JOURNAL OF VIROLOGY, June 1991, p. 3076–3082
0022-538X/91/063076-07$02.00/0

Vol. 65, No. 6

Identification of Polypeptides Encoded in Open Reading Frame 1b of the Putative Polymerase Gene of the Murine Coronavirus Mouse Hepatitis Virus A59

MARK R. DENISON,[1]* PHILIP W. ZOLTICK,[2] JULIAN L. LEIBOWITZ,[3] CATHERINE J. PACHUK,[2] AND SUSAN R. WEISS[2]

Departments of Pediatrics and Microbiology and Immunology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107[1]; Department of Microbiology, University of Pennsylvania Medical School, Philadelphia, Pennsylvania 19104[2]; and Department of Pathology and Laboratory Medicine, University of Texas Health Science Center, Houston, Texas 77030[3]

Received 17 December 1990/Accepted 1 March 1991

The polypeptides encoded in open reading frame (ORF) 1b of the mouse hepatitis virus A59 putative polymerase gene of RNA 1 were identified in the products of in vitro translation of genome RNA. Two antisera directed against fusion proteins containing sequences encoded in portions of the 3′-terminal 2.0 kb of ORF 1b were used to immunoprecipitate p90, p74, p53, p44, and p32 polypeptides. These polypeptides were clearly different in electrophoretic mobility, antiserum reactivity, and partial protease digestion pattern from viral structural proteins and from polypeptides encoded in the 5′ end of ORF 1a, previously identified by in vitro translation. The largest of these polypeptides had partial protease digestion patterns similar to those of polypeptides generated by in vitro translation of a synthetic mRNA derived from the 3′ end of ORF 1b. The polypeptides encoded in ORF 1b accumulated more slowly during in vitro translation than polypeptides encoded in ORF 1a. This is consistent with the hypothesis that translation of gene *A* initiates at the 5′ end of ORF 1a and that translation of ORF 1b occurs following a frameshift at the ORF 1a-ORF 1b junction. The use of in vitro translation of genome RNA and immunoprecipitation with antisera directed against various regions of the polypeptides encoded in gene *A* should make it possible to study synthesis and processing of the putative coronavirus polymerase.

The coronavirus mouse hepatitis virus (MHV) contains a positive-sense RNA genome of approximately 33 kb (27). Following uncoating of the RNA genome in the cytoplasm, replication is initiated by translation of the viral RNA-dependent RNA polymerase from the positive-sense genome RNA. This polymerase directs the synthesis of negative-strand RNA, from which six subgenomic mRNAs are in turn transcribed and subsequently translated into structural and nonstructural proteins. The coronaviruses transcribe a "nested set" of subgenomic mRNA species which have a common 3′ end as well as an identical 5′ leader sequence (11, 23, 34). A large body of data has been accumulated which supports a "leader-primed" mechanism for transcription of subgenomic RNAs (3, 33). According to this model, an approximately 72-nucleotide leader is synthesized by transcription from the 3′ end of the negative-strand RNA. This leader RNA dissociates from the template and then hybridizes to the conserved intergenic regions on the full-length negative strand to prime the transcription of both the subgenomic mRNAs and copies of the full-length genomic RNA (4). More recent evidence suggests that transcription of subgenomic mRNAs may occur on subgenomic negative-strand templates as well (30, 31).

The polymerase of the coronaviruses must possess many activities, including an ability to synthesize negative-strand RNA, leader RNA, subgenomic mRNAs, and full-length positive-strand virion RNA, as well as probable capping and polyadenylation activities. Both early and late membrane-associated polymerase activities have been described by Brayton et al. (6, 7), who suggested that these activities correspond to synthesis of negative-strand and positive-strand genomic and subgenomic RNA synthesis, respectively. However, Sawicki and Sawicki (29) observed the appearance of positive- and negative-strand RNAs at the same time after infection. Thus, it is still not clear how many viral polymerase complexes or activities are necessary during the MHV infectious cycle.

By analogy to other positive-stranded RNA viruses, it is presumed that the proteins containing polymerase activities must be translated from the input RNA genome immediately after uncoating in the cell cytoplasm. Based on in vitro translation experiments with both genomic and subgenomic RNAs (10, 14), it appears that each MHV mRNA is translated from its 5′ end. Thus, the polymerase is thought to reside in the 5′ gene *A* sequence. The 5′ putative polymerase gene *A* of another coronavirus, avian infectious bronchitis virus (IBV), has been completely sequenced and found to consist of two very large open reading frames (ORFs), F1 (ORF 1a) and F2 (ORF 1b), which potentially encode 441- and 300-kDa proteins, respectively (5). Translation is presumed to initiate at the beginning of F1, with the translation of F2 occurring via a translational frameshift. A sequence similar to that found between the *gag* and *pol* genes of retroviruses (18) has been identified near the beginning of F2 of IBV and has been found to direct efficient frameshifting in vitro (9). The second ORF of gene *A* of MHV-A59 (ORF 1b) has also been sequenced. MHV-A59 ORF 1b is 56% homologous to IBV F2 at the amino acid level and also contains a sequence near its 5′ end that is homologous to the IBV frameshift region and is capable of directing ribosomal frameshifting in vitro (8).

Despite the large amount of information available concerning the replication strategy of the coronaviruses and the

* Corresponding author.

sequence of the putative polymerase gene of RNA 1, identification and characterization of the polymerase proteins and their activities have been elusive. Full-length genomic RNA has been translated in vitro, yielding products of greater than 200 kDa (14, 25). A spectrum of polypeptides have been identified in the cell-free translation products of virion RNA, including the N-terminal products p28 and p220, both of which are derived from a precursor, p250, which is observed only in the presence of protease inhibitors (14). To date, no products of greater than 250 kDa have been clearly identified either by cell-free translation or in infected cells, nor have any proteins been isolated which could be clearly assigned to the second ORF (ORF 1b) of gene *A*. The data here suggest that rabbit reticulocyte lysates can support the translation of the entire gene *A* of MHV-A59. We used antibodies directed against portions of the proteins encoded in the 3'-terminal 2 kb of gene *A* (35) to identify several protein products encoded in ORF 1b of MHV-A59. These are p90, p74, p53, p44, and p32. The sequence of the ORF 1b region encoding these proteins contains a potential polymerase motif (GDD) (1) and thus may represent the first identified polymerase protein in a coronavirus.

MATERIALS AND METHODS

RNA preparation. MHV-A59 virion RNA was isolated from infected cells as described previously (14). Briefly, monolayers of 17Cl-1 cells were infected with MHV-A59 at a multiplicity of infection of 10, and the medium was harvested 10 to 12 h after infection, when ~70 to 80% of the cells were involved in syncytia. The supernatant medium was centrifuged briefly to remove cellular debris and then centrifuged at 4°C for 90 min at 80,000 × *g* to pellet the virus. This pellet was resuspended in NTE buffer (0.1 M NaCl, 0.01 M Tris [pH 7.4], 0.001 M EDTA) by sonication, layered on a 5 to 40% sucrose–NTE gradient, and centrifuged for 45 min in an SW41 Ti rotor at 105,000 × *g*, resulting in a discrete band of virus. The virus band was removed, repelleted, lysed in 0.5% sodium dodecyl sulfate (SDS) in NTE, and treated with proteinase K (200 μg/ml) prior to phenol-chloroform extraction and ethanol precipitation in the presence of 0.3 M sodium acetate. Approximately 1 μg of genome RNA per 10^7 cells was obtained by this method. Genome RNA was resuspended at a concentration of 100 ng/μl in NTE buffer in the presence of RNasin and stored at −70°C for use in in vitro translation reactions.

An mRNA transcript (96/2.9) representing the 3' portion of ORF 1b was prepared for in vitro translation. A 2.9-kb *Sma*I-*Sac*I fragment of MHV-A59 clone 96, obtained from P. Bredenbeek and W. Spaan (8, 27), was subcloned into *Sma*I- and *Sac*I-cleaved IBI-31 (Fig. 1B). This fragment contains 2,314 nucleotides from gene *A*, ORF 1b, and 514 nucleotides from gene *B*, ORF 2a. The resulting plasmid was cleaved at an *Eco*RI site downstream of the viral sequence. RNA was transcribed by using T7 RNA polymerase and purified for in vitro translation as described previously (10, 20).

In vitro translation. In vitro translations were performed in a rabbit reticulocyte lysate system, as described previously (14, 25, 26). The lysate was used as supplied (Promega), and mRNA was added at 200 ng per 25-μl reaction mix (8 to 12 μg/ml). Radioactive label, [^{35}S]methionine (NEN), was added at 1,000 μCi/ml unless otherwise indicated. For most experiments, translations were carried out for 90 min at 30°C and were terminated by the addition of 2× Laemmli buffer (22) or RIPA lysis buffer (1% Nonidet P-40 1% deoxycholate-sodium salt, 0.1% SDS, 150 mM NaCl, 10 mM Tris-HCl [pH 8.0], 0.5 mM phenylmethylsulfonyl flouride, aprotinin [20 μg/ml; Sigma]), or the reaction mixes were quick frozen in ethanol-dry ice and stored at −70°C.

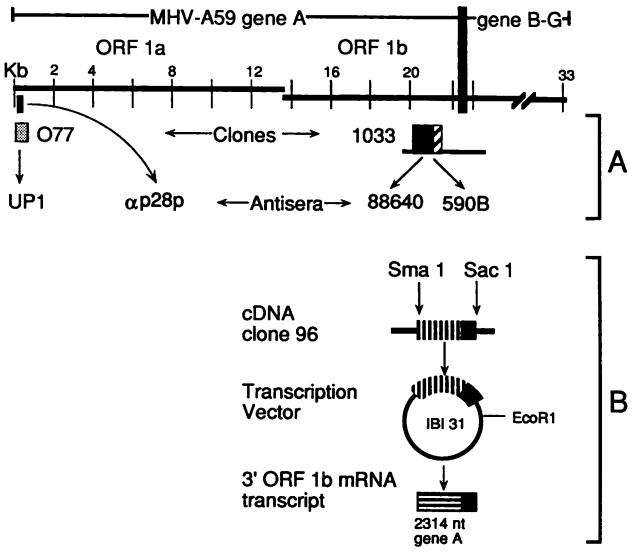


FIG. 1. Schematic of MHV-A59 gene *A* of RNA 1. (A) Locations of the clones O77 and 1033 on the MHV-A59 genome. These clones were used for synthesis of fusion proteins which were subsequently used for antiserum preparation in rabbits, as described in Materials and Methods. The antisera are indicated by arrows. αp28p denotes the antiserum prepared from a synthetic oligopeptide, as described in Materials and Methods. The size and location of this peptide are not shown to scale. (B) Synthesis of 96/2.9 mRNA transcript from clone 96. The locations of clone 96 and transcript 96/2.9 are shown to scale based on the gene map in panel A. nt, nucleotides.

Antibody preparation. Antisera directed against polypeptides encoded in gene *A* were raised in rabbits by using as immunogens procaryote-virus fusion proteins (35). Briefly, fragments of DNA from cloned cDNAs representing portions of MHV-A59 gene *A* were inserted into the bacterial plasmid pGE374 between "out of frame" and truncated *recA* and *lacZ* genes. The recombinant plasmids were used to transform *Escherichia coli* MC1061. Clones in which the reading frame of *lacZ* was restored and which therefore contained all three proteins fused in frame were selected on the basis of β-galactosidase activity. The tripartite fusion proteins were purified by their insolubility and used to raise antibodies in rabbits as described previously (35). The antisera relevant to the experiments described here (Fig. 1) are: (i) UP1, directed against the polypeptide encoded in a 0.3-kb *Bam*HI-*Pst*I fragment of clone O77, located 200 nucleotides from the beginning of ORF 1a, within the coding region of p28; (ii) 88640, directed against a polypeptide encoded in a 1.0-kb *Pst*I-*Pst*I fragment of clone 1033 located near the 3' end of ORF 1b; and (iii) 590B, directed against the polypeptide encoded in a 0.427-kb *Pst*I-*Sph*I fragment of clone 1033, near the extreme 3' end of ORF 1b (Fig. 1).

The antipeptide serum αp28p was raised in rabbits against a 14-amino-acid synthetic oligopeptide (amino acids 25 to 38 from the initiating methionine of MHV strain JHM ORF 1a), which is within the region coding for p28. This antiserum was a generous gift from Stanley Perlman of the University of Iowa. The hybridoma cell line secreting antibodies against the S protein (JA3.10) was obtained from John Fleming at the University of Wisconsin. The hybridoma cell line secret-

ing the antinucleocapsid antibody 1.16.1 was isolated after infection of a BALB/c mouse with MHV strain JHM (MHV-JHM) as described before (24). This antibody recognizes MHV-A59, MHV-JHM, and MHV-3 nucleocapsid protein in immunoprecipitation and Western immunoblot assays.

Immunoprecipitation. In vitro translation products were subjected to radio immunoprecipitation (RIP) with the above antisera. Samples of in vitro translation products (0.5×10^6 to 3.0×10^6 trichloroacetic acid [TCA]-precipitable counts) were diluted into 1 ml of RIPA buffer. For polyclonal antisera, typically 5 μl of serum was added directly to the sample in RIPA buffer. For the monoclonal antisera, 50 μl of hybridoma cell supernatant was used. The reaction mix was incubated overnight at 4°C with top-to-bottom inversion. Subsequently, 30 μl of protein A-Sepharose (Sigma) was added to the RIPA reaction mix and incubated for an additional 4 h at 4°C. The protein A-containing Sepharose beads were then pelleted in a microcentrifuge, the supernatant was removed, and the beads were washed four times with RIPA buffer. Finally, 50 μl of 2× Laemmli buffer was added to the beads, which were boiled for 4 min at 100°C prior to loading on gradient polyacrylamide-SDS gels for analysis.

Partial protease digests. Partial proteolysis was performed as described by Cleveland et al. (13). Proteins were separated on 5 to 18% SDS–polyacrylamide gradient gels. Bands of interest were identified either by autoradiography or by proximity to stained, nonradiolabeled markers and were then excised and placed into the wells of stacking gels over a 15% separating gel. Gel slices were overlaid with 50 μl of a 50-μg/ml solution of *Staphylococcus aureus* V8 protease (Sigma) in 0.1% SDS–1.0 mM EDTA–0.125 M Tris-HCl (pH 6.8)–10% glycerol and subjected to electrophoresis at 25 mA until the bromphenol blue dye front approached the bottom of the stacking gel. At that time the current was turned off, and proteolysis was allowed to proceed for 45 min, after which the current was reapplied at 25 to 30 mA for four hours. Gels were processed for fluorography in 22% PPO (2,5-diphenyloxazole) in dimethyl sulfoxide as described previously (14).

RESULTS

Translation and immunoprecipitation of MHV-A59 gene *A* ORF 1b products. Genome RNA was translated in a rabbit reticulocyte lysate, and the products were immunoprecipitated with antisera directed against fusion proteins containing sequences from the 3′ portions of ORF 1b. The location on the MHV-A59 genome of the DNA fragments that were used to produce fusion proteins and the antisera are shown in Fig. 1. In vitro translation products labeled with [^{35}S]methionine were immunoprecipitated with antisera 88640 and 590B and analyzed by polyacrylamide gel electrophoresis (PAGE) (Fig. 2). Polypeptides of 90 and 74 kDa were precipitated by both antisera. Another polypeptide of 53 kDa was precipitated by 88640 but not by 590B. Conversely, products of 44 and 32 kDa were precipitated by 590B but not by 88640. Based on the known sequence of ORF 1b of MHV-A59 and the known locations of the epitopes for 88640 and 590B on that sequence, p44 and p32 are probably overlapping C-terminal products. Because of their size, they must extend some distance into the region included in the clone used to produce 88640 (see Fig. 6). The failure of 88640 to precipitate p44 and p32 is most likely due to lack of antigenic determinants in the carboxy terminus of the clone used to produce 88640. This explanation is supported by

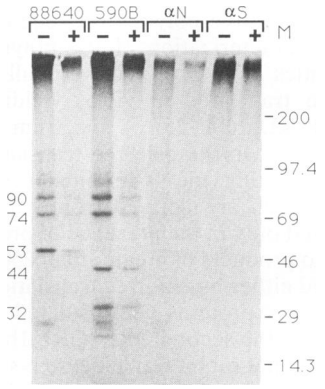


FIG. 2. RIP of in vitro translation products of MHV-A59 genome RNA. Purified genome RNA was translated in the presence of [^{35}S]methionine and immunoprecipitated with the antisera as indicated, with the RIP products analyzed by 5 to 18% gradient SDS–PAGE. Lanes: M, molecular mass markers; αN, monoclonal antinucleocapsid antibody; αS, monoclonal antispike antibody; 88640 and 590B, polyclonal antisera directed against sequences encoded in ORF 1b. Lanes − and + indicate the absence and presence of leupeptin, respectively. The numbers on the left indicate the mass (in kilodaltons) of specific polypeptides precipitated by the ORF 1b antisera.

sequence analysis, which shows a clustering of antigenic determinants (hydrophobicity, secondary structure, and antigenic index) in the 5′ end of the sequence used to produce 88640 and the 3′ end of the sequence used to produce 590B.

This pattern of proteins was observed in translation reactions with three separate MHV-A59 genome RNA preparations as templates. No polypeptides with mobilities similar to p90, p74, p53, p44, or p32 were precipitated by anti-N, anti-S, or preimmune serum. The addition of 2 mM leupeptin, a reversible serine-cysteine protease inhibitor, resulted in a 50% decrease in TCA-precipitable counts, but the pattern of products seen with 88640 and 590B was identical to that seen in the absence of leupeptin.

Comparison of immunoprecipitation products of in vitro translation of MHV-A59 ORF 1a and ORF 1b. To further confirm that the products identified originated from the 3′ end of ORF 1b, RIPs of in vitro translation products were carried out with the antiserum UP1, which is directed against a fusion protein made from a plasmid containing an 0.3-kb fragment from ORF 1a, within the p28 coding region (Fig. 1). As shown in Fig. 3, after reaction of in vitro translation products with UP1, the majority of the TCA-precipitable label was present in p28. A spectrum of other polypeptides were observed, ranging in size from 60 to 160 kDa, as observed previously (2, 14). A similar pattern was obtained after immunoprecipitation with the antiserum αp28p directed against a 14-amino-acid sequence within the p28 coding region. This was distinctly different from the pattern of products obtained from immunoprecipitation with either of the 3′ ORF 1b antisera, 88640 and 590B. In addition, as reported earlier (14), the cleavage of the ORF 1a product, p28, was blocked by the protease inhibitor leupeptin, whereas this was not true of the polypeptides encoded in ORF 1b.

Comparison of the ORF 1b polypeptides synthesized from genome RNA and from the ORF 1b synthetic mRNA transcript (96/2.9) by partial proteolysis. The relationships among the ORF 1b products was evaluated by partial protease digests of in vitro translation products (13, 14). Genome

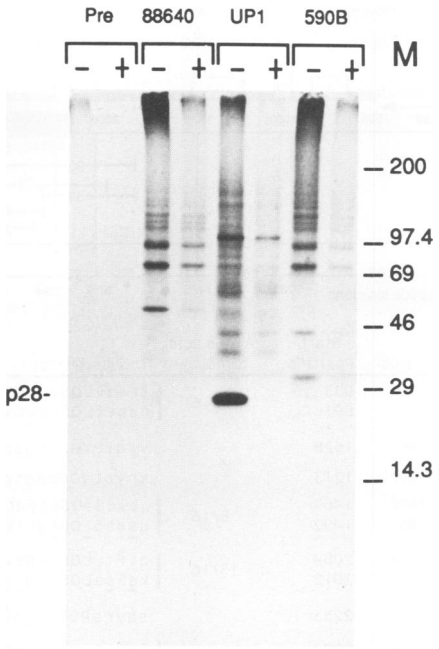


FIG. 3. Comparison of in vitro translation products of MHV genome RNA following immunoprecipitation with antisera directed against polypeptides encoded in ORF 1a and ORF 1b. Immunoprecipitation products were analyzed by 5 to 18% gradient SDS–PAGE. Pre, preimmune serum; 88640 and 590B are ORF 1b antisera, as in Fig. 1; UP1 is the ORF 1a antiserum. Lanes − and + indicate the absence and presence of leupeptin, respectively. Lane M, molecular mass markers (in kilodaltons).

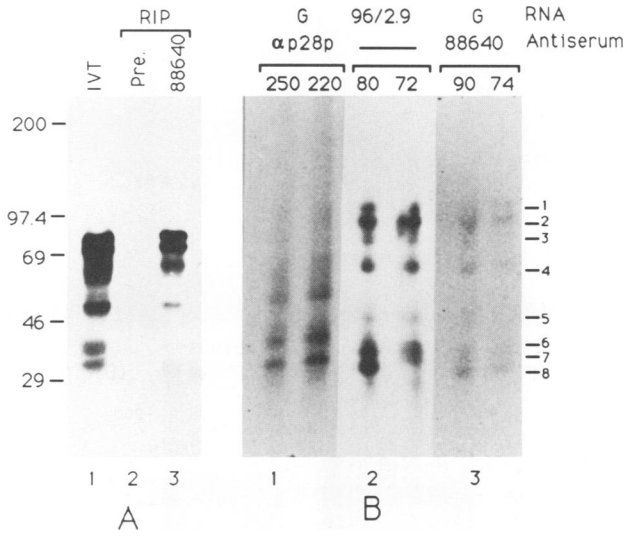


FIG. 4. Translation of 96/2.9 transcript RNA, immunoprecipitation with antiserum to ORF 1b, and comparison with genome RNA translation by partial proteolysis mapping. (A) ORF 1b mRNA transcript 96/2.9 was translated in vitro (IVT) and immunoprecipitated (RIP) with rabbit preimmune serum (Pre.) or ORF 1b antiserum (88640). (B) MHV-A59 genome RNA or the ORF 1b mRNA transcript 96/2.9 was translated in vitro and immunoprecipitated with the antisera indicated. The products were analyzed on preparative 5 to 18% gels, and specific polypeptide bands were excised, subjected to proteolysis by *S. aureus* V8 protease, and analyzed on a 15% gel. Panel 1, in vitro translation of genome RNA in the presence of leupeptin, RIP by αp28p. Polypeptides at 250 and 220 kDa were digested. Panel 2, translation of 96/2.9 mRNA transcript. Polypeptides at 80 and 72 kDa, as seen in panel A1, were digested. Panel 3, translation of genome RNA in the absence of leupeptin, with RIP by 88640. Major polypeptides at 90 and 74 kDa were digested (see Fig. 2). All digests were run on one gel, but the digests of the RIP products required significantly longer exposures (four times longer than for 96/2.9 mRNA). Digest fragment numbers for panels 2 and 3 are shown on the right.

ORF 1b products were compared with translation products of the mRNA transcript 96/2.9, which encodes the extreme 3′ 80 kDa of protein of ORF 1b. Genome RNA and the 96/2.9 transcript were translated, and the products were immunoprecipitated with antiserum 88640 and analyzed by 5 to 18% gradient SDS–PAGE. The 90- and 74-kDa products of genome RNA translation and the 80- and 72-kDa products of 96/2.9 RNA translation were excised from the gel and compared by partial proteolysis (Fig. 4). The products of ORF 1a, p220 and p250, which were immunoprecipitated from genome translation with the ORF 1a antipeptide serum αp28p were used as controls. Polypeptide bands excised from the gel were electrophoresed on 15% SDS gels with an overlay of *S. aureus* V8 protease, allowing partial proteolysis in the stacking gel (Fig. 4). Digests of the 80- and 72-kDa polypeptides from the ORF 1b synthetic RNA transcript resulted in a pattern with eight distinct fragments from the 80-kDa band and seven fragments from the 72-kDa band. The fragment not seen in the 72-kDa band digest had a molecular mass of significantly less than 14.3 kDa and probably represents the difference between the sizes of the 80- and 72-kDa bands (Fig. 4B). The digest patterns of the polypeptides derived from ORF 1a (p220 and p250) were similar to each other but were distinct from those derived from the products of the ORF 1b transcript 96/2.9 (Fig. 4B, panels 1 and 2). This was expected, since the products were derived from opposite poles of the gene. The digests of the 90- and 74-kDa ORF 1b polypeptides, immunoprecipitated from the products of translation of genome RNA by 88640, resulted in seven to eight fragments, all of which comigrated with those seen from the digests of the 80- and 72-kDa products of the ORF 1b synthetic mRNA transcript 96/2.9 (Fig. 5B, panels 2 and 3). Additional evidence to support this identity was found in the fact that the in vitro translation products of 96/2.9 were precipitated by 88640 (Fig. 4A) and 590B antisera and not by preimmune serum, antisera to regions of ORF 1b, or antibodies to MHV-A59 structural proteins (data not shown). This confirmed that the 90- and 74-kDa polypeptides translated from genome RNA were indeed derived from the 3′ end of ORF 1b.

Kinetics of accumulation of ORF 1b products. Accumulation of in vitro translation products of MHV-A59 genome RNA, the ORF 1b mRNA transcript 96/2.9, and the unrelated brome mosaic virus RNA (Promega) were measured over a time period from 0 to 90 min (Fig. 5). Each of the templates showed a similar rate of translation in the reticulocyte lysate, with the synthesis of 100 kDa of protein requiring approximately 9 min at 30°C. Figures 5C and D show the products of in vitro translation of genome RNA immunoprecipitated with antisera directed against sequences encoded in the 5′ end of ORF 1a (αp28p) and the 3′ end of ORF 1b (88640). When translation products were precipitated with αp28p, a series of increasingly large products appeared up to 30 min into translation, at which time p28 and p220 were first observed, as has been demonstrated previously (14). In contrast, the ORF 1b polypeptides p90,

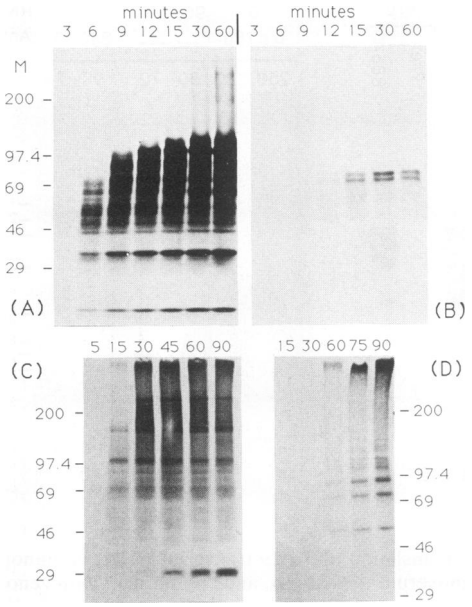
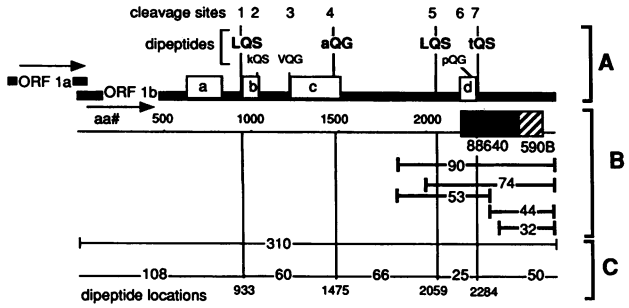


FIG. 5. Kinetics of appearance of MHV ORF 1a and ORF 1b products during in vitro translation of MHV-A59 genome RNA. Products were labeled with [^{35}S]methionine, and samples were removed for analysis by 5 to 18% gradient SDS–PAGE at the times (in minutes) shown above each of the lanes. (A) Translation of brome mosaic virus mRNA, used as a control. Molecular mass markers (in kilodaltons) are shown to the left. (B) Translation of 96/2.9 mRNA transcript. (C) Translation of genome RNA immunoprecipitated by αp28p antiserum. Markers are shown to the left. (D) Translation of genome RNA immunoprecipitated with 88640. Markers are shown to the right.



Site #	Virus	Start Position	Amino acid homology	Subsequence
1.	MHV	933	10/16	lrsavLQSvGaCVVCs
	IBV	891		raptt LQScGvCVVCn
2.	MHV	1028		vfglyKQSctgspyie
3.	MHV	1213		krycttVQGppgtgksh
4.	MHV	1469	8/16	qtvdSaQGseydfviy
	IBV	1492		dsetSlQGtglfklcn
5.	MHV	2054	13/16	ntFtrLQSleNvvYNl
	IBV	2012		ksFsaLQSidNlaYNm
6.	MHV	2253		shyrsPQGnpggnrvg
7.	MHV	2278	11/16	rgtiftQSrilssftp
	IBV	2350		tcypqlQSawtcgynm

D

FIG. 6. Products of MHV-A59 ORF 1b. (A) Schematic of potential 3C protease-like cleavage sites of ORF 1b. Cleavage site numbers refer to sequences shown in panel D. Dipeptides refer to the consensus cleavage site and the preceding (−1) amino acid. Amino acid (aa) homologies with IBV are in boldface. The location of potential SDD polymerase (A), metal binding (b), helicase (c), and GDD polymerase (d) motifs are shown. (B) Products identified by in vitro translation and immunoprecipitation with antisera directed against ORF 1b fusion proteins. Probable locations are based on immunoprecipitation patterns with antisera 88640 and 590B. The exact sequence relationship among products has not been determined. Note the location of the GDD sequence in relationship to the antiserum 88640 and the probable locations of p90 and p74. (C) Predicted ORF 1b proteins which would result from cleavages at the conserved sites shown in panel A if ORF 1b were translated as an intact polyprotein prior to cleavage. (D) Sequence of seven potential protease sites of ORF 1b. Boldface capital letters indicate direct amino acid matches with IBV. Underlining denotes homologous amino acid counterparts.

p74, and p53 were not detectable until 60 min. Based on the rates of in vitro protein synthesis on these templates, if any products had been the result of internal initiation at the start of ORF 1b, then the 3′ ORF 1b products should have been present within 15 to 20 min. These kinetics data strongly suggest that translation occurred at the expected rate after initiating at the 5′ end of ORF1a, with extension through the frameshift and into ORF 1b.

Prediction of protease sites within ORF 1b. Gorbalenya et al. (16) identified a region of F1 (ORF 1a) of IBV which has close homology to the 3C protease of the picornaviruses. A search for potential substrate sites for this putative 3C protease resulted in the identification of four potential 3C protease-like cleavage sites within F2 (ORF 1b) of IBV. Each of these regions contains the consensus dipeptide, consisting of a glutamine followed by glycine or serine [Q(G/S)], usually preceded by a leucine or another hydrophobic residue. These sites also have close homology to each other and to picornavirus 3C protease-like cleavage sites from −7 to +9 amino acids from the consensus cut site. We analyzed ORF 1b of MHV-A59 by using a protein subsequence match program (MacVector; IBI/Kodak) to search for sequences homologous to the IBV potential protease digestion sites. The search algorithm, which required an exact match in the consensus sequence dipeptide Q(G/S), identified seven conserved sites within ORF 1b of MHV-A59 (Fig. 6A and D). Five of these sites contained a hydrophobic residue in the −1 (P2) position with respect to the dipeptide. Two of the sites had a leucine in that position. Four of the sites identified in ORF 1b of MHV-A59 had between 9 and 11 amino acids that were identical or homologous to their counterparts in the corresponding IBV sequence.

DISCUSSION

In this report we present the first description of polypeptides translated from the second ORF of the putative polymerase gene of MHV-A59, ORF 1b. Specifically, polypeptides of 90, 74, 53, 44, and 32 kDa have been identified as products of in vitro translation of genome RNA. Several lines of evidence suggest that these polypeptides are indeed encoded in ORF 1b of gene *A*. They are immunoprecipitated by antisera raised against procaryote-virus fusion proteins containing sequences encoded in the 3′ end of ORF 1b but not by an antiserum directed against sequences in ORF 1a (αp28p) or by monoclonal antibodies directed against MHV structural proteins N and S. Furthermore, partial proteolysis of these products shows identity between the p90 and p74 genome translation products and the 80- and 72-kDa poly-

peptides resulting from in vitro translation of the ORF 1b mRNA transcript 96/2.9, which contains sequences from the 2.9 kb of the 3′ end of ORF 1b.

The kinetics of accumulation of these polypeptides during cell-free translation suggests that synthesis of the polypeptides encoded in ORF 1b occurs on full-length genomic RNA, presumably by initiation of translation at the 5′ end of ORF 1a, followed by translation of ORF 1a, with subsequent ribosomal frameshifting and translation of ORF 1b (9). The cell-free translation experiments shown in Fig. 5 can be used to calculate an in vitro amino acid incorporation rate of 1.6 to 2 residues per s. At this rate, the entire gene, encoding approximately 7,600 amino acids, should be translated within 60 to 70 min. If all translation is initiated at the 5′ end of the gene, 3′ ORF 1b products should not begin to appear prior to 45 to 60 min. Our experimental results are consistent with this prediction (Fig. 5D).

It was also noted that the proteins identified appeared as discrete bands, with accumulation over time but without identifiable precursors or subsequent processing of the proteins p90 and p74. If an intact polyprotein were made from the entire gene or from ORF 1b alone prior to generation of products, then precursors of 310 to 760 kDa might be expected. We did detect a diffuse, heterogeneous band of label in the region above 300 kDa which could represent a variety of precursor polyproteins, but it has not yet been possible to determine whether these are true precursors to the bands described.

Previous data from in vitro translation of the MHV genome indicate that proteolytic processing in the ORF 1a products begins quite soon after synthesis. Proteins translated from the N-terminal region of ORF 1a appear to be cleaved as soon as 160 kDa of protein has been synthesized (2, 14). Baker et al. concluded that an autoproteolytic activity is encoded in the sequence between 3.9 and 5.3 kb from the 5′ end of the genome of MHV-JHM, a closely related strain of MHV (2). This interpretation is supported by sequence analysis of IBV (16), which has revealed the presence of two potential protease activities within the first 9 kb of F1 (ORF 1a). One of these closely resembles the 3C protease of the picornaviruses, and the other has similarity to the cysteine protease of *Streptococcus pneumoniae*. The possibility that p28 cleavage occurs at a Tyr-Gly (32) and the demonstrated inhibition of p28 cleavage in vitro by leupeptin (Fig. 3) support the inference that a similar 3C-like protease may be active during translation of the gene. If MHV encodes a 3C protease-like activity in ORF 1a, it is possible that intramolecular self-cleavages and *trans*-cleavages would occur as soon as the proteases were synthesized and appropriate substrate sites were available. This type of cotranslational cleavage is well characterized in the picornaviruses, in which the first specific cleavages of P1 and P2 occur even before the P3 portion of the polyprotein is synthesized (28). The similarity of the IBV potential protease activities to those of the picornaviruses, the data from MHV-A59 genome in vitro translation, the presence of potential 3C protease-like cleavage sites, and the observed pattern of products from RIP with 88640 and 590B all suggest that larger precursors may be rapidly cleaved into products which represent the functional components of the polymerase gene. This further suggests that full-length ORF 1b or ORF 1a-ORF 1b polyproteins might not be identified, either in vitro or in vivo, except in the presence of a specific inhibitor or blocker of proteolysis.

We have identified several potential 3C protease-like cleavage sites in MHV-A59 ORF 1b, based on homology to picornavirus consensus dipeptide cleavage sites and homology to similar sites described in the sequence of ORF 1b of IBV. If these sites were utilized by a 3C-like protease from ORF 1a, then primary cleavage products of 108, 61, 66, 25, and 51 kDa would be expected from an ORF 1b polyprotein (Fig. 6). However, only one of the identified ORF 1b products, p74, can be potentially accounted for by cleavage at only one of these sites, specifically the QS at amino acid positions 2059 and 2060. Furthermore, none of the ORF 1b products demonstrated the same sensitivity in vitro to the protease inhibitor leupeptin as did the 5′ ORF 1a peptide p28 (Fig. 2 and 3), the cleavage of which was completely blocked by the concentrations used in the present experiments. Leupeptin is a reversible inhibitor of trypsinlike serine and many but not all cysteine proteases. Both of the potential protease activities identified in IBV are of the cysteine type, with similarities to those of poliovirus and other RNA viruses (16). Together these data suggest that although MHV-A59 may possess a protease activity in ORF 1a which effects the cleavage of p28, an entirely different protease activity may be necessary for the cleavage of ORF 1b products from a polyprotein precursor. The demonstration of precursors may require more specific, irreversible inhibitors of cysteine proteases or of other protease activities, such as those of the aspartic, metallo, or serine variety.

Functional proteins of the putative polymerase gene have not yet been isolated from any coronavirus, but based on sequence analysis, several potential activities have been proposed for MHV-A59 and the closely related strain MHV-JHM. These include a protease activity in ORF 1a of MHV-JHM (2) and potential polymerase (SDD) and helicase (8, 16) domains in ORF 1b of MHV-A59. The locations of these motifs with respect to potential 3C-like protease cleavage sites are shown in Fig. 6A. In addition, we have identified a possible "metal finger" nucleoside triphosphate-binding domain not previously recognized in MHV-A59, which is conserved in the sequence of IBV (16).

At least one of the ORF 1b products identified in this report, p90, must include a sequence of two asparagine residues preceded by a glycine (GDD), which begins at amino acid 2239 from the start of ORF 1b. This GDD tripeptide has been characterized as the active site of consensus RNA-dependent RNA polymerase sequences in several plus-strand RNA viruses (1). The GDD and flanking sequences in ORF 1b conform well to the hydrophobic-GDD-hydrophobic character of the RNA-dependent RNA polymerases described by Argos (1). In addition, this GDD sequence is likely to be an exposed site, as deduced from hydropathicity (15, 17, 21), surface probability (19), and secondary-structure predictions (12). Finally, the region flanking the A59 GDD sequence (GDDVIFS) demonstrates a match of six of seven amino acids with the poliovirus polymerase (GDDVIAS). It is not clear what role this motif plays in MHV replication, since this GDD sequence is not conserved in the sequence of IBV (8).

In conclusion, rabbit reticulocyte lysates appear to be capable of translating the entire 23 kb of the putative polymerase gene *A*. This is the largest described polyprotein gene of an RNA virus translated in vitro, and it was accomplished in a period of time well within the working span of the reticulocyte lysate. This should make it possible to study the entire scheme of coronavirus gene *A* synthesis and processing as more antibodies become available to different regions of the gene. We are presently using these antisera to probe infected cells. This should allow us to

identify the polypeptides important for viral polymerase activity during infection.

ACKNOWLEDGMENTS

We thank Bernadette Giangreco, J. DeVries, and Z. Bieleuka for expert technical assistance, S. Perlman and A. L. Olson for antibody αp28p, and S. Perlman, K. Holmes, and J. Gombold for helpful discussion.

This work was supported by Public Health Service grants R29 AI-26603 (M.D.), AI-17418, and NS-21954 (S.W.) and National Multiple Sclerosis Society grant NMSS RG 2203-A-5 (J.L.L.).

REFERENCES

1. **Argos, P.** 1988. A sequence motif in many polymerases. Nucleic Acids Res. **16**:9909–9916.
2. **Baker, S. C., C.-K. Shieh, L. H. Soe, M.-F. Chang, D. M. Vannier, and M. M. C. Lai.** 1989. Identification of a domain required for autoproteolytic cleavage of murine coronavirus gene A polyprotein. J. Virol. **63**:3693–3699.
3. **Baric, R. S., C. K. Shieh, S. A. Stohlman, and M. M. C. Lai.** 1987. Studies into the mechanism of MHV transcription, p. 137–149. *In* M. M. C. Lai and S. Stohlman (ed.), Coronaviruses. Plenum Publishing Corp., New York.
4. **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.
5. **Boursnell, M. F. 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. **Brayton, P. R., M. M. C. Lai, C. C. Patton, and S. A. Stohlman.** 1982. Characterization of two RNA polymerase activities induced by mouse hepatitis virus. J. Virol. **42**:847–853.
7. **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.
8. **Bredenbeek, P. J., C. J. Pachuk, A. F. H. Noten, J. Charite, W. Luytjes, S. R. Weiss, and W. J. M. Spaan.** 1990. The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59: a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. Nucleic Acids Res. **18**:1825–1832.
9. **Brierley, I., M. E. G. Boursnell, M. M. Binns, B. Billimoria, 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.
10. **Budzilowicz, C. J., and S. R. Weiss.** 1987. In vitro synthesis of two polypeptides from a nonstructural gene of coronavirus mouse hepatitis virus, strain A59. Virology **157**:509–515.
11. **Budzilowicz, C. J., S. P. Wilcznski, and S. R. Weiss.** 1985. Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. J. Virol. **53**:834–840.
12. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. Related Areas Mol. Biol. **47**:45–148.
13. **Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **262**:1102–1106.
14. **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.
15. **Engleman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15**:321–353.
16. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Coronavirus genome: prediction of putative functional domains in the nonstructural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res. **17**:4847–4861.
17. **Hopp, T. P., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78**:3824–3828.
18. **Jacks, T., and H. E. Varmus.** 1985. Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting. Science **230**:1237–1242.
19. **Janin, J., S. Wodak, M. Levitt, and B. Maigret.** 1978. Conformation of amino acid side-chains in proteins. J. Mol. Biol. **125**:357–386.
20. **Krieg, P. A., and D. A. Melton.** 1984. Functional messenger RNAs are produced by SP6 in vitro transcription of cloned cDNAs. Nucleic Acids Res. **12**:7057–7070.
21. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
23. **Lai, M. M. C., C. D. Patton, and S. A. Stohlman.** 1982. Further characterization of mRNAs of mouse hepatitis virus: presence of common 5′-end nucleotides. J. Virol. **41**:557–565.
24. **Leibowitz, J. L., J. R. DeVries, and M. Rodriquez.** 1987. Increased hepatotropism of mutants of MHV, strain JHM, selected with monoclonal antibodies, p. 321–331. *In* M. M. C. Lai and S. Stohlman (ed.), Coronaviruses. Plenum Publishing Corp., New York.
25. **Leibowitz, J. L., S. R. Weiss, E. Paavola, and C. W. Bond.** 1982. Cell-free translation of murine coronavirus RNA. J. Virol. **43**:903–913.
26. **Merrick, W. C.** 1983. Translation of exogenous mRNAs in reticulocyte lysates. Methods Enzymol. **101**:606–615.
27. **Pachuk, C. J., P. J. Bredenbeek, P. W. Zoltick, W. J. M. Spaan, and S. R. Weiss.** 1989. Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus, strain A59. Virology **171**:141–148.
28. **Palmenberg, A. C.** 1987. Picornaviral processing: some new ideas. J. Cell. Biochem. **33**:191–198.
29. **Sawicki, D. L., and S. G. Sawicki.** 1986. Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. J. Virol. **57**:328–334.
30. **Sawicki, S. G., and D. L. Sawicki.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J. Virol. **64**:1050–1056.
31. **Sethna, T. B., S. L. Hung, and D. A. Brian.** 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. Proc. Natl. Acad. Sci. USA **86**:5626–5630.
32. **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.
33. **Spaan, W. J. M., D. Cavanagh, and M. C. Horzinek.** 1988. Coronaviruses: structure and genome expression. J. Gen. Virol. **69**:2939–2952.
34. **Spaan, W. J. M., P. Rottier, M. C. Horzinek, and B. A. M. van der Zeijst.** 1981. Isolation and characterization of virus-specific mRNAs in cells infected with mouse hepatitis virus (MHV-A59). Virology **108**:424–434.
35. **Zoltick, P. W., J. L. Leibowitz, J. R. De Vries, G. M. Weinstock, and S. R. Weiss.** 1989. A general method for the induction and screening of antisera for cDNA-encoded polypeptides: antibodies specific for a coronavirus putative polymerase-encoding gene. Gene **85**:413–420.