JOURNAL OF VIROLOGY, Sept. 1989, p. 3729–3736
0022-538X/89/093729-08$02.00/0

Vol. 63, No. 9

Identification of a New Transcriptional Initiation Site and the Corresponding Functional Gene 2b in the Murine Coronavirus RNA Genome

CHIEN-KOU SHIEH,[1] HAN-JUNG LEE,[2] KYOKO YOKOMORI,[1] NICOLA LA MONICA,[1] SHINJI MAKINO,[1] AND MICHAEL M. C. LAI[1]*

Department of Microbiology, School of Medicine, University of Southern California, 2011 Zonal Avenue, Los Angeles, California 90033-1054,[1] and Institute of Molecular Biology, Academia Sinica, Nankang, Taipei 11529, Taiwan[2]

Received 21 February 1989/Accepted 17 May 1989

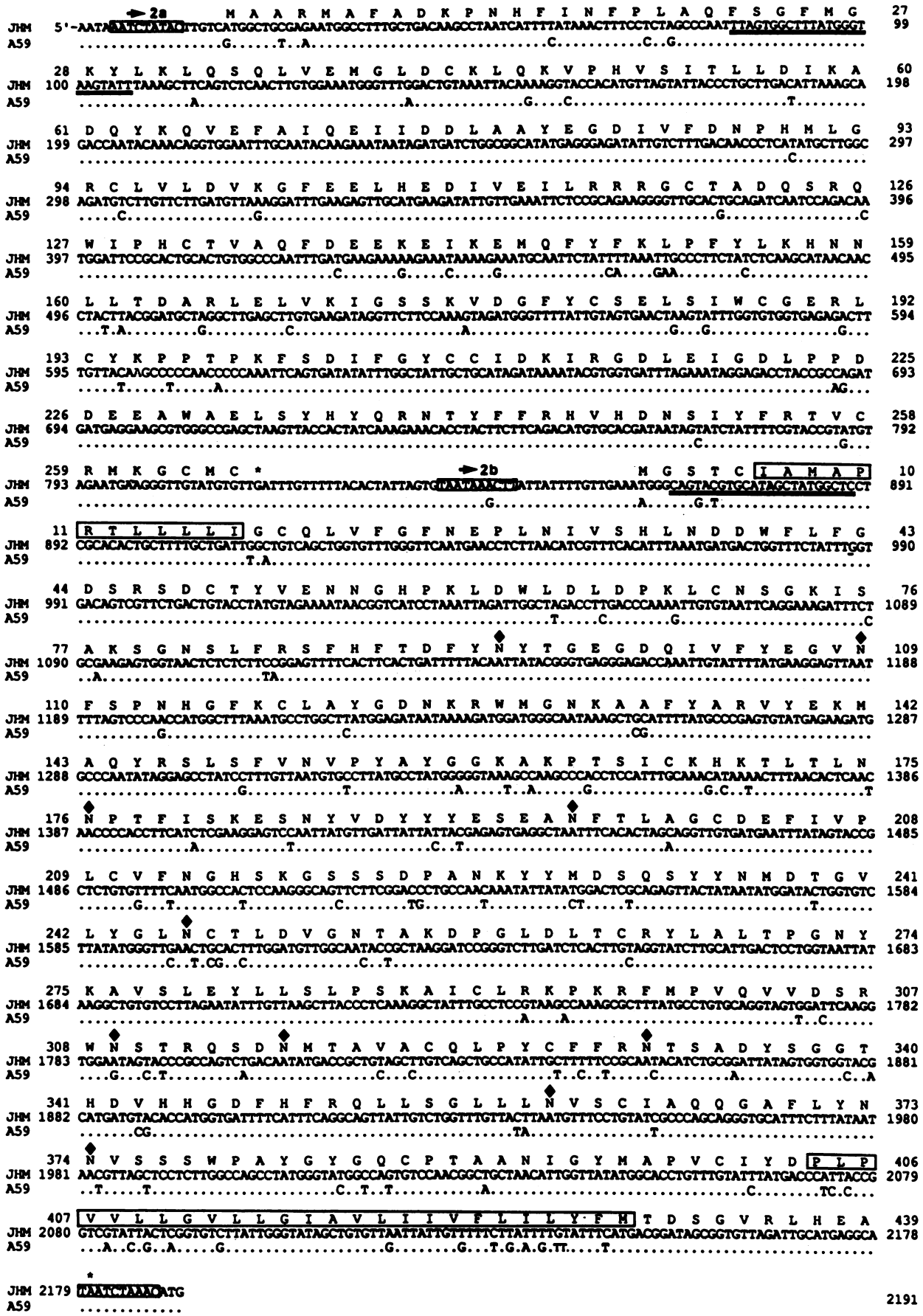
We have previously shown that some strains of the murine coronavirus mouse hepatitis virus (MHV) synthesize an additional mRNA species (mRNA 2b, previously called mRNA 2a) with a size intermediate between that of mRNAs 2 and 3, suggesting the presence of an optional transcriptional initiation site. This transcriptional start is dependent on the leader sequence of the virus strains. To study the mechanism of coronavirus transcriptional regulation, we have cloned and sequenced the region of the viral genome corresponding to the 5′ unique coding region of mRNA 2 of the JHM strain of MHV. In addition to the open reading frame (ORF) predicted to encode the viral nonstructural protein p30, a second complete ORF, with the potential to encode a 439-amino-acid polypeptide, was discovered. The transcriptional initiation sites of both mRNA 2a (formerly called mRNA 2) and mRNA 2b were determined by primer extension studies and RNA sequencing. The data indicated that transcription of mRNA 2a initiated at a site, UCUAUAC, that resembled the consensus intergenic sequence. In contrast, the start signal of the optional mRNA 2b, UAAUAAAC, deviated from the consensus sequence. mRNA 2b is a functional mRNA, as shown by in vitro translation studies of mRNA and ORF 2b and by the detection of an additional viral structural protein, gp65, in the JHM strain that synthesized this mRNA. Although the A59 strain of MHV was found to retain ORF 2b, it lacked the correct transcriptional and translational start signals for this gene. This study has therefore identified an optional gene product for murine coronaviruses and provided insights into the mechanism of regulation of MHV RNA transcription.

The murine coronavirus mouse hepatitis virus (MHV) is an enveloped virus containing a nonsegmented and positive-sensed RNA genome with a molecular weight of greater than 6×10^6 (18, 44). In infected cells, MHV synthesizes six major species of subgenomic mRNAs that are 3′ coterminal, each of which extends for various distances toward the 5′ end. This results in a nested-set structure of mRNAs (17). The 5′ unique portion of each mRNA is the functional part which is translated monocistronically (32, 36). These mRNAs code for three structural proteins, gp180/gp90 (E2), gp23 (E1), and pp50 (N), and an assortment of nonstructural proteins that have so far been poorly characterized.

All of the mRNAs of MHV, including genomic RNA, contain identical leader sequences of 72 to 77 nucleotides (16, 40). Although these leader nucleotides are not repeated in the internal region of genomic RNA, short stretches of consensus sequences similar to the 3′ end of the leader sequence have been found in all of the intergenic regions studied so far (6, 34, 35, 38). A considerable body of evidence suggests that MHV utilizes a unique mechanism of leader-primed transcription in which a leader RNA is transcribed from the 3′ end of the template RNA, is dissociated from the template, and then rejoins the template RNA at the downstream intergenic regions to serve as the primer for mRNA transcription (2, 3, 15, 26). This binding is probably mediated by the complementary sequences between the 3′ ends of the free-leader RNA and the template RNA at the intergenic regions, which include a consensus sequence UCUAAAC, common to the transcription initiation sites of most of MHV mRNAs (35). The free-leader RNA has been detected in MHV-infected cells (1, 3) and shown to be involved in mRNA transcription by the demonstration that the leader sequences of MHV mRNAs can be freely reassorted during mixed infection of two MHV strains (26). The 3′ end of the leader RNA contains a pentanucleotide sequence, UCUAA, that is repeated two to four times in different MHVs (24) and is also included in the intergenic consensus sequences (6, 25, 35). This pentanucleotide is probably part of the leader-binding sequence, and the number of repeats of this pentanucleotide in a given mRNA species is heterogeneous even in a homogeneous virus preparation (25). Thus, the MHV mRNA transcriptional start is the result of interaction between the 3′ end of the leader sequence and the intergenic sequences on the template RNA.

We have recently shown that the number of the UCUAA repeats at the 5′ end of the genomic RNA decreases upon serial in vitro passages of the virus (24). We have isolated JHM viruses with either two or three UCUAA repeats. Interestingly, the two-repeat virus, JHM(2), synthesizes an additional mRNA species, mRNA 2b (previously named mRNA 2a), with a size intermediate between that of mRNAs 2 and 3 (24). This finding suggests that this virus can efficiently recognize a cryptic transcriptional start site within the 5′ coding region of mRNA 2. In contrast, the A59 strain of MHV, another prototype MHV, does not synthesize this additional mRNA even though it contains only two UCUAA repeats (24). The published sequence of the 5′ unique portion of mRNA 2 of A59 shows the presence of a complete open reading frame (ORF) and an additional downstream incom-

* Corresponding author.


2a
M A A R M A F A D K P N H F I N F P L A Q F S G F M G
JHM 5'-AATA
A59
TTAGTGGCTTTATGGGT
TTGTCATGGCTGCGAGAATGGCCTTTGCTGACAAGCCTAATCATTTTATAAACTTTCCTCTAGCCCAAT
2b
M G S T C I A M A P
R T L L L L I G C Q L V F G F N E P L N I V S H L N D D W F L F G
D S R S D C T Y V E N N G H P K L D W L D L D P K L C N S G K I S
A K S G N S L F R S F H F T D F Y N Y T G E G D Q I V F Y E G V N
F S P N H G F K C L A Y G D N K R W M G N K A A F Y A R V Y E K M
A Q Y R S L S F V N V P Y A Y G G K A K P T S I C K H K T L T L N
N P T F I S K E S N Y V D Y Y Y E S E A N F T L A G C D E F I V P
L C V F N G H S K G S S S D P A N K Y Y M D S Q S Y Y N M D T G V
L Y G L N C T L D V G N T A K D P G L D L T C R Y L A L T P G N Y
K A V S L E Y L L S L P S K A I C L R K P K R F M P V Q V V D S R
W N S T R Q S D N M T A V A C Q L P Y C F F R N T S A D Y S G G T
H D V H H G D F H F R Q L L S G L L L N V S C I A Q Q G A F L Y N
N V S S S W P A Y G Y G Q C P T A A N I G Y M A P V C I Y D P L P
V V L L G V L L G I A V L I I V F L I L Y F M T D S G V R L H E A
TAATCTAAAC
ATG

plete ORF without the initiator AUG codon (21). This downstream ORF also lacks the consensus transcriptional start sequence present at the immediate upstream sites of all of the other mRNAs of MHV. These findings prompted us to examine the mRNA 2 sequence of the JHM strain. The results presented in this report reveal an additional functional gene that codes for a structural glycoprotein, gp65. The transcriptional start sequence of this new optional gene is also distinct from the known consensus sequence for other genes of MHV. These results provide further insights into the transcriptional regulation of MHV RNA.

MATERIALS AND METHODS

Virus and cells. The plaque-cloned JHM strain of MHV (23, 27) was used for cDNA cloning and sequencing. JHM(3) and JHM(2), which contain three and two repeats of UCUAA, respectively, at the 3′ end of the leader sequence, were isolated after 11 consecutive passages of the original plaque-cloned JHM strain (24). Viruses were propagated in the DBT cell line and purified on successive discontinuous and continuous sucrose gradients as described previously (27, 28).

Molecular cloning and DNA sequencing. A specific oligodeoxyribonucleotide (oligo 31; 5′-CAAGAGGGTAATAG TAAAAT-3′) complementary to the RNase T_1-resistant oligonucleotide no. 7 (29) was synthesized. This oligonucleotide had previously been mapped to the 5′ end of gene C (29). The procedures for RNA sequence analysis were identical to those described previously (35). cDNA cloning was performed by the general method of Gubler and Hoffman (9), using oligo 31 as primer for first-strand cDNA synthesis on purified genomic RNA. The double-stranded cDNAs were ligated to *Eco*RI linkers and inserted into the *Eco*RI site of plasmid pTZ18U (United States Biochemical Corp.). The recombinant DNAs were transformed into *Escherichia coli* MV1190 competent cells (8) and screened by colony hybridization (35), using the 5′-end-labeled oligo 31 as probe. Four overlapping cDNA clones were used to sequence the entire 2.2-kilobase (kb) of gene B by both chemical modification (30, 39) and dideoxynucleotide-chain termination (33) methods directly on plasmid DNA (7).

Isolation of intracellular mRNAs. The intracellular RNA of JHM(2) was extracted from infected DBT cells at 9 h postinfection (p.i.) as previously described (29). Virus-specific RNA was enriched by passing intracellular RNA through an oligo(dT)-cellulose (Collaborative Research, Inc.) column. The poly(A)-containing RNAs were mixed with ^{32}P-labeled mRNA of the same virus and separated by preparative electrophoresis on urea-agarose gels (27). The individual mRNA species were eluted and extracted by published procedures (20). The mRNAs were further purified on oligo(dT) columns to remove contaminating degraded genomic RNA.

Primer extension. Two specific oligodeoxyribonucleotides, oligo 61 (5′-AATACTTACCCATAAAGCCACTA-3′) and oligo 62 (5′-GAGCCATAGCTATGCACGTACTGC-3′), complementary to nucleotides 83 to 106 and 868 to 889, respectively, from the 5′ end of mRNA 2 (Fig. 1), were used. The oligonucleotides were 5′ end labeled with [γ-^{32}P]ATP (7,000 Ci/mmol; ICN Biochemicals) by polynucleotide kinase (Pharmacia, Inc.) (31). Primer extension reactions were performed as described previously (25) on purified mRNAs 2a and 2b separately. Reaction products were analyzed on 6% polyacrylamide gels containing 7.7 M urea.

RNA sequencing. Oligos 61 and 62 were used as primers to sequence both genomic RNA and mRNAs of JHM(2). The dideoxynucleotide-chain termination method adapted for RNA sequencing was used (46).

^{35}S and [^{3}H]glucosamine labeling of virion. The virus-infected DBT cells were switched to methionine-free medium at 7.5 h p.i. After 30 min of starvation for methionine, ^{35}S Translabel (100 μCi/ml; 1,193 Ci/mmol; ICN Biochemicals) was added, and the preparation was incubated for another 2 h. The medium was then replaced with normal minimum essential medium containing 10% tryptose phosphate broth and 2% fetal calf serum. For [^{3}H]glucosamine labeling, the virus-infected cells were switched at 1 h p.i. to minimal essential medium containing only 2% dialyzed fetal calf serum. D-[6-^{3}H(*N*)]glucosamine hydrochloride (100 μCi/ml; 27.0 Ci/mmol; Dupont, NEN Research Products) was added at 6 h p.i. Labeled virus was harvested at 14 h p.i. and purified as described previously (28, 29).

Construction of recombinant pT7(2b) clones. Plasmid p(2a2b) DNA, which contains 2.4 kb of MHV genome sequence covering the entire gene 2 region (N. La Monica, unpublished data), was cleaved with restriction enzymes *Pst*I and *Xba*I. The resulting 1.9-kb fragment was inserted into the transcription vector pT7-4, which contains the bacteriophage T7 promoter (42). This construct, designated pT7(2b), contains MHV sequences extending from the *Pst*I site located 476 bases upstream of the initiation codon of ORF 2b to the 3′ end of gene 2b.

In vitro transcription and translation. Recombinant plasmid pT7(2b) was linearized by digestion with restriction enzyme *Eco*RI and transcribed in vitro with T7 RNA polymerase as previously described (39). The resulting RNA was translated in mRNA-dependent rabbit reticulocyte lysate (Promega Biotec) in the presence of [^{35}S]methionine. Reactions were carried out in a final volume of 25 μl under conditions recommended by the manufacturer. Where specified, 2 μl of canine pancreatic microsomal membrane (Promega Biotec) was added to the reaction mixtures.

SDS-PAGE. For sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE), a linear-gradient SDS-PAGE gel (7.5 to 15% polyacrylamide) was used. The ^{35}S- and ^{3}H-labeled virions were suspended in 10 mM Tris hydrochloride (pH 7.5)–1 mM EDTA–1% SDS. Virus samples were then mixed with an equal volume of 2× sample buffer (0.12 M Tris hydrochloride [pH 6.8], 40% glycerol, 4% SDS, 3% dithiothreitol, 0.1% bromophenol blue) and incubated at 37°C for 10 min before being loaded onto the gel. Electrophoresis was performed at 55 V for 12 h. After electrophoresis, the gel was fixed with 40% methanol–7% glacial acetic acid for 30 min and treated for fluorography,

FIG. 1. DNA sequence of 2,191 bases in the gene B region of MHV strain JHM genomic cDNA clones. Translation of the two long ORFs is shown in single-letter amino acid code. The diverged nucleotide sequence of MHV strain A59 (21) is shown. The transcription start sequences are boxed. Positions of oligomers used in RNA sequencing and primer extension studies are double underlined. Potential sites for N-linked glycosylation are indicated (◆). The boxed amino acid sequences indicate the hydrophobic domains at the N terminus (amino acids 6 to 17) and C terminus (amino acids 404 to 429) of ORF 2b.

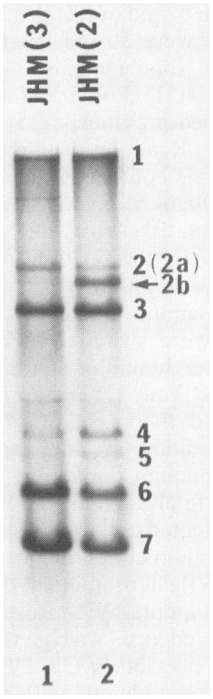


FIG. 2. Electrophoretic analysis of the virus-specific intracellular RNAs of MHV strains JHM(2) and JHM(3). ^{32}P-labeled intracellular viral RNAs were denatured with glyoxal-dimethyl sulfoxide and separated by electrophoresis on a 1% agarose gel. The additional mRNA in JHM(2) is marked (←). As indicated, mRNA 2 is renamed mRNA 2a.

using En^{3}Hance (Dupont, NEN Research Products). The gel was then dried and exposed to X-ray film.

RESULTS

Nucleotide sequence and predicted amino acid sequence of gene B of MHV strain JHM. Figure 1 shows the 2.2-kb sequence obtained from four overlapping cDNA clones covering gene B. This region likely represents the total gene B sequence, since the 2.2 kb roughly corresponds to the estimated molecular weight difference between mRNAs 2 and 3, which is approximately 7×10^5 (29). Furthermore, a consensus transcription start site was found upstream of the 2.2-kb sequence (see below), suggesting that this is the start of a gene. The sequence immediately downstream of this 2.2-kb sequence was identical to the published 5′-end sequence of gene C of JHM (34). The sequence upstream of it was an ORF that was continuous for at least 3 kb. This latter sequence is most likely part of gene A. Similar to findings for gene B of the A59 strain of MHV (21), two ORFs were found in this 2.2-kb region. The first ORF (ORF 2a) was 265 amino acids long and had a predicted molecular weight of 30,912, in agreement with results for the postulated gene product, p30, of mRNA 2 (36). In comparison with A59, the 5′-proximal AUG codon of ORF 2a of JHM was located 9 nucleotides further upstream of that of A59 (21). Since both of the first two AUG codons in JHM ORF 2a are in preferred context for translational initiation (14), it is not clear whether JHM will utilize the 5′-proximal AUG codon or bypass it (leaky scanning) (13) and initiate protein synthesis at the second AUG codon, which is the first initiator codon of the corresponding ORF in A59 (21). The second ORF (ORF 2b), which started 48 nucleotides downstream of the termination codon of ORF 2a, was 439 amino acids long and had a predicted molecular weight of 48,951. Within this predicted polypeptide were 10 potential N-glycosylation sites of the type Asn-X-Thr/Ser (X not being Pro) (Fig. 1). Thus, this protein is potentially a glycoprotein. There were two predicted hydrophobic regions at the N and C termini, suggesting that it may be a membrane protein. In comparison the second ORF of A59 appeared to be nonfunctional, since it does not have an initiator AUG codon (Fig. 1) (23). ORFs 2a and 2b of the JHM and A59 strains of MHV had 95 and 92% sequence homology, respectively.

Transcriptional initiation sites for mRNAs 2a and 2b. A sequence of UCUAUAC was located 5 nucleotides upstream of the initiation codon of ORF 2a (Fig. 1). This sequence was very similar to the consensus intergenic sequence, UCUAAAC, found at the transcriptional start sites of all of the MHV mRNAs (6, 35, 38). Surprisingly, no such sequence was found in the region between ORF 2a and ORF 2b. The only loosely related sequence, AAUAAACUU, was located 14 nucleotides upstream of the first AUG of ORF 2b. To determine the transcriptional initiation sites of mRNAs transcribed from these regions, we performed primer extension studies on the purified mRNAs of JHM(2), which synthesized an additional mRNA species, mRNA 2b, with a size intermediate between those of mRNAs 2 and 3 (Fig. 2) (24). We renamed the first two transcripts mRNAs 2a and mRNA 2b, respectively. These two mRNAs were eluted from the gel, further purified by oligo(dT) column chromatography, and used for primer extension. Two synthetic

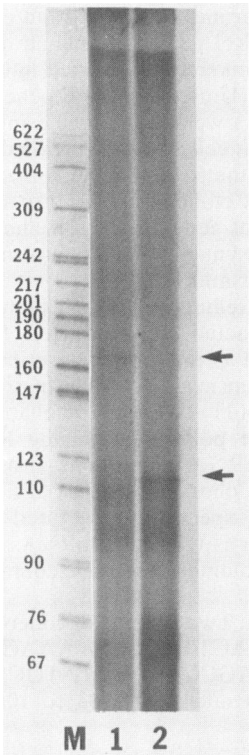


FIG. 3. Primer extension analysis of mRNAs 2a and 2b. Lane: 1, primer extension on purified mRNA 2a, using 5′-end-labeled oligo 61; 2, primer extension on mRNA 2b, using 5′-end-labeled oligo 62. The primer extension products obtained by reverse transcriptase are indicated (←). Marker DNA is pBR322 DNA digested with *Msp*I; molecular sizes (in bases) are shown on the left.

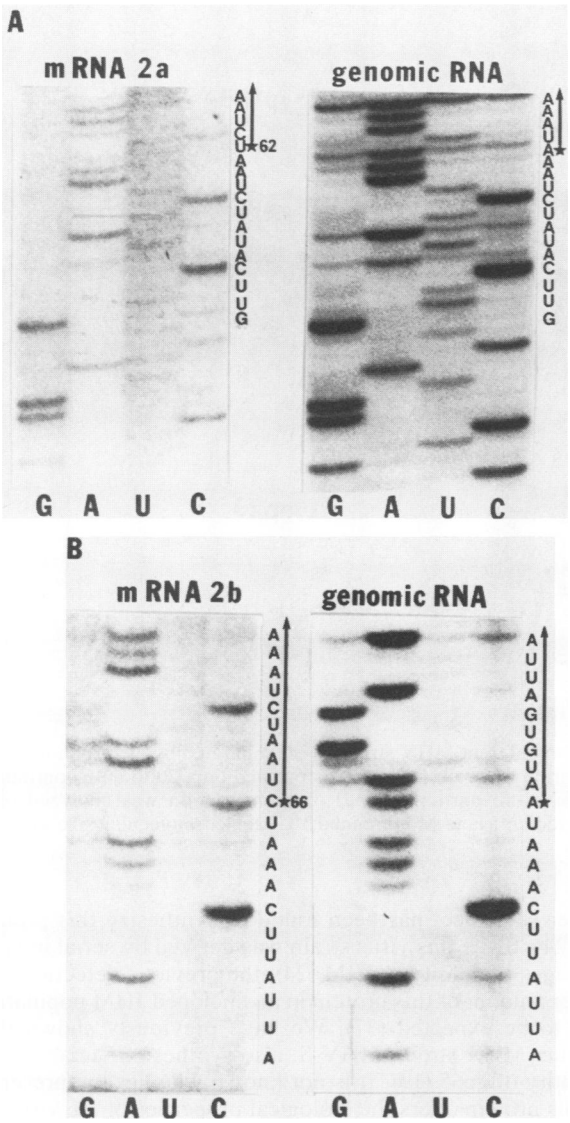


FIG. 4. Sequence comparison between the genomic RNA and mRNA at the transcription start regions. (A) RNA sequence obtained by using oligo 61 as primer for dideoxy-chain termination analysis on both genomic RNA and mRNA 2a of MHV strain JHM(2). (B) RNA sequence obtained by using oligo 62 as primer on both genomic RNA and mRNA 2b. The diverged sites between the genomic RNA and mRNAs are indicated (↓). Numbers on the right of the mRNA panels designate positions from the 5′ end of leader RNA.

oligodeoxyribonucleotides were used. Oligo 61, 5′-AATAC TTACCCATAAAGCCACTA-3′, is located 70 nucleotides downstream of the putative initiation site of ORF 2a. Oligo 62, 5′-GAGCCATAGCTATGCACGTACTGC-3′, is located 18 nucleotides downstream of the putative initiation site for ORF 2b. The oligo 61-primed extension product on mRNA 2a was about 165 nucleotides long, whereas the oligo 62-primed extension product on mRNA 2b was about 114 nucleotides long (Fig. 3). Since the leader sequence of JHM(2) is 72 nucleotides long (24), this result suggests that the two predicted sequences, UCUAUAC and AAUAAA CUU, are likely the transcriptional start sites for mRNAs 2a and 2b, respectively. To precisely define the leader-binding sequences, we also performed RNA sequencing of mRNAs 2a and 2b (Fig. 4). The sequences of both mRNAs 2a and 2b diverged from that of the genomic RNA immediately upstream of these two stretches of sequence, indicating that these two regions are indeed the initiation sites for transcription of mRNAs 2a and 2b, respectively. Both mRNAs contained a leader sequence with two UCUAA repeats, in agreement with the model of leader-primed transcription (15).

Identification of gene products of mRNA 2b. The presence of an additional ORF in the gene B region and the synthesis of a corresponding mRNA in some strains of MHV suggests that ORF 2b is a functional but optional gene. To detect the possible gene product of ORF 2b, we constructed a pT7 plasmid [pT7(2b)] containing the cDNA insert covering ORF 2b and used an in vitro transcription-translation system to identify the gene product. The primary translation product of the ORF 2b RNA was a protein of 45 kilodaltons, roughly equivalent to the predicted coding capacity of ORF 2b (Fig. 5A). The primary translation product became a protein of 65 kilodaltons when translation was performed in the presence of canine pancreatic microsomal membrane preparations (Fig. 5A, lane 4). This result suggests that the protein is glycosylated. Both proteins were precipitated by polyclonal antibodies prepared against JHM(2). Similar proteins were

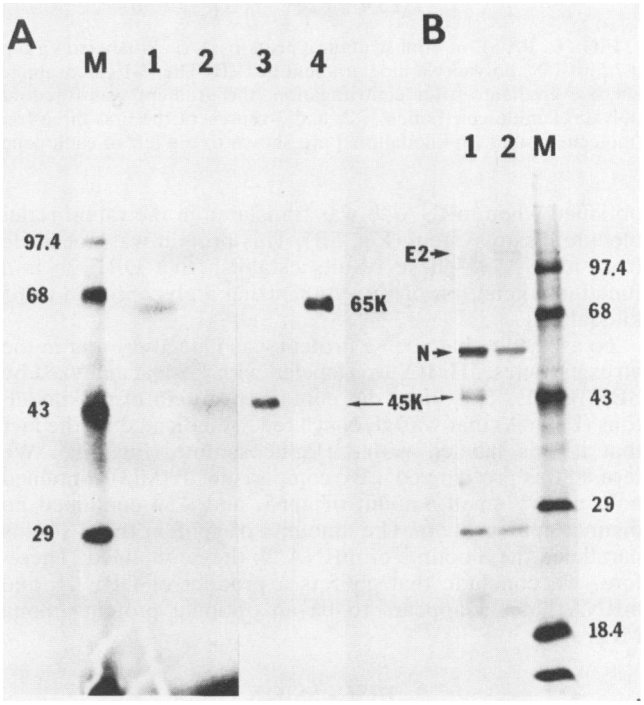


FIG. 5. SDS-PAGE analysis of in vitro translation products. (A) In vitro translation products of in vitro-transcribed ORF 2b RNA derived from JHM(2). Translation was performed in a rabbit reticulocyte lysate system, using [^{35}S]methionine. Lanes: M, ^{14}C-labeled protein size marker (Bethesda Research Laboratories; molecular sizes [in kilodaltons] are indicated in the margins); 1, immunoprecipitation of in vitro translation products synthesized in the presence of canine pancreatic microsomal membrane; 2, immunoprecipitation of in vitro translation products synthesized in the absence of membrane; 3, same as lane 2 but without immunoprecipitation; 4, same as lane 1 but without immunoprecipitation. (B) In vitro translation product of purified mRNAs 2b (lane 1) and 3 (lane 2) from JHM(2)-infected DBT cells. Translation products were immunoprecipitated with polyclonal antibody against JHM(2) virions.

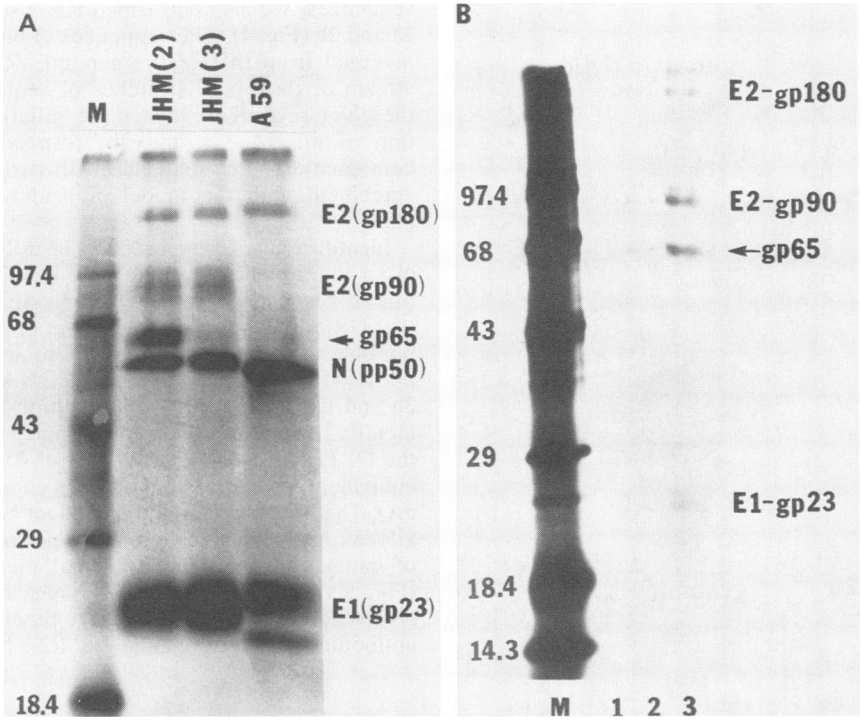


FIG. 6. PAGE of viral structural proteins. (A) ^{35}S-labeled viral proteins from different MHV strains were separated by electrophoresis on a 7.5 to 15% polyacrylamide gradient gel. (B) The [^{3}H]glucosamine-labeled virion of strain JHM(2) was purified on a 20 to 60% continuous sucrose gradient. After centrifugation, the gradient was fractionated into six 2-ml portions; 40 μl of each fraction was analyzed on a polyacrylamide gel. Lanes 1, 2, and 3 represent the first three fractions of gradient; lane M contained ^{14}C-labeled molecular size markers (molecular sizes [in kilodaltons] are shown to the left of each panel).

obtained when mRNA 2b was translated in the rabbit reticulocyte lysate system (Fig. 5B). This protein was not made from mRNA 3. These results establish that ORF 2b is a functional gene, probably synthesizing a glycoprotein of 65 kilodaltons.

To ascertain that such a protein was indeed present in the virus particles, JHM(2) was labeled with ^{35}S and analyzed by SDS-PAGE. This virus did contain a protein of 65 kilodaltons (Fig. 6A) that was glycosylated, as indicated by the fact that it was labeled with [^{3}H]glucosamine (Fig. 6B). We termed this protein gp65. By comparison, JHM(3) contained only a very small amount of gp65, and A59 contained no distinct gp65 protein. The amounts of gp65 in these strains paralleled the amounts of mRNA 2b they contained. Therefore, we conclude that gp65 is a product of ORF 2b and mRNA 2b and appears to be an optional protein among MHVs.

DISCUSSION

Significance of the optional virion protein gp65. The virus-specific proteins of the JHM strain of MHV have previously been studied (28, 37, 45). In addition to the proteins detected in strain A59 (41), a minor glycoprotein, gp65, was found in some strains of JHM (28, 37, 43). However, the real identity of this protein has been in doubt, since the MHV genome is not known to have any genetic region or mRNA species capable of coding for this additional protein (17, 36). The studies presented in this paper detected a new MHV gene that codes for this protein. The unique aspect of this protein is that it appears to be optionally expressed only in certain strains of MHVs. So far, only the JHM strain, with a unique leader sequence, has been found to synthesize this protein (22, 24). Since this virus strain was derived by serial in vitro passages of wild-type JHM (24), the previous detection of a minor amount of this protein in an uncloned JHM population was to be expected (37). We have previously shown that another MHV strain, MHV-S, also synthesizes an abundant quantity of gp65 (19). It is not known whether the presence of this protein alters the biological properties of the virus. A similar protein has been detected in bovine coronavirus (11), which, significantly, also synthesizes mRNA 2b (10). This protein is an envelope spike protein of bovine coronavirus and is thought to be responsible for the hemagglutination properties of the virus (12). Similar biological properties could not be detected for the MHV gp65 (K. Yokomori, unpublished observation). Thus, the possible function of this protein in MHV is not clear. It is interesting that the corresponding gene of strain A59 does not contain an initiator AUG codon (21), although a large part of this ORF is conserved between JHM and A59. This finding suggests that this gene is not required for viral infectivity. In another coronavirus, avian infectious bronchitis virus, no corresponding gene is predicted (5). However, hemagglutinin activity was detected in some strains of this virus (4). It has previously been shown that the predicted protein product of the ORF 2b in A59 shares some homology with the hemagglutinin of influenza C virus (21). Further studies of the viral protein should yield interesting insights into the evolution of coronaviruses.

Transcriptional control of gene 2b. Transcriptional control of gene 2b is particularly interesting. Several unique points associated with this gene are noteworthy.

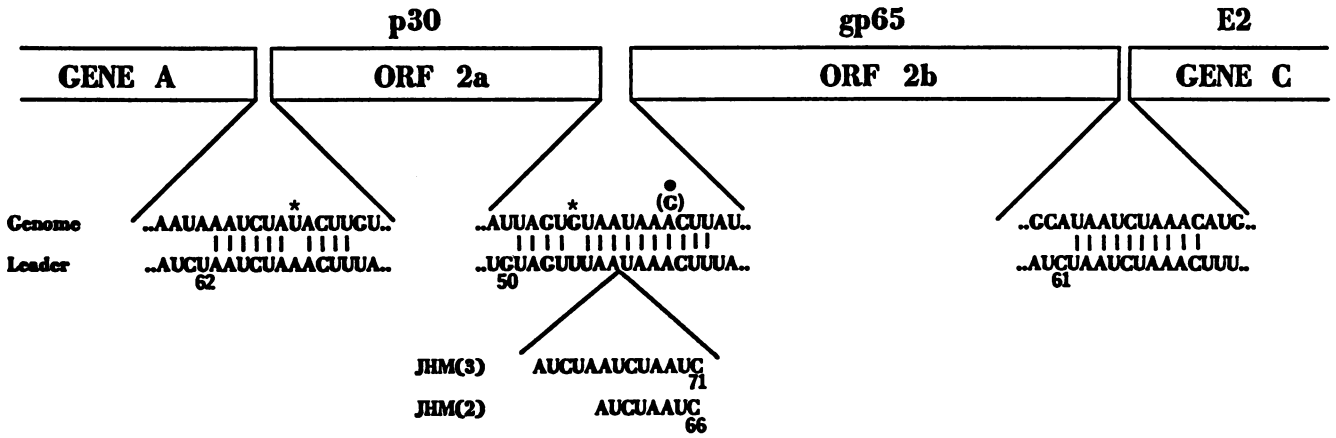


FIG. 7. Comparison of leader sequences with the transcription start sites for ORF 2a, ORF 2b, and gene C. Symbols: | , homologous nucleotides between the transcription start sites and the 3′ end of leader RNA; *, mismatches within this region. Numbers below the leader RNA designate positions from the 5′ end of mRNAs. JHM(3) and JHM(2) display the loop sequences that may be formed between the leader RNA to the transcription start site for ORF 2b. (G), Diverged A59 sequence at the initiation site of ORF 2b.

(i) The transcriptional initiation sites of all of the subgenomic mRNAs of MHV have been shown to contain a stretch of consensus sequence, UCUAAAC (6, 35, 38). This sequence is identical for all of the mRNAs except mRNA 6, which has one nucleotide substitution, UCCAAAC (6, 38). The transcription start site of mRNA 2a, UCUAUAC, determined in this study also has a single base substitution but in a different position. This raises the possibility that other loosely homologous sequences are utilized for the synthesis of minor mRNA species, as has been noted with some MHV strains (29). The transcription initiation site of mRNA 2b, UAAUAAAC, shows the most divergence from this consensus sequence. Whether this divergence is responsible for the optional nature of the transcriptional initiation from this site is not clear. It will be interesting to determine whether other similar sequences in gene 5, which also shows two ORFs (38), can initiate the synthesis of an additional mRNA for translation of the second ORF.

(ii) The optional transcription initiation of mRNA 2b provides another piece of evidence that the transcriptional initiation of coronavirus mRNAs is determined by the leader sequence. As we have shown previously (25), the 3′ end of the leader RNA of MHV contains a variable number of repeats of a UCUAA sequence, which is localized in the region complementary to the consensus transcription initiation sites of every mRNA and is thus considered to be essential for the transcription initiation of MHV mRNAs (35). Most interesting is that the JHM strain containing two UCUAA repeats utilizes the initiation site for mRNA 2b very efficiently, whereas that with three UCUAA repeats does not (24). This finding further supports the importance of this repeat sequence in transcriptional initiation of MHV mRNAs. Surprisingly, comparison of the 3′ end of leader sequence with the initiation points of mRNA 2b does not show complete complementarity within this repeat. Instead, the complementarity is in the sequences immediately upstream and downstream of these UCUAA repeats (Fig. 7). This sequence relationship departs from that at the transcriptional initiation sites of other subgenomic mRNAs, which may explain why other mRNAs are transcribed in JHM(2) and JHM(3), whereas mRNA 2b is efficiently transcribed only from JHM(2) (24). However, it is not clear why JHM(3) cannot initiate transcription of mRNA 2b efficiently. Perhaps the insertion loop created by the binding of leader RNA to this initiation site is too big for JHM(3), and hence interference of mRNA transcription occurs (Fig. 7).

(iii) In contrast to JHM(2), strain A59, which also has two UCUAA repeats, does not transcribe mRNA 2b. It should be noted that the A59 leader can initiate JHM mRNA 2b, as demonstrated by the synthesis of this mRNA by the recombinant MHV, RL-2, which has the A59 leader but the JHM 2b gene (22). Thus, the failure of A59 to synthesize mRNA 2b is most likely due to the single base substitution within the transcription initiation sequence (Fig. 1 and 7). The fact that ORF 2b of A59 has also lost its initiator AUG (21) may have interesting implications with respect to the evolution of MHVs. The preservation of this initiation site and the corresponding ORF may be under the same evolutionary pressure. This possibility can be tested by examining additional coronaviruses.

ACKNOWLEDGMENTS

We thank Lisa Soe, James Keck, and Lisa Welter for suggestions and advice during the course of this study.

This study was supported by Public Health Service research grants AI19244 and NS18146 from the National Institutes of Health. N.L.M. is a postdoctoral fellow of the National Multiple Sclerosis Society.

LITERATURE CITED

1. **Baric, R. S., C.-K. Shieh, S. A. Stohlman, and M. M. C. Lai.** 1987. Analysis of intracellular small RNAs of mouse hepatitis virus: evidence for discontinuous transcription. Virology **156:** 342–354.
2. **Baric, R. S., S. A. Stohlman, and M. M. C. Lai.** 1983. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chain. J. Virol. **48:**633–640.
3. **Baric, R. S., S. A. Stohlman, M. K. Razavi, and M. M. C. Lai.** 1985. Characterization of leader related-small RNAs in coronavirus-infected cells: further evidence for leader-primed mechanism of transcription. Virus Res. **3:**19–33.
4. **Bingham, R. W., M. H. Madge, and D. A. J. Tyrrell.** 1975. Haemagglutination by avian infectious bronchitis virus—a coronavirus. J. Gen. Virol. **28:**381–390.
5. **Boursnell, M. E. G., T. D. K. Brown, I. J. Foulds, P. F. Green, F. M. Tomley, and M. M. Binns.** 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J. Gen. Virol. **68:**57–77.
6. **Budzilowicz, C. J., S. Wilczynski, and S. R. Weiss.** 1985. Three

intergenic regions of nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. J. Virol. **53**:834–840.
7. **Chen, E. J., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
8. **Dagert, M., and S. D. Ehrlich.** 1979. Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. Gene **6**:23–29.
9. **Gubler, U., and B. J. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25**:263–269.
10. **Keck, J. G., B. G. Hogue, D. A. Brian, and M. M. C. Lai.** 1988. Temporal regulation of bovine coronavirus RNA synthesis. Virus Res. **9**:343–356.
11. **King, B., and D. A. Brian.** 1982. Bovine coronavirus structural proteins. J. Virol. **42**:700–707.
12. **King, B., B. J. Potts, and D. A. Brian.** 1985. Bovine coronavirus haemagglutinin protein. Virus Res. **2**:53–59.
13. **Kozak, M.** 1986. Regulation of protein synthesis in virus-infected animal cells. Adv. Virus Res. **31**:229–292.
14. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15**:8125–8148.
15. **Lai, M. M. C.** 1986. Coronavirus leader-RNA-primed transcription: an alternative mechanism to RNA splicing. BioEssays **5**:257–260.
16. **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.
17. **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 hepatotropic strain MHV-3. J. Virol. **39**:823–834.
18. **Lai, M. M. C., and S. A. Stohlman.** 1978. RNA of mouse hepatitis virus. J. Virol. **26**:236–242.
19. **Lai, M. M. C., and S. A. Stohlman.** 1981. Comparative analysis of RNA genome of mouse hepatitis virus. J. Virol. **38**:661–670.
20. **Langridge, L., P. Langridge, and P. L. Bergquist.** 1980. Extraction of nucleic acids from agarose gels. Anal. Biochem. **103**: 264–271.
21. **Luytjes, W., P. J. Bredenbeek, A. F. H. Noten, M. C. Horzinek, and W. J. M. Spaan.** 1988. Sequence of mouse hepatitis virus A59 mRNAs: indications for RNA recombination between coronaviruses and influenza C virus. Virology **166**:415–422.
22. **Makino, S., J. O. Fleming, J. G. Keck, S. A. Stohlman, and M. M. C. Lai.** 1987. RNA recombination of coronaviruses: localization of neutralizing epitopes and neuropathogenic determinants on the carboxyl terminus of peplomers. Proc. Natl. Acad. Sci. USA **84**:6567–6571.
23. **Makino, S., N. Fujioka, and K. Fujiwara.** 1985. Structure of the intracellular defective viral RNAs of defective interfering particles of mouse hepatitis virus. J. Virol. **54**:329–336.
24. **Makino, S., and M. M. C. Lai.** 1989. Evolution of the 5′-end of genomic RNA of murine coronaviruses during passages in vitro. Virology **169**:227–232.
25. **Makino, S., L. H. Soe, C.-K. Shieh, and M. M. C. Lai.** 1988. Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNA. J. Virol. **62**:3870–3873.
26. **Makino, S., S. A. Stohlman, and M. M. C. Lai.** 1986. Leader sequence of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. Proc. Natl. Acad. Sci. USA **83**:4204–4208.
27. **Makino, S., F. Taguchi, and K. Fujiwara.** 1984. Defective interfering particles of mouse hepatitis virus. Virology **133**:9–17.
28. **Makino, S., F. Taguchi, M. Hayami, and K. Fujiwara.** 1983. Characterization of small plaque mutants of mouse hepatitis virus, JHM strain. Microbiol. Immunol. **27**:445–454.
29. **Makino, S., F. Taguchi, N. Hirano, and K. Fujiwara.** 1984. Analysis of genomic and intracellular viral RNAs of small plaque mutants of mouse hepatitis virus, JHM strain. Virology **139**:138–151.
30. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
31. **Pedersen, F. S., and W. A. Haseltine.** 1980. A micromethod for detailed characterization of high molecular weight RNA. Methods Enzymol. **65**:680–687.
32. **Rottier, P. J. M., W. J. M. Spaan, M. Horzinek, and B. A. M. van der Zeijst.** 1981. Translation of three mouse hepatitis virus (MHV-A59) subgenomic RNAs in *Xenopus laevis* oocytes. J. Virol. **38**:20–26.
33. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
34. **Schmidt, I., M. Skinner, and S. G. Siddell.** 1987. Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. J. Gen. Virol. **68**:47–56.
35. **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.
36. **Siddell, S. G.** 1983. Coronavirus JHM: coding assignments of subgenomic mRNAs. J. Gen. Virol. **64**:113–125.
37. **Siddell, S. G., H. Wege, A. Barthel, and V. ter Meulen.** 1981. Coronavirus JHM: intracellular protein synthesis. J. Gen. Virol. **53**:145–155.
38. **Skinner, M. A., D. Ebner, and S. G. Siddell.** 1985. Coronavirus MHV-JHM mRNA 5′ has a sequence arrangement which potentially allows translation of a second, downstream open reading frame. J. Gen. Virol. **66**:581–592.
39. **Soe, L. H., C.-K. Shieh, S. C. Baker, M.-F. Chang, and M. M. C. Lai.** 1987. Sequence and translation of the murine coronavirus 5′-end genomic RNA reveals the N-terminal structure of the putative RNA polymerase. J. Virol. **61**:3968–3976.
40. **Spaan, W. J. M., 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.
41. **Sturman, L. S.** 1977. Characterization of a coronavirus. I. Structural proteins: effects of preparative conditions on the migration of protein in polyacrylamide gels. Virology **77**:637–649.
42. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**:1074–1078.
43. **Taguchi, F., P. T. Massa, and V. ter Meulen.** 1986. Characterization of a variant virus isolated from neural cell culture after infection of mouse coronavirus JHMV. Virology **155**:267–270.
44. **Wege, H., A. Müller, and V. ter Meulen.** 1978. Genomic RNA of the murine coronavirus JHM. J. Gen. Virol. **41**:217–227.
45. **Wege, H., H. Wege, K. Nagashima, and V. ter Meulen.** 1979. Structural polypeptides of the murine coronavirus JHM. J. Gen. Virol. **42**:37–47.
46. **Zimmern, D., and P. Kaesburg.** 1978. 3′-Terminal nucleotide sequence of encephalomyocarditis virus RNA determined by reverse transcriptase and chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **75**:4257–4261.