that different MHV mRNA species are present at different quantities in infected cells (23), this could provide a mechanism by which the rate of synthesis of the individual mRNAs is regulated.

We wish to thank Mahmood Kafaii and Gary Nelson for excellent technical assistance. We also thank Toni Ferry and Loretta Pedroza for assistance in manuscript preparation. This study was supported by U.S. Public Health Service Research Grant AI 19244 awarded by the National Institute of Allergy and Infectious Diseases and by National Science Foundation research Grant PCM-4507.

1. Lai, M. M. C. & Stohlman, S. A. (1978) *J. Virol.* **26,** 236–242.
2. Brayton, P. R., Ganges, R. G. & Stohlman, S. A. (1981) *J. Gen. Virol.* **56,** 457–460.
3. Wilhelmsen, K. C., Leibowitz, J. L., Bond, C. W. & Robb, J. A. (1981) *Virology* **110,** 225–230.
4. Sturman, L. S. & Holmes, K. V. (1983) *Adv. Virus Res.* **28,** 36–112.
5. Brayton, P. R., Lai, M. M. C., Patton, C. D. & Stohlman, S. A. (1982) *J. Virol.* **42,** 847–853.
6. Lai, M. M. C., Patton, C. D. & Stohlman, S. A. (1982) *J. Virol.* **44,** 487–492.
7. Lai, M. M. C., Brayton, P. R., Armen, R. C., Patton, C. C., Pugh, C. & Stohlman, S. A. (1981) *J. Virol.* **39,** 823–834.
8. Lai, M. M. C., Patton, C. D., Baric, R. S. & Stohlman, S. A. (1983) *J. Virol.* **46,** 1027–1033.
9. Baric, R. S., Stohlman, S. A. & Lai, M. M. C. (1983) *J. Virol.* **48,** 633–640.
10. Pedersen, F. S. & Haseltine, W. A. (1980) *Methods Enzymol.* **65,** 680–687.
11. Lamb, R. A. & Lai, C. J. (1980) *Cell* **21,** 475–485.
12. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–566.
13. McMaster, G. K. & Carmichael, G. G. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4835–4838.
14. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5202.
15. Friedman, T. (1979) *Am. J. Hum. Genet.* **31,** 19–28.
16. Lai, M. M. C., Patton, C. D. & Stohlman, S. A. (1982) *J. Virol.* **41,** 557–565.
17. Armstrong, J., Smeekens, S. & Rottier, P. (1983) *Nucleic Acids Res.* **11,** 883–891.
18. Hampe, A., Gobet, M., Even, J., Sherr, C. J. & Galibert, F. (1983) *J. Virol.* **45,** 466–472.
19. Bouloy, M., Plotch, S. J. & Krug, R. M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4886–4890.
20. Skinner, M. A. & Siddell, S. G. (1983) *Nucleic Acids Res.* **11,** 5045–5054.
21. Armstrong, J., Smeekens, S., Rottier, P., Spaan, W. & Van der Zeijst, B. A. M. (1984) in *Molecular Biology and Pathogenesis of Coronaviruses*, eds. Rottier, P. J. M., Van der Zeijst, B. A. M., Spaan, W. J. M. & Horzinek, M. C. (Plenum, New York), in press.
22. Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B. A. M. & Siddell, S. G. (1983) *EMBO J.* **2,** 1939–1944.
23. Leibowitz, J. L., Wilhelmsen, K. C. & Bond, C. W. (1981) *Virology* **114,** 39–51.