

5'-Proximal Hot Spot for an Inducible Positive-to-Negative-Strand Template Switch by Coronavirus RNA-Dependent RNA Polymerase[∇]

Hung-Yi Wu and David A. Brian*

Departments of Microbiology and Pathobiology, University of Tennessee, College of Veterinary Medicine, Knoxville, Tennessee 37996-0845

Received 21 August 2006/Accepted 7 January 2007

Coronaviruses have a positive-strand RNA genome and replicate through the use of a 3' nested set of subgenomic mRNAs each possessing a leader (65 to 90 nucleotides [nt] in length, depending on the viral species) identical to and derived from the genomic leader. One widely supported model for leader acquisition states that a template switch takes place during the generation of negative-strand antileader-containing templates used subsequently for subgenomic mRNA synthesis. In this process, the switch is largely driven by canonical heptameric donor sequences at intergenic sites on the genome that match an acceptor sequence at the 3' end of the genomic leader. With experimentally placed 22-nt-long donor sequences within a bovine coronavirus defective interfering (DI) RNA we have shown that matching sites occurring anywhere within a 65-nt-wide 5'-proximal genomic acceptor hot spot (nt 33 through 97) can be used for production of templates for subgenomic mRNA synthesis from the DI RNA. Here we report that with the same experimental approach, template switches can be induced *in trans* from an internal site in the DI RNA to the negative-strand antigenome of the helper virus. For these, a 3'-proximal 89-nt acceptor hot spot on the viral antigenome (nt 35 through 123), largely complementary to that described above, was found. Molecules resulting from these switches were not templates for subgenomic mRNA synthesis but, rather, ambisense chimeras potentially exceeding the viral genome in length. The results suggest the existence of a coronavirus 5'-proximal partially double-stranded template switch-facilitating structure of discrete width that contains both the viral genome and antigenome.

Template switching by RNA-dependent RNA polymerases (RdRps) is a mechanism that contributes to genetic recombination and sequence diversity in RNA viruses (24). Template switching during both positive- and negative-strand RNA synthesis have been documented (20, 30). Interestingly, some positive-strand RNA virus families in the *Nidovirus* order require a template switch for virus replication. Coronaviruses (for a leader of 65 to 90 nucleotides [nt], depending on the coronavirus species) (48) and arteriviruses (for a leader of 160 to 210 nt, depending on the arterivirus species) (41) appear to utilize a positive-to-positive-strand template switch (35, 36) during synthesis of antileader-containing negative-strand templates used in turn for the production of subgenomic mRNAs (sgmRNAs). The template switch in the end results in a leader on each sgmRNA that is identical to that on the viral genome. In toroviruses, only the largest of three sgmRNA species appears to gain an 18-nt leader in common with the genome, possibly by the same mechanism (42). The sgmRNAs for roniviruses do not share a leader sequence with the genome and therefore do not appear to use a discontinuous transcription step (10). In coronavirus-infected cells, sgmRNA-length negative strands (39) containing the antileader sequence (38) are found as components of sgmRNA-length double-stranded transcriptive intermediates (3, 34, 37). In coronaviruses there also exists the phenomenon of leader switching, wherein frequent RdRp template switching occurs near the 5' end of the genome (26), but whether this switch happens during positive- or negative-strand

synthesis has not been established, although models for both have been proposed (8, 26).

In a series of *in vivo* experiments designed to examine the causes of sequence similarity-induced, high-frequency, positive-to-positive-strand template switching associated with sgmRNA synthesis in a bovine coronavirus (BCoV) defective interfering (DI) RNA system (32, 43, 44), one set was designed to test whether a positive-to-negative-strand template switch could be similarly induced, thereby directly demonstrating by the nature of the product that the RdRp was undergoing negative-strand synthesis at the time of the switch. Such a template switch was found and is reported here. The data also reveal that the switch necessarily occurred *in trans* from the positive-strand DI RNA donor to the negative-strand viral antigenome acceptor and that an 89-nt-wide acceptor window, a hot spot on the viral antigenome (nt 35 through 123 from the 3' end), was used. Interestingly, the 89-nt acceptor hot spot is largely complementary to a previously described 65-nt acceptor hot spot on the positive-strand genome (nt 33 through 97 from the 5' end) used for a positive-to-positive-strand template switch (44). In addition, both hot spots overlap *cis*-acting signals for RNA replication (7, 33) (S. Raman and D. Brian, unpublished). The results together lead us to suggest that the coronavirus template switch-facilitating apparatus, perhaps a component of the transcription complex, is a partially double-stranded structure that contains both the viral genome and antigenome.

MATERIALS AND METHODS

Virus and cells. A DI RNA-free stock of the Mebus strain of BCoV (GenBank accession no. U00235) was used as a helper virus on human adenocarcinoma (HRT-18) cells as previously described (7). In some control experiments, the A59

* Corresponding author. Mailing address: Department of Microbiology, University of Tennessee, Knoxville, TN 37996-0845. Phone: (865) 974-4030. Fax: (865) 974-4007. E-mail: dbrian@utk.edu.

[∇] Published ahead of print on 17 January 2007.

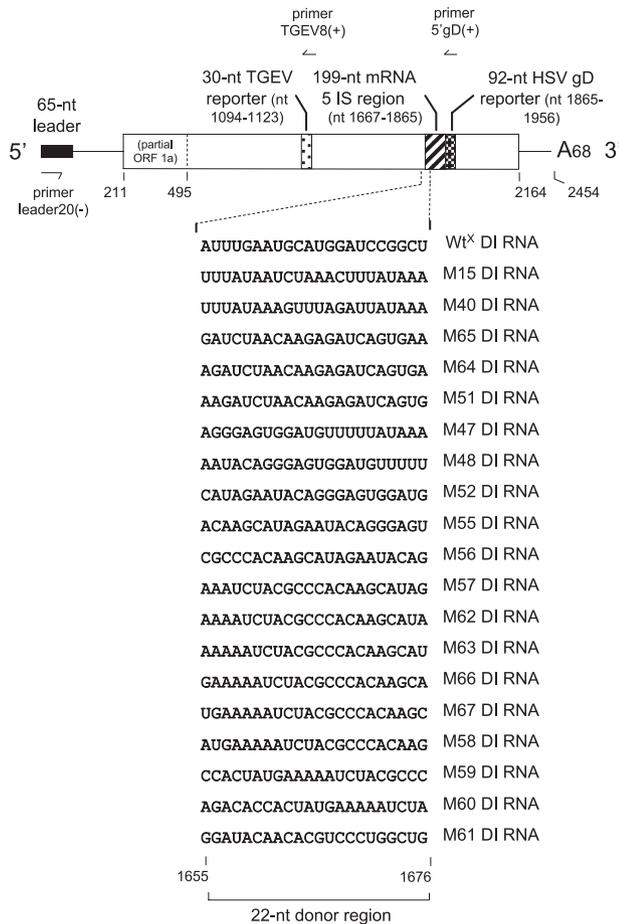


FIG. 1. Structure of reporter-containing BCoV DI RNA donor mutants used for testing positive-to-negative-strand RdRp template switching. Wt^X DI RNA is a cloned naturally occurring DI RNA modified to contain the BCoV intergenic sequence region for sgmRNA 5 and two reporter sequences (44). Shown is the 22-nt donor region (nt 1655 to 1676) within which mutations were made to produce M15 which directs a positive-to-positive-strand template switch (44), and mutants M40 to 67, as noted, tested here for positive-to-negative-strand template switching. The 65-nt BCoV leader sequence is illustrated by a filled rectangle. The 5' and 3' untranslated regions are identified. The binding regions for primers leader20(-), TGEV8(+), and 5'gD(+) are shown.

strain of mouse hepatitis virus (MHV-A59, GenBank accession no. NC 006852) was used as a helper virus on murine delayed brain tumor (DBT) cells for the amplification of the BCoV DI RNA as previously described (43).

Plasmid constructs and transfection with synthetic DI RNAs. Construction of pWt^X from pDrep1 (7) and its mutant 15 derivative (pM15) (Fig. 1 and 2) have

been described previously (44). Mutant 40 and all other mutants used in this study were constructed from pWt^X by the same overlap mutagenesis procedure used for making pM15 except for the use of synthetic oligonucleotides carrying the appropriate mutated 22-nt region (Fig. 1). Mutant DI RNAs, obtained as T7 RNA polymerase transcripts of their respective plasmid linearized at the MluI site [immediately downstream of the poly(A) tail] were transfected into BCoV-infected cells with the use of Lipofectin (Life Technologies) as previously described (44).

Sequence analysis of products of template switching. RNA (10 µg from a 35-mm dish containing ~8 × 10⁶ cells) was extracted at 24 h postinfection (hpi) with Trizol (Invitrogen) from cells that had been infected with passage 1 virus (VP1). VP1 was virus collected at 48 h postinfection (with helper virus) and 47 h posttransfection (with the respective T7 RNA polymerase-generated mutant DI RNA) (44). To determine the junction sequences on potential chimeric ambisense products from all 19 DI RNA mutants shown in Fig. 1, 2.5 µg of RNA in a 20-µl reverse transcription (RT) reaction mix was used with Superscript II RT (Invitrogen) and 0.1 µM primer BCoV1096(+) (5'-CGCACAACGTGCCA TGCCAC-3', which binds nt 1096 to 1115 in the BCoV genome, a sequence not in the DI RNA) to generate the cDNA product. Five microliters of the post-reaction RT mix was used in a 25-µl reaction mix containing 0.8 µM concentrations of each of the nested primers BCoV108(+) (5'-CCACTATGAAAAATC TACGCC-3', which binds nt 108 to 129 in both the BCoV genome and DI RNA) and 5'gD(+) (5'-GAGAGAGGCATCCGCCAAGGCATATTG-3', which binds nt 1866 to 1893 within the gD reporter sequence in all DI RNA mutants) under previously described conditions for 25 PCR cycles (44). PCR products were purified by native agarose gel electrophoresis, DNA from the resolved PCR band, obtained by suction-punching the gel with a micropipette, was cloned into the TOPO XL vector (Invitrogen), and the resulting plasmid DNA was sequenced with M13 universal sequencing primers (44). A minimum of three clones was sequenced for each product. In some experiments, as described in the text, primer BCoV631(+) (5'-GGCCACATGCTTGTAAACAGCGCAA CG-3', which binds nt 631 to 657 in the viral genome, a sequence not in the DI RNA) replaced BCoV108(+) in the above RT-PCR (data are discussed but not shown).

For the RT-PCR control reactions, two sets were done. In the first, reactions were carried out for M40 as described above and illustrated in Fig. 3A for M40, except for the differences noted here and summarized in Fig. 3C, lanes 2, 4, and 6 to 12. For lane 2, the RT reaction used 2.5 µg RNA from cells infected with VP1 obtained from M15-transfected cells at 24 hpi. For lane 4, the RT reaction used a mix of 2.5 µg RNA from BCoV-infected cells at 24 hpi, 1.0 ng of T7 RNA polymerase-generated M40 positive-strand transcript, and 0.1 ng of SP6 RNA polymerase-generated M40 negative-strand transcript. For lane 6, the RT reaction used 2.5 µg RNA from uninfected cells only. For lane 7, the RT reaction used 2.5 µg RNA from BCoV-infected cells at 24 hpi only. For lane 8, the RT reaction used a mix of 2.5 µg RNA from uninfected cells and 500 ng of T7 RNA polymerase-generated M40 positive-strand transcript. For lane 9, the RT reaction used 2.5 µg RNA from BCoV-infected cells at 24 hpi and 500 ng of T7 RNA polymerase-generated M40 positive-strand transcript. For lane 10, PCR only was done in the presence of 500 ng of MluI-linearized pM40 DNA. For lane 11, the RT reaction used 2.5 µg RNA from BCoV-infected cells and PCR was done after the addition of 500 ng of MluI-linearized pM40 DNA. For lane 12, the RT reaction used 2.5 µg RNA from BCoV-infected cells and primers 1096(+) and leader20(-), and PCR was done after the addition of 500 ng of MluI-linearized pM40 DNA. Leader20(-) primer (5'-GATTGTGAGCGATTGCGTG-3') binds nt 1 to 20 at the 3' end of the viral antigenome. In the second set of controls, reactions were carried out for M52, which causes generation of a

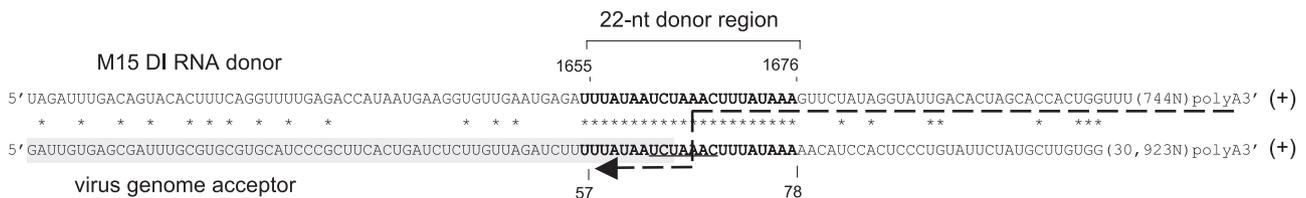


FIG. 2. Induction of a putative positive-to-positive-strand template switch by DI RNA mutant M15 that yields a negative-strand template for synthesis of leader-containing sgRNA. The putative RdRp pathway taken (dashed arrow) during the template switch from M15 which results in a previously characterized full-length (65-nt) leader on sgRNA (44) is shown. The 22-nt donor sequence in M15 DI RNA (at nt 1655 to 1676) is identical to viral genome nt 57 to 78 and contains the canonical core intergenic sequence, UCUAAC. Note that the positive-strand donor and positive-strand acceptor molecules are drawn in parallel orientation. Note also that the template switch could occur anywhere within the region of perfect sequence identity, identified by the continuous stretch of asterisks, and yield the same negative-strand anti-sgmRNA product.

chimeric molecule (lane 13; see Fig. 5B). In this control, RNA from MHV-A59-infected DBT cells was used under conditions described here and summarized in lane 14. MHV-A59-infected DBT cells were shown previously to support the replication of BCoV DI RNA (43). For lane 14, the RT reaction used a mix of 2.5 μ g RNA from DBT cells infected with VP1 obtained from MHV-A59-infected and M52 DI RNA-transfected DBT cells extracted at 24 hpi and 2.5 μ g RNA from BCoV-infected HRT cells. In this case, DBT cells were infected with MHV-A59, at a multiplicity of infection of 1, and transfected at 1 hpi with 500 ng of M52 DI RNA as previously described (43), and VP1 supernatant virus was harvested at 24 hpi.

RT-PCR monitoring of DI RNA and helper virus-derived sgmRNA 7. The same cell-extracted RNA samples used for chimeric RNA analyses were used to monitor the presence of DI RNA and helper virus-derived sgmRNA 7 as evidence of DI RNA amplification and helper virus replication. To identify DI RNA, primer TGEV8(+) (5'-CATGGCACCATCCTTGCAACCCAGA-3'), which binds nt 1 to 25 in the 30-nt TGEV reporter sequence, was used for RT and also with primer leader20(-) for PCR. To identify BCoV sgmRNA 7 in BCoV-infected cells, primer BCoVn29482(+) (5'-GGTTGAACATTCTAG ATTGGTCGGAC-3'), which binds nt 29,482 to 29,509 in the BCoV genome (nt 87 to 114 in the N gene), was used for RT and then also with primer leader20(-) for PCR. To identify MHV-A59 sgmRNA 7 in MHV-A59-infected cells, primer MHV-N(+) (5'-GTTGGGTTGAGTAGTTGCAGTC-3'), which binds nt 29,821 to 29,842 in the MHV-A59 genome (nt 153 to 174 in the N gene), was used for RT and then also with primer MHV-A59leader20(-) (5'-GAGTGAT TGGCGTCCGTACGTACC-3'), which binds nt 6 to 29 in the 3' end of the MHV-A59 antileader for PCR. RT-PCR products were resolved by electrophoresis on a 1% agarose gel along with size markers and stained with ethidium bromide (EtBR).

Oligonucleotide primers used in this study. Note that in all cases the sign + or - in the primer name indicates the polarity of the RNA (or DNA) bound by the primer.

RESULTS

Induction of an RdRp template switch from a positive-strand DI RNA donor to the helper virus negative-strand antigenome acceptor. A previous study (44) had established that an RdRp template switch, presumably positive-strand-to-positive-strand during negative-strand RNA synthesis from a donor site within a BCoV DI RNA (Wt^x DI RNA) (Fig. 1), to an acceptor site within the 5'-proximal 65-nt genomic leader occurred with sufficient frequency to generate an sgmRNA with a 33-nt (mini) leader. By using engineered sequences of 22 nt in length at this internal site in the DI RNA that fully matched sequences within a 5'-proximal region of the viral genome, a 65-nt-wide genomic acceptor hot spot for the switches was identified that mapped between nt 33 and 97 (44). The 22-nt donor and acceptor sequences did not need to include the canonical core intergenic sequence (UCUAAAC) mapping at nt 63 through 69 in the viral genome for successful template switching. One example of an induced positive-to-positive-strand template switch is illustrated by M15, which carried a

22-nt donor sequence identical to nt 57 to 78 in the viral genome (Fig. 2).

To seek direct evidence for similarity-assisted template switching during negative-strand synthesis and at the same time test the possibility of a switch in *trans*, an attempt was made to induce a switch from the positive-strand DI RNA donor to the negative-strand viral antigenome. A chimeric product bearing the predicted sequence would directly demonstrate such a pathway. To perform the experiment, a 22-nt DI RNA donor sequence matching nt 57 to 78 from the 3' end of the viral antigenome was placed into the DI RNA between nt positions 1655 and 1676 to create M40 (Fig. 1 and Fig. 3A). This sequence is complementary to nt 57 to 78 from the 5' end of the viral genome described above to make M15.

A chimeric ambisense molecule resulting from an RdRp template switch from the positive-strand M40 DI RNA donor template to the negative-strand viral antigenome acceptor template (as depicted in Fig. 3A) would consist of the 5'-terminal 800 nt of M40 DI RNA negative strand (i.e., the sequence copied from the 3'-terminal 800 nt of M40 DI RNA positive strand) and the 3'-terminal 30,954 nt of the viral genome (i.e., the sequence copied from the 5'-terminal 30,954 nt of the viral antigenome). Alternatively, if the negative strand of another M40 DI RNA (rather than the viral antigenome) were to function as the acceptor, the resulting chimeric molecule would consist of the 5'-terminal 800 nt of M40 DI RNA negative strand and the 3'-terminal 2,376 nt of M40 RNA positive strand. Northern analysis of RNA at 24 hpi from cells infected with M40 DI RNA-containing VP1 failed to detect molecules of either 31,754 nt in length (a chimera of the first type) or 3,176 nt in length (a chimera of the second type) (data not shown). However, by using a viral genome-specific primer [BCoV 1096(+), which binds nt 1096 to 1115 in the viral genome] for RT followed by a nested primer set [primer BCoV 108(+), which binds viral genomic (and DI RNA) nt 108 to 129, and primer gD(+), which binds the DI RNA-specific herpes simplex virus (HSV) gD reporter sequence] to detect an RNA chimera of the first type, a product of 290 nt was found (Fig. 3A and C, lanes 1, 3, and 5). The 290-nt product was cloned, and the sequence of the upper strand, which reflects the RdRp product, revealed its predicted origin, its chimeric nature, and the RdRp template switch site (Fig. 3B). Note that the template switch could have occurred anywhere within the 22-nt region of identity between the donor and acceptor sites and yield the same fusion site between the negative-

depicted are the binding sites for primers 5'gD(+) and BCoV108(+) used in the nested PCR. The DNA product made from primer 5'gD(+) is represented by a small dashed arrow. Immediately below the third open arrowhead is shown the upper strand of the predicted 290-nt long chimeric PCR product. (B) DNA sequencing results from the cloned 290-nt RT-PCR product shown in lane 5 of panel C. (C) Summary of experimental and control RT-PCR conditions. The conditions are detailed further in Materials and Methods and in Results. The products of the reactions were visualized by native agarose gel electrophoresis and EtBr staining. Lanes 1 through 12 refer to RT-PCR conditions for M40 DI RNA. The 290-nt M40 chimeric product is shown in lanes 1, 3, and 5. Lanes 13 and 14 refer to RT-PCR conditions for M52. The 268-nt M52 chimeric product is shown in lane 13. Molecular length DNA markers are not shown. (D) Comparison of the BCoV antigenome (top) and MHV-A59 antigenome (bottom) in the region from which the BCoV DI RNA M40 and M52 mutant sequences (boldface type) are derived. Asterisks identify identical bases. (E) RT-PCR verification of M52 DI RNA amplification and helper virus replication. The RT-PCR-amplified 1.2-kb product from DI RNA in BCoV-infected cells (lane 1, upper band) and from MHV-infected cells (lane 2, upper band), and the RT-PCR-amplified 180-nt sgmRNA 7 product from BCoV-infected cells (lane 1, lower band) and 247-nt sgmRNA 7 product from MHV-infected cells (lane 2, lower band) are shown.

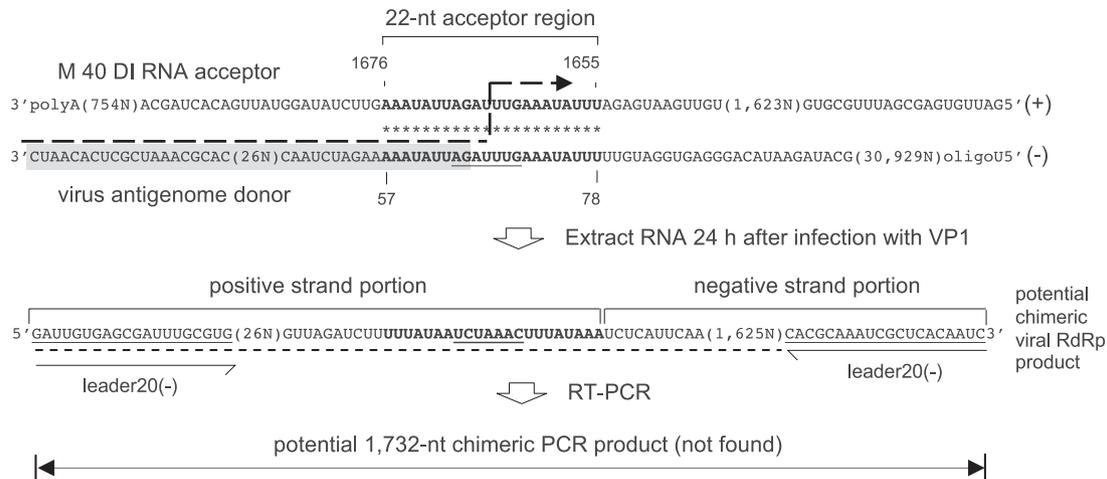


FIG. 4. Failure to induce an RdRp template switch from the negative-strand viral antigenome to the internal site on the positive-strand DI RNA. In this case, the template switch would be *in trans* from the negative-strand viral antigenome (nt 57 to 78, numbering from the 3' end) to the internal acceptor region (nt 1655 to 1676, numbering from the 5' end) on the positive-strand M40 DI RNA (i.e., the reverse direction of that depicted in Fig. 3A). Note that, in this case, the negative-strand donor and the positive-strand acceptor molecules are drawn in an unnatural parallel orientation for ease of illustration. The same VP1 RNA preparation as used for the RT-PCR analysis shown in Fig. 3C, lane 5, was tested by RT-PCR with probe leader20(-) for both RT and PCR. With a negative-to-positive-strand template switch as depicted, a 1,732-nt product would be expected. No product was observed in an EtBr-stained gel (data not shown).

and positive-strand sequence stretches in the ambisense RdRp product.

Since the RNA analyzed by RT-PCR in all cases came from cells infected with progeny virus containing packaged progeny DI RNAs (i.e., VP1) and not from the cells that had been originally transfected with synthetic transcripts, it is unlikely that the T7 RNA polymerase-generated transcripts in the presence of viral genome and antigenome or contaminating plasmid DNA delivered to the parent cells by transfection would have served as templates for an RT-PCR-generated 290-nt chimeric product. Nevertheless, controls for RT-PCR-generated chimeric products were carried out as described in Materials and Methods and as summarized in Fig. 3C, lanes 2, 4, 6 to 12, and 14, and none were found. Of special interest are the control reactions depicted in Fig. 3C, lanes 4, 12, and 14. In lane 4, RNA molecules mimicking those in M40-transfected, BCoV-infected cells were mixed for the RT-PCRs. That is, positive- and negative-strand viral genomic RNA obtained from infected cells and positive- and negative-strand DI RNAs produced by T7 RNA polymerase and SP6 RNA transcripts of pM40, respectively, were mixed before the RT reaction. Since DI RNAs exist at ~200 molecules and viral genomes at ~20 molecules per cell at 24 hpi from VP1 infection as measured by Northern analysis (44), then ~1 pg of positive-strand DI RNA and ~0.1 pg of negative-strand DI RNA exist in 2×10^6 cells, the number of cells yielding 2.5 μ g of RNA used in the RT reaction. A mixture containing these concentrations along with 2.5 μ g RNA from BCoV-infected cells was tested, as well as mixtures containing 10-fold-increasing amounts of DI RNAs up to 1.0 ng of positive-strand RNA and 0.1 ng of negative-strand RNA (the amounts used in lane 4), and none yielded an RT-PCR product. In lane 12, the RT reaction used 2.5 μ g RNA from BCoV-infected cells and primers 1096(+) and leader20(-) to generate negative- and positive-strand cDNA from the 5'-proximal region of the viral genome, respectively,

before the addition of 500 ng of MluI-linearized M40 plasmid DNA and subsequent PCR. In lane 14, DI M52 RNA was amplified intracellularly with MHV-A59 helper virus under conditions previously shown to amplify BCoV DI RNA (43) and then extracted and mixed with BCoV genomic and antigenomic RNA prior to the RT reaction. M52 was chosen for this control, since unlike M40, it shares very little sequence homology with the analogous region of the MHV-A59 antigenome (only 4 nt of 22 as shown in Fig. 3D) and therefore was unlikely to yield an RdRp-generated chimeric product in the DBT cells. In this case, no 268-nt RT-PCR product was found, although RT-PCR showed MHV-A59 to have replicated (Fig. 3E, lane 2, lower panel) and the DI RNA to have amplified (Fig. 3D, lane 2, upper panel). Thus, intracellular-amplified M40 DI RNA could be mixed with intracellular-amplified viral genome prior to the *in vitro* RT-PCRs and no RT-PCR product was generated. It should be noted that chimeric DNA molecules that might have arisen by polymerase halt-mediated linkage of primers (13, 19) during PCR in the presence of RT-generated cDNA molecules made by mispriming on M40-generated sgmRNAs or sgmRNA-negative strands, or on the viral genome or antigenome, were controlled for by the reactions in Fig. 3C, lanes 12; that is, positive- and negative-strand pM40 DNA molecules were present along with cDNA made from the virus genome and antigenome and no RT-PCR products were found. These results together indicate that the 290-nt chimeric product shown in Fig. 2B, lanes 1, 3, and 5, resulted from RT-PCR recognition of a viral RdRp-generated molecule.

The validity of this conclusion is reinforced by the results of four other experiments. (i) With the same RT-PCR protocol as used for M40 (Fig. 3A) but with RNA from VP1 of M15 DI RNA-transfected, BCoV-infected cells, no chimeric product was obtained (Fig. 3C, lane 2). (ii) With RNA from VP1 of M40 DI RNA-transfected, BCoV-infected cells (i.e., the same

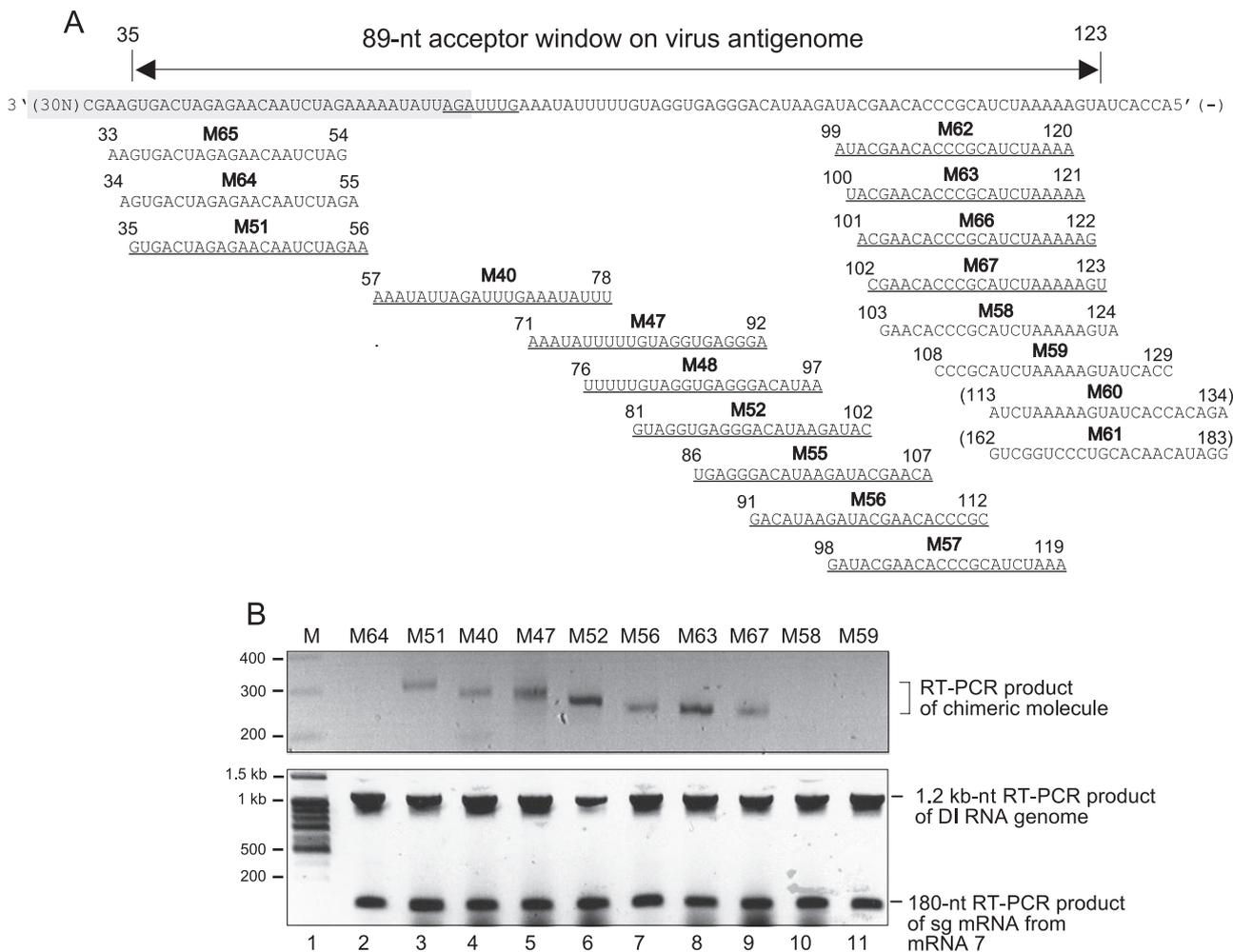


FIG. 5. BCov antigenome acceptor window for induced positive-to-negative-strand template switching. (A) Mapping of the template switching window. Sequences in the viral antigenome that were placed in the DI RNA donor site (nt 1655 to 1676) to form the indicated DI RNA mutants are shown. An underlined sequence identifies a DI RNA mutant that induced a template switch, as indicated by an RT-PCR product and as verified by the sequence of the cloned chimeric product. A sequence not underlined identifies a DI RNA mutant that failed to induce a template switch. On the viral antigenome, the antileader sequence is shaded and the 3'-proximal 89-nt acceptor window is identified. (B) RT-PCR identification of chimeric molecules made with the primer sets described in the legend to Fig. 3A. Top panel: results from an EtBr-stained agarose gel of RT-PCR products of selected mutants that together span and flank the window are shown. Sequence analysis of the cloned RT-PCR product from each band confirmed its chimeric nature and predicted template switch site (shown in Table 1). M, molecular length DNA markers in nt. Bottom panel: a separate gel showing the 1.2-kb RT-PCR products from DI RNA (upper band) and the 180-nt products from sg mRNA 7 (lower band) made with the specific primer sets described in Materials and Methods.

RNA from which the 290-nt RT-PCR product shown in Fig. 3C, lane 3, was obtained) but with primer leader20(-) for both RT and PCR, a procedure that would detect a 1,732-nt chimeric ambisense product resulting from a template switch by the RdRp moving in the opposite direction of that shown in Fig. 3A (Fig. 4), no RT-PCR product was obtained (data not shown). Note that, in this case (Fig. 4), the RdRp would need to switch from a 3'-proximal antigenomic donor site (nt 57 to 78) to an internal positive-strand DI RNA acceptor site (nt 1655 to 1676) to produce a chimeric product. (iii) Independent confirmation of recombinant chimera formation from M40 DI RNA was observed when a second downstream primer specific for the viral antigenome was used [i.e., when primer BCoV631(+)] (specific for the viral genome) replaced primer BCoV108(+)] (which binds a site found in both the DI RNA

and viral genome) in the RT step described in the legend to Fig. 3A]. Cloning and sequencing of the expected 680-nt product in this case confirmed that it too had resulted from RT-PCR recognition of a chimeric molecule produced by an RdRp switch to the viral antigenome at the predicted site (data not shown). (iv) When a series of DI RNA mutants in which other 22-nt sequences that matched sites within the antigenome were used, only those within a discrete window were observed to cause a template switch, thereby identifying a template-switching hot spot (data described below). The sites outside this window, therefore, serve as internal controls for artificial RT-PCR-generated chimeras.

These results therefore indicate that, by creating a 22-nt matching sequence between the positive-strand DI RNA donor template and the viral negative-strand antigenome accep-

tor template, an RdRp switch onto the antigenome was induced that generated a chimeric product of the predicted sequence. This template switch could only have occurred during negative-strand synthesis on the donor template, and the viral antigenome was necessarily acting as the acceptor template in *trans*. It remains to be determined whether the negative strand of a separate DI RNA can also function as an acceptor. Whether the large ambisense molecule undergoes replication remains to be determined as well, but its presence was below the threshold of detection by Northern analysis, suggesting that it is synthesized at low levels *de novo* by template switching in cells infected with VP1 virus carrying the packaged progeny M40 DI RNA. It should be noted that intracellular BCoV genomic RNA is observed as a relatively low abundance molecule at 24 hpi, approximately 20 molecules per cell, and is itself only faintly visible by Northern analysis (17).

Nucleotides 35 to 123 from the 3' end of the viral antigenome function as the acceptor window for template switching from the positive-strand DI RNA to the viral antigenome. To map the RdRp acceptor window on the viral antigenome, 17 additional DI RNA mutants named M47, M48, M51, M52, and M55 to M67, designed to examine acceptor sites between 3'-proximal nt 33 through 183 on the viral antigenome, were tested by using the method described for M40 in the legend to Fig. 3A. That is, in each mutant, a 22-nt sequence matching a site between nt 33 and 183 on the viral antigenome was made and tested. As illustrated in Fig. 5A and B and summarized in Table 1, template acceptor sites between nt 35 and 123 from the 3' end of the viral antigenome were found as determined by cloning and sequencing the chimeric RT-PCR products. When RT-PCR assays from the same RNA samples were undertaken to monitor levels of the DI RNA (Fig. 5B, bottom panel, upper band) and sgRNA 7 from the helper virus (Fig. 5B, bottom panel, lower band), little variation was found between transfections. These results show that the variation in the amounts of chimeric molecules found were a function of the rate of template switching and not of variations in transfection efficiency, DI RNA replicating ability, or the level of helper virus replication. Therefore, there is an acceptor window of 89 nt in the 3'-proximal region of the viral antigenome within which the RdRp can be induced to switch in *trans* from the DI RNA donor template. Template switching was not induced at tested sites outside this window.

DISCUSSION

In experiments presented here, we sought to directly test whether an induced similarity-assisted RdRp template switch, resembling that thought to produce discontinuous negative-strand templates for sgRNA synthesis, occurs during negative-strand synthesis. We reasoned that a template switch induced in *trans* from a positive-strand DI RNA to the negative-strand helper virus antigenome would yield a partially negative-strand and partially positive-strand chimeric product whose sequence would identify the RdRp pathway used for its synthesis. Ambisense products of such a switch were found and, surprisingly, an acceptor hot spot for the switch on the viral antigenome was also identified that coincided roughly with the previously identified acceptor hot spot on the viral

TABLE 1. Junction sequences in ambisense viral RdRp products resulting from mutant DI RNA-induced positive-to-negative-strand template switching

DI RNA mutant	Junction sequence in chimeric RdRp product (5'-3') ^a	Junction site ^b	No. of clones sequenced
M15	No product		
M65	No product		
M64	No product		
M51	AGAAC/ CACUGAUCUCUUGUUGAUCUUUUUAU	1676/35	3
M40	AGAAC/ UUUAUAAUCUAAACUUUAUAAAACAU	1676/57	6
M47	AGAAC/ UUUAUAAAAACAUCACUCCUUGUAUU	1676/71	3
M48	AGAAC/ AAAAACAUCACUCCUUGUAUU CUAUG	1676/76	3
M52	AGAAC/ CAUCCACUCCUUGUAUUCUAUGCUUGU	1676/81	3
M55	AGAAC/ ACUCCUUGUAUUCUAUGCUUGU GGGCG	1676/86	3
M56	AGAAC/ CGUAUUCUAUGCUUGU GGGCGUAGAU	1676/91	3
M57	AGAAC/ CUAUGCUUGU GGGCGUAGAUUUUCAU	1676/98	3
M62	AGAAC/ UAUGCUUGU GGGCGUAGAUUUUCAU	1676/99	3
M63	AGAAC/ AUGCUUGU GGGCGUAGAUUUUCAUAG	1676/100	3
M66	AGAAC/ UGCUGU GGGCGUAGAUUUUCAUAGU	1676/101	3
M67	AGAAC/ GCUUGU GGGCGUAGAUUUUCAUAGU	1676/102	3
M58	No product		
M59	No product		
M60	No product		
M61	No product		

^a Sequences were revealed by sequencing the upper strand of the RT-PCR product. The sequence to the left of the slash is the negative strand and that to the right of the slash is the positive strand. Boldface type indicates the sequence complementary to that placed in the DI RNA between nt 1655 and 1676 to form the indicated mutant.

^b The numbers to the left and right of the slash refer, respectively, to the position of the nucleotides immediately to the left and to the right of the slash in the second column. The number to the left indicates the nt position in the DI RNA (beginning from the 5' end of the DI RNA) and the number to the right indicates the nt position in the viral antigenome (beginning from the 3' end of the viral antigenome).

genome for high-frequency positive-to-positive-strand template switching (44). In addition, the window overlaps *cis*-acting DI RNA replication signaling elements (7, 33), at least one of which, the negative-strand form of stem-loop III (Raman and Brian, unpublished), is a higher-order structure. We therefore postulate that this region of the 5' end of the genome, possibly as a sometimes open structure, as depicted in Fig. 6A, constitutes a multifunctional structure that works to facilitate template switching.

Although internal (i.e., not at the genome termini) positive-to-negative-strand RdRp template switching observed here was in an experimentally induced system and the resulting chimeric molecules have no obvious function, we believe the results have twofold significance. The first is general in that it identifies a pathway whereby RdRps of positive-strand RNA viruses may contribute to genome sequence diversity. This, to our knowledge, is only the second report of such a phenomenon in a positive-strand RNA virus, the first being that of an internal spontaneous template switch in the turnip crinkle virus (6). A single template switch from a strand of one polarity to that of another would of course be a dead-end for the virus. Two switches in close proximity, on the other hand, would potentially yield a nonlethal palindrome in the genome.

The second realm of significance appears specific to coronaviruses (and perhaps other members in the *Nidovirus* order). The acceptor hot spot for an internal positive-to-negative template switch, along with the complementary hot spot on the genome for positive-to-positive-strand template switching (44), identifies a remarkably active 5'-proximal region of the genome for RdRp template switching. This region, therefore, (i)

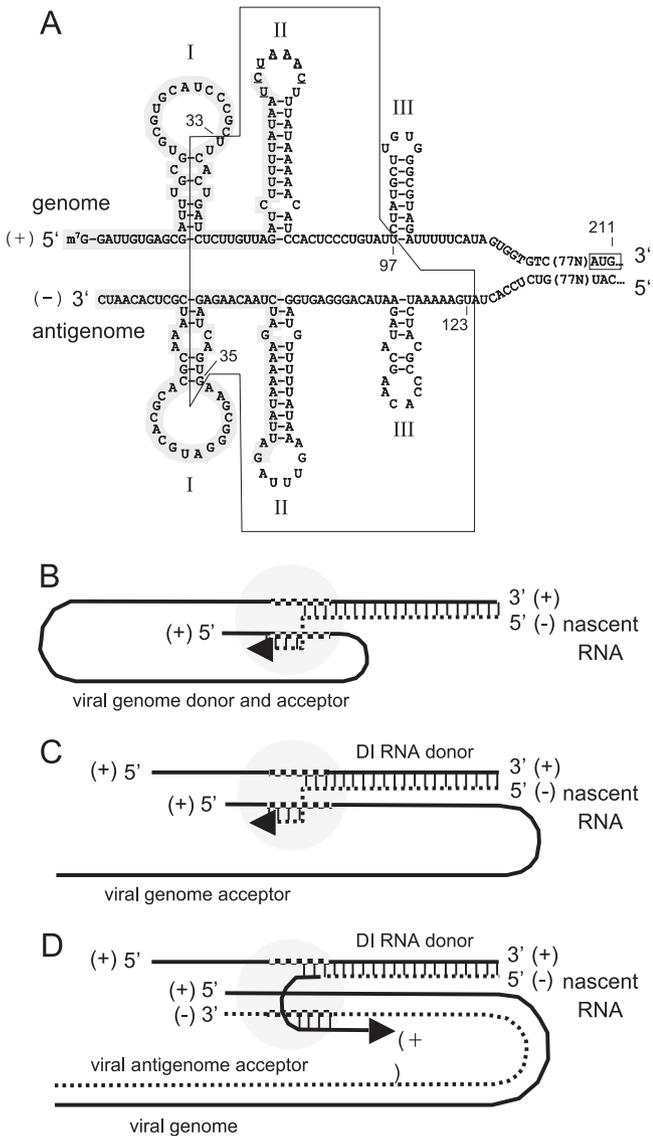


FIG. 6. Model proposed for the coronavirus 5'-terminal double-stranded template switch-facilitating structure and pathway for the RdRp template switch in *trans*. (A) Model for the putative double-stranded template switch-facilitating structure. Mapped acceptor hot spots (boxed) for template switches onto the bovine coronavirus genome and antigenome drawn as an open structure in the context of stem-loops I, II, and III as predicted and structure mapped in the positive strand (8, 33). In the negative strand, stem-loops II and III are predicted (33), and stem-loop I is drawn as a near-mirror image of stem-loop I in the positive strand. The 65-nt leader and antileader sequences are shaded, the nucleotides making up the leader-associated intergenic core sequence (UCUAAAC) in the loop of stem-loop II are underlined, and the start codon for open reading frame 1 is boxed. It should be noted that the stem-loop I region for group 2 coronaviruses (represented here as nt 1 to 51 for BCoV) has recently been alternatively represented as two stem-loops by Kang et al. (21). (B) Proposed model as redrawn from Zúñiga et al. (48) for positive-to-positive-strand template switching from donor to acceptor sites in *cis* during synthesis of negative-strand templates for sgRNA synthesis. In this model, only the positive-strand viral genomic template is shown. The matching donor and acceptor regions are checked. Base-paired regions between the positive-strand genome and nascent negative strand are indicated by vertical lines. The complete transcription complex with proteins and RNA is depicted by the shaded oval. (C) Proposed model as drawn by Wu et al. (44) for positive-to-positive-strand tem-

probably represents a structure involved in the formation of the antileader-containing templates for sgRNA synthesis (44, 48) and (ii) possibly represents a structure involved in high-frequency genomic 5' end recombination (22), of which "leader switching" (8, 26) may be one manifestation. Interestingly, both of these template-switching activities, and others that include the addition of UCUAA elements near the leader in genomic MHV (25, 27), heterogeneous leader-mRNA fusion sites under the influence of the 9-nt UUUUAUAAAC element mapping just downstream of the MHV genomic leader (45, 46), and the formation of novel sgRNAs with aberrant leader-body junctions (12), take place within the identified hot spots. Interestingly, on this point, the template switching may not be exclusively in one direction. Since, in M40 DI RNA, two heptameric regions of high similarity exist between the potential donor sequence in M40 DI RNA (nt 1655 to 1676) and the leader-associated acceptor region in the viral genome (or DI RNA) (nt 57 to 78), we thought there could also possibly be a positive-to-positive-strand template switch and synthesis of sgRNA from M40 DI RNA (the alignments are not shown). This possibility was tested, and a small amount of two sgRNA species, one with a full-length leader of 65 nt and another with a 53-nt minileader, were found (data not shown), indicating that sgRNA synthesis from M40 DI RNA does occur and that it conforms to a pattern described earlier for other DI RNA mutants (44).

What strikes us about the template switch-facilitating structure is that the antigenome appears to be part of it. This view is based in part on the results of several studies that led to the conclusion that proximity between donor and acceptor templates contributes importantly to the frequency of sequence similarity-induced template switching. Examples include cross-over hot spots for RdRps (19, 29) and RTs (1, 2, 11, 16, 28). Thus, in our view, the antigenome must be nearby and possibly associated with the genome in a double-stranded or at least partially double-stranded structure at the time of the template switch. This view is consistent with the finding that the antigenome is not found free in the cytoplasm of infected cells but is in a form that copurifies with the genome in membrane-protected complexes (37), possibly in a transcriptive intermediate (34). As with positive-to-positive-strand template switching in the BCoV DI RNA (44), switching here appears driven by a sequence similarity-induced mechanism (5, 30). What other structural features might mechanistically contribute to the switch is, of course, not known at this time. Based on precedents in other RdRp template switching systems, however, higher-order structures that promote or enhance template switching via interactions with proteins (23, 30, 31) might be

plate switching in *trans* from a DI RNA donor to the viral genome acceptor. In this model, only the positive-strand templates are shown. (D) Proposed model for positive-to-negative-strand template switching in *trans* from a DI RNA donor to the viral antigenome acceptor (to explain the data in this report). In this model, both the full-length viral genome and antigenome templates are shown. The switch generates a chimeric ambisense product that is potentially longer than the viral genome. Note that in this model the 3' terminus of the completed viral antigenome along with the 5' terminus of the viral genome are components of the transcription complex.

involved. In that regard, stem-loops II and III (Fig. 6A) are candidates for such structures. For these to function as higher-order structures, however, the 5' end of the genome and 3' end of the antigenome would probably need to be in an open form, as depicted in Fig. 6A. Another potentially important factor is a discontinuous donor template that might take part in a "forced" template switching mechanism (9, 30). Such a role has been suggested for the coronaviral endonuclease nsp 15 or NendoU (14, 15, 40), which in one study demonstrated endonuclease activity for double-stranded, nonmethylated RNA (18) but in another, demonstrated activity for only single-stranded RNA (4).

If the antigenome is a component of the transcription complex, then one model for the complex might be that depicted in Fig. 6D, wherein negative-strand synthesis is initiated in *trans* at the 3' end of positive-strand DI RNA by an RdRp that is somehow associated with both the 3' end of the viral antigenome and the 5' end of the viral genome within a moving transcription complex. This model accommodates the data described in this report in that the 3' end of the antigenome is available within the transcription complex for positive-to-negative-strand template switching in *trans*. If the model in Fig. 6D is correct, then it might be expected that positive-to-positive-strand template switching could also occur in *trans* within the transcription complex, for example, between the DI RNA and viral genome, as depicted in Fig. 6C, and as described by Wu et al. (44) and Zhang et al. (47). The model is a modification of that drawn by Zúñiga et al. (Fig. 6B) (48) in which a moving transcription complex carries the 5' end of the viral genome and facilitates high-frequency positive-to-positive-strand template switching in *cis*. A mechanism requiring the presence of the antigenome 3' terminus to be part of the transcription complex would, additionally, ensure that conditions for genome replication have been met before sgRNA synthesis begins. An acceptor hot spot on the viral antigenome might also function as an acceptor site for negative-to-negative-strand template switching during positive-strand synthesis that results in "leader switching" as proposed earlier (26, 45).

Further experimentation using methods other than those described here are needed to determine whether the antigenome is a component of the coronaviral transcription complex. It will also be important to determine the precise RNA and protein composition of the transcription complex to understand what factors within the structure direct high-frequency template switching.

ACKNOWLEDGMENTS

We thank Kimberley Nixon, Aykut Ozdarendeli, Gwyn Williams, and Ruey-Yi Chang (Dong-Hua University, Taiwan) for valuable discussions.

This work was supported by Public Health Service grant AI4367 from the NIAID and by funds from the University of Tennessee, College of Veterinary Medicine Center of Excellence Program.

REFERENCES

- Andersen, E. S., R. E. Jeeninga, C. K. Damgaard, B. Berkhout, and J. Kijms. 2003. Dimerization and template switching in the 5' untranslated region between various subtypes of human immunodeficiency virus type 1. *J. Virol.* **77**:3020–3030.
- Balakrishnan, M., B. P. Roques, P. J. Fay, and R. A. Bambara. 2003. Template dimerization promotes an acceptor invasion-induced transfer mechanism during human immunodeficiency virus type 1 minus-strand synthesis. *J. Virol.* **77**:4710–4721.
- Baric, R. S., and B. Yount. 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. *J. Virol.* **74**:4039–4046.
- Bhardwaj, K., S. Jingchuan, A. Holzenburg, L. A. Guarino, and C. C. Kao. 2006. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J. Mol. Biol.* **361**:243–258.
- Brian, D. A., and W. J. M. Spaan. 1997. Recombination and coronavirus defective interfering RNAs. *Semin. Virol.* **8**:101–111.
- Carpenter, C. D., and A. E. Simon. 1994. Recombination between plus and minus strands of turnip crinkle virus. *Virology* **201**:419–423.
- Chang, R. Y., M. A. Hofmann, P. B. Sethna, and D. A. Brian. 1994. A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. *J. Virol.* **68**:8223–8231.
- Chang, R. Y., R. Krishnan, and D. A. Brian. 1996. The UCUAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. *J. Virol.* **70**:2720–2729.
- Coffin, J. M. 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *J. Gen. Virol.* **42**:1–26.
- Cowley, J. A., C. M. Dimmock, and P. J. Walker. 2002. Gill-associated nidovirus of *Penaeus monodon* prawns transcribes 3'-coterminal subgenomic mRNAs that do not possess 5'-leader sequences. *J. Gen. Virol.* **83**:927–935.
- Delviks, K. A., and V. K. Pathak. 1999. Effect of distance between homologous sequences and 3' homology on the frequency of retroviral reverse transcriptase template switching. *J. Virol.* **73**:7923–7932.
- Fischer, F., C. F. Stegen, C. A. Koetzner, and P. S. Masters. 1997. Analysis of a recombinant mouse hepatitis virus expressing a foreign gene reveals a novel aspect of coronavirus transcription. *J. Virol.* **71**:5148–5160.
- Frohman, M. A., and G. R. Martin. 1990. Detection of homologous recombinants, p. 228–236. *In* M. A. Innis (ed.), *PCR protocols*. Academic Press, New York, NY.
- Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder. 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**:17–37.
- Guarino, L. A., K. Bhardwaj, W. Dong, J. Sun, A. Holzenburg, and C. C. Kao. 2005. Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J. Mol. Biol.* **353**:1106–1117.
- Havert, M. B., and D. D. Loeb. 1997. *cis*-Acting sequences in addition to donor and acceptor sites are required for template switching during synthesis of plus-strand DNA for duck hepatitis B virus. *J. Virol.* **71**:5336–5344.
- 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.
- Ivanov, K. A., T. Hertzog, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr. 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc. Natl. Acad. Sci. USA* **101**:12694–12699.
- Jarvis, T. C., and K. Kirkegaard. 1992. Poliovirus RNA recombination: mechanistic studies in the absence of selection. *EMBO J.* **11**:3135–3145.
- Jarvis, T. C., and K. Kirkegaard. 1991. The polymerase in its labyrinth: mechanisms and implications of RNA recombination. *Trends Genet.* **7**:186–191.
- Kang, H., M. Feng, M. E. Schroeder, D. P. Giedroc, and J. L. Leibowitz. 2006. Putative *cis*-acting stem-loops in the 5' untranslated region of the severe acute respiratory syndrome coronavirus can substitute for their mouse hepatitis virus counterparts. *J. Virol.* **80**:10600–10614.
- Keck, J. G., S. A. Stohlman, L. H. Soe, S. Makino, and M. M. Lai. 1987. Multiple recombination sites at the 5'-end of murine coronavirus RNA. *Virology* **156**:331–341.
- Kim, M. J., and C. Kao. 2001. Factors regulating template switch in vitro by viral RNA-dependent RNA polymerases: implications for RNA-RNA recombination. *Proc. Natl. Acad. Sci. USA* **98**:4972–4977.
- Lai, M. M. 1992. Genetic recombination in RNA viruses. *Curr. Top. Microbiol. Immunol.* **176**:21–32.
- Makino, S., and M. M. Lai. 1989. Evolution of the 5'-end of genomic RNA of murine coronaviruses during passages in vitro. *Virology* **169**:227–232.
- Makino, S., and M. M. Lai. 1989. High-frequency leader sequence switching during coronavirus defective interfering RNA replication. *J. Virol.* **63**:5285–5292.
- Makino, S., L. H. Soe, C. K. Shieh, and M. M. Lai. 1988. Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. *J. Virol.* **62**:3870–3873.
- Mikkelsen, J. G., S. V. Rasmussen, and F. S. Pedersen. 2004. Complementarity-directed RNA dimer-linkage promotes retroviral recombination in vivo. *Nucleic Acids Res.* **32**:102–114.
- Nagy, P. D., and J. J. Bujarski. 1993. Targeting the site of RNA-RNA recombination in brome mosaic virus with antisense sequences. *Proc. Natl. Acad. Sci. USA* **90**:6390–6394.
- Nagy, P. D., and A. E. Simon. 1997. New insights into the mechanisms of RNA recombination. *Virology* **235**:1–9.
- Nagy, P. D., C. Zhang, and A. E. Simon. 1998. Dissecting RNA recombination in vitro: role of RNA sequences and the viral replicase. *EMBO J.* **17**:2392–2403.
- Ozdarendeli, A., S. Ku, S. Rochat, G. D. Williams, S. D. Senanayake, and D. A. Brian. 2001. Downstream sequences influence the choice between a

- naturally occurring noncanonical and closely positioned upstream canonical heptameric fusion motif during bovine coronavirus subgenomic mRNA synthesis. *J. Virol.* **75**:7362–7374.
33. **Raman, S., P. Bouma, G. D. Williams, and D. A. Brian.** 2003. Stem-loop III in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. *J. Virol.* **77**:6720–6730.
 34. **Sawicki, D., T. Wang, and S. Sawicki.** 2001. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J. Gen. Virol.* **82**:385–396.
 35. **Sawicki, S. G., and D. L. Sawicki.** 2005. Coronavirus transcription: a perspective. *Curr. Top. Microbiol. Immunol.* **287**:31–55.
 36. **Sawicki, S. G., and D. L. Sawicki.** 1998. A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* **440**:215–219.
 37. **Sethna, P. B., and D. A. Brian.** 1997. Coronavirus genomic and subgenomic minus-strand RNAs copartition in membrane-protected replication complexes. *J. Virol.* **71**:7744–7749.
 38. **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.
 39. **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.
 40. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Rozanov, W. J. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **331**:991–1004.
 41. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci. USA* **96**:12056–12061.
 42. **van Vliet, A. L., S. L. Smits, P. J. Rottier, and R. J. de Groot.** 2002. Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J.* **21**:6571–6580.
 43. **Wu, H. Y., J. S. Guy, D. Yoo, R. Vlasak, E. Urbach, and D. A. Brian.** 2003. Common RNA replication signals exist among group 2 coronaviruses: evidence for in vivo recombination between animal and human coronavirus molecules. *Virology* **315**:174–183.
 44. **Wu, H. Y., A. Ozdarendeli, and D. A. Brian.** 2006. Bovine coronavirus 5'-proximal genomic acceptor hotspot for discontinuous transcription is 65 nucleotides wide. *J. Virol.* **80**:2183–2193.
 45. **Zhang, X., and M. M. Lai.** 1996. A 5'-proximal RNA sequence of murine coronavirus as a potential initiation site for genomic-length mRNA transcription. *J. Virol.* **70**:705–711.
 46. **Zhang, X., and M. M. Lai.** 1994. Unusual heterogeneity of leader-mRNA fusion in a murine coronavirus: implications for the mechanism of RNA transcription and recombination. *J. Virol.* **68**:6626–6633.
 47. **Zhang, X., C. L. Liao, and M. M. Lai.** 1994. Coronavirus leader RNA regulates and initiates subgenomic mRNA transcription both in *trans* and in *cis*. *J. Virol.* **68**:4738–4746.
 48. **Zúñiga, S., I. Sola, S. Alonso, and L. Enjuanes.** 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.* **78**:980–994.