# A cis-Acting Function for the Coronavirus Leader in Defective Interfering RNA Replication

RUEY-YI CHANG, MARTIN A. HOFMANN,† PHIROZE B. SETHNA,‡ AND DAVID A. BRIAN\*

Department of Microbiology, The University of Tennessee, Knoxville, Tennessee 37996-0845

Received 27 June 1994/Accepted 30 August 1994

To test the hypothesis that the 65-nucleotide (nt) leader on subgenomic mRNAs suffices as a 5'-terminal cis-acting signal for RNA replication, a corollary to the notion that coronavirus mRNAs behave as replicons, synthetic RNA transcripts of a cloned, reporter-containing N mRNA (mRNA 7) of the bovine coronavirus with a precise 5' terminus and a 3' poly(A) of 68 nt were tested for replication after being transfected into helper virus-infected cells. No replication was observed, but synthetic transcripts of a cloned reporter-containing defective interfering (DI) RNA differing from the N mRNA construct by 433 nt of continuous 5'-proximal genomic sequence between the leader and the N open reading frame did replicate and become packaged, indicating the insufficiency of the leader alone as a 5' signal for replication of transfected RNA molecules. The leader was shown to be a necessary part of the cis-acting signal for DI RNA replication, however, since removal of terminal bases that destroyed a predicted intraleader stem-loop also destroyed replicating ability. Surprisingly, when the same stem-loop was disrupted by base substitutions, replication appeared only minimally impaired and the leader was found to have rapidly reverted to wild type during DI RNA replication, a phenomenon reminiscent of high-frequency leader switching in the mouse hepatitis coronavirus. These results suggest that once a minimal structural requirement for leader is fulfilled for initiation of DI RNA replication, the wild-type leader is strongly preferred for subsequent replication. They also demonstrate that, in contrast to reported natural mouse hepatitis coronavirus DI RNAs, the DI RNA of the bovine coronavirus does not require sequence elements originating from discontinuous downstream regions within the polymerase gene for replication or for packaging.

Coronaviruses possess a positive-strand polyadenylated RNA genome of approximately 30 kb that undergoes replication entirely within the cytoplasm of infected cells (reviewed in reference 25). During virus replication, not only is a minusstrand copy of the full-length genome generated (the antigenome) but so too are minus-strand counterparts of the six to eight 3'-coterminal subgenomic mRNAs (15, 18, 42, 43). The facts that the subgenomic minus strands (anti-mRNAs) are components of metabolically active replicative intermediates (39) and that subgenomic mRNAs and anti-mRNAs amplify at à rate roughly inversely proportional to their length (18, 43) led to the suggestion that the subgenomic mRNAs might serve as templates for synthesis of subgenomic minus strands that in turn serve as full-length templates for mRNA synthesis, i.e., amplification by the mechanism of replication (18, 43). Subgenomic mRNA amplification via replication, or possibly by transcription from internal promoters on the subgenomic minus strands (5), would in theory facilitate coronavirus replication since it would relieve demands on the low-abundance antigenome for mRNA production, a molecule already needed as a template for genome replication and for making the initial mRNA transcripts (47).

The behavior of mRNAs as replicons would suggest that the 65-nucleotide (nt) leader, encoded only at the genome 5' terminus but found as a common 5'-terminal sequence on all subgenomic mRNAs, suffices as a 5'-terminal cis-acting signal for replication. Assuming that a transfected mRNA molecule

would become translocated to the replicational complex, the sufficiency of the 5' leader as a promoter ought to be testable. In experiments described here, we tested the replicability of a transfected coronavirus subgenomic mRNA but found no evidence of replication, indicating that under these conditions the 5' leader alone is insufficient as a replication signal. These experiments, therefore, leave unresolved the role of the subgenomic mRNAs in mRNA amplification. However, by examining the requirements for the leader on a cloned defective interfering (DI) RNA, whose transfected transcripts do undergo replication, we demonstrate that the 5'-terminal portion of the leader which forms a potential stem-loop is required in cis for replication. We were unable to test the requirement of the stem-loop per se since leaders with stem-loops modified by site-directed mutagenesis underwent rapid reversion to the wild-type leader.

# MATERIALS AND METHODS

Virus and cells. Bovine coronavirus (BCV) Mebus was plaque purified three times, and a stock at passage 5 was used as previously described (18, 27). The human rectal tumor cell line HRT-18 was used in all experiments (18, 27).

Construction of reporter-containing cDNA clones of N mRNA and DI RNA. pNrep2, a reporter-containing clone of the BCV N mRNA, was constructed by ligating three fragments from separate sources, sequentially, into the pGEM3Zf(-) vector (Promega, Madison, Wis.) and modifying the 5' end of the insert by in vitro mutagenesis (Fig. 1). (i) The unmodified 5' end of pNrep2 was derived from a cDNA clone (pBL7) of the (almost complete) leader, the remaining 5' UTR of N mRNA, and the 5' end of the N open reading frame (ORF) (through the XmnI site, 48 nt downstream from the N start codon). To prepare pBL7, cytoplasmic RNA from cells at 6 h

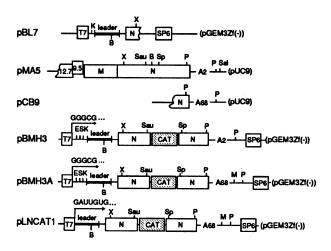
<sup>\*</sup> Corresponding author. Phone: (615) 974-4030. Fax: (615) 974-4007. Electronic mail address: Brian@utkvx.utk.edu.

<sup>†</sup> Present address: Institute of Virology and Immunoprophylaxis, CH-3147 Mittelhausern, Switzerland.

<sup>‡</sup> Present address: Division of Molecular Genetics and Molecular Biology, Burroughs Wellcome Co., Research Triangle Park, NC 27709.

8224 CHANG ET AL. J. VIROL.

#### A. Precursor constructs



#### B. Constructs used for testing replication

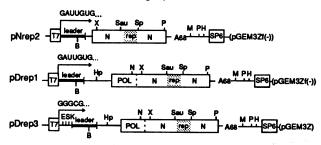


FIG. 1. Plasmids used for the study of BCV mRNA and DI RNA replication. (A) Precursor constructs. Plasmids used in the construction of reporter-containing mRNA (pNrep2), reporter-containing DI RNA (pDrep1), and reporter-containing DI RNA with a modified 5' terminus for making 5'-terminal deletions (pDrep3) are shown. (B) Constructs used for testing replication. Structures of plasmids pNrep2, pDrep1, and pDrep3 are shown. Abbreviations: B, BgIII; E, EcoRI; H, HindIII; Hp, HpaI; K, KpnI; M, MluI; N, NdeI; P, PstI; S, SacI; Sal, SaII; Sau, SauI; Sp, SpeI; X, XmnI; CAT, 660-nt CAT gene; rep, 30-nt TGEV N-specific oligonucleotide reporter. The first bases synthesized by the T7 RNA polymerase promoter during transcript synthesis are indicated. The 65-nt leader is indicated by a heavy solid line.

postinfection and the XmnI site-containing oligonucleotide 5'-CCAGAACGATTTCCAAAGGACGCTCT-3' (primer 1, complementary to nt 34 to 59 in the N ORF on plus-strand RNA) were used for cDNA synthesis and the EcoRI and KpnI site-containing oligonucleotide 5'-cggaattcggtacctTTGCGAG CGATTTGCGTGCG-3' (primer 2, the minus-strand [antileader]-binding primer) and primer 1 were used for DNA amplification by PCR. (In primer 2, the 5'-most capital letter is now known to be base 3 in the BCV leader and base 6 in the leader is now known to be U, not C [13, 17].) The 149-nt PCR product was digested with EcoRI and XmnI, filled in with Klenow enzyme (regenerating the XmnI site), and ligated into EcoRI-digested and filled in pGEM3Zf(-) vector, forming pBL7. (ii) The middle section of pNrep2 (the 1,233-nt XmnI-PstI fragment) was derived from clone pMA5 (27), a 2,815-nt clone in the pUC9 vector that contains the 3' 2,815 nt of the BCV genome (i.e., all of the 9.5-kDa protein, M, and N genes, the 3' UTR, and a 3' A tail of 2 nt). Before its use in subcloning, however, the 660-nt chloramphenicol acetyltransferase (CAT) gene including 119 nt of flanking sequence with a BamHI linker at each end (P. L. Biochemicals) was cloned

into the unique BglII site in MA5, forming pBNC5 (not shown). By joining the 2,400-nt CAT-containing XmnI-SalI fragment of pBNC5 with the 129-nt KpnI-XmnI leader-containing fragment of pBL7 and cloning this into KpnI and SalI-linearized pGEM3Zf(-), pBMH3 was formed. (iii) The 3' end of pNrep2 was derived from the 3' end of the pCB9, a 1,100-nt 3'-end clone of the BCV genome (27), with a poly(A) tail of 68 nt, that had been modified to contain a MluI site at the 3' end of the poly(A) tail. To insert the MluI site, an MluI site-containing oligonucleotide, 5'-A<sub>10</sub>CGCGTC<sub>10</sub>-3' (primer 3), which binds to the minus strand of pCB9 at the base of the poly(A) tail, its complement (primer 4), oligonucleotide 5'-TGGGAATCTTGACGAGCCCCAGAAGG3' (primer 5), which binds to minus-strand RNA beginning 266 nt upstream of the PstI site in the N ORF, and the plus-strand-binding pUC9 "reverse" (P. L. Biochemicals) primer were used for PCR-directed mutagenesis. The PstI-digested PCR product, a 422-nt fragment containing a poly(A)<sub>68</sub> tail and MluI site, was ligated into the 4.8-kb PstI-digested fragment of pBMH3 to create pBMH3A. (iv) The 5' end of pBMH3A was modified by the loop-out mutagenesis method of Kunkel (24) to contain the proper 5' terminus of the BCV leader (5'-GAUUGUG...) (13, 17) immediately downstream from the T7 promoter. The resulting clone was pLNCAT1. (v) To assemble the final pNrep2 construct, the CAT gene in pLNCAT1 was replaced with the in-frame 30-nt transmissible gastroenteritis virus (TGEV) Nspecific oligonucleotide. For this, the 26-mer oligonucleotide 5'-TCTGGGTTGCCAAGGATGGTGCCATG-3' (representing nt 338 to 363 in the TGEV N ORF [21], called TGEV oligonucleotide 7) and its complement (TGEV oligonucleotide 8) were annealed and ligated into the Bg/II-cut, filled-in pMA5. The TGEV oligonucleotide-containing SauI-SpeI fragment was used to replace the analogous fragment in pLNCAT1, thus creating pNrep2.

pDrep1 was constructed by cDNA cloning the 5' 557 nt of the DI RNA (most of which is shown in Fig. 2B) with the N-binding primer 1 (described above) and the *Eco*RI and *Kpn*I site-containing antileader-binding primer 5'-cggaattcggtaccGA TTGTGAGCGATTTGCGTGCG-3' (primer 6) and replacing the 78-nt *Bgl*II-*Xmn*I fragment of pNrep2 with the cloned 499-nt *Bgl*II-*Xmn*I fragment of the DI RNA (Fig. 1).

pDrep3 was constructed by modifying the 5' end of pDrep1 (Fig. 1). (i) By using the *NdeI* site-containing oligonucleotide 5'-CCTCCAAATCATATGGACGTGTATTC-3' (primer 7) that binds to the positive strand of nt 246 to 271 downstream from the DI ORF start codon, antileader-binding primer 6, and virion RNA, a blunt-ended fragment was generated by PCR with Pfu DNA polymerase, and this was cloned into *SmaI*-linearized pGEM3Z to form pGDI5 (not shown). (ii) The 2,109-nt *HpaI-HindIII* fragment from pDrep1 was used to replace the *HpaI-HindIII* fragment of pGDI5, yielding pDrep2 (not shown). (iii) From this, the 15-nt *KpnI-KpnI* fragment within the multiple cloning region was removed, and the resultant clone was named pDrep3.

In vitro mutagenesis of cloned reporter-containing DI RNA. To prepare a set of leader 5' deletions of different lengths, pDrep3 was linearized with EcoRI and digested with Bal 31 exonuclease (0.2 U/ $\mu$ g of DNA at 30°C), and fragments collected at 2 through 15 min were blunt ended with mung bean nuclease and digested with HindIII, which cuts within the multiple-cloning region downstream of the poly(A) tail. Gelpurified fragments were ligated into pDrep3 that had been linearized with EcoRI, blunt ended with mung bean nuclease, and cut with HindIII. Clones with 5'-terminal leader deletions of 4, 8, 13, and 50 nt were used for replication studies.

To disrupt stem-loop 1 in pDrep1 transcripts, pLM1 (in

which nt 5 through 11 were replaced with CACTCGC) and pLM2 (in which nt 30 through 37 were replaced with GCGA GTG) were generated. For this, a modification of the PCRbased gene splicing by overlap extension technique (19) was used. For making pLM1, the gel-purified 232-nt PCR product from a pDrep1-templated reaction with primer 8 (5'-GTTGA TCTTCGACATTGTGACC-3'), which binds to plus-strand nt 204 through 225, and primer LM1(-) (5'-CACTATAGATT CACTCGCATTTGCGTGC-3'), which (except for mutated [underlined] bases) binds to minus-strand nucleotides -7 through 21, was used with the gel-purified 738-nt product of a reaction involving leader plus-strand-binding primer (5'-AC AAGAGATCAGTGAAGCGG-3') (which binds to nt 29 through 48) and the NdeI site-containing vector-binding primer 9 (5'-GAGTGCACCATATGCGGTGT-3') (which binds upstream of the T7 polymerase promoter) in an overlap PCR. In the overlap, the 232- and 738-nt fragments were used together with primers 8 and 9 to form a product of 915 nt, from which the 660-nt NaeI vector-HpaI insert fragment was used to replace the analogous region in pDrep1, generating pLM1.

pLM2 was similarly generated, except that the overlap PCR involved the 460-nt fragment from a reaction with the LM2(-) primer (5'-CGTGCATCCGCGACTGTGATCTCTTG-3', where mutated bases are underlined) and the NdeI-containing primer 7, the 915-nt fragment from a reaction with primers 8 and 9, and primers 9 and 7 themselves to generate a 1,170-nt fragment. From this, the 660-nt NaeI-HpaI fragment was used to replace the analogous region in pDrep1, generating pLM2, or the analogous region in pLM1, generating pLMC, the compensatory double mutant.

Assay for replication: transfection and Northern blotting analysis. Transfections were done by the method of Felgner et al. (10) essentially as described by Rice et al. (37). Briefly, 5 µg of MluI-linearized DNA, blunt ended with mung bean nuclease, was transcribed with 40 U of T7 RNA polymerase in a 100-µl reaction mixture. Capping with <sup>7m</sup>GpppG, when done, followed the procedure of the manufacturer (Promega). The transcription reaction mixture was treated with 5 U of RNasefree DNase (Promega), and the RNA was chromatographed through a Biospin 6 column (Bio-Rad). Cells in 35-mm dishes at 50 to 80% confluency (approximately 106 cells) were infected with BCV at a multiplicity of approximately 5 PFU per cell by incubating cells with inoculum at 37°C for 1 h, refeeding with 2 ml of growth medium containing 10% fetal bovine serum, and incubating at 37°C for 1 h. For transfection, each dish of cells was rinsed three times with Ca2+- and Mg2+-free phosphate-buffered saline (PBS) and treated for 10 min at 25°C with 200 μl of PBS containing 10 μg of lipofectin (Bethesda Research Laboratories) and 1 µg (for the experiments in Fig. 3A and C) or 200 ng (for all other experiments) of transcript. Cells were rinsed with 2 ml of Dulbecco's modified Eagle's medium and incubated at 37°C with 2 ml of medium containing 10% fetal bovine serum until cytoplasmic RNA extraction was done at the indicated times.

For serial passage of progeny virus, supernatant fluids were harvested at the indicated times and 500 µl was used directly to infect confluent cells in a 35-mm dish. Cytoplasmic RNA was extracted for analysis at 48 h postinfection.

The yield of cytoplasmic RNA from one 35-mm dish was approximately 10  $\mu$ g; 2.5  $\mu$ g was used per well in a formaldehyde-agarose gel. Denaturing RNA electrophoresis, RNA blotting, and hybridization were done as previously described (18, 43). From 1 to 100 ng of transcript was loaded per lane when used as marker, probe-specific activities ranged from 1.5  $\times$  10<sup>6</sup> to 3.5  $\times$  10<sup>6</sup> cpm/pmol, and film exposure times ranged from 6 h to 1 week. Quantitation in Northern (RNA) analyses

was done with either the radioanalytic imaging system (AM-BIS, Inc., San Diego, Calif.), or the imaging densitometer (Bio-Rad).

Assay for leader reversion: RNase protection assay. To generate minus-strand RNA probes of the pLM1 and PLMC leader sequences, pLM1 and pLMC were truncated to yield probes of 275 nt. For this, the *HpaI-HindIII* fragment [containing the entire Pol-N ORF and poly(A) tail] was removed from each, the *HindIII* site was blunt ended by fill-in with Klenow enzyme, and the clones were religated to form pLM1D and pLMCD. <sup>32</sup>P-labeled riboprobes were generated with SP6 polymerase after linearization of the plasmids with *PvuII*, which cuts within the vector.

For the RNase protection assay, essentially the method of Vidal and Kolakofsky (46) was used. Approximately 60,000 cpm of riboprobe purified by electrophoresis was coprecipitated in ethanol with 5 µg of cytoplasmic RNA or 100 ng of in vitro generated RNA transcripts. Dried RNA was dissolved in 9 μl of 0.1% N-lauroylsarcosine in water, heated for 1 min at 90°C, quick chilled, and made 0.1% N-lauroylsarcosine, 0.3 M NaCl, 20 mM Tris-HCl [pH 7.4], and 2 mM EDTA ( $1\times$ annealing buffer) by addition of 1 µl of 10× annealing buffer. After annealing for 30 min at 70°C, 50 µl of 0.1% sodium dodecyl sulfate and (SDS)-40  $\mu g$  of RNase A per ml in 2.5× annealing buffer was added. After 60 min at 30°C, 7 µl of 1% SDS and 2.1 µl of proteinase K (20 mg/ml) were added, and the mixture was incubated for 30 min at 30°C, phenol-chloroform extracted after the addition of 5 µg of tRNA, and ethanol precipitated. RNAs were dissolved in 10 µl of 80% formamide plus dyes, heated for 2 min at 90°C, and separated on a 6% polyacrylamide sequencing gel at 5 µl per lane.

Enzymes. Enzymes were purchased from New England Biolabs, Promega Biotec, and Stratagene, La Jolla, Calif.

Other methods. Molecular biological manipulations were carried out by standard methods (38). All constructs were confirmed by sequencing through junction regions (14, 38). RNA structural analyses were done with the aid of the Microgenie program (Beckman Instruments), which uses the RNA-folding algorithm of Tinoco et al. (44).

Nucleotide sequence accession number. The GenBank/EMBL accession number is U00735.

## **RESULTS**

The 5' cis-acting signal for RNA replication includes more than the leader. To directly test the hypothesis that the leader is sufficient as a 5'-terminal signal for coronavirus RNA replication (18, 43), a clone of BCV N mRNA (40) with a precise 5'-terminal sequence (13, 17) on the 65-nt leader, a 3' poly(A) tail of 68 nt, and an in-frame 30-nt reporter sequence within the N ORF (construct pNrep2) was used to prepare capped RNA transcripts in vitro, which were then tested for replication by transfection into infected cells. This system was based on the successful use of trans-acting RNA replicational machinery in the analysis of replicating transfected subgenomic (DI) RNA molecules of picornaviruses (12), togaviruses (29), and coronaviruses (33, 45). In our system, replication was monitored by Northern analysis with a strand-specific probe for the reporter sequence. When synthetic capped transcripts of the N mRNA construct (pNrep2) were transfected into infected cells and cytoplasmic RNA extracts were examined, no increase in the abundance of plus strands was observed and minus strands were not found (data not shown). Likewise, no evidence of replication was found when the CAT gene (not in frame with the N ORF) or a 36-nt portion of the CAT gene (in frame with the I ORF but not the N ORF) was

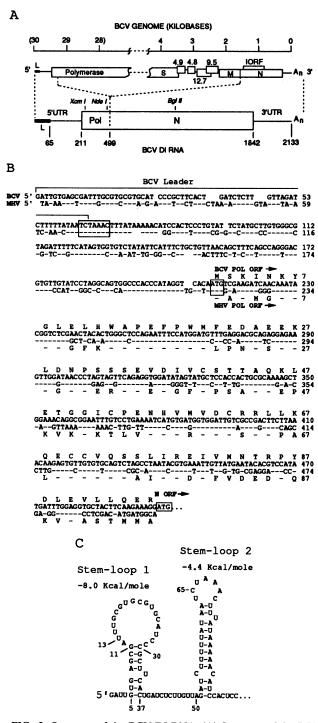


FIG. 2. Structure of the BCV DI RNA. (A) Structure of the BCV DI RNA relative to the BCV genome. Abbreviations: L, leader; S, spike protein gene; M, membrane protein gene; N, nucleocapsid protein gene; IORF, internal ORF. The total length of the BCV genome is unknown; therefore, parentheses indicate the estimated 28-, 29-, and 30-kb position markers. Base 499 of the DI RNA is the first base in the N ORF which forms a contiguous ORF with the first 288 nt of the polymerase gene. Unique *XcmI* and *NdeI* sites within the Pol portion of the ORF and the *BgIII* site within the N portion of the ORF into which the reporter sequence was inserted are shown. (B) The aligned 5' ends of the BCV DI RNA (which is the same as the BCV genome to base 499 [16]) and MHV genome. The BCV DI RNA nucleotide sequence and corresponding ORFs are shown above, and

used in place of the TGEV oligonucleotide reporter sequence (data not shown).

To determine what sequence in addition to the 5' leader might be required to promote RNA replication, the 5' 557-nt terminus of a DI RNA, known from Northern blot experiments to contain the leader and at least a portion of the N mRNA (reference 18 and data not shown) was cloned and tested for its ability to provide replicational competence to the reportercontaining mRNA (described below). The naturally occurring DI RNA is intermediate in size between the N and M mRNAs and was discovered during a study on persistently infected cells (18). Sequence analysis of the cloned 5'-terminal 557 nt of the DI RNA demonstrated the first 498 nt to be identical to the 5' terminus of the BCV genome (16) and base 499 to be the beginning of the N ORF (40). Thus, the natural BCV DI RNA is apparently a simple recombinant between the genomic 5'-terminal 498 nt and the N mRNA (Fig. 2A), but this remains to be confirmed through sequence analysis of the 3'-proximal portion of the natural DI RNA. A 288-nt ORF, identified as a portion of the Pol ORF in Fig. 1B and 2A and B, begins at nt 211 in the BCV DI RNA and is contiguous with the whole of the N ORF. Sequence comparison between the BCV and mouse hepatitis virus (MHV) JHM genomic 5'-terminal 498 nt (28) (Fig. 2B) shows a nucleotide sequence identity of 64% and the putative polymerase ORFs beginning at the same relative position on the genome. For the 96-amino-acid portion of the polymerase proteins, there is a 48% amino acid sequence similarity between BCV and MHV, compared with a similarity of 60, 87, and 70%, respectively, for the S, M, and N proteins of the two viruses (1, 27).

By using the BglII site within the leader and the XmnI site 49 nt downstream from the start of the N ORF (Fig. 1B), the 77-nt 5' UTR in the reporter-containing N mRNA construct (pNrep2) was replaced with the 5'-terminal 498-nt sequence upstream from the start of N ORF in the DI RNA to form pDrep1. When synthetic uncapped transcripts from MluIlinearized pDrep1 were transfected into BCV-infected cells, replication was observed, as demonstrated by a 11-fold increase in abundance of plus-strand RNA through 18 h posttransfection (Fig. 3A, lanes 11 to 14) and by the appearance of minus-strand RNA that peaked at 6 h posttransfection (lanes 18 to 21). Both the abundance of minus strands relative to plus strands and the kinetics of minus-strand accumulation and subsequent decay reflected patterns observed for genomic and subgenomic minus-strand RNAs in BCV-infected cells (18). Successful pDrep1 replication under conditions in which there was no pNrep2 replication could not be explained by a difference in RNA stabilities, since both had a similar decay rate following transfection into uninfected cells, a two- and threefold decrease, respectively, between 2 and 12 h (Fig. 3B). It was further demonstrated that transcripts used for transfection need not be capped, since capped and uncapped transcripts of pDrep1 replicated equally well (Fig. 3C). Thus, capping was not done in experiments described below. One notable feature of pDrep1 is that it underwent efficient packaging, as evidenced by virion transmission for four serial passages (Fig. 3D).

Taken together, these experiments indicate that the leader sequence of 65 nt is not sufficient as a 5' cis-acting signal to

those of MHV are shown below. The position of the 65-nt BCV leader is indicated. The consensus intergenic sequence and the start codons for the DI and N ORFs are boxed. (C) Predicted secondary structures at the 5' end of the BCV DI RNA (and genome).

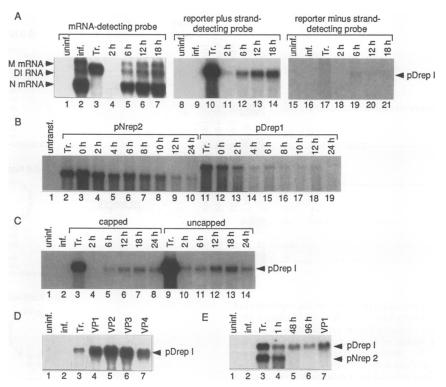


FIG. 3. RNA-blotting evidence for replication and packaging of the reporter-containing BCV DI RNA (pDrep1). (A) Evidence of pDrep1 replication. Denaturing RNA electrophoresis and RNA blotting were done as described in Materials and Methods. In lanes 1 to 7, RNA was probed with a 5'-end-labeled oligonucleotide that binds to plus strands of the N mRNA (nt 35 to 60 in the N ORF, named primer 1 in Materials and Methods), and all larger mRNAs in the nested set. Lanes: 1, RNA from uninfected cells; 2, RNA from infected cells at 26 h postinfection; 8, 10, and 17, pDrep1 transcript marker; 4 to 7, cytoplasmic RNA from cells infected and then transfected (times indicate hours posttransfection); 8 to 14, RNA probed with TGEV oligonucleotide 8, which detects the positive strand of the reporter sequence; 15 to 21, RNA probed with TGEV oligonucleotide 7, which detects the negative strand of the reporter sequence. (B) Evidence of pNrep2 and pDrep1 stability in uninfected cells. Transcripts were transfected, and cytoplasmic RNA was extracted at the times indicated. Tr., transcript marker. RNA extracts were analyzed to detect plus-strand reporter sequence. (C) Evidence that both uncapped and capped transcripts of pDrep1 replicate. Tr., transcript marker. RNA extracts were analyzed to detect plus-strand reporter sequence. (D) Evidence that pDrep1 is packaged. Virus produced by cells transfected with pDrep1 was used to infect new cells, and cytoplasmic RNA from these (VP1) was analyzed at 48 h postinfection. Supernatant virus from this and subsequent infections was likewise passaged (VP2 through VP4). Tr., transcript marker. RNA extracts were analyzed to detect plus-strand reporter sequence. (E) Evidence that transfected mRNA transcripts (pNrep2) fail to replicate when cotransfected with DI RNA (pDrep1) transcripts. Tr., pDrep1 and pNrep2 transcript markers; uninf., uninfected; inf., infected. RNA extracts were analyzed to detect plus-strand reporter sequence.

promote replication of a transfected molecule and that some or all of the additional 433-nt sequence between the leader and the start of N in the DI RNA is required in cis for replication. To reexamine whether the N mRNA would replicate in the presence of replicating DI RNA, evidence of pNrep2 replication was sought after pNrep2 and pDrep1 were transfected together (Fig. 3E, lanes 4 to 7). Still no replication of pNrep2 was evident, indicating again the insufficiency of the leader alone to provide the 5' cis-acting signal for replication of a transfected mRNA molecule. It should be noted that although the amount of RNA taken up by cells, as determined by measurements at 1 or 2 h posttransfection, varied among experiments and was not entirely controllable (Fig. 3A to C and E) (see below), evidence of replication in all transfection experiments was based both on an abundant or increased amount of plus-strand RNA after long-term incubation (24, 48, or 96 h) and on the presence of plus-strand RNA at 48 h postinfection from passage 1 progeny virus (VP1).

To date, the 3' limit of the 5'-terminal cis-acting signal for pDrep1 replication has not been identified. A limited number of deletion mutations located within the 210-nt 5' UTR and downstream of the 65-nt leader have all prevented replication

(data not shown), suggesting that the entire 5' UTR may be a part of the signal. Interestingly, frameshift mutations within the Pol-N fusion ORF at the unique *XcmI*, *NdeI*, (Fig. 2A), and six other sites (6) that lead to premature termination of translation of the DI RNA ORF also disrupted replication ability, suggesting a mechanistic linkage between translation and replication (data not shown).

5'-terminal deletions of the leader disrupt a predicted intraleader stem-loop and also destroy replicating ability. To determine whether the leader is required in cis for replication of the DI RNA, 5'-terminal deletions were made and the mutants were tested for replication (Fig. 4A). To generate a nested set of deletions with Bal 31, pDrep1 was modified to pDrep3 by insertion of a 22-nt sequence containing EcoRI, SacI, and KpnI restriction endonuclease sites between the T7 promoter and the first base of the leader. By using EcoRI-linearized DNA to prepare Bal 31 digestions, clones containing deletions of 4, 8, 13, and 50 nt were generated. However, it should be noted that for each of these, transcripts began with a 5'-terminal gggcg sequence that was derived from the multiple cloning region of the vector. That is, the pD4 clone has nt 1 to 4 of the leader (GAUU) replaced with the gggcg

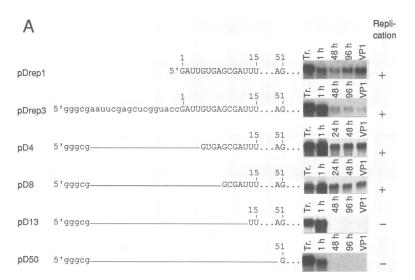
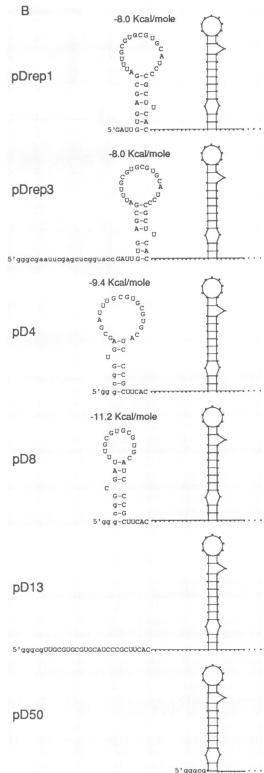


FIG. 4. Effects of 5'-terminal leader deletions on the replication of pDrep1 and on the predicted stem-loop 1 structure. (A) Deletion sites and evidence of replication. Lowercase letters represent bases derived from the multiple-cloning region of the vector. Deleted bases are indicated by a solid line. Tr., transcript marker. DI RNAs were identified by RNA-blotting analysis with probe that detects plus-strand reporter sequence. (B) Predicted stem-loop 1 structures in the deletion mutants.

sequence, and so on (Fig. 4A). It should also be noted that the predicted stem-loop I of pDrep1 with a free energy -8.0kcal/mol (-33.5 kJ/mol) and a bulging U at position 34 remains unchanged in pDrep3 but undergoes modifications to become a stem-loop of -9.4 kcal/mol (-39.3 kJ/mol) with a bulging U at original base position 6 for pD4 and one of -11.2kcal/mol (-46.9 kJ/mol) with a bulging C at original base position 10 for pD8. For pD13 and pD50, the predicted stem-loop disappears (Fig. 4B). Whereas transcripts of pD4 and pD8 replicated almost as well as those from pDrep1, transcripts from pD13 and pD50 showed no evidence of replication (Fig. 4A). Thus, although the 5'-terminal portion of the leader up through 8 nt was dispensable, the tolerated limit for this truncation lies somewhere between nt 8 and 13 from the 5' end. Whether the replicating molecules resulting from transfection with D4 and D8 are revertants with a wild-type leader (as described below for stem-loop mutants) was not tested.

Intriguingly, pDrep3 with a 5' extension of 22 nt also replicated but accumulated at a lower rate (Fig. 4A; data not shown). Whereas RNA-blotting experiments demonstrated the additional sequence to be missing in progeny DI RNA propagated from passaged virus (data not shown), 5'-terminal extensions were found in a small fraction of naturally occurring BCV mRNAs (17) and similar extensions have been shown to be tolerated in replicating molecules of other plus-strand RNA viruses (4).

Leader with stem-loop 1 disrupted by base substitutions is rapidly replaced by wild-type leader. The concurrent loss of replicating ability and stem-loop 1 suggested that this stem-loop or its variants (Fig. 4B) might be an important structural element in the 5' cis-acting signal for replication. To test this, stem-loop 1 in pDrep1 was disrupted by site-directed mutagenesis to each half of the stem independently, thus creating



pLM1 and pLM2, which were then tested separately for replication. By putting both mutations in the same construct, pLMC, a stem-loop was recreated (without a bulging U, however), and transcripts of this were likewise tested for replica-

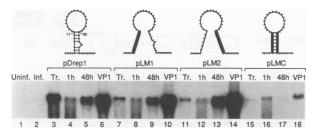


FIG. 5. Effects of substituted bases within stem-loop 1 on the replication of pDrep1 RNA. Substituted regions in pLM1, pLM2, and pLMC are shown by a heavy solid line. Tr., transcript marker. DI RNAs were identified in RNA-blotting analysis with probe that detects plus-strand reporter sequences.

tion. Surprisingly, the replicating ability of pLM1 and pLM2 did not appear impaired relative to that of pDrep1 (Fig. 5, lanes 3 to 14), although the compensatory mutant pLMC showed some impairment (lanes 15 to 18). All three mutants became packaged and replicated during passage 1 of progeny virus on fresh cells (lanes 10, 14, and 18).

To determine whether the leaders on replicating molecules resulting from transfection with mutants had retained their mutant genotype or had possibly reverted to the wild type, an RNase protection assay was done on each mutant at 1 and 48 h posttransfection and at 48 h after passage of progeny virus on fresh cells (Fig. 6). Surprisingly, all had reverted to the wild-type leader within 48 h of transfection and maintained this genotype in the progeny of packaged DI RNA (Fig. 6A, lanes 8 to 10; Fig. 6B, lanes 9, 10, 13, and 14).

### **DISCUSSION**

In experiments described above, we directly tested the hypothesis stated previously (18, 42, 43) that coronavirus subgenomic mRNAs behave as replicons in the presence of viral polymerase(s) and other cofactors. Specifically, we addressed the question whether the 5'-terminal untranslated sequence (a leader of 65 nt plus an additional 12 nt upstream from the start codon) of the N mRNA (mRNA 7; the shortest, most 3'-terminal species [40]) is sufficient to provide replicational competence to the mRNA molecule when introduced into infected cells by transfection. An important part of our experimental design was to be certain that the 5' terminus on the transfected transcripts was identical to the terminus of BCV mRNAs, namely, 5' GA.... Since the 5' terminus of another mammalian coronavirus, the porcine TGEV, also begins with GA... (9), we thought that the 5'-terminal G might be critical for the initiation of mRNA replication. Precise termini are known to be critical for the replication of single-stranded RNAs of other animal viruses (7, 30). In our experiments no evidence of mRNA replication from transfected molecules was found, thereby confirming and extending the studies of Makino et al., whose nonreplicating synthetic mRNA-like transcripts of MHV begin with 5' U. . . (33), a base that aligns with the second position of the BCV leader (17) (Fig. 2B).

The 5'-terminal 498 nt on the BCV DI RNA (including the 65-nt 5'-terminal leader sequence) did impart replicational competence to the transfected reporter-containing N mRNA when it replaced the 5'-terminal 77 nt of the BCV N mRNA. The mechanistic contribution of the additional 433 nt in the DI RNA to the replicational ability of the transfected molecule remains to be determined. Three functional elements that

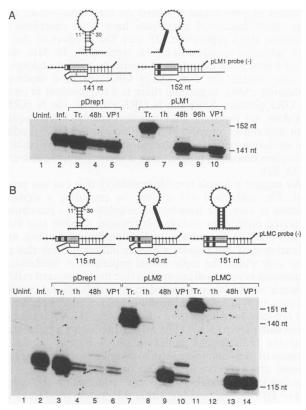


FIG. 6. Evidence of leader reversion for leaders with stem-loop 1 substitutions mutations. Radiolabeled minus-strand riboprobes of pLM1 leader (A) or pLMC leader (B) were used in RNase protection assays with cytoplasmic RNA from infected and transfected cells. Tr., transcript marker. In panel A, the size of the expected protected fragment resulting from leader reversion to wild type is 141 nt. In panel B, the size of the expected protected fragment resulting from leader reversion to wild type is 115 nt.

could be supplied by the additional sequence, thereby making the 5'-terminal replication signal complete, are suggested by studies in other RNA replication systems. (i) The minus-strand complement of the additional sequence may be providing a promoter or a portion of a promoter for plus-strand synthesis not supplied by the antileader alone. While promoters for some animal RNA viruses can be less than 20 nt long (29, 30), raising the possibility that the entire promoter for coronavirus plus-strand synthesis resides entirely within the 65-nt antileader, the full promoter may in fact be larger or even extend to discontinuous downstream regions. Since virtually all mutations we have made downstream of the leader destroyed replicating ability, we are unable to map the 3' terminus of the replication signal or to define the full length of the promoter. It is clear, however, that sequences mapping within discontinuous downstream regions of the polymerase gene and shown to be required for the replication of MHV JHM DI RNAs (22, 32) are not part of the BCV DI RNA replication signal, since the homologous sequences are not in the BCV DI RNA. Moreover, such a sequence appears not to be a universal requirement for MHV DI RNA replication, since a synthetic MHV DI RNA with a structure very similar to pDrep1 (36) can replicate. (ii) The 5'-terminal plus-strand sequence extending downstream of the leader might be binding cellular or viral factors, protein, or RNA, in a complex necessary for the initiation of plus-strand synthesis on the minus-strand template. Such functional complexes have been described for poliovirus RNA replication (2). (iii) The product of the DI ORF might be required in cis for replication. To date, most out-of-frame mutations made at unique restriction endonuclease sites within the 2-kb Pol-N ORF (6) have destroyed replicating ability, suggesting there is a requirement in cis for the ORF product. Since the N ORF exists on the N mRNA that does not replicate after transfection, the ORF requirement may reflect a qualitative need for the Pol ORF product. The cis-acting requirement for coronavirus DI translation in replication remains a puzzle, since it seems not to be universal (8, 23, 32).

We suggest a fourth testable possibility that has less precedent. The additional 433 nt may be providing a signaling function to carry the transfected molecule into a membraneprotected replication complex. It remains possible that coronavirus mRNAs do undergo replication during the virus multiplication cycle but require a precise initiation event that can occur only within the replicational apparatus. A fundamental assumption in our experiments was that the transfected mRNA molecule would be transported to the replicational complex and presented in a way that enables replication. This, however, may not have happened, even when pNrep2 RNA was cotransfected with pDrep1 RNA (Fig. 3E). Evidence for a membraneprotected replication complex in which roughly equimolar amounts of plus- and minus-strand counterparts exist (for both genome and subgenomic species) has been found (41). The targeting function could conceivably reside within the unique part of DI RNA or possibly within the protein product of the truncated polymerase ORF on the replicating DI RNA.

Although the 65-nt leader is insufficient as a 5'-terminal cis-acting signal for replication of the transfected mRNA molecule, it is clear from our experiments that it is a necessary part of the signal for DI RNA replication. Since the terminal deletions of 13 and 50 nt (but not of 8 nt or less) correlated with a loss in the predicted stem-loop 1 (-8 kcal/mol), we postulated that this secondary structure is critical for the cis-acting replication signal. Surprisingly, substitution mutations that disrupted and then restored stem-loop 1 appeared to have little or no effect on the replication of the DI RNA, presumably because of the rapid reversion to the wild-type leader (summarized in Fig. 7). This appears to have resulted from a high-frequency recombination event (26) that is reminiscent of the leader-switching phenomenon described for MHV DI RNA (34) and suggests that once a minimal structural requirement for leader is fulfilled for initiation of DI RNA replication, the wild-type leader is strongly preferred and is selected for subsequent replication. These results, however, leave unanswered the exact role played by the 5'-terminal portion of the leader in initiating replication.

It has been proposed that leader switching on the MHV DI RNA might arise from a leader-priming mechanism in which free leader anneals to the intergenic consensus sequence on the template just downstream of the genomic leader to prime replication (3, 31, 34). This is a variation of the leader-priming model for subgenomic transcription, in which the leader used for priming is provided in *trans* (20, 23, 35). If this is the mechanism of leader switching, one possible role for the leader 5' terminus might be a necessary interaction with the leader-priming machinery. Further analysis of this must await a replicating DI RNA in which leader reversion (switching) does not occur. Such a DI RNA has been reported for MHV (34), and it is one in which a 9-nt sequence (UUUAUAAAC), occurring immediately downstream of the UCUAAAC sequence, has been deleted. An identical sequence is found in an

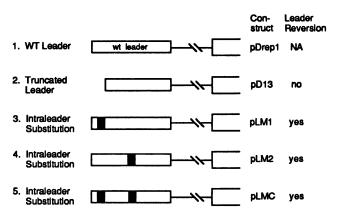


FIG. 7. Mutations and leader reversion. Mutated regions in the 65-nt leader are drawn to scale. For transcripts of pD13 (shown) and pD50 (not shown), no replication was observed, and hence it is presumed that no reversion to wild-type (WT) leader occurred. For transcripts of pLM1, pLM2, and pLMC, leader reversion was demostrated by an RNase protection assay (see the text). Solid regions of leader represent regions of base substitutions. NA, not applicable.

analogous region of the BCV DI RNA, but we do not know if its deletion will prevent leader reversion and yet allow replication.

Finally, perpetuation of a DI RNA may be aided in part by its ability to become packaged (11, 45). In this regard, the BCV DI RNA may differ mechanistically from the MHV DI RNAs (11, 45), for which the packaging signal is described as being a portion of the 3' terminus of the polymerase gene that forms part of the DI RNA mosaic structure. This cannot be the case for the BCV DI RNA, or for the BCV subgenomic mRNAs which also become efficiently packaged (18), since neither possesses a sequence homologous to the identified packaging signal in MHV. The BCV DI packaging signal remains to be characterized.

#### **ACKNOWLEDGMENTS**

We thank Sylvie Rochat for instruction on the RNase protection assay, Rajesh Krishnan for many helpful discussions, and David Hacker for critical reading of the manuscript.

This work was supported primarily by Public Health Service grant AI 14367 from the National Institutes of Health and in part by funds from the University of Tennessee College of Veterinary Medicine Center of Excellence Program for Livestock Diseases and Human Health. M.A.H. was supported by a postdoctoral fellowship from the Swiss National Science Foundation.

# REFERENCES

- Abraham, S., T. E. Kienzle, W. Lapps, and D. A. Brian. 1990. Deduced sequence of the bovine coronavirus spike protein and identification of the internal proteolytic cleavage site. Virology 176:296-301.
- Andino, R., G. E. Rieckhof, P. L. Achacoso, and D. Baltimore. 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5' end of viral RNA. EMBO J. 12:3587–3598.
- 3. 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.
- Ball, L. A., and Y. Li. 1993. cis-acting requirements for the replication of flock house virus RNA 2. J. Virol. 67:3544–3551.
- Brian, D. A., R.-Y. Chang, P. B. Sethna, and M. A. Hofmann. 1994.
   Role of subgenomic minus-strand RNA in coronavirus replication.
   Positive-strand RNA viruses. Arch. Virol. 9(Suppl.):173–180.
- 6. Chang, R.-Y., and D. A. Brian. Unpublished data.
- 7. Collins, P. L., M. A. Mink, and D. S. Stec. 1991. Rescue of

- synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. Proc. Natl. Acad. Sci. USA 88:9663–9667.
- DeGroot, R. J., R. G. Van der Most, and W. J. M. Spaan. 1992. The fitness of defective interfering murine coronavirus DI-a and its derivatives is described by nonsense and frameshift mutations. J. Virol. 66:5898-5905.
- 9. Dombrowski, S. M., and D. A. Brian. Unpublished data.
- Felgner, P. L., T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielsen. 1987.
   Lipofectin: a highly efficient, lipid-mediated DNA-transfection procedure. Proc. Natl. Acad. Sci. USA 84:7413-7417.
- Fosmire, J. A., K. Hwang, and S. Makino. 1992. Identification and characterization of a coronavirus packaging signal. J. Virol. 66: 3522-3530.
- Hagino-Yamagishi, K., and A. Nomoto. 1989. In vitro construction of poliovirus defective interfering particles. J. Virol. 63:5386–5392.
- Hofmann, M. A., and D. A. Brian. 1991. A PCR-enhanced method for sequencing the 5' end of messenger RNAs. PCR Methods Applic. 1:43-45.
- Hofmann, M. A., and D. A. Brian. 1991. Sequencing PCR DNA amplified directly from a bacterial colony. BioTechniques 11:30– 31
- Hofmann, M. A., and D. A. Brian. 1991. The 5' end of coronavirus minus-strand RNAs contain a short poly(U) tract. J. Virol. 65: 6331–6333.
- 16. Hofmann, M. A., R.-Y. Chang, S. Ku, and D. A. Brian. Leader-mRNA junction sequences are unique for each subgenomic mRNA species in the bovine coronavirus and remains so throughout persistent infection. Virology 196:163–171.
- Hofmann, M. A., S. D. Senanayake, and D. A. Brian. A translationattenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. Proc. Natl. Acad. Sci. USA 90:11733-11737.
- 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.
- Horton, R. M., Z. Cai, S. N. Ho, and L. R. Pease. 1990. Gene splicing by overlap extension: tailor made genes using the polymerase chain reaction. BioTechniques 8:528-535.
- Jeong, Y. S., and S. Makino. 1994. Evidence for coronavirus discontinuous transcription. J. Virol. 68:2615–2623.
- Kapke, P. A., and D. A. Brian. 1986. Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. Virology 151:41-49.
- 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.
- Kim, Y.-N., M. M. C. Lai, and S. Makino. 1993. Generation and selection of coronavirus defective interfering RNA with large open reading frame by RNA recombination and possible editing. Virology 194:244–253.
- Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA 82:488– 492.
- Lai, M. M. C. 1990. Coronavirus: organization, replication, and expression of genome. Annu. Rev. Microbiol. 44:303–333.
- Lai, M. M. C. 1992. RNA recombination in animal and plant viruses. Microbiol. Rev. 56:61-79.
- Lapps, W., B. G. Hogue, and D. A. Brian. 1987. Sequence analysis
  of the bovine coronavirus nucleocapsid and matrix protein genes.
  Virology 157:47-57.
- 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.
- Levis, R., B. G. Weiss, M. Tsiang, H. Huang, and S. Schlesinger. 1986. Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. Cell 44:137-145.
- Li, X., and P. Palese. 1992. Mutational analysis of the promoter required for influenza virus virion RNA synthesis. J. Virol. 66: 4331–4338.
- Liao, C.-L., and M. M. C. Lai. 1992. RNA recombination in a coronavirus: recombination between viral genomic RNA and transfected RNA fragments. J. Virol. 66:6117–6124.
- 32. 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 discontinuous sequence for replication. J. Virol. 67:6110–6118.
- 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.
- Makino, S., and M. M. C. Lai. 1989. High-frequency leader sequence switching during coronavirus defective interfering RNA replication. J. Virol. 63:5285-5292.
- 35. Makino, S., S. Stohlman, and M. M. C. Lai. 1986. Leader sequence of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. Proc. Natl. Acad. Sci. USA 83:4204-4208.
- Masters, P., C. A. Koetzner, C. A. Kerr, and Y. Heo. 1994.
   Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in coronavirus mouse hepatitis virus. J. Virol. 68:328-337.
- Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers. 1989.
   Transcription of Yellow Fever RNA from full-length cDNA templates produced by in vitro ligation. New Biol. 1:285-296.
- 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.
- Sawicki, S. G., and D. L. Sawicki. 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in mRNA synthesis. J. Virol. 64:1050-1056.
- Senanayake, S. D., M. A. Hofmann, J. L. Maki, and D. A. Brian. 1992. The nucleocapsid protein gene of the bovine coronavirus is bicistronic. J. Virol. 66:5277-5283.
- 41. Sethna, P. B., and D. A. Brian. Unpublished data.
- Sethna, P. B., M. A. Hofmann, and D. A. Brian. 1991. Minusstrand copies of replicating coronavirus mRNAs contain antileaders. J. Virol. 65:320–325.
- Sethna, P. B., S.-L. Hung, and D. A. Brian. 1989. Coronavirus subgenomic minus-strand RNA and the potential for mRNA replicons. Proc. Natl. Acad. Sci. USA 86:5626–5630.
- 44. Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla. 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. 246:40-41.
- 45. van der Most, R. G., P. J. Bredenbeek, and W. J. M. Spaan. 1991. A domain at the 3' end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. J. Virol. 65:3219-3226.
- Vidal, S., and D. Kolakofsky. 1989. Modified model for the switch from Sendai virus transcription to replication. J. Virol. 63:1951– 1059
- 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.