Analysis of Intracellular Small RNAs of Mouse Hepatitis Virus: Evidence for Discontinuous Transcription

RALPH S. BARIC,[1] CHIEN-KOU SHIEH, STEPHEN A. STOHLMAN, AND MICHAEL M. C. LAI[2]

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

We have previously shown the presence of multiple small leader-containing RNA species in mouse hepatitis virus (MHV)-infected cells. In this paper, we have analyzed the origin, structure, and mechanism of synthesis of these small RNAs. Using cDNA probes specific for leader RNA and genes A, D, and F, we demonstrate that subsets of these small RNAs were derived from the various viral genes. These subsets have discrete and reproducible sizes, varying with the gene from which they are derived. The size of each subset correlates with regions of secondary structure, whose free energy ranges from -1.6 to -77.1 kcal/mol, in each of the mRNAs examined. In addition, identical subsets were detected on the replicative intermediate (RI) RNA, suggesting that they represent functional transcriptional intermediates. The biological significance of these small RNAs is further supported by the detection of leader-containing RNAs of 47, 50, and 57 nucleotides in length, which correspond to the crossover sites in two MHV recombinant viruses. These data, coupled with the high frequency of RNA recombination during MHV infection, suggest that the viral polymerase may pause in or around regions of secondary structure, thereby generating pools of free leader-containing RNA intermediates which can reassociate with the template, acting as primers for the synthesis of full-length or recombinant RNAs. These data suggest that MHV transcription uses a discontinuous and nonprocessive mechanism in which RNA polymerase allows the partial RNA products to be dissociated from the template temporarily during the process of transcription. © 1987 Academic Press, Inc.

INTRODUCTION

Mouse hepatitis virus (MHV), a member of Coronaviridae, contains a linear single-stranded and positive-sense RNA of 5.4×10^6 Da (Lai and Stohlman, 1978). The genomic RNA is enclosed in a helical nucleocapsid structure constructed from multiple copies of nucleocapsid protein (N) (Sturman, 1977). Virions are enveloped and contain two virus-specific glycoproteins of 180 kDa (E2) and 23 kDa (E1) (Sturman et al., 1980).

Upon entry into the host cell, the genomic RNA is translated into an early polymerase which directs the synthesis of a full-length negative-sense viral RNA (Brayton et al., 1982; Lai et al., 1982). In turn, the negative-sense RNA is transcribed by a late or altered early polymerase into the genomic RNA and six subgenomic mRNAs (Brayton et al., 1984). These mRNAs range from 0.6 to 5.4×10^6 Da in length and are arranged in the form of a nested set from the 3'-end of the genome (Lai et al., 1981). The 5'-ends of each mRNA and the genomic RNA appear to contain an identical leader sequence of about 72 nucleotides, which are encoded only at the 5'-end of the genome (Lai et al., 1983, 1984;

Spaan et al., 1983). The uv transcriptional mapping studies suggest that the mRNAs are not derived from cleavage of a large precursor RNA (Jacobs et al., 1981), nor is there any evidence indicating that nuclear factors are required for MHV replication (Brayton et al., 1981; Wilhelmsen et al., 1981). These data suggest that conventional eukaryotic RNA splicing is not involved in MHV transcription. Analysis of MHV replicative intermediate and replicative form RNAs suggests that a free leader RNA may be involved in subgenomic mRNA synthesis (Baric et al., 1983). In addition, discrete small leader-related RNAs of various sizes have been detected in MHV-infected cells (Baric et al., 1985). We have also isolated a ts mutant of MHV which synthesizes only small leader RNAs, but not mRNAs at the nonpermissive temperature (Baric et al., 1985). These data suggest that the MHV mRNAs are synthesized discontinuously and may utilize a single or multiple free leader RNAs as primers for the transcription of different subgenomic mRNAs. Additional proof that free *trans*-acting leader RNA(s) function in MHV transcription came from the demonstration of reassortment of leader RNAs between two different strains of MHV (Makino et al., 1986b). However, the exact leader RNA species which function in MHV transcription has not been identified.

RNA recombination occurs at very high frequencies during mixed infection with two heterologous strains

---

[1] Present address: University of North Carolina, Department of Parasitology and Laboratory Practice, 348 Rosenau Hall 201H, Chapel Hill, NC 27514.

[2] To whom requests for reprints should be addressed.

of MHV (Lai et al., 1985; Makino et al., 1986a). These data, coupled with the presence of discrete large leader-containing RNAs which range from 84 to 1000 nucleotides in length in MHV-infected cells (Baric et al., 1985), suggest that discontinuous RNA intermediates may be dissociated and reassort between viral RNA templates to generate recombinant viruses by a copy-choice mechanism (Makino et al., 1986a). Therefore, the larger leader-containing RNAs present in MHV-infected cells may represent functional intermediates of RNA transcription and recombination. In this paper, we have analyzed the origin, structure, and probable mechanism of synthesis of these RNAs. The data suggest that MHV RNA transcription may pause at sites corresponding to hairpin loops in the RNA template, and they support a mechanism of discontinuous RNA transcription in which RNA intermediates can be dissociated and reassociated with the RNA template intermittently during the course of transcription.

MATERIALS AND METHODS

Virus and cells

The A59 strain of MHV was propagated in either DBT or L2 cells at 37° in Dulbecco's modified MEM medium supplemented with 10% fetal calf serum containing 100 μg/ml penicillin and 50 μg/ml streptomycin. Infection was performed at a m.o.i. of 1 to 5, as previously described (Baric et al., 1983).

Preparation of MHV intracellular RNA

RNA was extracted from infected cells by the phenol/chloroform method between 5 and 7 hr postinfection (70% CPE) (Baric et al., 1983). Following ethanol precipitation, the RNA was washed once in 70% ethanol to remove excess salts and analyzed by electrophoresis on polyacrylamide gels as previously described (Baric et al., 1985).

Isolation of MHV replicative intermediate (RI) RNA and free RNAs

Intracellular RNA was extracted from infected cells as described above and applied to 15–30% sucrose gradients made in NTE buffer (0.1 M NaCl, 0.01 M Tris–HCl, pH 7.4, and 0.001 M EDTA) containing 0.1% SDS. RNA was sedimented at 45,000 rpm in an SW55Ti rotor for 2 hr at 18° and 0.2-ml fractions were collected. In duplicate gradients, 28, 18, and 4 S RNA were included as size markers. To isolate MHV RI RNAs, the 18–40 S RNA fractions were pooled and precipitated with ethanol. It has previously been shown that the RI RNAs are contained within this size fraction (Baric et al., 1983, 1985; Sawicki and Sawicki, 1986). The 4–

10 S RNA fractions were collected as the free small RNAs.

Nick translations of MHV cDNA clones and filter hybridization

The cDNA clones of MHV genomic RNA used in this study are summarized in Fig. 1. Clones F82 and C96 represent overlapping regions in the 5'-end of the genomic RNA (Shieh et al., 1987). F82 is a 3.8-kb clone representing the very 5'-end of the genome including the leader RNA sequences, while clone C96 represents internal sequences in gene A. Two other clones representing the internal sequences of MHV genome were also used: Clone D63 contains entire sequences of genes D and E and part of genes C and F (Shieh, unpublished), and clone PHN42 spans the entire gene F (kindly provided by Dr. Heiner Niemann, Giessen, Germany). An internal XbaI fragment (0.4 kb), designated D63X, of clone D63, was used as the probe for gene D. This fragment includes nucleotides 47–375 in the mRNA 4 coding sequence as reported (Skinner and Siddell, 1985). The construction of subclones derived from PHN 42 followed the procedures described by Maniatis et al. (1982). Subclone E1-500 was derived by blunt-end ligating a 509 nucleotide DdeI fragment, located between nucleotides 90 and 609 in the gene F sequence (Armstrong et al., 1984), into plasmid pT7-2 at the SmaI site (pT7-2 was purchased from Amersham). Subclone E1-300 was derived from a 246-nucleotide fragment, located between nucleotides 534 and 780 in the E1 RNA (Armstrong et al., 1984), and was inserted into plasmid pT7-2 at the HindIII/SmaI site. E1-L was excised from PHN 42 by an EcoRI/ScaI restriction digestion. It contains the first 84 nucleotides of the mRNA 6, including the 72-nucleotide leader RNA sequence. This fragment was used directly for nick translation without subcloning. Nick translations of cDNA clones (0.5–1.0 μg) were performed in 50 mM Tris–HCl (pH 7.8), 5 mM MgCl$_2$, 5 mM DTT, 50 μg/ml bovine serum albumin (BSA), 10 mM each of dATP, dGTP, and TTP, 100 μCi [α-^{32}P]dCTP (3200 Ci/mmol), 5% glycerol, 1 μg/ml DNase I (Worthington), and 2 U of DNA polymerase I (Boehringer Mannheim Biochem) for 2 hr at 15°. Unincorporated nucleotides were removed by several precipitations with 80% ethanol. Hybridizations were performed at 42° for 36 hr in 50% formamide, 10 mM Na phosphate (pH 6.5), 5× SSC (1× SSC = 0.15 M NaCl, 0.015 M Na citrate, pH 7.0), 10× Denhardt's solution (1× Denhardt's solution: 0.02% each of Ficoll, polyvinylpyrrolidone, and BSA) and 125 μg/ml salmon sperm DNA (Baric et al., 1985). Approximately 40–60 × 10^6 cpm/μg of DNA were hybridized to each blot. After hybridization, the filters were

washed, dried, and exposed to XAR films at −80° in the presence of an intensifying screen.

Polyacrylamide gel electrophoresis and electroblotting

Equivalent amounts of intracellular RNA (25 or 50 μg) were separated by electrophoresis on 8% polyacrylamide gels in 0.1 M Tris–borate buffer (pH 8.4) containing 1 mM EDTA and 6 M urea. For analysis of small RNAs ranging between 30 and 150 nucleotides in length, the samples were electrophoresed at 1000 V (25–30 mA) until the bromphenol blue dye marker had migrated approximately 20 cm from the origin. To analyze larger RNAs ranging between 100 and 300 nucleotides in length, electrophoresis was at 1000 V until the xylene cyanol dye marker had migrated 35–40 cm from the origin. Following electrophoresis, urea was removed by twice washing the gel in ice-cold TAE buffer (0.04 M Tris–HCl, pH 7.4, 0.02 M Na acetate, and 0.001 M EDTA) for 15 min each, and the RNA was electroblotted to Zeta-probe paper (Bio-Rad) in cold TAE. Electroblotting was performed at 4° in a circulating chamber at 390 mA for 20–24 hr and an additional 2 hr at 1 A to ensure the transfer of larger RNAs above 500 nucleotides in length (Baric et al., 1985; Stellwag and Dahlberg, 1980). Following electrotransfer, the paper was gently washed in TAE and baked at 80° for 2 hr. Prehybridization was performed for 36 hr at 42° in 50% formamide, 10 mM Na phosphate (pH 6.5), 5× SSC, 10× Denhardt's solution and 125 μg/ml salmon sperm DNA.

Precise sizes of the small RNAs were determined by electrophoresis of M13 mp 7 sequencing T-ladders in lanes adjacent to intracellular RNA preparations. Since electrotransfer oftentimes decreases the resolution of individual bands, complete sets of M13 sequencing A, T, C, G ladders were also electrophoresed with each set of RNA preparations, fixed, dried, and separately exposed to XAR-5 film. The M13-T ladders electrotransferred with RNA preparations were aligned with the A, T, C, G ladders to determine the exact sizes of small RNAs.

Computing stability of MHV RNA secondary structure

Predictions for regions of secondary structure in MHV RNA were made using the Zucker RNA folding program through Bionet (Zucker and Stiegler, 1981). Stability of individual hairpin loops within the MHV sequence was calculated from the thermodynamics of adding a base pair to a double-stranded helix at 25° as reported by Tinoco et al. (1973) and modified by Salser (1977). More specifically, the energies for stacking of internal GU

pairs were computed as follows: GU next to GC, −1.3 kcal/mol; GU next to AU or GU, −0.3 kcal/mol. In addition, −1.0 kcal/mol was subtracted from loop structures containing U nucleotides (Tinoco et al., 1973).

RESULTS

Construction of specific cDNA probes

To determine the structure and origin of the small RNAs present in MHV-infected cells, we constructed several cDNA probes specific for different regions of the MHV genome. We chose three genes for this study, namely, genes A, D, and F. Gene A probably encodes an RNA polymerase, while genes D and F encode a nonstructural protein p14 and the E1 protein, respectively. Since gene A represents the 5′-most region of the genome, the small RNAs originating from this gene represent products of RNA replication. The other two genes represent two internal genes which have been well characterized (Armstrong et al., 1984; Skinner and Siddell, 1985). The map positions of the probes used in this study are depicted in Fig. 1.

The probe specific for leader region was derived from a gene F-specific clone (pHN42) which contains the entire leader sequence. The fragment (E1-L) containing the first 84 nucleotides of MHV mRNA6 was used as the leader-specific probe in subsequent experiments (see Fig. 1b). The clones E1-500 and E1-300 were used to serve as the probes for the 5′- and 3′-ends of the gene F, respectively (Fig. 1b). The E1-300 clone would not detect any leader-containing RNAs smaller than 534 nucleotides.

Clones specific for genes A and D were derived from cDNA clones as described under Materials and Methods.

Analysis of small RNAs originating within the leader sequence

To determine the structure and origin of the small RNAs present in MHV-infected cells, we first used a leader-specific cDNA probe to identify small RNAs terminated within or around the leader RNA sequence. These small RNAs should represent RNA species derived from the 5′-end of genomic or subgenomic mRNAs. These RNA species have been detected previously (Baric et al., 1985), but their size and structure have not been precisely determined. Intracellular RNA was extracted from MHV-infected cells and separated by electrophoresis on polyacrylamide gels. Following transfer to Zeta-probe paper, the blot was probed with a leader-specific cDNA probe (E1-L) (Fig. 1b). As shown in Fig. 2, several distinct small RNAs ranging from 47 to 84 nucleotides were detected, similar to our previous
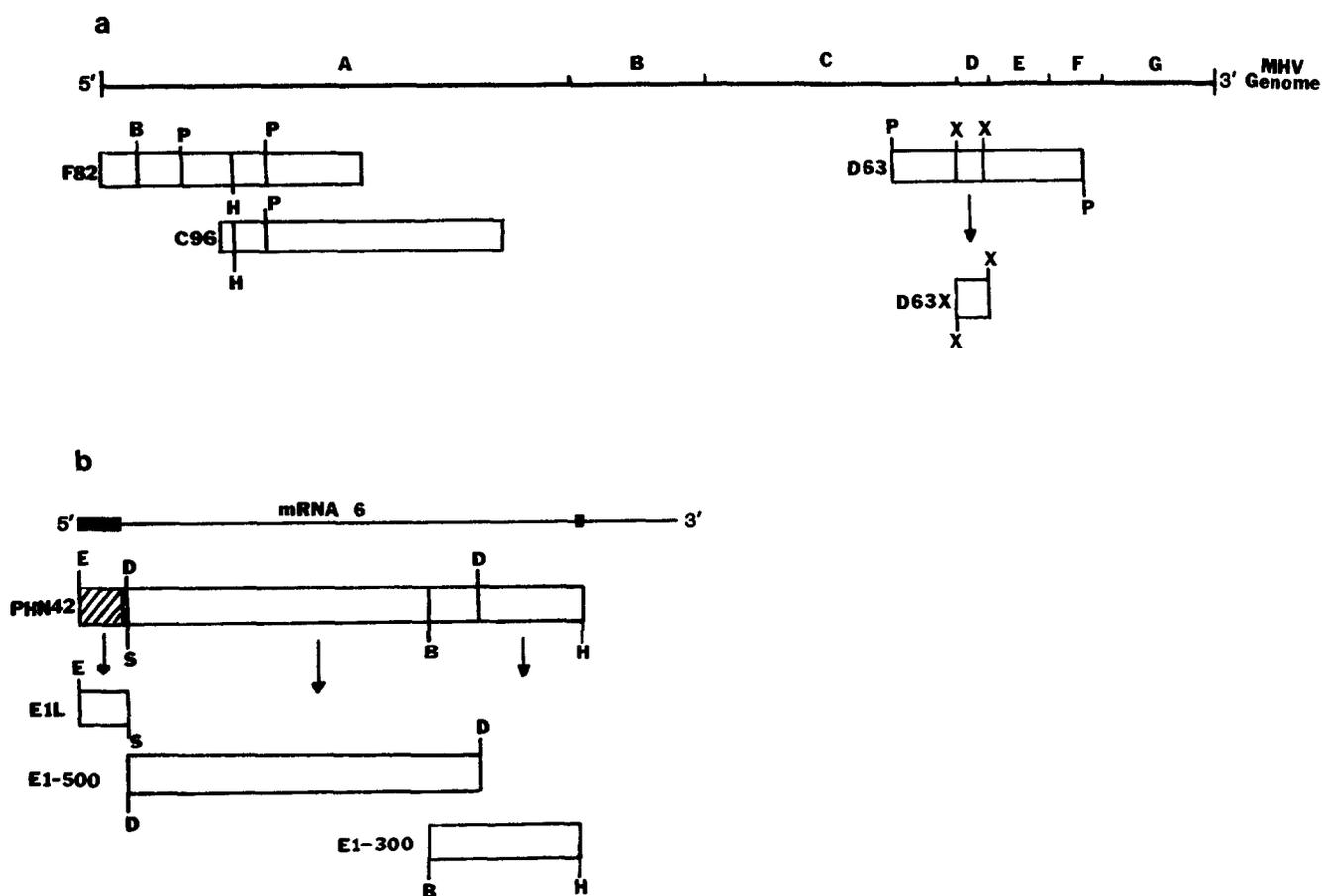
FIG. 1. The structure of the MHV cDNA probes. Clone F82 contains the complete 5'-end of the genome and overlaps with clone C96. These clones were used to detect small RNAs originating during RNA replication (a). Clones D63X, E1-500, and E1-300 were used to detect small RNAs originating during transcription of mRNA 4 (gene D) and mRNA 6 (gene F) (b). Probe E1-L was used to detect small leader RNAs containing the leader sequence. The exact sizes of these clones are described under Materials and Methods. Abbreviations used for cDNA fragments: B, BamHI; D, DdeI; E, EcoRI; H, HindIII; P, PstI; S, ScaI; X, XbaI.

report (Baric et al., 1985). The sizes of these RNA species were precisely determined from the accompanying T-ladder of M13 sequences. Their sizes were reproducible, as evident from the RNA patterns of two preparations derived from different cell lines, although the relative amounts of the various RNA species varied. The four smallest RNA species of 47, 50, 57, and 65 nucleotides correspond in size to the lengths between the 5'-end of the leader RNA and a region of two weak overlapping hairpin loops around nucleotides 42–56 and 52–72 (Fig. 3; Table 1) (Shieh et al., 1987). Thus, these RNA species are likely to represent RNA transcriptional products terminating within the leader sequences. The larger leader-containing RNAs of 74 and 77 nucleotides, and occasionally an RNA of 84 nucleotides, also correspond to the region of a hairpin loop at nucleotides 80–143 from the 5'-end of the genome and a postulated transcriptional termination signal (UUUAUAAA) for the MHV leader RNA synthesis (Shieh et al., 1987) (Fig. 3). These are the major small leader-containing RNA species.

Larger RNA species were also reproducibly detected by the leader-specific probe, consistent with previously published results (Baric et al., 1985); however, under the conditions used, they were not well resolved. These leader-containing RNAs range from 110 to more than 1000 nucleotides in length. To determine the structure of these larger leader-containing RNA species, the RNA blots were hybridized with cDNA clone F82, representing 3.8 kb of the sequences at the 5'-end of the genomic RNA (Shieh et al., 1987) (Fig. 1a). Several distinct RNA species ranging from 136 to larger than 281 nucleotides were resolved (Fig. 4A). It is notable that the two independently isolated preparations of intracellular RNA gave identical RNA bands. When the clone C96, representing the internal region of gene A (approximately 3.8 kb from the 5'-end) (Fig. 1a), was used, no distinct small RNA bands were detected (Fig. 4B). The hybridization with C96 was generally weaker than that with F82, since the C96 probe would not detect any RNA species smaller than 3800 nucleotides (Shieh et al., 1987), which were detectable with F82. These
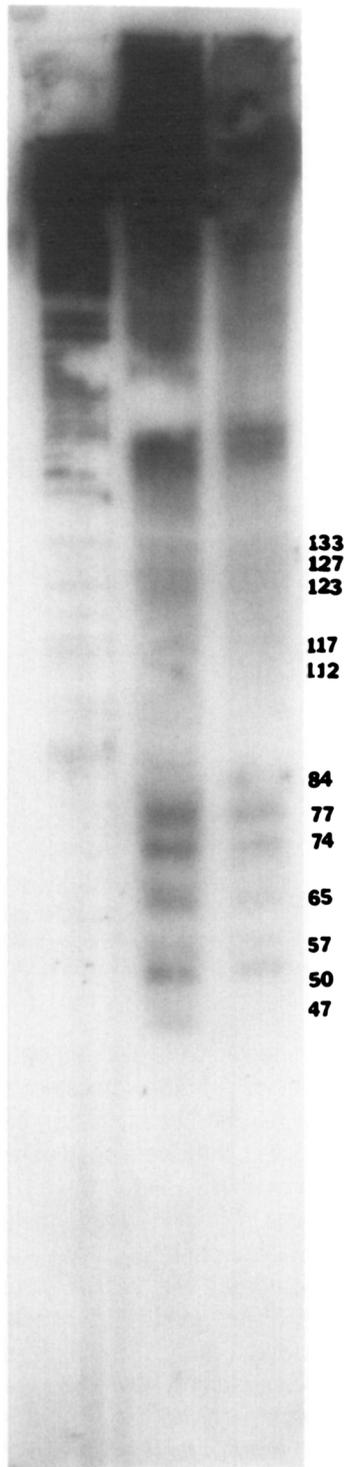
FIG. 2. Analysis of MHV leader-containing RNAs. Intracellular RNA extracted from MHV-infected cells was separated by electrophoresis on 8% polyacrylamide gels containing urea and was transferred to Zeta-probe paper. The nick-translated leader-specific probe (E1-L) was hybridized to the blot. Lane 1, M13 sequencing T ladder. Lane 2, MHV-infected L2 cells. Lane 3, MHV-infected DBT cells. The sizes of T bands were determined from the accompanying A, G, C, T ladders of M13 sequencing (data not shown).
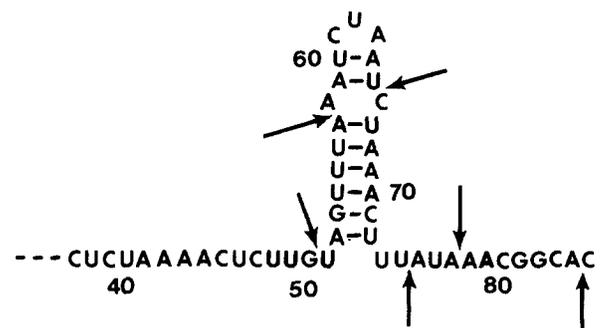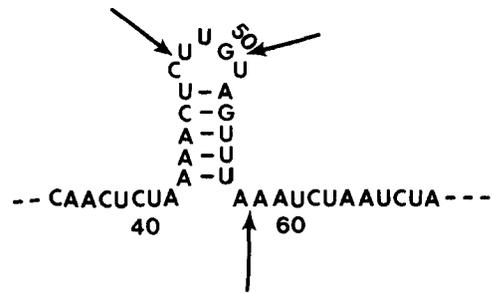
FIG. 3. The predicted secondary structures of the MHV leader RNA. The sequence of the 5'-end leader region of A59 genomic RNA is derived from the 5'-end sequence of the JHM strain (Shieh et al., 1987). In contrast to JHM, only two UCUAA repeats (nucleotides 60–69) are present in A59 (Spaan et al., 1983; Lai et al., 1984). The two potential overlapping hairpin loops are shown separately. Arrows indicate the termination sites of the small RNAs detected in A59-infected cells (Fig. 2). The free energy of these hairpin loops is shown in Table 1.

data suggest that these RNAs represent the transcriptional intermediates derived from the 5'-end of the genomic RNA, and that transcription from the 5'-end of the genome terminates or pauses at distinct sites. The length of these RNAs correspond to the lengths between the 5'-end and the sites of potential hairpins (Table 1; see Discussion).

Analysis of small RNAs originating from internal genes

The leader-containing RNAs larger than 110 nucleotides in length detected by the leader-specific cDNA probe were more heterogeneous (Fig. 2) (Baric et al., 1985), suggesting that multiple RNA species in this size range were present. However, the probe representing the 5'-end sequences of gene A detected discrete species (Fig. 4A), suggesting that the larger leader-containing RNAs may also represent RNA species derived from other regions of the genome during subgenomic RNA synthesis. Therefore, cDNA probes

TABLE 1

STABILITY OF MHV HAIRPIN LOOP STRUCTURES PRESENT
WITHIN THE MHV RNA

Origin	Hairpin loop locations[a]	Free energy[b] (kcal/mol)	Size of RNA intermediates detected[c]
Leader	42–56	−0.8	47, 50, 57
Leader	52–72	−1.6	47, 50, 57, 65, 74, 77, 84
mRNA1	80–143	−18.7	136, 138, 141
mRNA1	165–188	−6.9	188, 191
mRNA1	148–320	−77.1	162, 188, 191, 241, 260, 271, 281
mRNA4	106–131	−13.0	117, 135
mRNA4	105–162	−34.8	117, 135, 160, 168, 173
mRNA4	168–204	−12.3	168, 173, 199, 205
mRNA4	211–235	−6.6	218
mRNA6	105–122	−8.2	118, 123
mRNA6	114–141	−16.1	127, 143/145
mRNA6	123–167	−26.4	127, 143/145, 150, 158, 164, 169
mRNA6	193–226	−6.2	None
mRNA6	242–285	−12.0	241–260, 270–281

[a] Locations are expressed in numbers of nucleotides from the 5'-end of respective mRNAs.

[b] Free energies for each region of secondary structure were calculated as described under Materials and Methods. The free energy for the first 280 nucleotides of the mRNA 6 (gene F) was calculated at −90.1 kcal/mol, while a free energy of −65.4 kcal/mol was predicted from the first 200 nucleotides of the sequence of mRNA4 (gene D). Free energy of the first 250 nucleotides of the 5'-end of the genome was estimated at −65.4 kcal, while free energy of the first 90 nucleotides of the leader sequence was −17.4 kcal. The sequence of gene A was obtained from the sequence of JHM strain (Shieh et al., 1987). The sequences of mRNA4 and 6 were as published (Armstrong et al., 1984; Skinner and Siddell, 1985). Some regions are listed more than once because of several possible secondary structures.

[c] Expressed in number of nucleotides. The data were derived from Figs. 2, 4, and 5. In the case of overlapping secondary structures, small RNAs of appropriate size were placed in each region of hairpin loop structure, since it could not be determined which loop structure induced pausing at this site.

specific for the 5'-coding sequences of genes D and F were tested. The probe representing the 5'-end of the gene F (E1-500) detected RNA species of 118, 123, 127, 143/145, 150, 158, 164, 169, and two groups of poorly resolved RNAs ranging between 241–260 and 270–281 nucleotides in length (Fig. 5A). These RNA species were different from those detected with the gene A probe (Fig. 4A). In addition, they were not detected with a probe (E1-300) representing the 3'-end of the gene between nucleotides 534 and 780 (Fig. 5C). The gene D-specific cDNA probe (D63X) (Fig. 1a) also detected a different set of small RNA species of 117, 135, 160, 168, 205, and 218 nucleotides in length (Fig. 5B). Smaller amounts of RNA of 173 and 199 nucleotides in length were also detected. These data

suggest that transcription of subgenomic mRNAs also terminates or pauses at many sites.

Computer models of RNA folding suggest the presence of hairpin loops in regions of nucleotides 105–122, 114–141, 123–167, 193–226, and 242–285 within the sequence of mRNA6 (gene F) (Fig. 6b) and in the regions of nucleotides 106–131, 105–162, and 168–204 and 211–235 within mRNA4 (gene D) (Fig. 6a). The predicted stability of these hairpin structures is shown in Table 1. These data show that most of the small RNAs detected in MHV-infected cells have termination sites within the hairpin loop structures, suggesting a possible correlation between the generation of small RNAs and the presence of secondary structure in the template or product RNAs.

Presence of small RNA species on the replicative intermediate (RI) RNA

If the small RNA species detected in MHV-infected cells represent specific products released from the template RNA during transcriptional pausing, rather than degradation products of the mRNAs, then the small RNA species would be expected to be present in the RI structure which is involved in RNA transcription. Conversely, if RNA transcription proceeds in a continuous manner, the RNA species on the RI structure should be heterogeneous, without predominant RNA species. To distinguish between these two possibilities, 18–40 S RI RNA (Baric et al., 1983; Sawicki and Sawicki, 1986) was isolated and analyzed for the presence of small leader-containing RNA using two probes, F82 and E1-500, representing genes A and F, respectively. As shown in Fig. 7, the small RNA species present on the RI RNA are the same as those present in the whole cell lysates. Furthermore, the 4 S RNA fractions also contain an identical set of small RNAs, indicating that some of the small RNAs were dissociated from the RI RNA. These results suggest that the small RNA species are true transcriptional intermediates.

DISCUSSION

Previous findings in our laboratory demonstrate that RNA recombination could occur at a very high frequency during a mixed infection with two strains of MHV (Makino et al., 1986a). This high frequency is reminiscent of RNA reassortment described for viruses with segmented RNAs (Fields, 1981) and suggests that MHV replication might involve the generation of free and discontinuous RNA intermediates. These intermediates could participate in RNA recombination by a copy-choice mechanism. The detection of multiple discrete species of leader-containing RNAs, which
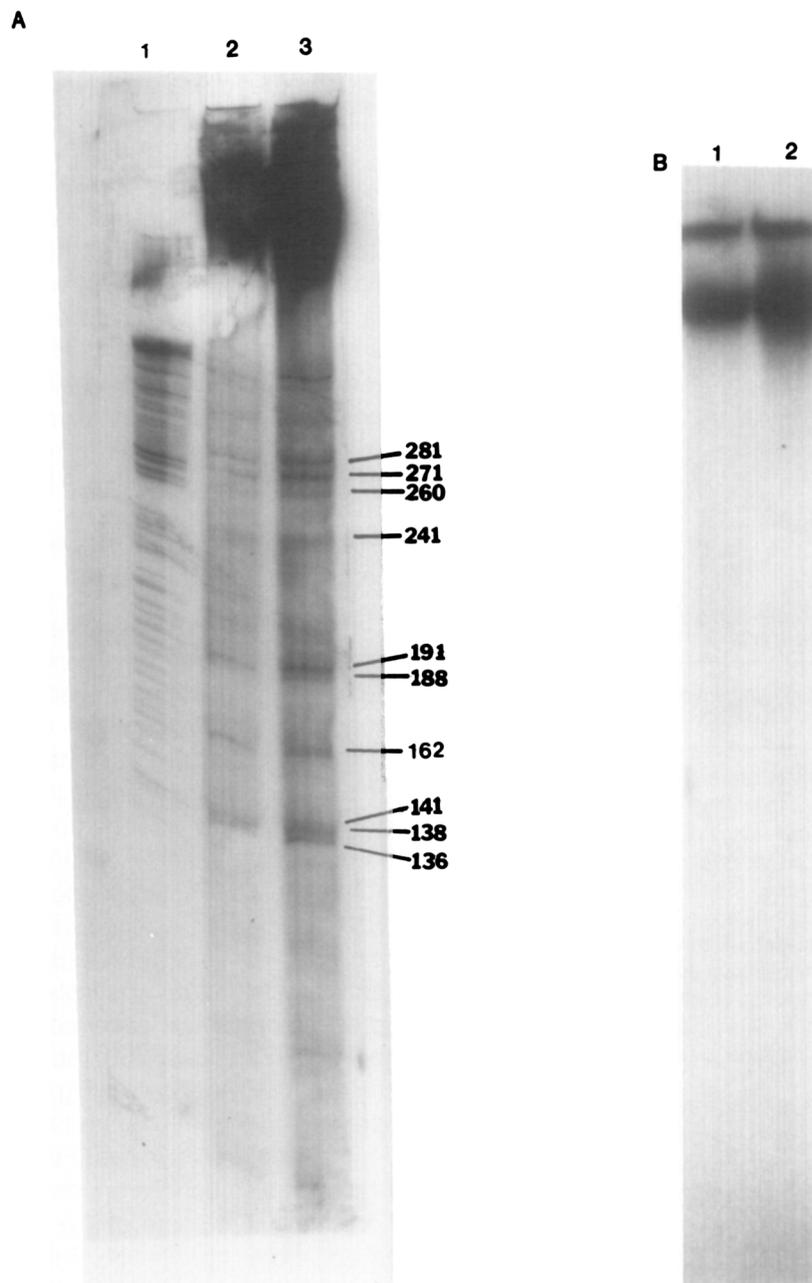
FIG. 4. Analysis of MHV small RNAs originating at the 5'-end of gene A. Intracellular RNA was separated by electrophoresis under conditions which separate RNAs between 100 and 300 nucleotides in length (Materials and Methods). Following transfer to Zeta-probe paper, RNAs were probed with the nick-translated clones F82 (A) or C96 (B). (A) Lane 1, M13 sequencing T-ladder; lane 2, A59-infected L2 cells; lane 3, A59-infected DBT cells. (B) Lane 1, A59-infected L2 cells; lane 2, A59-infected DBT cells.

range from 47 to more than 1000 nucleotides in length, in MHV-infected cells further supports this model (Baric et al., 1985). In this report, we examined the origin, structure, and mechanism of synthesis of these RNAs. The data suggest that these RNA species are likely to represent transcriptional pausing products of MHV RNA synthesis. The pausing sites correspond, in general, to the potential hairpin loops present in the template or product RNAs. This transcriptional pausing mech-

anism is reminiscent of the transcriptional pausing of Qβ phage RNA, in which RNA polymerase pauses in regions of hairpin loop structures (Mills et al., 1978). Pausing also occurs during DNA-directed DNA synthesis (Huang and Hearst, 1981; Sherman and Gefter, 1976), DNA-directed RNA synthesis (Maizels, 1973; Rosenberg et al., 1978) and in reverse transcriptase reactions (Efstratiadis et al., 1975; Haseltine et al., 1976). Similar to the mechanism of Qβ RNA transcrip-
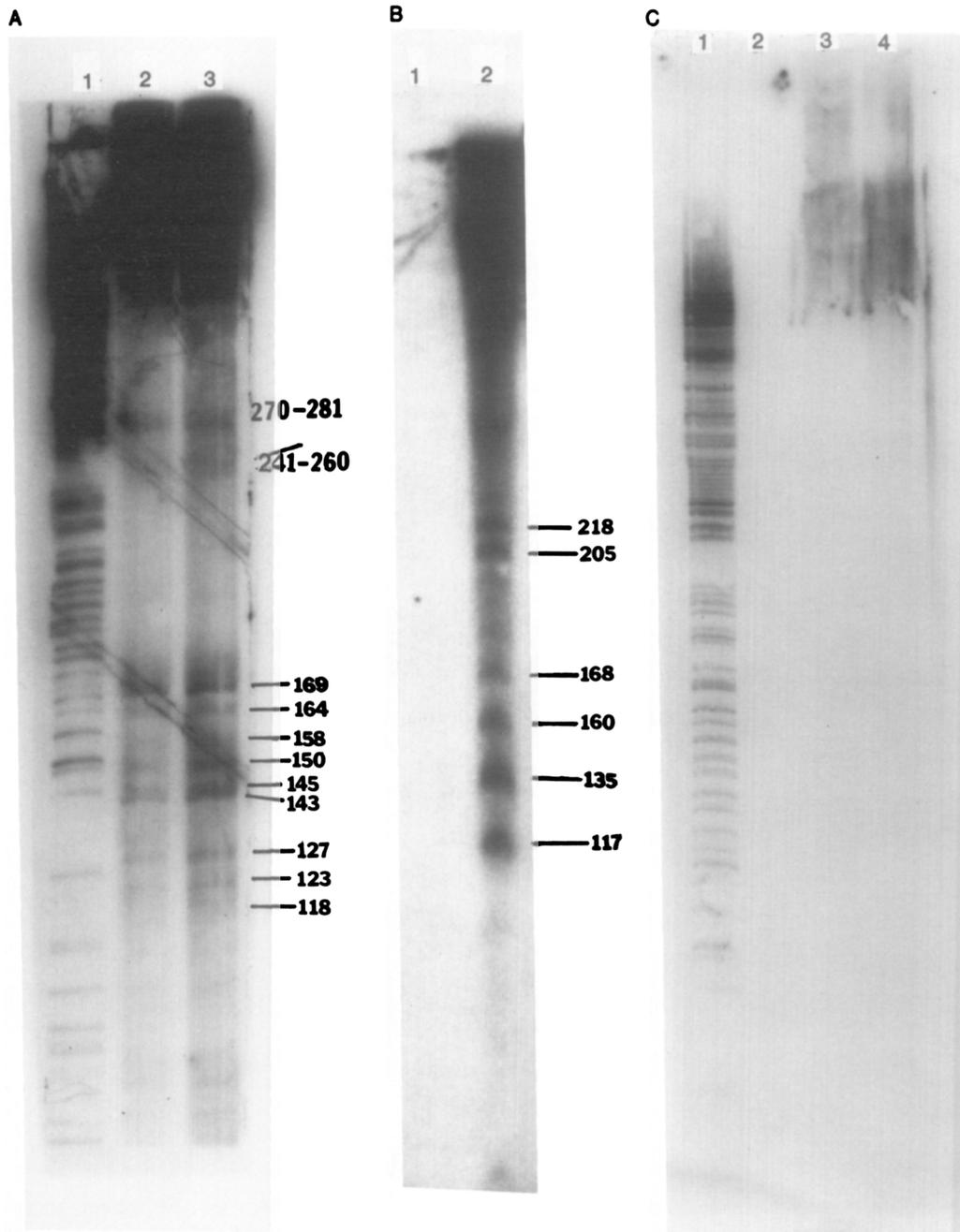
FIG. 5. Analysis of small RNAs originating within genes D and F. Intracellular RNA was separated and transferred as described in Fig. 4. Nick-translated probes specific for gene F (A, C) or gene D (B) were used. Similar amounts of RNA were used for each blot. (A) RNAs detected with E1-500. Lane 1, M13 T-ladder; lane 2, A59-infected L2 cells; lane 3, A59-infected DBT cells. (B) RNAs detected with D63X. Lane 1, uninfected DBT cells; lane 2, A59-infected DBT cells. (C) RNAs detected with E1-300. Lane 1, M13 T-ladder; lane 2, uninfected DBT cells; lane 3, A59-infected DBT cells; lane 4, A59-infected L-2 cells.

tion, the majority of MHV small RNAs terminate at the 3′-side of hairpin loop structures in the product or template strands, which have free energies ranging from −1.6 to −77.1 kcal/mol (Fig. 6 and Table 1). The stabilities of these hairpin loops are comparable to those of Qβ (−9.8 to −23.8 kcal) and tRNA (−4.5 to −15.0 kcal) hairpin loops, whose existence in the RNA has been demonstrated by various physical and biochem-

ical methods (Auron et al., 1982; Gehrke et al., 1983; Kramer and Mills, 1981). Whether the pausing RNA intermediates of Qβ or other pausing transcriptional model systems are dissociated from the RNA template is not known; however, the data presented in this paper clearly show that at least some of the MHV RNA intermediates are separated from the template RNA strand. It has previously been shown that the nascent
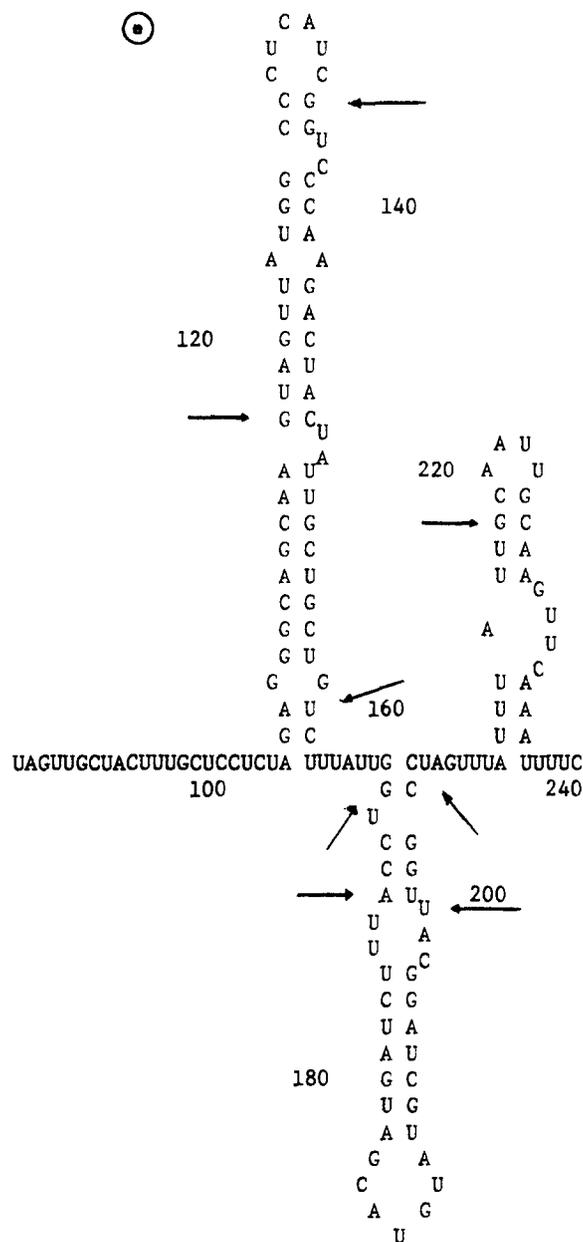
```
        ⊙            C  A
                     U    U
                     C    C
                     C  G  ◄────
                     C  G U
                     G  C C
                     G  C      140
                     U  A
                     A    A
                     U  G
                     U  A
              120    G  C
                     A  U
                     U  A
        ────►        G  C U
                     A  U A          A  U
                     A  U     220  A    U
                     A  U          C  G
                     C  G    ────► G  C
                     G  C          U  A
                     A  U          U  A G
                     C  G                 U
                     G  C          A      U
                     G  U
                     G    G        U  A C
                     A  U  ╲ 160   U  A
                     G  C          U  A
UAGUUGCUACUUUGCUCCUCUA  UUUAUUG  CUAGUUUA  UUUUC
        100               G  C              240
                        ╱ U
                       ╱  C  G
                          C  G
        ────►          A  U U ◄──── 200
                       U     A
                       U     A
                       U  G C
                       C  G
                       U  A
                       A  U
              180      G  C
                       U  G
                       A  U
                       G    A
                     C        U
                       A    G
                          U
```

FIG. 6. Predicted regions of secondary structure at the 5′-ends of mRNA 4 (gene D) and mRNA 6 (gene F). The sequences of mRNA 4 and mRNA 6 were derived from Skinner and Siddell (1985) and Armstrong et al. (1984), respectively. Arrows mark the possible termination points of the small RNAs detected in MHV-infected cells. The free energies of the hairpin loops are shown in Table 1. In cases of overlapping secondary structures, only the more stable structure is shown. (a) mRNA 4; (b) mRNA 6.

RNA chains of Qβ and MS2 phages are probably bound to the template strand only by short duplex regions (Weissman, 1974). It is also possible that the MHV RNA polymerase is a nonprocessive enzyme, thus releasing some of the RNA intermediates after pausing.

Currently, there is no direct evidence to prove that these free small RNAs actually rebind to the RNA template and participate in continuing transcription; however, the high frequency of RNA recombination (Makino et al., 1986a) suggests that such rejoining does occur.

The recent isolation of several MHV recombinants with multiple crossovers further strengthened this probability (Keck et al., 1987). Most interestingly, the crossover sites in the two recently isolated recombinants, A-1 and A-5, have been mapped within a region of the leader sequences (nucleotides 35–60 from the 5′-end) (Keck et al., 1987), where three small RNA species were detected (Fig. 2). The isolation of these recombinants strongly suggests the functional roles of the pausing RNA intermediates. It is noteworthy that, during the
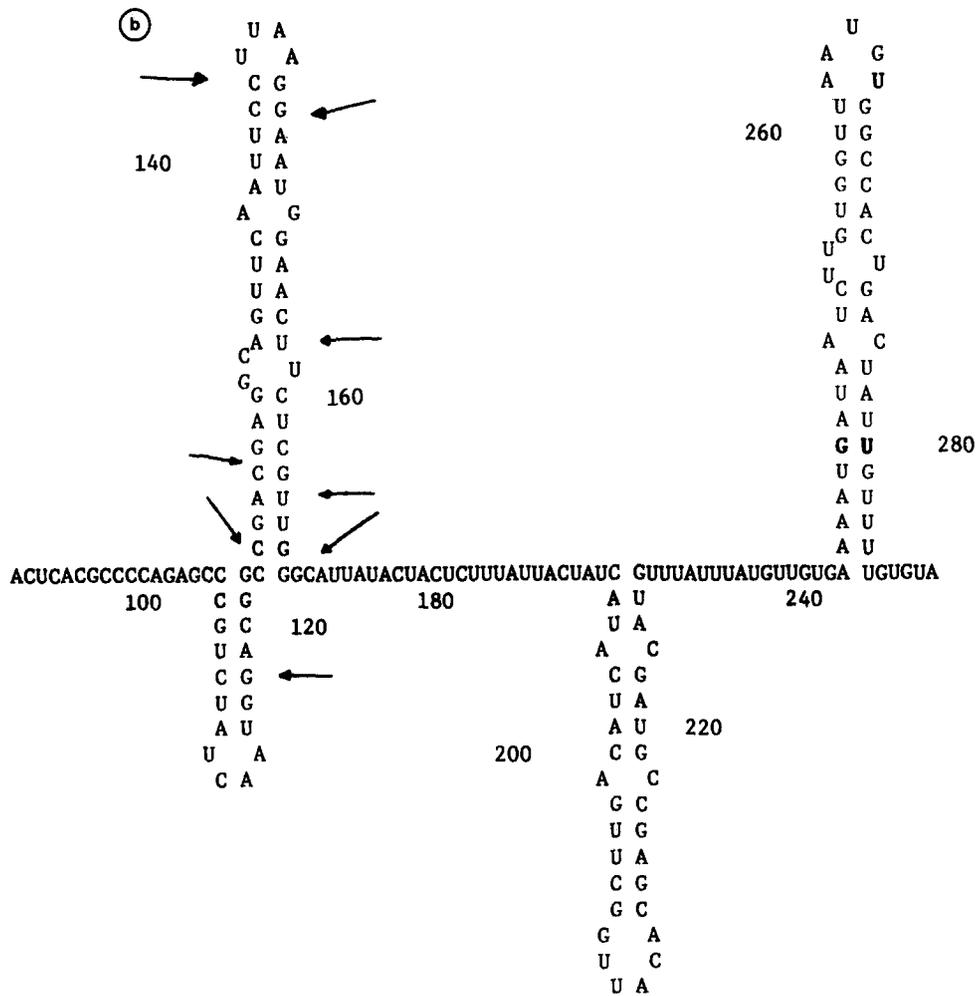
(b)

```
            U  A                                           U
           U    A                                        A     G
      →    C  G                                          A     U
           C  G   ←                                      U  G
           U  A                               260        U  G
  140      U  A                                          G  C
           A  U                                          G  C
           A    G                                        U  A
           C  G                                        U G  C
           U  A                                       U      U
           U  A                                        U C  G
           G  C                                         U  A
          A U      ←                                   A    C
         C   U                                          A  U
        G                                               U  A
          G  C   160                                    A  U
           A  U                                         G  U    280
       →   G  C                                         U  G
           C  G                                         A  U
          A  U    ←                                     A  U
        \  G  U        /                                A  U
           C  G   ↙
ACUCACGCCCCAGAGCC GC GGCAUUAUACUACUCUUUAUUACUAUC GUUUAUUUAUGUUGUGA UGUGUA
     100      C  G        180              A  U        240
             G  C   120                    U  A
             U  A                          A    C
             C  G   ←                      C  G
             U  G                          U  A
             A  U                          A  U   220
          U     A             200          C  G
          C  A                           A    C
                                           G  C
                                           U  G
                                           U  A
                                           C  G
                                           G  C
                                         G     A
                                         U     C
                                           U  A
```

FIG. 6—*Continued.*

transduction of the proto-oncogene c-*fps* by a retrovirus, right-handed recombination occurred near or in stable hairpin loop structures (Huang *et al.*, 1986). These data, coupled with the fact that reverse transcriptase "pauses" in regions of secondary structure (Haseltine *et al.*, 1976), suggest a common mechanism of recombination for retroviruses and coronaviruses, involving paused transcriptional intermediates. In addition, it has previously been shown that short oligonucleotides can act as primers for RNA synthesis *in vitro* (Minkley and Pribnow, 1973; Niyogi and Stevens, 1965). Thus, it seems likely that the free RNAs in MHV-infected cells could also be used for chain elongation. Recently, our laboratory has shown that the MHV leader RNA can act in *trans* to participate in RNA transcription (Makino *et al.*, 1986b).

The four leader-containing RNA species of 65–84 nucleotides in length are the most likely candidates for the primers of subgenomic mRNA transcription, since they fall within the size range of the leader sequences

present in the MHV subgenomic mRNAs (Lai *et al.*, 1984; Spaan *et al.*, 1983). It is not clear, however, whether multiple leader RNAs are involved in priming mRNA transcription. According to our present model of leader-primed transcription, these leader RNAs would have to be cleaved before serving as primer (Shieh *et al.*, 1987). These RNAs are likely generated due to the hairpin loop structures in nucleotides 52–72 and 80–143, and the intervening AU-rich sequences at the 5'-end of the genome (Fig. 3) (Shieh *et al.*, 1987).

It should be noted that the pausing or termination sites of these small RNAs were determined by assuming that they were generated from correct transcriptional initiation sites. Although the extremely small amount of these RNAs makes it impossible to resolve this issue definitively, several reasons argue for the interpretation that they were initiated from the same site: first, the leader-primed transcription mechanism of MHV (Baric *et al.*, 1983; Shieh *et al.*, 1987) indicates that the same primer is used for the transcription of
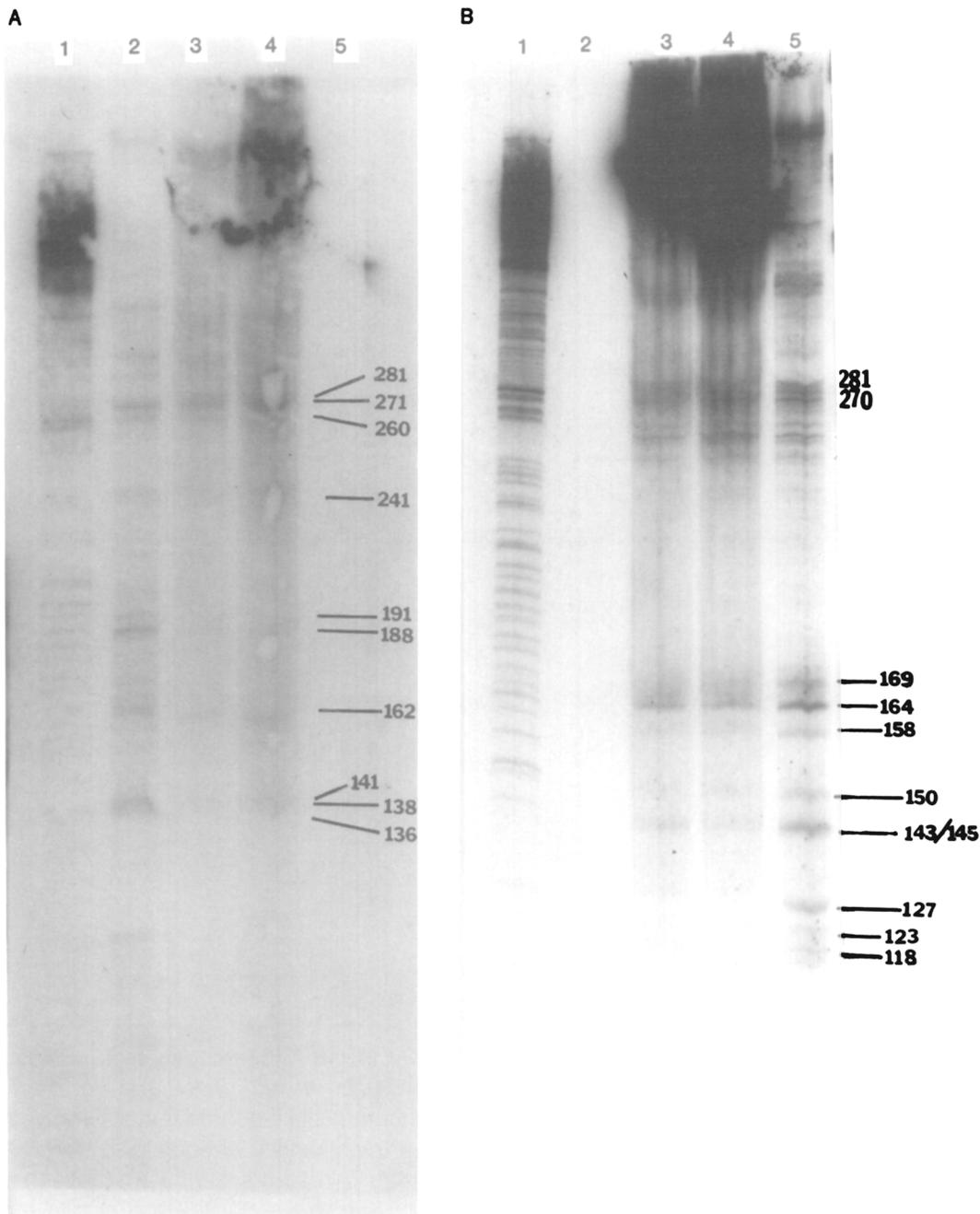
FIG. 7. Analysis of small RNAs present on RI or as free RNAs. Intracellular RNA was extracted from MHV-infected cells and separated on 10–30% sucrose gradients. The 18–40 S RNA, representing the RI, and 4 S RNA fractions were collected separately and analyzed on 8% polyacrylamide gels, transferred to Z-probe paper, and probed with nick-translated F82 (A) and E1-500 (B) separately. (A) Lane 1, M13 T-ladder; lane 2, 4 S RNA fractions; lane 3, 18–40 S RNA; lane 4, total intracellular RNA from A59-infected cells; lane 5, uninfected DBT cells. (B) Lane 1, M13 T-ladder; lane 2, uninfected DBT cells; lane 3, total intracellular RNA from A59-infected cells; lane 4, 18–40 S RNA; lane 5, 4 S RNA fractions.

various mRNAs; second, these RNA species are very reproducible in different RNA preparations and are present on replicative intermediate RNA; third, RNA recombinants have been isolated whose recombination sites may be in an area corresponding to the termination sites of these small RNAs (Keck *et al.,* 1986), suggesting that these RNAs have biological functions;

and finally, correct initiation was noted in a similar study with QB phages (Mills *et al.,* 1978).

The biochemical and biological data presented here and elsewhere (Baric *et al.,* 1985; Makino *et al.,* 1986a) are, thus, consistent with a mechanism of MHV transcription which is discontinuous and nonprocessive (Fig. 8). In this model, the viral polymerase pauses in
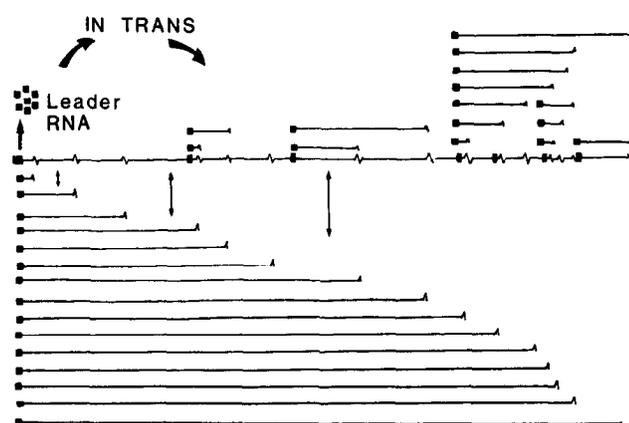
FIG. 8. Proposed model for MHV discontinuous transcription. In this model, MHV positive-sense RNA transcription initiates at the 3'-end of the negative-strand and "pauses" in regions of secondary structure in either the product or template strands. These intermediate RNAs can dissociate from the negative-strand template and reassociate to function as primers for the synthesis of full-length RNAs.

regions of secondary structure, producing discrete nascent RNA intermediates on replicative intermediate RNA. Either a portion of or all of these small RNAs dissociate from the template strand and then reassociate to act as primers for continuing RNA elongation. Further proof of this model will require studies using an in vitro transcription system.

ACKNOWLEDGMENTS

REFERENCES

ARMSTRONG, J., NIEMANN, H., SMEEKENS, S., ROTTIER, P., and WARREN, G. (1984). Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. Nature (London) 308, 751–752.

AURON, P. E., WEBER, L. D., and RICH, A. (1982). Comparison of transfer ribonucleic acid structures using cobra venom and S₁ endonucleases. Biochemistry 21, 4700–4706.

BARIC, R. S., STOHLMAN, S. A., and LAI, M. M. C. (1983). Characterization of replicative intermediate RNA of mouse hepatitis virus: Presence of leader RNA sequences on nascent chains. J. Virol. 48, 633–640.

BARIC, R. S., STOHLMAN, S. A., RAZAVI, M. K., and LAI, M. M. C. (1985). Characterization of leader-related small RNAs in coronavirus-infected cells: Further evidence for leader-primed mechanism of transcription. Virus Res. 3, 19–33.

BRAYTON, P. R., GANGES, R. G., and STOHLMAN, S. A. (1981). Host cell nuclear function and murine hepatitis virus replication. J. Gen. Virol. 56, 457–460.

BRAYTON, P. R., LAI, M. M. C., PATTON, C. D., and STOHLMAN, S. A. (1982). Characterization of two RNA polymerase activities induced by mouse hepatitis virus. J. Virol. 42, 847–853.

BRAYTON, P. R., STOHLMAN, S. A., and LAI, M. M. C. (1984). Further characterization of mouse hepatitis virus RNA-dependent RNA polymerases. Virology 133, 197–201.

EFSTRATIADIS, A., MANIATIS, T., KAFATOS, F. C., JEFFREY, A., and VOURNAKIS, J. N. (1975). Full length and discrete partial reverse transcripts of globin and chorion mRNAs. Cell 4, 367–378.

FIELDS, B. N. (1981). Genetics of reovirus. Curr. Top. Microb. Immunol. 91, 1–24.

GEHRKE, J., AURON, P. E., QUIGLEY, G. J., RICH, A., and SONENBERG, N. (1983). 5'-Conformation of capped alfalfa mosaic virus ribonucleic acid 4 may reflect its independence of the cap structure or of cap-binding protein for efficient translation. Biochemistry 22, 5157–5164.

HASELTINE, W. A., KLEID, D. G., PANET, A., ROTHENBERG, E., and BALTIMORE, D. (1976). Ordered transcription of RNA tumor virus genomes. J. Mol. Biol. 106, 109–131.

HUANG, C. C., HAY, N., and BISHOP, J. M. (1986). The role of RNA molecules in transduction of the proto-oncogene c-fps. Cell 44, 935–940.

HUANG, C. C., and HEARST, J. E. (1981). Fine mapping of secondary structures of fd phage DNA in the region of the replication origin. Nucleic Acids Res. 9, 5587–5599.

JACOBS, L., SPAAN, W. J. M., HORZINEK, M. C., and VAN DER ZEIJST, B. A. M. (1981). Synthesis of subgenomic mRNAs of mouse hepatitis virus is initiated independently: Evidence from U.V. transcriptional mapping. J. Virol. 34, 401–406.

KECK, J., STOHLMAN, S. A., SOE, L., MAKINO, S., and LAI, M. M. C. (1987). Multiple recombination sites at the 5'-end of murine coronavirus RNA. Virology 156, 331–341.

KRAMER, F. R., and MILLS, D. R. (1981). Secondary structure formation during RNA synthesis. Nucleic Acids Res. 9, 5109–5124.

LAI, M. M. C., BARIC, R. S., BRAYTON, P. R., and STOHLMAN, S. A. (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.

LAI, M. M. C., BARIC, R. S., MAKINO, S., KECK, J. G., EGBERT, J., LEIBOWITZ, J. L., and STOHLMAN, S. A. (1985). Recombination between non-segmented RNA genomes of murine coronaviruses. J. Virol. 56, 449–456.

LAI, M. M. C., BRAYTON, P. R., ARMEN, R. C., PATTON, C. D., PUGH, C., and STOHLMAN, S. A. (1981). Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from hepatotropic strain MHV-3. J. Virol. 39, 823–834.

LAI, M. M. C., PATTON, C. D., and STOHLMAN, S. A. (1982). Replication of mouse hepatitis virus: Negative-stranded RNA and replicative form RNA are of genomic length. J. Virol. 44, 487–492.

LAI, M. M. C., and STOHLMAN, S. A. (1978). RNA of mouse hepatitis virus. J. Virol. 26, 236–242.

LEIBOWITZ, J. L., WILHELMSEN, K. C., and BOND, C. W. (1981). The virus specific intracellular RNA species of two murine coronaviruses: MHV-A59 and MHV-JHM. Virology 114, 39–51.

MAIZELS, N. M. (1973). The nucleotide sequence of the lactose messenger ribonucleic acid transcribed from the UV5 promotor mutant of Escherichia coli. Proc. Natl. Acad. Sci. USA 70, 3585–3589.

MAKINO, S., KECK, J. G., STOHLMAN, S. A., and LAI, M. M. C. (1986a). High frequency RNA recombination of murine coronaviruses. J. Virol. 57, 729–737.

MAKINO, S., STOHLMAN, S. A., and LAI, M. M. C. (1986b). 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.

MANIATIS, T., FRITSCH, E. F., and SAMBROOK, J. (1982). "Molecular Cloning—A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

MILLS, D. R., DABKIN, C., and KRAMER, F. R. (1978). Template-determined, variable rate of RNA chain elongation. *Cell* **15**, 541–550.

MINKLEY, E. G., and PRIBNOW, D. (1973). Transcription of the early region of bacteriophage T7: Selective initiation with dinucleotides. *J. Mol. Biol.* **77**, 255–277.

NIYOGI, S. K., and STEVENS, A. (1965). Studies of the ribonucleic acid polymerase from *Escherichia coli*. IV. Effect of oligonucleotides on the ribonucleic acid polymerase reaction with synthetic polyribonucleotides as templates. *J. Biol. Chem.* **240**, 2593–2598.

ROSENBERG, M., COURT, D., SHIMATAKE, H., BRADY, C., and WULFF, D. L. (1978). The relationship between function and DNA sequence in an intercistronic regulatory region in phage lambda. *Nature (London)* **272**, 414–423.

SALSER, W. (1977). Globin mRNA sequences: analysis of base pairing and evolutionary implications. *Cold Spring Harbor Symp. Quant. Biol.* **42**, 985–1002.

SAWICKI, S. G., and SAWICKI, D. L. (1986). Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *J. Virol.* **57**, 328–334.

SHERMAN, L. A., and GEFTER, M. L. (1976). Studies on the mechanism of enzymatic DNA elongation by *Escherichia coli* DNA polymerase II. *J. Mol. Biol.* **103**, 61–76.

SHIEH, C.-K., SOE, L., MAKINO, S., CHANG, M. F., STOHLMAN, S. A., and LAI, M. M. C. (1987). The 5'-end sequence of murine coronavirus genome: Implications for multiple fusion sites in leader-primed transcription. *Virology* **156**, 321–330.

SKINNER, M. A., EBNER, D., and SIDDELL, S. G. (1985). Coronavirus MHV-JHM. mRNA5 has a sequence arrangement which potentially allows translation of a second down-stream open reading frame. *J. Gen. Virol.* **66**, 581–592.

SKINNER, M. A., and SIDDELL, S. G. (1985). Coding sequence of coronavirus MHV-JHM mRNA 4. *J. Gen. Virol.* **66**, 593–596.

SPAAN, W., DELIUS, H., SKINNER, M., ARMSTRONG, J., ROTTIER, P., SMEEKENS, S., VAN DER ZEIJST, B. A. M., and SIDDELL, S. G. (1983). Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* **2**, 1839–1844.

STELLWAG, E. J., and DAHLBERG, A. E. (1980). Electrophoretic transfer of DNA, RNA and protein onto diazobenzyloxymethyl (DBM) paper. *Nucleic Acids Res.* **8**, 299–317.

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.

STURMAN, L. S., HOLMES, K. V., and BEHNKE, J. (1980). Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J. Virol.* **33**, 449–462.

TINOCO, I., BORER, P. N., DENGLER, B., LEVINE, M. D., UHLENBECK, O. C., CROTHERS, D. M., and GRALLA, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biol.* **246**, 40–41.

WEISSMAN, C. (1974). The making of a phage. *FEBS Lett.* **40**, S10–S18.

WILHELMSEN, K. C., LEIBOWITZ, J. L., BOND, C. W., and ROBB, J. A. (1981). The replication of murine coronaviruses in enucleated cells. *Virology* **110**, 225–230.

ZUCKER, M., and STIEGLER, P. (1981). Optimal computer folding of large RNA sequence using thermodynamics and auxillary information. *Nucleic Acids Res.* **9**, 133–148.