

Identification of a Domain Required for Autoproteolytic Cleavage of Murine Coronavirus Gene A Polyprotein

SUSAN C. BAKER, CHIEN-KOU SHIEH, LISA H. SOE, MING-FU CHANG, DAVID M. VANNIER, AND
MICHAEL M. C. LAI*

Department of Microbiology, School of Medicine, University of Southern California, Los Angeles, California 90033

Received 21 February 1989/Accepted 26 May 1989

The 5'-most gene of the murine coronavirus genome, gene A, is presumed to encode viral RNA-dependent RNA polymerase. It has previously been shown that the N-terminal portion of this gene product is cleaved into a protein of 28 kilodaltons (p28). To further understand the mechanism of synthesis of the p28 protein, cDNA clones representing the 5'-most 5.3 kilobases of murine coronavirus mouse hepatitis virus strain JHM were sequenced and subcloned into pT7 vectors from which RNAs were transcribed and translated *in vitro*. The sequence was found to encode a single long open reading frame continuing from near the 5' terminus of the genome. Although p28 is encoded from the first 1 kilobase at the 5' end of the genome, translation of *in vitro*-transcribed RNAs indicated that this protein was not detected unless the product of the entire 5.3-kilobase region was synthesized. Translation of RNAs of 3.9 kilobases or smaller yielded proteins which contained the p28 sequence, but p28 was not cleaved. This suggests that the sequence in the region between 3.9 and 5.3 kilobases from the 5' end of the genomic RNA is essential for proteolytic cleavage and contains autoproteolytic activity. The p28 protein could not be cleaved from the smaller primary translation products of gene A, even in the presence of the larger autocleaving protein. Cleavage of the p28 protein was inhibited by addition of the protease inhibitor ZnCl₂. This study thus identified a protein domain essential for autoproteolytic cleavage of the gene A polyprotein.

Mouse hepatitis virus (MHV), a member of the *Coronaviridae* family, is an enveloped virus which contains a single-stranded infectious RNA genome with a molecular weight of more than 6×10^6 (19, 40). Upon infection of a susceptible cell, the viral genomic RNA is first translated, producing viral RNA-dependent RNA polymerase. This polymerase transcribes virion genomic RNA into a negative-strand RNA, which is, in turn, transcribed into a positive-sense genomic RNA and six subgenomic mRNAs (4, 5, 18, 19; for a review see reference 14). These mRNAs form a nested set structure; that is, all have a common 3'-end sequence and poly(A) tail but extend for various lengths in the 5' direction. An additional feature of these mRNAs is that all of them possess an identical 5'-end leader sequence of approximately 72 nucleotides. This leader sequence is derived from the 5' end of the genomic RNA (15, 17, 34). Current evidence favors a unique leader-primed transcription mechanism for the synthesis of these mRNAs. In this model, leader RNA is synthesized, dissociates from the template RNA, and rebinds to the negative-strand RNA template at the intergenic regions where there is a complementary sequence between the leader and RNA template. The leader RNA then serves as a primer for transcription of each mRNA (2, 30).

Translation studies indicate that MHV mRNAs are functionally monocistronic and only the 5' unique region of each mRNA is translated (21, 27, 31). The translation products include three virus structural proteins which are the nucleocapsid protein, encoded by mRNA 7, and two glycoproteins, E1 and E2, encoded by mRNAs 6 and 3, respectively (35, 37). Several small, nonstructural proteins are encoded by mRNAs 2, 4, and 5 (31, 32). The function of these nonstructural proteins has yet to be determined. The largest mRNA, which is of genomic length, is thought to encode

viral RNA-dependent RNA polymerase. Sequencing of the entire genome of another coronavirus, avian infectious bronchitis virus (IBV), suggests that this gene has a capacity to encode two very large proteins of greater than 300 kilodaltons (kDa) (3). *In vitro* translation studies of murine coronavirus genomic RNA indicated that a large polyprotein of 250 kDa is synthesized and cleaved to 28- and 220-kDa proteins (8, 33).

We have been interested in studying the RNA polymerase(s) responsible for the complex transcription strategy of coronaviruses. Previous sequencing and translation studies had confirmed that the p28 protein represents the N-terminal portion of the presumed RNA polymerase precursor (8, 33). To further elucidate the structure and mechanism of synthesis of the potential RNA-dependent RNA polymerase of MHV, cDNA clones representing the 5'-end 2.0- to 5.3-kilobase (kb) region of genomic RNA were sequenced and translated *in vitro*. Sequence information obtained from these cDNA clones predicts an extension of a single long open reading frame which starts at nucleotide 215 of the viral genomic RNA (33). Translation studies of *in vitro*-synthesized RNA derived from these cDNA clones showed that this protein may contain an autoproteolytic activity which is responsible for cleavage of the p28 protein. The protein domain essential for this proteolytic activity is located in the region from 3.9 to 5.3 kb from the 5' end of the genomic RNA.

MATERIALS AND METHODS

Virus and cells. The plaque-cloned JHM strain of MHV (24) was used throughout this study and was propagated on DBT cells, a mouse astrocytoma cell line (12), at low multiplicities of infection. Virus was harvested and purified from the medium, and viral RNA was prepared as previously described (23).

cDNA clones. cDNA clones representing the 5'-terminal 6

* Corresponding author.

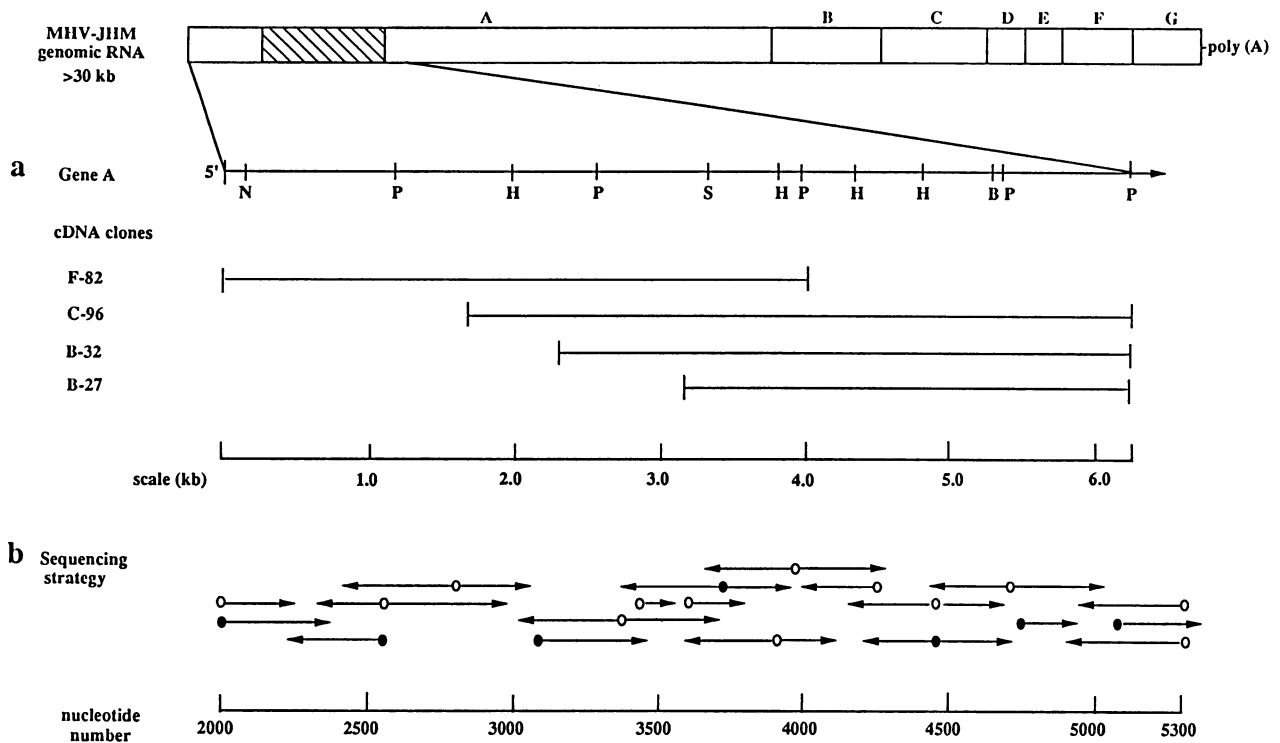


FIG. 1. Diagram of MHV-JHM genomic RNA, cDNA clones, and strategy used for sequencing of the 2.0- to 5.3-kb region (boxed) of gene A clones. (a) Restriction map and structure of the MHV-JHM gene A cDNA clones sequenced. (b) Direction and extent of sequence information obtained from individual restriction sites. Arrows starting with open circles indicate Maxam and Gilbert sequencing with 3'-end-labeled DNA. Abbreviations: B, *Bgl*II; H, *Hind*III; N, *Nar*I; P, *Pst*I; S, *Sph*I. Lengths are expressed in kilobase pairs.

kb of the MHV-JHM genome were derived from genomic RNA extracted from sucrose gradient-purified virus. These clones were generated by priming for first-strand cDNA synthesis with specific oligodeoxyribonucleotides as previously described (30, 33). The procedures for cDNA cloning were adapted from those of Gubler and Hoffman (11) and have been described previously (33). Figure 1 shows the positions of the cDNA clones used to obtain the sequence of the 5'-end 2.0 to 5.3 kb of the MHV-JHM genomic RNA.

DNA sequencing. Sequencing was performed as previously described (33). Both the Sanger dideoxyribonucleotide chain termination method (28) and the Maxam and Gilbert chemical modification method (25, 26) were used. Sequence analyses were obtained with the University of Wisconsin sequence analysis program.

Construction of pT7 recombinant vectors. Construction of plasmid pT7F82N has been previously described (33). This construct, representing the MHV-JHM genomic sequence from the *Nar*I site (nucleotide 187) to the first *Hind*III site (nucleotide 1989), was renamed pT7-NH and used as the parent plasmid for construction of recombinant vectors representing further extensions of the gene A sequence. A recombinant plasmid containing an additional 1.8 kb of the gene A sequence was constructed by excision of the 1.8-kb *Hind*III-to-*Hind*III fragment from cDNA clone C96 (30) and its ligation into the *Hind*III site of pT7-NH. This plasmid was designated pT7-N2H. An additional construct extending the gene A sequences to 5.3 kb was prepared by excising the *Sph*I-to-*Bgl*II fragment from clone C96 and inserting it into the *Sph*I and *Bam*HI (found in the polylinker region of pT7)

sites of pT7-N2H, and the resulting plasmid was designated pT7-NBgl.

In vitro transcription and translation. Recombinant pT7 plasmids were linearized by restriction enzyme digestion and transcribed in vitro with T7 RNA polymerase as previously described (33, 38). The resulting RNA was translated in an mRNA-dependent rabbit reticulocyte lysate (Promega Biotech) under conditions optimized for MHV RNA translation (8). Briefly, 200 ng of capped, in vitro-synthesized RNA was translated in the presence of 180 mM potassium acetate-1.5 mM magnesium acetate-1 mM amino acids minus methionine- 35 S]methionine (1,200 mCi/ml; Dupont, NEN Bioproducts) for 90 min at 30°C in a total reaction volume of 25 μ l. Translation products were analyzed by polyacrylamide gel electrophoresis as described by Maizel (22), with 5 to 15% gradient polyacrylamide gels containing 1% sodium dodecyl sulfate (SDS). Following electrophoresis, the gels were fixed for 30 min in 40% methanol-7% acetic acid, dried, and exposed to Kodak X-ray film at -70°C.

Anti-p28 serum. A synthetic peptide representing amino acids 78 to 93 (NH₃-R-D-I-F-V-D-E-D-P-Q-K-V-E-A-S-T-COOH) of the p28 protein (33) was synthesized by the Microchemistry Laboratory at the University of Southern California Comprehensive Cancer Center. The peptide (2 mg) was mixed with complete Freund adjuvant and injected subcutaneously into several sites along the back of a 3- to 4-kg New Zealand White rabbit. The rabbit was boosted 2 weeks later by intramuscular injection of the peptide (2 mg) in incomplete Freund adjuvant and again after another 2 weeks with the same peptide (2 mg) in phosphate-buffered

saline. The rabbit was bled from the ear vein 1 week after the final injection, and serum was tested by enzyme-linked immunosorbent assay (10) for anti-peptide titer.

Immunoprecipitation of in vitro-translated products. Immunoprecipitation was performed by the method of Kessler (13). In vitro-translated products were diluted to 1 ml in RIPA buffer (50 mM Tris hydrochloride [pH 7.4], 0.3 M NaCl, 4 mM EDTA, 0.5% Triton X-100, 0.1% SDS) and incubated with 3 μ l of anti-p28 immunoglobulin G for 16 h at 4°C. Antibody-antigen complexes were collected by binding to 100 μ l of a 10% (wt/vol) solution of *Staphylococcus aureus* cells (Pansorbin; Calbiochem-Behring, La Jolla, Calif.) for 30 min at 4°C and subsequently washed three times with 1 ml of RIPA buffer. The immunoprecipitated proteins were then eluted from the bacteria by being boiled for 2 min in electrophoresis sample buffer (0.1 M β -mercaptoethanol, 1% sodium dodecyl sulfate, 0.08 M Tris hydrochloride [pH 6.8], 10% glycerol). The bacteria were removed by centrifugation, and proteins were analyzed by electrophoresis on SDS-5 to 15% polyacrylamide gels (22).

RESULTS

DNA sequencing and analysis. To understand the structure and biochemical properties of the probable RNA polymerase of MHV, we first sequenced the 5' end of the gene that encodes this protein. The sequence of the 5'-most 2.0-kb region that encodes the p28 protein has been reported previously (33). We extended the sequence for an additional 3.3 kb, covering a region containing a probable proteolytic domain (see below). The strategy for sequencing this region is shown in Fig. 1. The nucleotide sequence was verified by repeated sequencing of both strands. The dideoxy-chain termination (28) and chemical modification methods (25, 26) were used. Figure 2 shows MHV gene A sequences from 2.0 to 5.3 kb from the 5' end of the genome, continuing from the previously published gene A sequence (33). Analysis of open reading frames of the sequence revealed an extension of the single long ORF of gene A. The two alternative reading frames are closed by frequent termination codons, and no other frame remains open for more than 50 amino acids.

Translation and processing of MHV gene A protein products. In vitro translation studies of MHV genomic RNA have revealed two primary protein products, p28 and p220 (8, 9, 33). It has previously been shown that p28 is an amino-terminal cleavage product of the MHV gene A protein (8, 33). To obtain more information on the translation and processing of gene A protein products, we subcloned various fragments of cDNA clones containing gene A sequences into pT7 vectors (Fig. 3). These plasmids represent gene A sequences starting from nucleotide 187 (the *Nar*I site), which is just upstream of the initiator AUG (nucleotide 215), and extend to 2.0, 3.9, and 5.3 kb from the 5' end of the genome. RNA was transcribed from linearized plasmids and translated in the presence of [³⁵S]methionine in rabbit reticulocyte lysates as described in Materials and Methods. The protein products were analyzed by polyacrylamide gel electrophoresis (Fig. 4). In vitro translation of MHV-JHM genomic RNA resulted in synthesis of the p28 and p220 proteins (Fig. 4, lane A), in agreement with previously published results (8, 33). In addition, several proteins of intermediate sizes were also detected, probably representing degradation or premature termination products of p220 or products from inappropriate translation initiation. The identity of p28 was confirmed by immunoprecipitation with an antiserum specific to a peptide of p28. This antiserum

precipitated only p28 and not the larger proteins (Fig. 4, lane E). Therefore, p28 represents the specific N-terminal cleavage product of MHV gene A (8, 33). In contrast to the genomic RNA translation products, translation of an RNA representing the first 2.0 kb of gene A yielded only a primary translation product of approximately 65 kDa (Fig. 4, lane B). Translation of an RNA extending 3.7 kb into the gene A coding region also resulted in translation of a primary protein product of approximately 135 kDa (Fig. 4, lane C). No p28 was detected in either of these two translation reactions, even though both primary translation products were predicted to contain the p28 sequence. However, translation of an RNA extending to 5.3 kb of the gene A sequence yielded a translation product of approximately 160 kDa and the p28 protein (Fig. 4, lane D). The identity of p28 was confirmed by its precipitation by the specific antiserum against the peptide of p28 (Fig. 4, lane H). No other translation products were precipitated, indicating that cleavage of p28 from the primary translation product was specific. The size of p28 translated from the 5.3-kb RNA is identical to that from the MHV genomic RNA, suggesting that the cleavage sites are identical in both cases. These data indicated that translation of a region between 3.9 and 5.3 kb downstream from the 5' end of the genomic RNA is required for cleavage of the N-terminal portion, p28, from the primary gene A protein. The estimated molecular masses of the primary translation products of the three in vitro-transcribed RNAs were 66, 136, and 186 kDa, respectively. These are in close agreement with the estimated molecular masses of the translation products detected on the gels: 65 kDa for pT7-NH, 135 kDa for pT7-N2H, and 160 plus 28 kDa for pT7-NBgl.

Interestingly, the 65-kDa protein from the pT7-NH-generated translation reaction was precipitated by p28-specific serum (Fig. 4, lane F), indicating that this uncleaved protein retains the p28 epitope. In contrast, no specific protein product could be precipitated from the pT7-N2H translation reaction (Fig. 4, lane G), indicating that the p28 epitope is either unavailable for binding or absent from that protein. As predicted, p160 was not precipitated, since it does not contain the p28 sequence. The above data indicated that correct cleavage of the p28 protein occurred only after translation of a region more than 3.9 kb from the 5' end of the viral RNA. This suggests that the region between 3.9 and 5.3 kb contains autoproteolytic activity or that it induces a conformational change in the protein which then allows cleavage of p28 (see Discussion). To determine whether the potential proteolytic activity present in the larger protein may cleave the smaller in vitro translation product, p65, intermolecularly, pT7-NH RNA and MHV genomic RNA were cotranslated in the same reticulocyte lysate. No cleavage of the 65-kDa polypeptide was noted under this condition, although the protease activity was present and acted on the large protein, as shown by the appearance of the p28 protein from translation of genomic RNA (Fig. 5, lanes A and C). The 65-kDa polypeptide retained the p28 epitope and was precipitated by antiserum specific to the p28 protein (Fig. 5, lanes E and F). Nor was the 65-kDa polypeptide cleaved in cotranslation experiments with pT7-NBgl RNA or by incubation with the translation products from either MHV-genomic RNA or pT7-NBgl RNA (data not shown). These results indicated that either the cleavage site for correct processing of the p28 protein is not available in the 65-kDa polypeptide or MHV gene A protease can act only intramolecularly.

To identify the primary translation product of pT7-NBgl RNA, we performed in vitro translation with protease inhib-

596	S L M V N F S R E V V D M C M D M A L L F M H D V K V A T K Y V K K	629
2001	-CATTGATGGTTAACTTTAGCCGAGAGGTAGTGCATGTGTATGGACATGGCTTTATTGTTTCATGCATGATGTTAAAGTGCCCACTAAGTATGTAAAG	2102
630	V T G K V A V R F K A L G I A V V R K I T E W F D L A V D T A A S A	663
2103	GTTACTGGCAAAGTAGCAGTGGCTTTAAAGCGTTAGGCATTGCTGTTGTCAGAAAAATTACTGAATGGTTTGACTTAGCCGTGGACACTGCTGCTAGTGGC	2204
664	A G W L C Y Q L V N G L F A V A N G V I T F I Q E V P N Y Q E F I N	697
2205	GCTGGATGGCTTTGCTACCAGCTGGTAAATGGCTTATTTCAGTAGCCAAATGGTGTATAAACCTTTATACAGGAGGTGCCCTAATCATCAAGAAATTTATTAAC	2306
698	N Q H F N S H L H P P E L V K N F V D K F K T F F K V L I D S M S V	731
2307	AATCAGCATTCAATTCACACCTGCATCCGCTGAGCTTGTCAAGAAATTTGTTGACAAAGTCAAGACATTTTCAAGGCTTGTATGCACCTATGTCGGTC	2408
732	S I L S G L T V V K T A S N R V C L A G S K V Y E V V Q K S L P A Y	765
2409	TCTATCTTGTCTGGACTTACTGTTGTCAAGACTGCCCAATAGGGTATGCTTGTCTGGCAGTAAGGTTTATGAAGTTGTGCAGAAATCTCTGCTGCATAT	2510
766	I M P V V G C S E A T C L V G E I E P A V F E D D V V D V V K A P L T	799
2511	ATTATGCTGTGGGTCAGCAAGCCACTTGTGTTGGTGGAGATGAACTTGCAGCTTCTTGAAGATGATGTTGTTGATGGTAAAGCCCCCTTAAACA	2612
800	Y Q G C C K P P S S F E K I C I V D K L Y M A K C G D Q F Y P V V	833
2613	TATCAAGGCTGTGTAAGCCACCCTCTTCTTCGAGAGATTGTATTGTGGATAAATATATATGCCCAAGTGGTGGATCAATCTCACTGTGGTGGT	2714
834	D N D T V G V L D Q C W R F P C A G K K V V F N D K P K V K E V P S	867
2715	GATAACGACACTGTTGGCTGTAGATCAGTGTGGAGGTTCCCTGCGGGCAGAAAGTCGTGTTAAACGCAAGCTAAAGTAAAGGAGGTACCTCC	2816
868	T R K I K I I F A L D A T F D S V L S K A C S E F E V D K D V T L D	901
2817	ACGCGAAGATTAAAGATTATCTTCGCTCGGATCGACCTTTGATAGTGTCTCTCGAAGGCAATGTTGAGAGTTGATAAAGATGTTACATTGGAT	2918
902	E L L D V V L D A V E S T L S P C K E H G V I G T K V C A L L K G W	935
2919	GAGCTGCTTGAATGTTGCTTACGCGAGTTGAGAGTACGCTCAGCCCTTGAAGGAGCAGTGTGATAGGCACAAAAGTTTGTGCTACTTAAAGTTGG	3020
936	W T I M S I F L M K E A K K L L P S R M Y V L S A P D E D C V A T D	969
3021	TGGACGATTATGCTATCTTTTGTGAAGGAGGCGAAGAAGTATTGCCCTCCCGTATGATGTCCTCTCTGCTCTGATGAAGACTGCGGTTGCAACAGAT	3122
970	V V Y A D E N Q D D D A D D P V L V L V A D T Q E E D G V A R E Q V D	1003
3123	GTTGTATATGACAGATAAACAAGATGATGATGCTGACGATCTGTAGTCTTGTGCGCGATACCAAGAAGAGGACGGCGTTGCGAGGAGCAGGTTGAT	3224
1004	S A D S E I C V A H T G G Q E M T E P D V V G S Q T P I A S A E E T	1037
3225	TCGGCTGATTCGAAAATTTGTTGCGCACACTGGTGTCAAGAAATGACTGAGCCGTGATGTCGTCGGATCTCAAACCTCCATCGCCTCTGCTGAGGAAACC	3326
1038	E V G E A C D R E G I A E V K A T G V C A D A L D A C P D O A V E A F D	1071
3327	GAACTCGGTGAGGATCGCACAGGGAAGGATGCTGAGGTCAAGCAACTGTGTGCTGATGTTAGATGCTGCCCGATCAAGTGAAGGACCTTGTAT	3428
1072	I E K V E D S I L S E L Q T E L N A P A D K T Y E D V L A F D A I Y	1105
3429	ATTGAAAAGTTGAAGACTCTATCTTAAGTGAAGTTCAAAACCGAAGTAAATGCGCCCGGCAAGACCTATGAGGATGCTTGGCATTGATGCCATATAC	3530
1106	S E T L S A F Y A V P S D E T H F K V C G F Y S P A I E R T N C W L	1139
3531	TCAGAGACGTTGTCTGCATCTATGCTGTGCGGAGTGTAGACGCACCTTAAAGTGTGGATTCTATTGCGCAGCTATAGAGCCTACTAATTTGTGGCTG	3632
1140	R S T L I V M Q S L P L E F K D L G M Q K L W L S Y K A G Y D Q C F	1173
3633	CGTCTACTTTGATAGTAAGCAGAGTTTACCTTTGGAATTTAAAGACTTGGGAGTGCAAAAGCTTGTCTTACAAAGCTGCTATGATGCTGCTT	3734
1174	V D K L V K S A P K S I I L P Q G G Y V A D F A Y F F L S Q C S F K	1207
3735	GTGGACAAACTAGTTAAGAGCGCCCAAGCTATATATCTTCCCAAGGTGGCTATGTCGAGATTTGCTATTTTTCTCAAGCCAGTGTAGCTCAAA	3836
1208	V H A N W R C L K C G M E L K L Q L G L D A V F F Y G D V V S H M C K	1241
3837	GTTCACTGTAAGTGGCGTGTCTAAAGTGGCAGTGAAGTTCAGGCTGAGCGCGGTTTCTATGGAGAGGTTGTGTCTATATGTTGAAG	3938
1242	C G N S M T L L S A D I P Y T F D F G V R D D K F C A F Y T P R K V	1275
3939	TGTGGTAAATGACATGACCTTGTGCTGCAGATATACCTTACACATTGATTTTGGAGTGCAGATGACAAAATTTGTGCTTTTACAGCCGAGAAAAGTC	4040
1276	F R A A C A V D V N D C H S M A V V D G K Q I D G K V V T K F N G D	1309
4041	TTTAGGGCTGCTTGTGGCGTTGATGTTAATGATGCACTCTATGGCTGTTGGATGGCAAGCAAAATGATGGTAAAGTTGTACCAAATTTAATGGTGAC	4142
1310	K F D F M V G H G M T F S M S P F E I A Q L Y G S C I T P N V C F V	1343
4143	AAATTTGATTTATGGTGGCTATGGTATGACATTTAGTATGCTCCCTTTGAGATCGCCAGCTTATGGTTCAATGATATAACCAAAATGTTGTTTGT	4244
1344	K G D V I K V L R R V G A E V I V N P A N G R M A H G A G V A G A I	1377
4245	AAAGAGATGTTATAAAGTTTTGCGCAGAGTGGTGTCAAGTCAATGCTCAACCTGCTAATGGGCGCATGGCTCATGGTGGCGGTTGTCAGGTGCTATA	4346
1378	A K A A G K A F I N E T A D M V K A Q G V C Q V G G C Y E S T G G K	1411
4347	GCTAAGCAGCGGCAAGCTTTTATTAATGAGACTGCGCATATGGTGAAGGCTCAGGCGGTTTGCAGGTTGGTGGATGCTATGATCTACCGGTTGATAA	4448
1412	L C K K V L N I V G P D A R G H G N E C Y S L L E R A Y Q H I N K V	1445
4449	TTATGTAAGAAAGTCTAACATTGTAGGACAGATGCGCGAGGGATGGCAACGAATGCTATTCACCTTTTAGAGCGTCTTATCAGCATATAAATAGTGT	4550
1446	D N V V T T L I S A G I F S V P T D V S L T Y L L G V V T K N V I L	1479
4551	GACAATGTTGTCACCACCTTAATTTGCGGCTGGTATATTAGTGTGCTACTGATGTTCCCTTAACTTATCTACTTGGTGTAGTGACGAAGAATGCTATCTT	4652
1480	V S N N Q D D F D V I E K C Q V T S V A G T K A L S F Q L A K N L C	1513
4653	GTCAGTAATAACAGGATGATTTGATGTGATAGAGAAGTGTCAAGTGACCTCCGTGCTGGTACCAAGCGCTATCAITTCATTTGAGCCAAAATTTGTGC	4754
1514	R D V K F V T N A C S S L F S E S S F V S S Y D V L Q E V E A L D M	1547
4755	CTGATGTAAGTTTGTGACGAATGCATGATGCTGCTTTAGTGAATCAAGCTTTGTCTCAAGCTATGATGTTGTCAGGAAGTTGAAGCGCTGGACATG	4856
1548	I Y N W M M M L V S L C R L I W T V C P Q T G V L S I N L I V L M V	1581
4857	ATATCAATGATGATGATGCTCGTCTTTGTCAGGCTAATATGGACTGTCTGCGCACAGACTGGCTCTTGTCAATAAATTTGATGTTGATGGTGG	4958
1582	L R T I K Y F E C P G E V F V S S Q G K K F G Y V Q N G S F K E A S	1615
4959	TTAAGAACCATTAAGTATTTGAATGTCGGGCGAGGTTTTGTATCTAGCCAGGCAAAAAGTTTGGTTATGTTCAAGTGGTTCATTTAAGGAGGGCAGT	5060
1616	V S Q I R A L L A N K V D V L C T V D G V N F R S C C V A E G E V F	1649
5061	GTTAGCCAAATAAGGGCTTTGCTGCTAATAGTGTGATGCTGTTGCTGATGTTGATGTTAATCTCCGCTCTGCTGTTGTCAGAGGTTGAAGTTTTT	5162
1650	G K T L G S V F C T D I N T K V R C S A I H K G K V F F Q Y S G L	1683
5163	GGCAAGCATTAGGTTGATTTTTTGTGATGGTAAATGTAAACAAAGTGTAGGTGATGCTTCAAGGGAAGGTTTTCTTCAAGTATAGTGGTCTG	5264
1684	S A A D L A A V K D A F-	1695
5265	CTGCGGCGAGTCTCGCTGCTGTTAAAGATGCCTTT--	5300

FIG. 2. DNA sequence of the 5'-end 2.0 to 5.3 kb of MHV-JHM genomic cDNA clones. A translation of the long open reading frame is shown in the single-letter amino acid code.

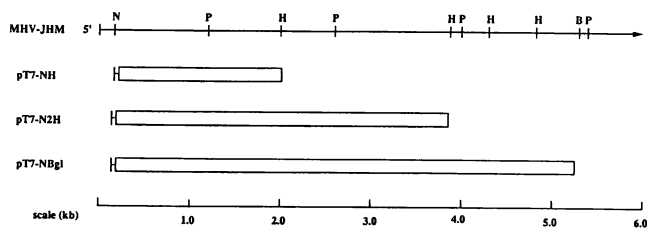


FIG. 3. Restriction map of recombinant plasmids used to transcribe RNA. The open box denotes the long open reading frame of gene A. Plasmids were constructed as described in Materials and Methods and linearized by restriction enzyme digestion, and capped RNA was transcribed by T7 RNA polymerase. For abbreviations, see the legend to Fig. 1.

itors. As shown previously (8), cleavage of p28 from the primary translation product (250 kDa) of MHV genomic RNA could be inhibited by addition of 1 mM ZnCl₂ to the translation reaction (Fig. 6, lane B). When pT7-NBgl RNA was translated with 1 mM ZnCl₂, a protein of 185 kDa, but

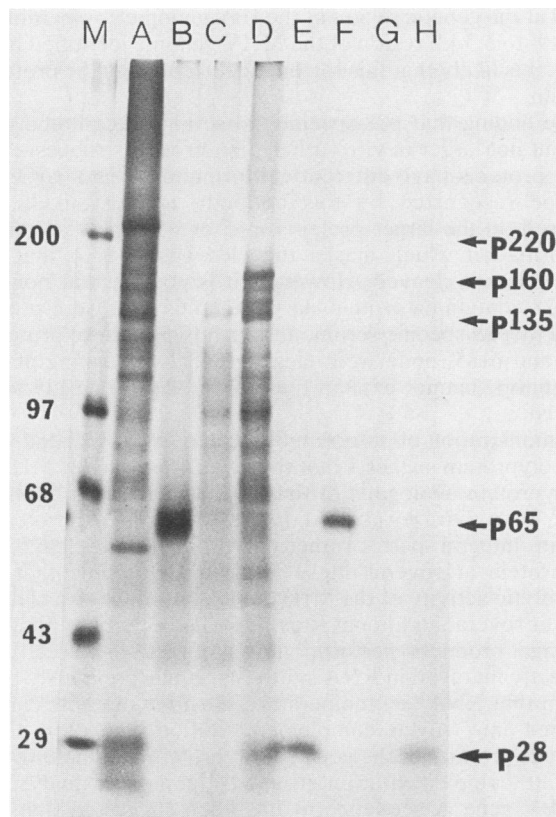


FIG. 4. Translation and immunoprecipitation of MHV-JHM gene A polypeptides. Capped RNA was synthesized from linearized plasmids pT7-NH, pT7-N2H, and pT7-NBgl with T7 RNA polymerase. The RNA was translated in vitro in rabbit reticulocyte lysates in the presence of [³⁵S]methionine. The protein products were analyzed on a 5 to 15% gradient SDS-polyacrylamide gel. Lanes A, B, C, and D, respectively, show translation products of MHV-JHM genomic RNA, pT7-NH-generated RNA, pT7-N2H-generated RNA, and pT7-NBgl-generated RNA. Lanes E, F, G, and H show the products of immunoprecipitation with antisera to p28. Lane M contained ¹⁴C-labeled marker polypeptides; molecular masses are given in kilodaltons on the left.

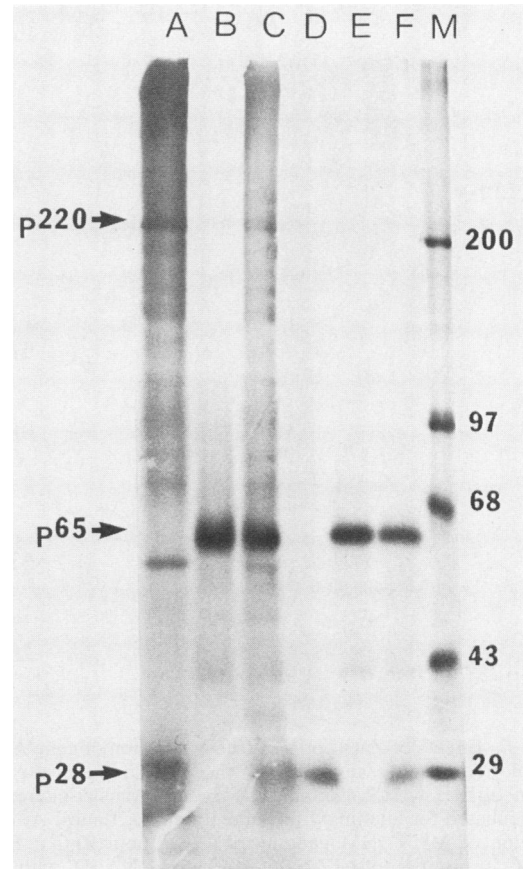


FIG. 5. Cotranslation of MHV-JHM genomic RNA and pT7-NH RNA. MHV-JHM genomic RNA and pT7-NH-generated RNA were translated in vitro separately (lanes A and B, respectively) or together (lane C). The product of the in vitro translation reactions were then immunoprecipitated with antiserum to p28 (lanes D to F). Lane M contained ¹⁴C-labeled marker polypeptides; molecular masses are given in kilodaltons on the right.

no p28, was obtained (Fig. 6, lane D). This result is in agreement with the predicted molecular weight of the primary translation product of this open reading frame. Neither p250 from the MHV genomic RNA translation reaction nor p185 from the pT7-NBgl translation reaction was precipitated by antiserum specific to the p28 protein (Fig. 6, lanes F and H, respectively). This result suggests that these primary translation products maintain a conformation in which this linear peptide epitope is masked or unavailable to bind the antibody.

DISCUSSION

Theoretical considerations suggest that the 5'-most gene of the RNA genome of MHV, gene A, codes for MHV-specific RNA-dependent RNA polymerases (33). The size and sequences of this genetic region have not been completely determined, and only the amino-terminal gene A protein, p28, has been detected in virus-infected cells (9). By analogy to the complete sequence of the genomic RNA of the related coronavirus IBV (3, 7), it has been suggested that the gene that codes for MHV RNA polymerase is probably greater than 20 kilobases long. Recent sequencing work on the MHV genome indeed supports this estimate (C.-K.

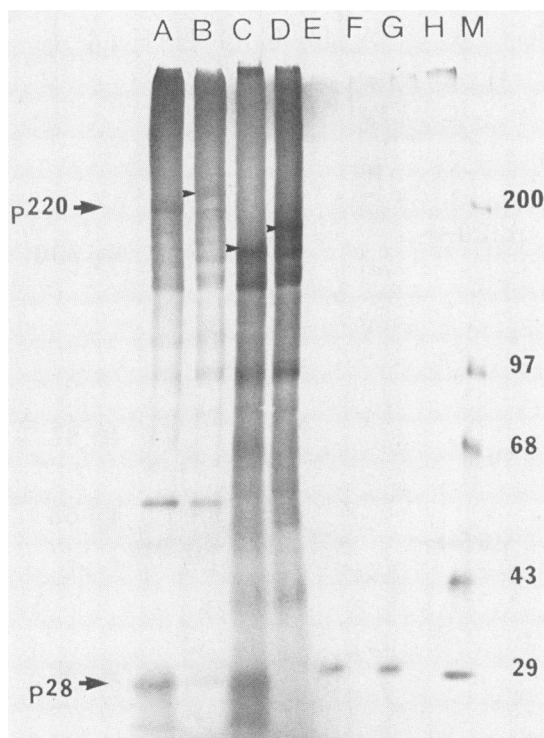


FIG. 6. Effects of $ZnCl_2$ on cell-free translation of gene A RNA. Zinc chloride (1 mM) was added to the reticulocyte lysate before addition of MHV-JHM genomic RNA, and translation reactions were incubated for 90 min with [^{35}S]methionine. Lanes: A, control translation of MHV-JHM genomic RNA without $ZnCl_2$; B, with $ZnCl_2$; C, translation of pT7-NBgl-generated RNA; D, with $ZnCl_2$. The *in vitro* translation products immunoprecipitated with antiserum to p28 are shown in lanes E to H, respectively. The arrows indicate primary (lanes B and D) and cleaved (lanes A and C) translation products. The sizes of ^{14}C -labeled molecular mass markers (lane M) are given in kilodaltons.

Shieh, unpublished data). This large genetic region is probably expressed as a single protein, since there is only one mRNA species that corresponds to gene A. According to the nested-set structure of coronavirus mRNAs (16), only one translation initiation site at the 5' end of each mRNA is utilized, and thus, only one functional protein is probably expressed from this entire gene. Significantly, the sequence of the IBV genome reveals two overlapping open reading frames in this region, which are likely translated into a polyprotein by a ribosomal frameshifting mechanism (6). The exceptionally large size of this gene product suggests that the protein has many different functional domains. This possibility is consistent with the very complex nature of coronavirus RNA synthesis, which employs a unique mechanism of leader-primed transcription (1, 2, 23, 30). Thus, it is likely that the viral RNA polymerase possesses several different enzymatic activities required for carrying out this complex transcription and replication process. Alternatively, the primary gene product of this large gene may undergo posttranslational processing into smaller proteins. This possibility is comparable to other positive-stranded RNA viruses (29, 36), in which several nonstructural proteins are generated from posttranslational cleavage of a single polyprotein. Indeed, it has already been shown that the N terminus of the potential RNA polymerase of MHV is cleaved co- or posttranslationally, resulting in the appear-

ance of a 28-kDa protein (8, 33). The data presented in this report show that the protein made from gene A may contain autoproteolytic activity which is responsible for cleavage of p28 from the primary translation product. It is not clear whether p28 is part of RNA polymerase. Conceivably, it may be a subunit of the polymerases responsible for activities involved in the synthesis of negative-strand, leader, or positive-strand RNA of MHV.

The data presented in this paper are compatible with the presence of autoproteolytic activity in the protein region corresponding to 3.9 to 5.3 kb from the 5' end of the genome. However, the data did not rule out the possibility that this protein region is needed only for maintaining a protein conformation required for the protease activity which is located at an upstream site or for making the p28 cleavage site accessible to proteases. Thus, the exact location of the protease domain cannot be determined. Experiments such as site-specific mutagenesis are needed to further establish the localization of the protease activity. However, it has recently been shown that the IBV gene A product contains a region with remote homology to a *Streptococcus pneumoniae* protease (39; A. E. Gorbalenya, E. V. Koonin, A. P. Donchenko, and V. M. Blinov, *Nucleic Acids Res.*, in press). Interestingly, this putative protease domain is located at the genetic region of the IBV genome comparable to the 3.9- to 5.3-kb region of the MHV genome identified here. Thus, it is likely that this region indeed contains the protease domain.

The finding that p28-specific antiserum precipitated only p65 but not larger *in vitro* translation products suggests that these proteins have different conformations. The particular epitope recognized by this antibody is most likely not exposed on the larger proteins. p65 or p135 may assume a conformation which masks the cleavage site so that the protein is not cleaved. However, it is notable that none of the p28-containing proteins larger than 65 kDa were precipitated by p28-specific serum, although two of these proteins, p250 and p185, underwent cleavage. Thus, masking of this p28 epitope cannot explain the failure of p65 or p135 to be cleaved.

Demonstration of autoproteolytic activity associated with this polyprotein suggests that the gene A product of MHV is a polyprotein analogous to that of picornaviruses, alphaviruses, or flaviviruses (29, 36). In these viruses, the proteases are an integral part of the polyprotein and cleave the polyprotein at several sites. Thus, it is possible that the proteolytic activity of the MHV gene A product also cleaves itself at several additional sites. It is likely that the different cleavage products perform various functions directly or indirectly involved in RNA synthesis. It has previously been shown that RNA (-) temperature-sensitive mutants can be grouped into seven complementation groups (20). Each cleavage product of the gene A polyprotein may belong to a separate complementation group. By computer analysis of the IBV gene A sequence, it has been suggested that this protein contains different domains for RNA polymerase, helicase, nucleotide-binding activities, and several proteases (Gorbalenya et al., in press). Indeed, the complex mechanism of MHV RNA synthesis may require all of these enzymatic activities. So far, study of the mechanism of coronavirus RNA synthesis has been hampered by the paucity of RNA polymerase and an inability to identify gene A protein products synthesized in MHV-infected cells. The approach demonstrated in this paper may provide a useful method for detection of various enzymatic activities associated with these proteins.

ACKNOWLEDGMENTS

We thank Lynn Williams for synthesis of the synthetic peptide and John Fleming and Ligaya Pen for performing enzyme-linked immunosorbent assays on the anti-p28 serum.

This work was supported by Public Health Service research grants A119244 and NS18146 from the National Institutes of Health. S.C.B. is a postdoctoral fellow of the Arthritis Foundation, and L.H.S. is a postdoctoral fellow of the Bank of America-Giannini Foundation.

LITERATURE CITED

1. 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 chains. *J. Virol.* **48**:633-640.
2. Baric, R. S., S. A. Stohlman, M. K. Razavi, and M. M. C. Lai. 1983. Characterization of leader-related small RNAs in coronavirus-infected cells: further evidence for leader-primed mechanism of transcription. *Virus Res.* **3**:19-33.
3. 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.
4. Brayton, P. R., M. M. C. Lai, C. D. Patton, and S. A. Stohlman. 1982. Characterization of two RNA polymerase activities induced by mouse hepatitis virus. *J. Virol.* **42**:847-853.
5. Brayton, P. R., S. A. Stohlman, and M. M. C. Lai. 1984. Further characterization of mouse hepatitis virus RNA dependent RNA polymerases. *Virology* **133**:197-201.
6. Briery, I., M. E. G. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. K. Brown, and S. C. Inglis. 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* **6**:3779-3785.
7. Brown, T. D. K., M. E. G. Boursnell, M. M. Binns, and F. M. Tomley. 1986. Cloning and sequencing of 5' terminal sequences from avian infectious bronchitis virus genomic RNA. *J. Gen. Virol.* **67**:221-228.
8. Denison, M. R., and S. Perlman. 1986. Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. *J. Virol.* **60**:12-18.
9. Denison, M. R., and S. Perlman. 1987. Identification of a putative polymerase gene product in cells infected with murine coronavirus A59. *Virology* **157**:565-568.
10. Fleming, J. O., M. D. Trousdale, F. A. K. El-Zaatari, S. A. Stohlman, and L. P. Weiner. 1986. Pathogenicity of antigenic variants of murine coronavirus JHM selected with monoclonal antibodies. *J. Virol.* **58**:869-875.
11. Gubler, U., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene* **25**:263-269.
12. Hirano, N., K. Fujiwara, S. Hino, and M. Matsumoto. 1974. Replication and plaque formation of mouse hepatitis virus (MHV-2) in mouse cell line DBT culture. *Arch. Gesamte Virusforsch.* **44**:298-302.
13. Kessler, S. W. 1981. Use of protein A bearing staphylococci for the immunoprecipitation and isolation of antigens from cells. *Methods Enzymol.* **73**:442-459.
14. Lai, M. M. C. 1988. Replication of coronavirus RNA, p. 115-136. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), *RNA genetics*. Vol. 1. CRC Press, Inc., Boca Raton, Fla.
15. 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.
16. 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.
17. Lai, M. M. C., C. D. Patton, R. S. Baric, and S. A. Stohlman. 1983. Presence of leader sequences in the mRNA of mouse hepatitis virus. *J. Virol.* **46**:1027-1033.
18. Lai, M. M. C., C. D. Patton, and S. A. Stohlman. 1982. Replication of mouse hepatitis virus: negative-stranded RNA and replicative form RNA are of genome length. *J. Virol.* **44**:487-492.
19. Lai, M. M. C., and S. A. Stohlman. 1978. RNA of mouse hepatitis virus. *J. Virol.* **26**:236-242.
20. Leibowitz, J. L., J. R. DeVries, and M. V. Haspel. 1982. Genetic analysis of murine hepatitis virus strain JHM. *J. Virol.* **42**:1080-1088.
21. Leibowitz, J. L., S. R. Weiss, E. Paavola, and C. W. Bond. 1982. Cell-free translation of murine coronavirus RNA. *J. Virol.* **43**:905-913.
22. Maizel, J. 1971. Polyacrylamide gel electrophoresis of viral proteins. *Methods Virol.* **5**:176-246.
23. Makino, S., S. A. Stohlman, and M. M. C. Lai. 1986. Leader sequences 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.
24. Makino, S., F. Taguchi, and K. Fujiwara. 1984. Defective interfering particles of mouse hepatitis virus. *Virology* **133**:9-17.
25. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560-564.
26. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
27. Rottier, P. J. M., W. J. M. Spaan, M. C. Horzinek, and B. A. M. van der Zeijst. 1981. Translation of three mouse hepatitis virus strain A59 subgenomic RNAs in *Xenopus laevis* oocytes. *J. Virol.* **38**:20-26.
28. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
29. Semler, B. L., R. J. Kuhn, and E. Wimmer. 1988. Replication of the poliovirus genome, p. 23-48. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), *RNA genetics*. Vol. I. CRC Press, Inc., Boca Raton, Fla.
30. 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.
31. Siddell, S. G. 1983. Coronavirus JHM: coding assignments of subgenomic mRNAs. *J. Gen. Virol.* **64**:113-125.
32. Siddell, S. G., H. Wege, A. Barthel, and V. ter Meulen. 1981. Coronavirus JHM: intracellular protein synthesis. *J. Gen. Virol.* **53**:145-155.
33. 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.
34. 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.
35. Stohlman, S. A., and M. M. C. Lai. 1979. Phosphoproteins of murine hepatitis viruses. *J. Virol.* **32**:672-675.
36. Strauss, J. H., and E. G. Strauss. 1988. Replication of the RNAs of alphaviruses and flaviviruses, p. 71-90. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), *RNA genetics*. Vol. I. CRC Press, Inc., Boca Raton, Fla.
37. 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.
38. 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.
39. Tai, J. Y., A. A. Kortt, T. Y. Liu, and S. D. Elliot. 1976. Primary structure of streptococcal proteinase. III. Isolation of cyanogen bromide peptides: complete covalent structure of the polypeptide chain. *J. Biol. Chem.* **251**:1955-1959.
40. Wege, H., A. Muller, and V. ter Meulen. 1978. Genomic RNA of the murine coronavirus JHM. *J. Gen. Virol.* **41**:217-227.