JOURNAL OF VIROLOGY, Jan. 1995, p. 272–280 Vol. 69, No. 1
0022-538X/95/$04.00+0

The Effect of Two Closely Inserted Transcription Consensus Sequences on Coronavirus Transcription

MYUNGSOO JOO AND SHINJI MAKINO*

Department of Microbiology, The University of Texas at Austin, Austin, Texas 78712-1095

Received 5 May 1994/Accepted 5 October 1994

Insertion of an intergenic region from the murine coronavirus mouse hepatitis virus into a mouse hepatitis virus defective interfering (DI) RNA led to transcription of subgenomic DI RNA in helper virus-infected cells. Using this system, we studied how two intergenic regions in close proximity affected subgenomic RNA synthesis. When two intergenic regions were separated by more than 100 nucleotides, slightly less of the larger subgenomic DI RNA (synthesized from the upstream intergenic region) was made; this difference was significant when the intergenic region separation was less than about 35 nucleotides. Deletion of sequences flanking the two intergenic regions inserted in close proximity did not affect transcription. No significant change in the ratio of the two subgenomic DI RNAs was observed when the sequence between the two intergenic regions was altered. Removal of the downstream intergenic region restored transcription of the larger subgenomic DI RNA. The UCUAAAC consensus sequence was needed for efficient suppression of the larger subgenomic DI RNA synthesis. These results demonstrated that the downstream intergenic sequence was suppressing subgenomic DI RNA synthesis from the upstream intergenic region. We discuss possible mechanisms to account for the regulation of this suppression of subgenomic DI RNA synthesis and the ways in which they relate to the general regulation of coronavirus transcription.

Many positive-strand RNA viruses express their genes in infected cells through subgenomic-length mRNAs. Viruses belonging to the alphavirus superfamily, which includes many animal and plant viruses (1, 4, 6), synthesize a subgenomic mRNA in order to express virus-specific gene(s). Transcription of the subgenomic mRNA starts at the promoter sequence present in the genomic-size negative-strand RNA. This type of subgenomic transcriptional regulation represents an area of considerable research into members of the alphavirus superfamily (3, 5, 7, 22, 31, 34, 40, 45). The coronaviruses are another positive-strand RNA virus group that depend on subgenomic mRNAs for gene expression. Several different species of coronavirus subgenomic mRNAs, which together make up a 3′-coterminal nested-set structure, are synthesized in virus-infected cells (15, 17, 21, 42, 43). The 5′ end of coronavirus genomic RNA and the 5′ ends of its subgenomic mRNAs bear a coronavirus signature that the alphavirus superfamily lacks, i.e., the leader sequence (16, 18, 41). The presence of the leader sequence on coronavirus subgenomic RNAs indicates that coronavirus transcription may be a unique mechanism that differs from that of the other positive-strand RNA viruses.

Mouse hepatitis virus (MHV) is the coronavirus prototype. It contains a single-stranded, positive-sense RNA of approximately 31 kb (19, 20, 32). In MHV-infected cells, seven to eight species of virus-specific mRNAs are synthesized; they are named mRNA 1 to mRNA 7 in decreasing order of size (17, 21). The 5′ end of the MHV genomic RNA and subgenomic mRNAs starts with a 72- to 77-nucleotide (nt) leader sequence (16, 41). The genomic RNA encodes only one copy of the leader sequence, found at its 5′ end. By an unknown mechanism, leader sequences fuse with the subgenomic mRNA body sequences. The mRNA body sequences begin from a consensus sequence (UCUAAAC or a very similar sequence) in the intergenic region, which is located upstream of each MHV gene (16, 41). The intergenic region preceding gene 7 (gene 7 encodes the nucleocapsid protein) carries the same 18-nt sequence found at the 3′ region of the genomic leader sequence (39).

MHV subgenomic RNAs are not detected in MHV virions; therefore, at some point in the virus replication cycle, subgenomic-size RNAs are thought to be synthesized from a genomic-size RNA and a leader sequence is attached to the 5′ end of each mRNA. Virions of some other coronaviruses do appear to contain a small amount of subgenomic RNAs with unknown biological function (10, 38, 49). Subgenomic negative-strand RNAs, each of which corresponds to subgenomic mRNAs, are present in coronavirus-infected cells (37, 38). These subgenomic negative-strand RNAs contain the antileader sequence at their 3′ end (37). MHV-infected cells contain subgenomic-size replicative intermediate RNAs, indicating that RNA elongates on a subgenomic-size RNA template (36). Therefore, there are at least two stages in coronavirus transcription: transcription of subgenomic-size RNA from the genomic-size template RNA (primary transcription) and synthesis of RNA on the subgenomic-size RNA template (secondary transcription) (11, 47). Although, we do not know the polarities of the primary and secondary transcription templates, we do know that MHV transcription involves discontinuous transcription (12, 29, 48) and that all activities necessary for MHV transcription are present at least during the first 6 h of infection (11).

A system that exploits defective interfering (DI) RNAs of MHV for the study of coronavirus transcription is established (25). In this system, subgenomic DI RNA is synthesized from an inserted intergenic region during DI RNA replication (25). One study in which this system was used demonstrated that the sequences flanking the intergenic region preceding gene 7 do not play a role in subgenomic DI RNA transcription (24). However, studies of an MHV mutant virus and bovine coronavirus subgenomic mRNAs raise the possibility that a se-

* Corresponding author. Mailing address: Department of Microbiology, The University of Texas at Austin, ESB 304, 24th at Speedway, Austin, Texas 78712-1095. Phone: (512) 471-6876. Fax: (512) 471-7088. Electronic mail address: makino@mail.utexas.edu.

quence(s) outside of the intergenic consensus sequence affects coronavirus transcription. An MHV mutant virus, MHV-S no. 8, which was isolated from cells persistently infected with MHV-S, possesses the 3′ half of the genomic leader sequence inserted in the 5′ region of gene 7 (44). This insertion results in the presence of two consensus sequences separated by 0.1 kb within gene 7. Interestingly, the amount of "larger" mRNA 7, which is synthesized from the upstream consensus sequence, is only 5% of that of the "smaller" mRNA 7, which is synthesized from the downstream consensus sequence (44). Perhaps, in MHV-S no. 8, the insertion of the 3′ half of the genomic leader sequence into gene 7 inhibits transcription of the larger mRNA 7 (44). Hofmann et al. (9) demonstrated that in bovine coronavirus, a subgenomic mRNA is not synthesized from the predicted intergenic consensus sequence; instead, a subgenomic mRNA is synthesized from another sequence, located 15 nt downstream of the predicted intergenic consensus sequence. Their data led us to speculate that subgenomic mRNA transcription from the upstream consensus sequence is inhibited by the presence of a downstream cryptic transcription consensus sequence. From these studies of MHV-S no. 8 and bovine coronavirus, we hypothesized that two coronavirus intergenic consensus sequences that are located in close proximity may interact in such a way that the presence of a downstream consensus sequence may inhibit transcription of subgenomic mRNA from an upstream consensus sequence. We examined this possibility and present new aspects of coronavirus transcription regulation.

MATERIALS AND METHODS

Viruses and cells. The plaque-cloned A59 strain of MHV (MHV-A59) (17) was used as a helper virus. Mouse DBT cells (8) were used for growth of viruses.

DNA construction. The small *Sph*I-*Eco*RV fragment of MHV cDNA clone MT1/31, that of MT1/24, and that of MT1/18 (24) were inserted into the large *Sph*I-*Kpn*I fragment of MT1/174, yielding MS36, MS29, and MS23, respectively. MHV cDNA clones MT1/110, MT1/88, and MT1/49 (24) were incubated with two oligonucleotides in PCR buffer at 93°C for 30 s, 55°C for 30 s, and 72°C for 100 s for 25 cycles; the oligonucleotides were oligonucleotide 1497 (5′-CCG CATTGGTACCAATCTAA-3′), which has a *Kpn*I site and binds to the antigenomic sense of these clones at nt 2329 to 2348 from the 5′ end, and oligonucleotide 2243 (5′-TATCTACGGTACCTTTCT-3′) also with a *Kpn*I site but binding to the genomic sense of these clones at nt 471 to 454 from the 3′ end; the PCR buffer consisted of 0.05 M KCl, 0.01 M Tris-HCl (pH 8.3), 0.0025 M $MgCl_2$, 0.01% gelatin, 0.17 mM each deoxynucleoside triphosphates, and 5 U of *Taq* polymerase (Promega). Each PCR product was digested with *Kpn*I and inserted into the *Kpn*I site of MT1/174, generating MS124, MS102, and MS63, respectively. MS23ΔUF was constructed by removing a 0.8-kb *Kpn*I-*Spe*I fragment of MS23. The 0.83-kb long MT1/18 *Spe*I-*Eco*RV fragment was inserted into the large MT1/18 *Spe*I-*Kpn*I fragment, yielding MS23ΔDF. Under the same PCR conditions as described above, MT1/174 was incubated with oligonucleotide 2241 (5′-TCTAAGGTACCAGGATGTCTTTT-3′), which contains a *Kpn*I site and binds to the antigenomic sequence of the MT1/174 intergenic region, and oligonucleotide 1189 (5′-GTTGGATATCTGCTTGGGC-3′), which contains an *Eco*RV site and binds to the genomic sequence 0.17 kb downstream of the intergenic sequence. The 0.18-kb long *Kpn*I-*Eco*RV PCR fragment was inserted into the *Kpn*I-*Eco*RV site of MT1/18. The 1.25-kb long MT1/24 (24) *Sph*I-*Eco*RV fragment was inserted into the *Kpn*I-*Sph*I site of this newly constructed plasmid, yielding MS23ΔDIG. MJWT (13) was incubated with oligonucleotides 1497 and 1189 under the same PCR conditions as described above. The 0.18-kb *Kpn*I-*Eco*RV PCR fragment was inserted into the corresponding sites of MT1/18. The 0.83-kb long *Spe*I-*Eco*RV fragment of MT1/18 was inserted into the *Spe*I-*Kpn*I site of this new plasmid to produce MS23D1R. A similar procedure to the one used for construction of MS23D1R was applied to MS23DC2G construction, in which oligonucleotide 10021 (5′-CACCGTATTGGTACCAAT GTAAAC-3′), which binds to the antigenomic sequence of PR6 (25) at nt 3089 to 4012 from the 5′ end, was used, instead of oligonucleotide 1497. MS23 was incubated with oligonucleotide 10024 (5′-GAGCTCGGTACCAATCTAATCT AAACTATTATAATTAATCTAATCTAAACTTTAAGGATGTC-3′), which binds to the antigenomic sequence of MS23 at 686 to 624 nt from the 3′ end, and oligonucleotide 1189 under the same PCR conditions described above. The *Kpn*I-*Eco*RV PCR fragment was inserted into the large *Kpn*I-*Eco*RV site of MT1/18, yielding MS23AU. For each mutant the entire region obtained by insertion of the PCR product was sequenced to confirm the presence of the specific mutations and the absence of extraneous mutations.

RNA transcription and transfection. Plasmid DNAs were linearized by *Xba*I digestion and transcribed with T7 RNA polymerase as previously described (26). The lipofection procedure was used for RNA transfection as previously described (25).

Preparation of virus-specific intracellular RNA and Northern blotting. Virus-specific RNAs in virus-infected cells were extracted as previously described (30). For each sample, 1.5 μg of intracellular RNA was denatured and electrophoresed through a 1% agarose gel containing formaldehyde, and the separated RNA was blotted onto nylon filters as described previously (25). The nylon filters were soaked in a prehybridization buffer, and Northern (RNA) blot hybridization was performed (11). A gel-purified 0.25-kb *Nru*I-*Msc*I fragment from MT1/174 was labeled by the random-priming procedure (35) and used as a probe. This probe corresponds to nt 18 to 262 from the 3′ end of the MHV DI cDNA.

Primer extension. The oligonucleotides were 5′-end labeled with [γ-^{32}P]ATP by using polynucleotide kinase (35). Poly(A)-containing RNAs were used for primer extension analysis as described previously (27). Reaction products were analyzed on 6% polyacrylamide gels containing 7 M urea.

PCR and direct sequencing of the PCR products. Primer extension products were purified from the gel and amplified by PCR under the same conditions as described above. The gel-purified reverse transcriptase-PCR products were separated by agarose gel electrophoresis. Direct PCR sequencing was performed by the procedure established by Winship (46).

RESULTS

Effect of two proximally inserted intergenic regions on MHV subgenomic DI RNA transcription. In an attempt to gauge the effect of a downstream intergenic region on its upstream neighbor, we measured the synthesis of two subgenomic DI RNAs that were transcribed from two proximally inserted intergenic consensus sequences. We constructed a series of MHV DI cDNAs, each of which contained two intergenic regions in the same parental clone, MHV DI RNA-derived cDNA clone MT1/174 (24) (Fig. 1). MT1/174 has a 174-nt inserted sequence, which includes both the 18-nt intergenic region, AAUCUAAUCUAAACUUUA, and 156 nt located immediately downstream of the intergenic region; this intergenic region and its flanking sequence derive from between genes 6 and 7. All clones contained an insertion upstream of the 18-nt intergenic region of MT1/174 (Fig. 1). The inserted sequence consisted of the 18-nt intergenic region preceding gene 7 attached to different lengths of downstream sequence and a few nucleotides from the 3′ end, which were generated by construction procedures (Fig. 1). Therefore, these newly constructed DI cDNA clones bore two 18-nt long intergenic regions; most of the sequences between the two intergenic regions derived from downstream of the intergenic region between genes 6 and 7. We named these DI cDNAs according to the distance between the first nucleotide of the upstream intergenic region and the first nucleotide of the downstream intergenic region. In MS124, the first nucleotide of the downstream intergenic region is located 124 nt from the first nucleotide of the upstream intergenic region.

DI RNAs synthesized in vitro from the mutant clones were transfected into monolayers of MHV-A59-infected DBT cells (25). After overnight incubation, the culture fluid was harvested and passaged once more to prepare passage 1 virus sample. This passage 1 sample was used as an inoculum for the analysis of intracellular RNA species. Virus-specific intracellular RNA was extracted at 7 h postinfection and analyzed by Northern blotting with a probe that specifically hybridizes with all MHV RNAs (Fig. 2). Genomic DI RNA and subgenomic DI RNA were synthesized from all the DI RNAs carrying dual intergenic sites. Most of the mutants made two subgenomic DI RNAs of approximately the same size; these were difficult to separate on agarose gels. Therefore, we used primer extension analysis to examine the quantities of the two subgenomic DI RNAs in the mutants with similar-size subgenomic DI RNAs. The 5′-end-labeled oligonucleotide 10007 (5′-TCTTTCTGAT

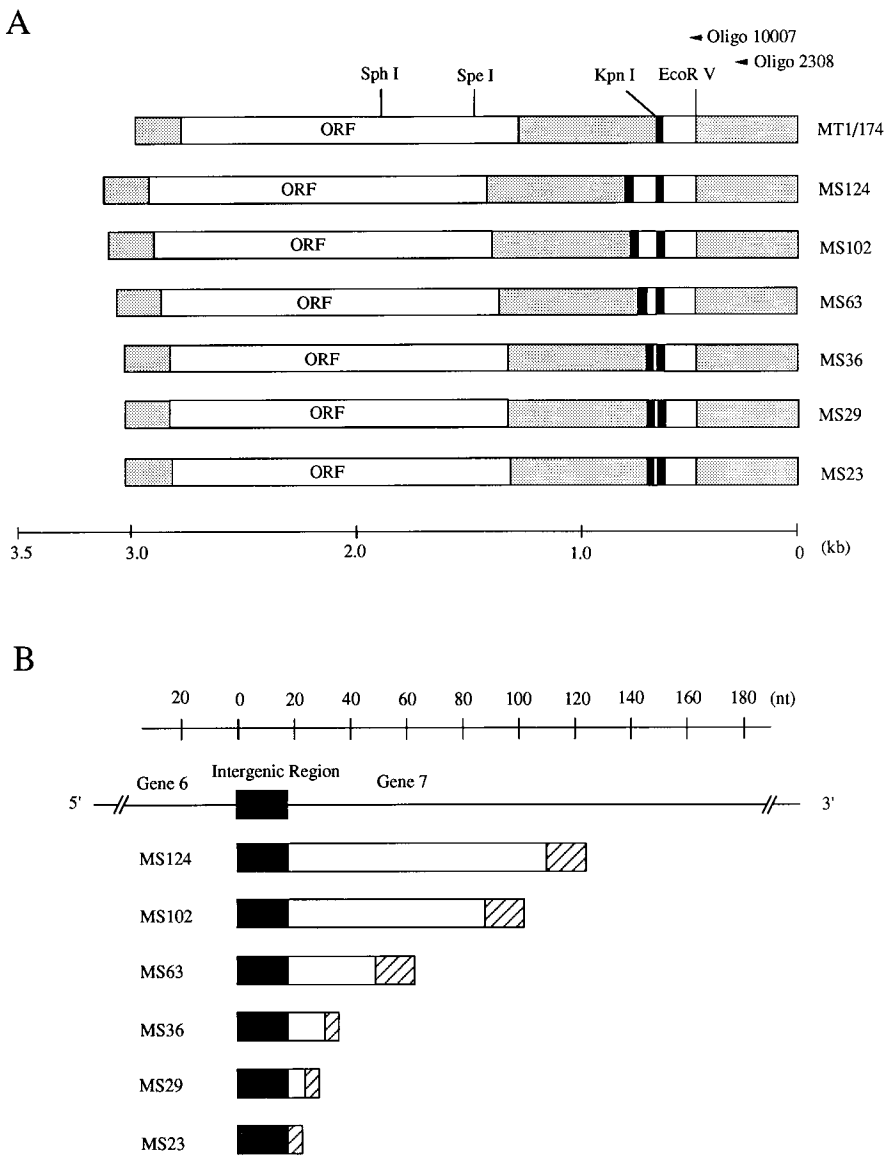

FIG. 1. Schematic diagram of the structure of MT1/174 and DI RNAs containing two intergenic sequences. (A) The DI-specific open reading frame is labeled ORF. The black boxes and white boxes represent the 18-nt intergenic region (AAUCUAAUCUAAACUUUA) and its downstream gene 7 sequence, respectively. Restriction enzyme sites used for construction of the mutants and oligonucleotides used for primer extension are shown. (B) Schematic diagram of the specific inserted sequences. All mutant DI RNAs had the same structure as MT1/174, except that each mutant contained an inserted sequence, shown by a boxed region. The black box, white box, and hatched box represent the 18-nt intergenic region, its downstream gene 7 sequence, and a non-MHV sequence, respectively.

ATCTGCT-3′), which specifically binds to genomic and subgenomic DI RNAs but not to helper virus mRNAs, hybridized with intracellular RNA species, and the hybridized primer was extended with reverse transcriptase. Primer-binding sites on genomic DI RNA are shown in Fig. 1. Primer extension products were then analyzed by electrophoresis on sequencing gels (Fig. 3). The amount of primer extension products applied to the gels was adjusted so that all of the products from the small subgenomic DI RNA had equivalent radioactivity. Large amounts of a primer extension product corresponding to a smaller subgenomic DI RNA, synthesized from the downstream intergenic region, were made in MS23-, MS29-, and MS36-replicating cells (Fig. 3A). We frequently observed a minor primer extension product, which migrated as if it were 5 nt longer than the major primer extension product (Fig. 3, open triangle). The amount of a primer extension product corresponding to a larger subgenomic DI RNA, produced from the upstream intergenic region, was significantly reduced (Fig. 3A). Densitometric analysis of the autoradiograms revealed that the molar ratios of the larger primer extension product to the major smaller primer extension product were approximately 0.04, 0.03, and 0.02 in MS36, MS29, and MS23, respectively; this ratio increased as the distance between the two intergenic regions increased.

The structures of the primer extension products were examined to confirm that these primer extension products indeed represented subgenomic DI RNA species. All MHV subgenomic mRNAs have the leader sequence at their 5′ end; therefore, we used oligonucleotide 10007 and another oligonucleotide, oligonucleotide 95 (5′-GATTGGCGTCCGTAC

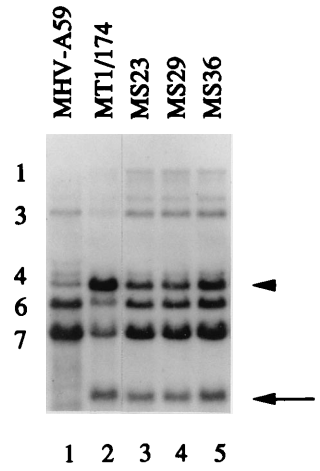

FIG. 2. Northern blot analysis of genomic DI RNA and subgenomic DI RNA. Numbers 1, 3, 4, 6, and 7 (on the left) represent major MHV-specific mRNA species. The arrowhead indicates genomic DI RNAs. Several subgenomic DI RNAs, which were not resolved by agarose gel electrophoresis, are shown by an arrow.

GTA-3′), which specifically binds to the 3′ end of the antileader sequence, for priming the synthesis of PCR products from the primer extension products. Direct sequencing of PCR products demonstrated that the larger and the smaller primer extension products contained the expected structure of the larger and the smaller subgenomic DI RNA, respectively (data not shown). All the subgenomic DI RNAs contained two repeats of UCUAA at the leader-body junction site. The minor primer extension products, which migrated slightly more slowly than the major primer extension product of the smaller subgenomic DI RNAs, exhibited structures that were similar to the smaller subgenomic DI RNAs, except that they contained three UCUAA repeats. This result was consistent with the observation that the leader fusion site on a given species of MHV subgenomic mRNA is heterogeneous and that heterogeneity is due to a variation in the number of UCUAA pentanucleotide repeats at the leader-body fusion site (28). This heterogeneity may be caused by imprecise leader RNA binding to the template RNA (28). These analyses demonstrated that transcription of the larger subgenomic DI RNA was inhibited when two intergenic regions were inserted within 23 to 36 nt of each other.

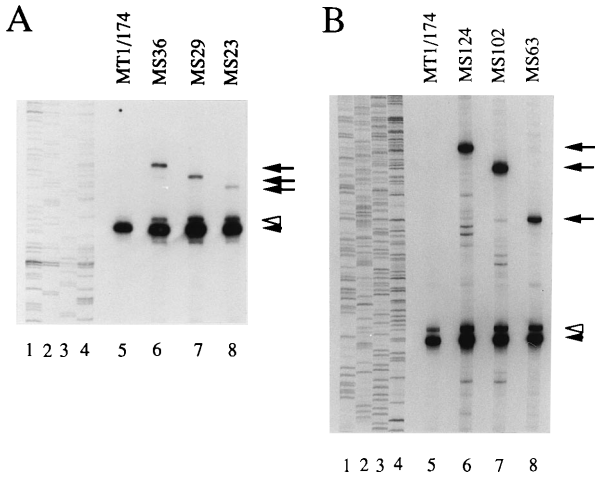

FIG. 3. Primer extension analysis of subgenomic DI RNAs. The 5′-end-labeled oligonucleotide 10007 was hybridized with intracellular RNAs and extended with reverse transcriptase. The products were electrophoresed on sequencing gels. Panels A and B represent two independently electrophoresed gels. Lanes 1 to 4 show a sequence ladder size marker. The arrows and arrowheads indicate the larger subgenomic DI RNA primer extension products and the smaller subgenomic DI RNA primer extension products, respectively. The minor bands indicated by the open arrowheads are the smaller subgenomic DI RNAs which contained three UCUAA repeats at the leader-body junction (see the text).

We examined the effect of the length of the "gap" between the two intergenic sites on inhibition of the larger subgenomic DI RNA transcription. The amounts of the larger and smaller subgenomic DI RNAs of MS124, MS102, and MS63 were compared (Fig. 3B). Primer extension analysis and direct sequencing of PCR products of the primer extension products demonstrated that the molar ratios of the larger primer extension product to the major smaller primer extension product were 0.77, 0.88, and 0.40 in MS124, MS102, and MS63, respectively, demonstrating that more of the larger subgenomic DI RNAs (shown by the arrows in Fig. 3B) were synthesized in MS124-, MS102-, and MS63-replicating cells than had been made in MS36-, MS29-, and MS23-replicating cells (Fig. 3A). The larger subgenomic DI RNA was only slightly less abundant than the smaller DI RNA in MS124 and MS102 (shown by the arrows in Fig. 3B). The larger subgenomic DI RNA in MS63-replicating cells (Fig. 3B) was noticeably less abundant than the smaller subgenomic DI RNA, whereas this difference was not as significant as the difference seen in MS36-, MS29-, and MS23-replicating cells (Fig. 3A). When two intergenic regions were separated by about 60 nt, a distinct inhibition of transcription of the larger subgenomic DI RNA occurred. Of the DI RNAs that we analyzed, the degree of inhibition increased as the distance between the two intergenic regions decreased; MS23, with the shortest sequence between the two intergenic sites, demonstrated the greatest inhibition of the larger subgenomic DI RNA transcription.

Effect of sequences flanking the two inserted intergenic regions on subgenomic DI RNA transcription. We looked for an effect of the sequences flanking the two inserted intergenic regions on subgenomic DI RNA transcription by analyzing two MS23-derived DI cDNAs, MS23ΔDF and MS23ΔUF. Downstream of the MS23 intergenic region, we deleted 156 nt to create MS23ΔDF. Upstream of the MS23 intergenic region, we removed 0.8 kb to make MS23ΔUF (Fig. 4A). The area downstream of the *Eco*RV site in MS23 is a part of an MHV DI RNA *cis*-acting RNA replication signal (14, 23); therefore, we did not test the effects of deletion of further downstream. Synthesis of subgenomic DI RNA was monitored by primer extension analysis. For the analysis of MS23ΔDF subgenomic DI RNAs, we used oligonucleotide 2308 (5′-CTTTCTCGCGA GGGGTTTAGATTAGACCACCGAGT-3′) (Fig. 1), which hybridized 0.25 kb from the 3′ end of MHV genomic RNA, as a primer, and for the analysis of MS23ΔUF subgenomic DI RNAs, we used oligonucleotide 10007. Primer extension analysis and direct sequencing of the PCR products of the primer extension products demonstrated that the transcription of the large subgenomic DI RNA was inhibited in both DI RNAs (Fig. 4B). The molar ratios of the larger subgenomic DI RNA to the major smaller subgenomic DI RNA species were approximately 0.01, 0.015, and 0.02 in MS23ΔDF, MS23ΔUF, and MS23, respectively. These studies demonstrated that the flanking sequences of the two inserted intergenic regions did not affect the inhibition of transcription of the larger subgenomic DI RNA. We sometimes observed an additional band migrating more slowly than the primer extension product from

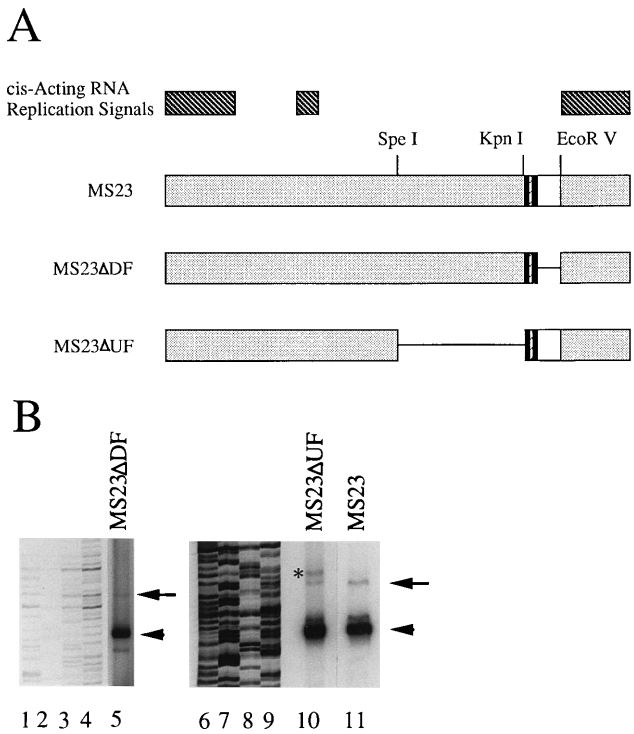

FIG. 4. Schematic diagram of the structure of MS23, MS23ΔDF, and MS23ΔUF (A) and primer extension analysis of subgenomic DI RNAs (B). (A) The black box, white box, and hatched box (between two black boxes) represent the 18-nt intergenic region, its downstream gene 7 sequence, and a non-MHV sequence, respectively. The locations of MHV DI RNA *cis*-acting RNA replication signals, which are necessary for MHV DI RNA replication (14), are also shown. (B) Oligonucleotide 2308 was used as primer for the analysis of subgenomic DI RNA of MS23ΔDF, and oligonucleotide 10007 was used for analysis of subgenomic DI RNA of MS23ΔUF and MS23. The arrows and arrowheads indicate the primer extension products of the larger and smaller subgenomic DI RNAs, respectively. The asterisk indicates a band that most probably corresponds to a premature termination product of the genomic DI RNA primer extension product (see text). Lanes 1 to 4 show a sequence ladder size marker.

the MS23ΔUF large subgenomic DI RNA (Fig. 4B, asterisk in lane 10). After purifying and attempting to sequence this band by PCR with oligonucleotides 10007 and 95, we failed to obtain a product, indicating that this band was not a primer extension product of the subgenomic DI RNA. It probably represented a premature termination product of the genomic DI RNA primer extension product.

Investigation of a possible effect of the downstream intergenic region on inhibition of transcription from the upstream intergenic consensus sequence. To directly test the effect of the downstream intergenic region on inhibition of transcription from the upstream intergenic region, we used a new construct, MS23ΔDIG. This MS23-derived DI cDNA lacked the entire 18-nt downstream intergenic region; in its place, MS23ΔDIG carried an 8-nt insertion from a non-MHV sequence, which was generated through the DNA construction procedure (Fig. 5). Northern blot analysis of MS23ΔDIG intracellular RNA species demonstrated that the amount of MS23ΔDIG subgenomic DI RNA was comparable to that of MT1/174 (data not shown), which synthesizes a high level of subgenomic DI RNA. Primer extension analysis of MS23ΔDIG with oligonucleotide 10007 showed an abundance of product corresponding to the larger subgenomic DI RNA, whereas no primer extension product made from the smaller subgenomic DI RNA was seen (Fig. 6A). These data clearly demonstrated that the presence of a downstream intergenic region inhibited transcription from the upstream intergenic region.

Effects on subgenomic DI RNA transcription of sequence changes in the downstream intergenic region and in the sequence between the two intergenic consensus sequences. Nucleotide substitutions or deletions within the intergenic region affect the efficiency of transcription of that subgenomic DI RNA (13, 25). Nucleotide substitutions or deletions in a downstream intergenic region might affect transcription from a upstream intergenic region. To test this possibility, we constructed two MS23-derived DI cDNAs, MS23D1R and MS23DC2G. MS23D1R had a structure similar to that of MS23, except that MS23D1R lacked one repeat of the UCUAA sequence from the downstream intergenic sequence (Fig. 5). The amount of a subgenomic DI RNA synthesized from an intergenic region lacking one UCUAA repeat is about three-eighths of that synthesized from the complete 18-nt intergenic region (25). In MS23DC2G, the downstream UC UAAAC consensus sequence of the MS23D1R was altered to UGUAAAC. Nucleotide substitution in the UCUAAAC consensus sequence to UGUAAAC significantly decreases the efficiency of subgenomic DI RNA transcription (13); the amount of subgenomic DI RNA synthesized from a UGUAAAC sequence is approximately one-third of that synthesized from the UCUAAAC consensus sequence (13).

Northern blot analysis revealed that subgenomic DI RNAs were efficiently transcribed in MS23D1R- and MSDC2G-replicating cells (Fig. 7). By densitometric analysis of autoradiograms, the molar ratios of subgenomic DI RNA to genomic DI RNA in MT1/174-, MS23D1R-, and MSDC2G-replicating cells were, on average, 0.8, 0.5, and 0.5, respectively. We quantitated the larger and smaller subgenomic DI RNAs by primer extension analysis with oligonucleotide 10007 as the primer. For the experiment in Fig. 6A, the amount of primer extension

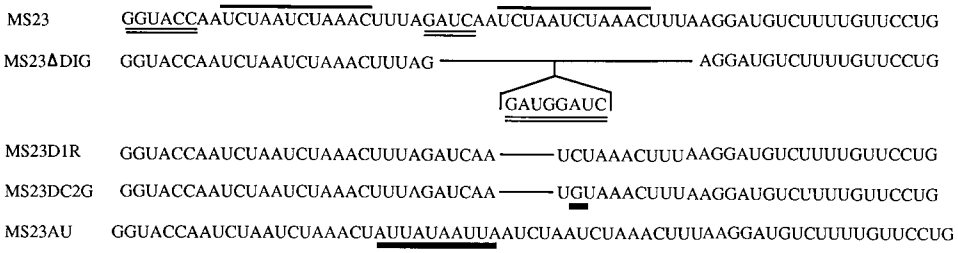

FIG. 5. Sequence at the boundary of the inserted intergenic regions of MS23, MS23ΔDIG, MS23D1R, MS23DC2G, and MS23AU. Non-MHV sequences are shown by double underlines. Bold underlines indicate mutated sequences. Deleted nucleotides are shown by thin lines; tandem repeats of UCUAA are shown by two lines above the MS23 sequence.

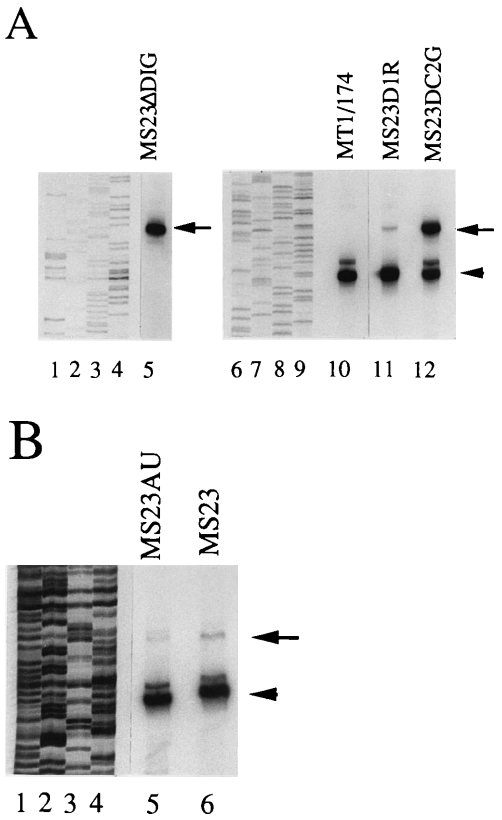

FIG. 6. Primer extension analysis of subgenomic DI RNAs. The 5′-end-labeled oligonucleotide 10007 was used as a primer. Lanes 1 to 4 and 6 to 9 in panel A and lanes 1 to 4 in panel B are size markers. The arrows and arrowheads indicate the primer extension products of the larger and smaller subgenomic DI RNAs, respectively.

products applied to gels was adjusted so that the smaller subgenomic DI RNAs showed roughly equivalent radioactivity. Synthesis of the larger subgenomic DI RNA was inhibited in MS23D1R, although the molar ratio of the larger subgenomic DI RNA to the smaller subgenomic DI RNA of MS23D1R was approximately 0.06, which is slightly larger than that of MS23 (Fig. 6A). The presence of an intergenic region lacking one UCUAA sequence inhibited transcription of the larger subgenomic DI RNA only slightly less than did the "wild-type" configuration of two consensus intergenic sequences. In MS23DC2G DI RNA-replicating cells, similar amounts of the larger and smaller subgenomic DI RNAs were synthesized (Fig. 6A, lane 12), demonstrating that the mutated consensus sequence did not significantly inhibit the larger subgenomic DI RNA transcription. Northern blot analysis and primer extension analysis revealed that for MS23DC2G, the molar ratio of the smaller subgenomic DI RNA to the genomic DI RNA was approximately half that of that in MS23D1R; these data suggested that subgenomic DI RNA transcription from UGUAAAC was reduced. We estimated the molar ratio of the larger subgenomic DI RNA to genomic DI RNA in the same mutant, MS23DC2G, to be 0.25, which was lower than the 0.8 molar ratio of the subgenomic DI RNA to the genomic DI RNA of MT1/174; transcription of the larger subgenomic DI RNA in MS23DC2G was less efficient than transcription of MT1/174 subgenomic DI RNA. Therefore, synthesis of the larger subgenomic DI RNA in MS23DC2G-replicating cells was inhibited by the presence of the downstream UGUAAAC sequence, whereas the level of inhibition was not as high as that found in MS23D1R-replicating cells.

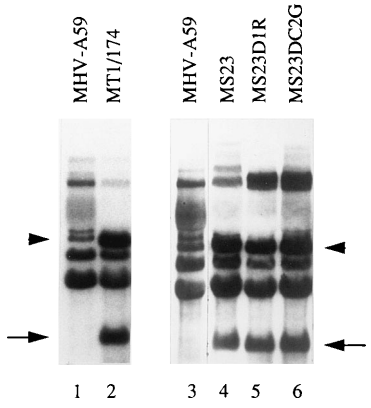

FIG. 7. Northern blot analysis of genomic and subgenomic DI RNA. The arrowheads and arrows indicate genomic and subgenomic DI RNAs, respectively.

The actual sequence of the nucleotides between the two intergenic regions might affect the transcriptional efficiency of the larger subgenomic DI RNA; we tested this by changing the nucleotides between the two repeated UCUAA sequences of the intergenic regions in clone MS23, making clone MS23AU (Fig. 5). Primer extension analysis of MS23AU with oligonucleotide 10007 as a primer revealed that transcription of the larger subgenomic DI RNA was inhibited (Fig. 6B). In cells replicating MS23AU, the molar ratio of the larger subgenomic DI RNA to the smaller subgenomic DI RNA was about half of that in MS23-replicating cells; we saw this difference consistently in repeated experiments. Frequently, a minor band migrated somewhat more slowly than the primer extension product of the larger subgenomic DI RNA of MS23AU, but this band was not stable in repeated experiments. We tried unsuccessfully to make a PCR product from the purified minor band by using oligonucleotide 10007 and oligonucleotide 95; this band probably was not a primer extension product of the subgenomic DI RNA but, rather, represented premature termination of the genomic DI RNA primer extension product. This experiment indicated that inhibition of transcription from the upstream intergenic region did not depend on the nucleotide sequence between the two intergenic regions, because there was significant inhibition after mutation of the "gap" between the intergenic sites. That mutation did have an effect, however, which we saw as a small increase in the inhibition of transcription from the upstream intergenic site.

These studies clearly demonstrated that a downstream intergenic consensus sequence inhibited transcription of subgenomic DI RNA from an upstream intergenic region located nearby.

DISCUSSION

In the present study, transcription of the larger of two subgenomic DI RNAs, which was synthesized from the upstream intergenic region of two closely inserted intergenic regions, was inhibited; its inhibition was caused by the presence of the downstream intergenic consensus sequence. This conclusion was based on the analyses of MS23 and MS23-derived mutants. Our study of mutants MS23ΔDF and MS23ΔUF clearly demonstrated that the flanking sequences of the two closely inserted intergenic regions did not play a role in the inhibition of transcription from the upstream intergenic

region. Inhibition of the larger subgenomic DI RNA transcription occurred even when the sequence between the two intergenic regions was changed, as in MS23AU. The presence of the downstream intergenic region was indeed responsible for the transcriptional inhibition at the upstream intergenic region; we demonstrated this with mutant MS23ΔDIG, which lacked the downstream intergenic region. Furthermore, the strength of transcriptional activity, initiated at the downstream intergenic region, seemed to be important for the inhibition of subgenomic DI RNA transcription from the upstream intergenic region; mutant MS23DC2G, which underproduced the smaller subgenomic DI RNA, had a reduced effect on the inhibition of the larger subgenomic DI RNA transcription.

In the case of the gene 6 to 7 intergenic sequence, transcription efficiency is not affected by sequences flanking the intergenic sequence (24). A contrasting exception to that, presented here, was when the flanking sequence was itself an intergenic region and affected transcription by inhibiting the intergenic region that it flanked. Recently, we placed a transcription consensus sequence in the middle of a 0.4-kb fragment; the fragment was located at a fixed position but was derived from various regions of MHV, and we found that transcription of subgenomic DI RNAs varied among the DI RNA constructs (12a). These data indicated that flanking sequences of the inserted intergenic region affected subgenomic DI RNA transcription efficiency. It is possible that each of the naturally occurring MHV intergenic sites is regulated in an analogous but slightly different way that depends on flanking sequences.

The novel finding presented here is that subgenomic DI RNA synthesis from the upstream intergenic region was inhibited by the presence of the downstream transcription consensus sequence. Our finding explains well why a bovine coronavirus subgenomic mRNA is not synthesized from the predicted intergenic consensus sequence but is synthesized from another sequence, located 15 nt downstream of this intergenic consensus sequence (9). The conclusions from our study also explain why in MHV-S no. 8-infected cells the amount of the larger mRNA 7, which is synthesized from the upstream consensus sequence, is significantly smaller than that of the smaller mRNA 7, which is synthesized from the downstream consensus sequence. The transcriptional regulation of MHV DI RNAs is most probably governed by the same mechanism as coronavirus transcription, and the data shown in the present study therefore shed light directly on the understanding of the actual coronavirus transcription mechanism.

We wanted to explain our data in view of the different existing coronavirus transcription models, and we envisioned how they might extend those models. The model proposed by Sawicki and Sawicki suggested that the intergenic region functions as a transcription attenuator and that subgenomic-size negative-strand RNA, which is used as the template for the subgenomic mRNA transcription, is synthesized from the input genomic positive-strand RNA (36). Accordingly, RNA polymerase should detach from the positive-strand genomic-size RNA template along with nascent subgenomic-size negative-strand RNA. This means that a reduced amount of RNA polymerase would be available for synthesis of the larger subgenomic negative-strand RNA. As a result, the level of the larger subgenomic mRNA should be consistently lower than that of the smaller subgenomic mRNA. We found that the level of the larger subgenomic DI RNA in MS124- and MS102-replicating cells was slightly lower than that of the smaller subgenomic DI RNA; these data seem to be consistent with the coronavirus transcription model proposed by Sawicki and Sawicki (36). If the intergenic sequence functions as a transcription attenuator and if two intergenic regions are separated by a short distance, RNA polymerase which resumes transcription from the downstream intergenic region (on the positive-strand RNA) may not be attenuated as strongly at the upstream intergenic region (on the positive-strand RNA). This may be the reason why the level of larger subgenomic DI RNA was significantly lower than that of the smaller subgenomic DI RNA in MS23-, MS29-, and MS36-replicating cells.

Another coronavirus transcription model, the leader-primed transcription model, proposes that a free leader RNA is transcribed from the 3′ end of the genomic-size, negative-strand template RNA and rejoins the template RNA at the downstream intergenic regions to serve as the primer for mRNA transcription (2, 15). If coronavirus primary transcription uses leader-primed transcription, it is not easy to explain why free leader RNAs did not efficiently prime synthesis of the larger subgenomic RNA, because it would seem that according to the leader-primed model, equal amounts of the larger and smaller subgenomic RNAs should be synthesized. However, if one assumes that coronavirus transcription has some similarity to a ρ-independent transcription termination mechanism, in which elongation of RNA is terminated by the formation of a stable hairpin structure of nascent RNA molecules (33), it is still possible to explain the present data by the leader-primed transcription model. It is possible that a stable RNA structure can be formed by the nascent larger subgenomic DI RNA, which should contain the 5′-end leader sequence and the downstream proximal intergenic consensus sequence.

Our data fit an alternative mechanism that we hypothesize here. A scanning transcription factor composed of either a virus-specific protein(s) or a combination of a virus-specific protein(s) and a host protein(s) reads the genomic-size negative-strand RNA from the 5′ end toward the 3′ end. The transcription factor recognizes an intergenic region on the template RNA, generates a stable transcription complex, and initiates subgenomic-size positive-strand RNA transcription; after transcription, the factor resumes scanning of the genomic-size negative-strand RNA in search of the next intergenic region. The intergenic region flanking sequences may help to form the transcription complex. Should an upstream intergenic region (on the positive-strand RNA) be located close to a downstream intergenic region (on the positive-strand RNA), the upstream intergenic region, because of its proximity, may also be involved in the transcriptional complex of the downstream intergenic region. In this case, only the downstream intergenic sequence would be positioned optimally for positive-strand subgenomic RNA transcription. After completion of transcription from the downstream intergenic region, the upstream intergenic region, because of its association with the downstream transcription complex, would still not be free to be recognized by the scanning transcription factor and would be passed over for transcription by the same transcription complex of which it is a part.

Formation of a transcription complex probably requires less than 0.1 kb of the flanking sequence located upstream of the intergenic region; this part of our model is based on the fact that significant inhibition of the larger subgenomic DI RNA was not observed in MS102 and MS124. However, it should be noted that there is an inhibition of the larger mRNA 7 synthesis in strain MHV-S no. 8. MHV-S no. 8 has a greater-than-0.1-kb sequence between its two transcription consensus sequences (44). These two observations are compatible if an upstream intergenic region becomes located "functionally" less than 0.1 kb from the downstream intergenic region; this could happen by virtue of an intervening secondary or tertiary structure; in this case, intergenic sequences separated by a primary sequence of more than 0.1 kb might be rendered

proximal and the upstream sequence could be incorporated within the transcriptional complex of the downstream intergenic region.

The structure of the transcription complex may be affected by the sequence and/or structure of upstream flanking sequence. Perhaps this is why the molar ratio of the larger subgenomic DI RNA to the smaller subgenomic DI RNA of MS23AU was consistently lower than that of MS23. Furthermore, the reduced inhibition of transcription of the larger subgenomic DI RNA in MS23DC2G may be explained by the hypothesis that the transcription factor does not efficiently recognize the mutated consensus sequence, UGUAAAC, from which subgenomic DI RNA transcription activity is reduced significantly (13). If free leader RNA recognizes the transcription complex and primes subgenomic mRNA transcription and if the efficiency of primary transcription is determined by the structure of the transcription complex on the negative-strand RNA template, then our scanning transcription factor hypothesis is consistent with the leader-primed transcription model.

ACKNOWLEDGMENTS

We thank Young-Nam Kim for performing some of the preliminary work.

This work was supported by Public Health Service grants AI29984 and AI32591 from the National Institutes of Health.

REFERENCES

1. **Ahlquist, P., E. G. Strauss, C. M. Rice, J. H. Strauss, J. Haseloff, and D. Zimmern.** 1985. Sindbis virus proteins nsP1 and nsP2 contain homology to nonstructural proteins from several RNA plant viruses. J. Virol. **53:**536–542.
2. **Baker, S. C., and M. M. C. Lai.** 1990. An in vitro system for the leader-primed transcription of coronavirus mRNAs. EMBO J. **9:**4173–4179.
3. **French, R., and P. Ahlquist.** 1988. Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus. J. Virol. **62:**2411–2420.
4. **Goldbach, R., O. LeGall, and J. Wellink.** 1991. Alpha-like viruses in plants. Semin. Virol. **2:**19–25.
5. **Goulden, M. G., G. P. Lomonosoff, J. W. Davies, and K. R. Wood.** 1990. The complete nucleotide sequence of PEBV RNA2 reveals the presence of a novel open reading frame and provides insights into the structure of tobraviral subgenomic promoters. Nucleic Acids Res. **18:**4507–4512.
6. **Haseloff, J., P. Goelet, D. Zimmern, P. Ahlquist, R. Dasgupta, and P. Kaesberg.** 1984. Striking similarities in amino acid sequence among nonstructural proteins encoded by RNA viruses that have dissimilar genomic organization. Proc. Natl. Acad. Sci. USA **81:**4358–4362.
7. **Hertz, J. M., and H. V. Huang.** 1992. Utilization of heterologous alphavirus junction sequences as promoters by Sindbis virus. J. Virol. **66:**857–864.
8. **Hirano, N., K. Fujiwara, S. Hino, and M. Matsumoto.** 1974. Replication and plaque formation of mouse hepatitis virus (MHV-2) in mouse cell line DBT culture. Arch. Gesamte Virusforch. **44:**298–302.
9. **Hofmann, M. A., R.-Y. Chang, S. Ku, and D. A. Brian.** 1993. Leader-mRNA junction sequences are unique for each subgenomic mRNA species in the bovine coronavirus and remain so throughout persistent infection. Virology **196:**163–171.
10. **Hofmann, M. A., P. B. Sethna, and D. A. Brian.** 1990. Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. J. Virol. **64:**4108–4114.
11. **Jeong, Y. S., and S. Makino.** 1992. Mechanism of coronavirus transcription: duration of primary transcription initiation activity and effect of subgenomic RNA transcription on RNA replication. J. Virol. **66:**3339–3346.
12. **Jeong, Y. S., and S. Makino.** 1994. Evidence for coronavirus discontinuous transcription. J. Virol. **68:**2615–2623.

12a. **Jeong, Y. S., and S. Makino.** Unpublished data.

13. **Joo, M., and S. Makino.** 1992. Mutagenic analysis of the coronavirus intergenic consensus sequence. J. Virol. **66:**6330–6337.
14. **Kim, Y.-N., Y. S. Jeong, and S. Makino.** 1993. Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication. Virology **197:**53–63.
15. **Lai, M. M. C.** 1990. Coronavirus: organization, replication, and expression of genome. Annu. Rev. Microbiol. **44:**303–333.
16. **Lai, M. M. C., R. S. Baric, P. R. Brayton, and S. A. Stohlman.** 1984. Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. Proc. Natl. Acad. Sci. USA **81:**3626–3630.
17. **Lai, M. M. C., P. R. Brayton, R. C. Armen, C. D. Patton, C. Pugh, and S. A. Stohlman.** 1981. Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from hepatotropic strain MHV-3. J. Virol. **39:**823–834.
18. **Lai, M. M. C., C. D. Patton, R. S. Baric, and S. A. Stohlman.** 1983. Presence of leader sequences in the mRNA of mouse hepatitis virus. J. Virol. **46:**1027–1033.
19. **Lai, M. M. C., and S. A. Stohlman.** 1978. RNA of mouse hepatitis virus. J. Virol. **26:**236–242.
20. **Lee, H.-J., C.-K. Shieh, A. E. Gorbalenya, E. V. Eugene, N. La Monica, J. Tuler, A. Bagdzhadzhyan, and M. M. C. Lai.** 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology **180:**567–582.
21. **Leibowitz, J. L., K. C. Wilhelmsen, and C. W. Bond.** 1981. The virus-specific intracellular RNA species of two murine coronaviruses: MHV-A59 and MHV-JHM. Virology **114:**39–51.
22. **Levis, R., S. Schlesinger, and H. V. Huang.** 1990. Promoter for Sindbis virus RNA-dependent subgenomic transcription. J. Virol. **64:**1726–1733.
23. **Lin, Y.-J., and M. M. C. Lai.** 1993. Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontiguous sequence for replication. J. Virol. **67:**6110–6118.
24. **Makino, S., and M. Joo.** 1993. Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. J. Virol. **67:**3304–3311.
25. **Makino, S., M. Joo, and J. K. Makino.** 1991. A system for study of coronavirus mRNA synthesis: a regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. J. Virol. **65:**6031–6041.
26. **Makino, S., and M. M. C. Lai.** 1989. High-frequency leader sequence switching during coronavirus defective interfering RNA replication. J. Virol. **63:**5285–5292.
27. **Makino, S., C.-K. Shieh, L. H. Soe, S. C. Baker, and M. M. C. Lai.** 1988. Primary structure and translation of a defective interfering RNA of murine coronavirus. Virology **166:**550–560.
28. **Makino, S., L. H. Soe, C.-K. Shieh, and M. M. C. Lai.** 1988. Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. J. Virol. **62:**3870–3873.
29. **Makino, S., S. A. Stohlman, and M. M. C. Lai.** 1986. Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of the free leader RNA in transcription. Proc. Natl. Acad. Sci. USA **83:**4204–4208.
30. **Makino, S., F. Taguchi, N. Hirano, and K. Fujiwara.** 1984. Analysis of genomic and intracellular viral RNAs of small plaque mutants of mouse hepatitis virus, JHM strain. Virology **139:**138–151.
31. **March, L. E., T. W. Dreher, and T. C. Hall.** 1988. Mutational analysis of the core and modular sequences of the BMV RNA3 subgenomic promoter. Nucleic Acids Res. **16:**981–995.
32. **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 virus, strain A59. Virology **171:**141–148.
33. **Platt, T.** 1981. Termination of transcription and its regulation in the tryptophan operon of E. coli. Cell **24:**10–23.
34. **Raju, R., and H. V. Huang.** 1991. Analysis of Sindbis virus promoter recognition in vivo, using novel vectors with two subgenomic mRNA promoters. J. Virol. **65:**2501–2510.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **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.
37. **Sethna, P. B., M. A. Hofmann, and D. A. Brian.** 1991. Minus-strand copies of replicating coronavirus mRNAs contain antileaders. J. Virol. **65:**320–325.
38. **Sethna, P. 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.
39. **Shieh, C.-K., L. H. Soe, S. Makino, M.-F. Chang, S. A. Stohlman, and M. M. C. Lai.** 1987. The 5′-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. Virology **156:**321–330.
40. **Smirnyagina, E., Y.-H. Hsu, N. Chua, and P. Ahlquist.** 1994. Second-site mutations in the brome mosaic virus RNA3 intercistronic region partially suppress a defect in coat protein mRNA transcription. Virology **198:**427–436.
41. **Spaan, W., H. Delius, M. Skinner, J. Armstrong, P. Rottier, S. Smeekens, B. A. M. van der Zeijst, and S. G. Siddell.** 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. EMBO J. **2:**1939–1944.
42. **Stern, D. F., and S. I. T. Kennedy.** 1980. Coronavirus multiplication strategy. I. Identification and characterization of virus-specified RNA. J. Virol. **34:**665–674.
43. **Stern, D. F., and S. I. T. Kennedy.** 1980. Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. J. Virol. **36:**440–449.
44. **Taguchi, F., T. Ikeda, S. Makino, and H. Yoshikura.** 1994. A murine coronavirus MHV-S isolate from persistently infected cells has a leader and

two consensus sequences between the M and N genes. Virology **198:**355–359.
45. **van der Kuyl, A. C., L. Neelman, and J. F. Bol.** 1991. *cis*-acting elements involved in replication of alfalfa mosaic virus RNAs *in vitro*. Virology **183:**687–694.
46. **Winship, P. R.** 1989. An improved method for directly sequencing PCR material using dimethyl sulfoxide. Nucleic Acids Res. **17:**1266.
47. **Yokomori, K., L. R. Banner, and M. M. C. Lai.** 1992. Coronavirus mRNA transcription: UV light transcriptional mapping studies suggest an early requirement for a genomic-length template. J. Virol. **66:**4671–4678.
48. **Zhang, X., C.-L. Liao, and M. M. C. Lai.** 1994. Coronavirus leader RNA regulates and initiates subgenomic mRNA both in *trans* and in *cis*. J. Virol. **68:**4738–4746.
49. **Zhao, X., K. Shaw, and D. Cavanagh.** 1993. Presence of subgenomic mRNAs in virions of coronavirus IBV. Virology **196:**172–178.