JVI Journals.ASM.org

Dependence of Coronavirus RNA Replication on an NH_2-Terminal Partial Nonstructural Protein 1 in *cis*

Yu-Pin Su, Yi-Hsin Fan, David A. Brian

Department of Biomedical and Diagnostic Sciences, University of Tennessee College of Veterinary Medicine, Knoxville, Tennessee, USA

ABSTRACT

Genomes of positive (+)-strand RNA viruses use *cis*-acting signals to direct both translation and replication. Here we examine two 5′-proximal *cis*-replication signals of different character in a defective interfering (DI) RNA of the bovine coronavirus (BCoV) that map within a 322-nucleotide (nt) sequence (136 nt from the genomic 5′ untranslated region and 186 nt from the nonstructural protein 1 [nsp1]-coding region) not found in the otherwise-identical nonreplicating subgenomic mRNA7 (sgmRNA7). The natural DI RNA is structurally a fusion of the two ends of the BCoV genome that results in a single open reading frame between a partial nsp1-coding region and the entire N gene. (i) In the first examination, mutation analyses of a recently discovered long-range RNA-RNA base-paired structure between the 5′ untranslated region and the partial nsp1-coding region showed that it, possibly in concert with adjacent stem-loops, is a *cis*-acting replication signal in the (+) strand. We postulate that the higher-order structure promotes (+)-strand synthesis. (ii) In the second examination, analyses of multiple frame shifts, truncations, and point mutations within the partial nsp1-coding region showed that synthesis of a PEFP core amino acid sequence within a group A lineage betacoronavirus-conserved NH_2-proximal WAPEFPWM domain is required in *cis* for DI RNA replication. We postulate that the nascent protein, as part of an RNA-associated translating complex, acts to direct the DI RNA to a critical site, enabling RNA replication. We suggest that these results have implications for viral genome replication and explain, in part, why coronavirus sgmRNAs fail to replicate.

IMPORTANCE

***cis*-Acting RNA and protein structures that regulate (+)-strand RNA virus genome synthesis are potential sites for blocking virus replication. Here we describe two: a previously suspected 5′-proximal long-range higher-order RNA structure and a novel nascent NH_2-terminal protein component of nsp1 that are common among betacoronaviruses of group A lineage.**

What constitutes the *cis*-acting requirements for coronavirus RNA replication has remained an intriguing question since it was discovered that the subgenomic mRNAs (sgmRNAs) of coronaviruses (used primarily to synthesize viral structural proteins) are both (i) 5′ and 3′ coterminal with the genome for at least ~70 and 1,670 nucleotides (nt), respectively, lengths greater than those of many viral RNA polymerase promoters (1–3), and (ii) are present in sgmRNA-length replication-intermediate-like double-stranded RNA structures that are involved in sgmRNA synthesis (4–6) yet fail to replicate when transfected, as synthetic transcripts, into virus-infected cells (Fig. 1) (7). If replication of the coronavirus sgmRNAs normally occurs during infection, it might be expected that they would replicate following their transfection into virus-infected cells, since all *trans*-acting factors required for viral RNA replication are present. In coronaviruses, the 5′ two-thirds of the single-stranded positive (+)-strand ~30-kb coronavirus genome is used as mRNA for synthesis of overlapping polyproteins 1a (~4,000 amino acids [aa]) and 1ab (~7,000 aa), which are proteolytically processed into the 16 replicase proteins that make up the replication/transcription complex, whereas the 3′ one-third of the genome is transcribed into a 3′ nested set of sgmRNAs that are coterminal with the genome but are translated separately (3, 8, 9). One model widely used to explain the origin of the sgmRNA-length replication-intermediate-like double-stranded RNAs was proposed by Sawicki et al. (4, 6, 10, 11). In this model, (i) the genome is envisioned as the only template for negative (−)-strand RNA synthesis, and (ii) an RNA-dependent RNA polymerase (RdRp) template-switching event takes place during (−)-strand synthesis from the viral genome template at intergenic donor core sequence (also termed transcription-regulating sequence) sites (UCUAAAC in bovine coronavirus [BCoV] and mouse hepatitis virus [MHV]) to the 5′-proximal leader acceptor core sequence (UCUAAAC) on the genome (i.e., a discontinuous transcription step) to create a sgmRNA-length (−)-strand RNA. In this model, the sgmRNA-length (−)-strand RNA (5, 12, 13) then functions as a template for synthesis of new sgmRNA. The term proposed for the sgmRNA-length, partially double-stranded structure hence became "transcriptive intermediate" (4, 11) rather than "replicative intermediate," as was initially used (5, 6, 12), to more clearly identify the viral genome as the only template for sgmRNA (−)-strand synthesis. This model for sgmRNA synthesis from the genome was tested first by reverse genetics in an arterivirus (14), a fellow member of the *Nidovirales* order with a very similar pattern of sgmRNA generation, and second in the coronavirus (15), and the results with both viruses are consistent with the Sawicki model. More recently, it has been learned that the

Received 13 March 2014 **Accepted** 21 May 2014

Published ahead of print 28 May 2014

Editor: S. Perlman

Address correspondence to David A. Brian, dbrian@utk.edu.

Y.-P.S. and Y.-H.F. contributed equally to this study.

doi:10.1128/JVI.00738-14

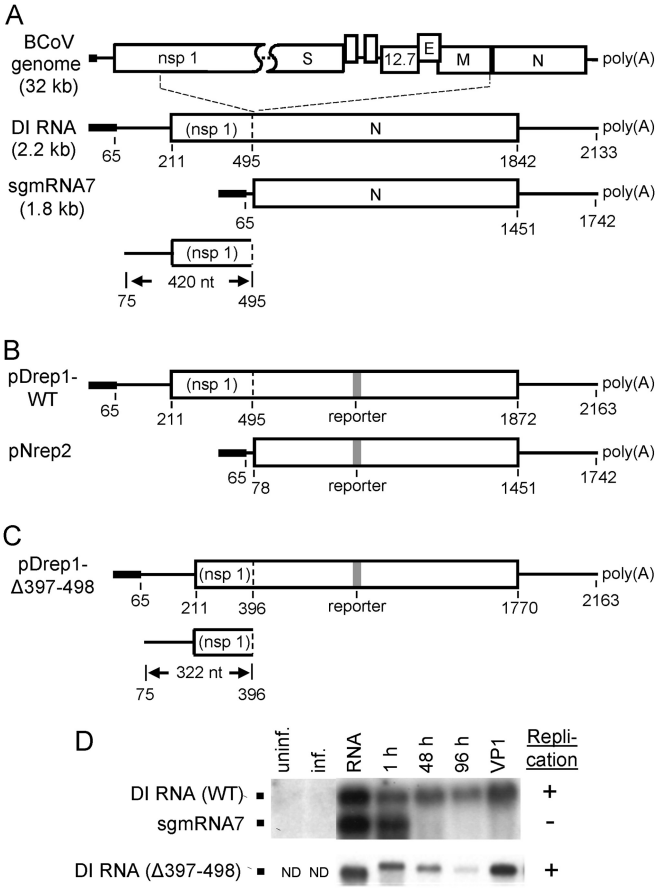
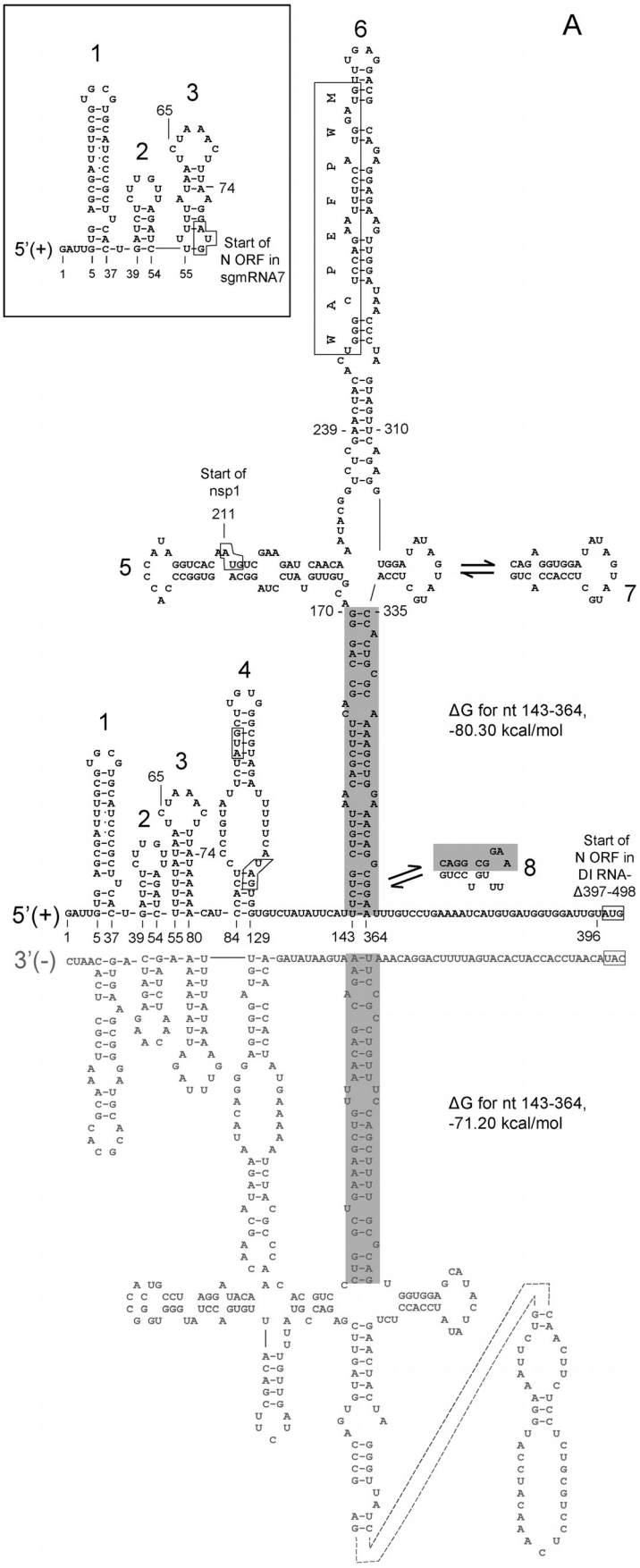


FIG 1 Three hundred twenty-two-nucleotide sequence difference between the minimalized replication-competent BCoV DI RNA and the replication-incompetent sgmRNA7. (A) Schematic representation of the parent BCoV genome, the naturally occurring replication-competent BCoV DI RNA, and replication-incompetent sgmRNA7. Note that the naturally occurring DI RNA and sgmRNA7 are identical at the ends but differ by a contiguous 420-nt 5′-proximal sequence. (B) When cDNAs of the DI RNA and sgmRNA7 were cloned and an in-frame 30-nt reporter for Northern blot analysis was inserted within the N gene, they were named pDrep1-WT and pNrep2, respectively (7). (C) The 420-nt sequence in the naturally occurring DI RNA was shortened from its 3′ end to 322 nt, and replication competence was retained (26). (D) Northern blot analyses showing the replication patterns of reporter-containing DI RNA and sgmRNA7. (Reprinted from references 7 and 26). (Upper) DI RNA (WT) (as represented by transcripts of pDrep1-WT) and sgmRNA7 RNA (as represented by transcripts of pNrep2) were cotransfected into BCoV (helper virus)-infected cells, and RNA abundance was measured by hybridization with a reporter-specific ^{32}P-radiolabeled probe (7). From the Northern blot it can be seen that the DI RNA replicates following transfection into helper virus-infected cells and gets packaged, whereas the sgmRNA7 does not. (Lower) DI RNA (Δ397-498) (as represented by transcripts of pDrep1Δ397-498) was transfected into BCoV-infected cells, and Northern blot analyses were carried out as described for the upper panel (26). Lanes: uninf., uninfected; inf., infected; RNA, sample of the nonpolyadenylated RNA used for transfection of cells infected 1 h earlier; 1 h, 48 h, 96 h, times posttransfection; VP1, first virus passage, RNA extracted from VP1 virus-infected cells at 48 h postinfection; ND, not determined. Replication was considered positive if there had been accumulation of DI RNA over time in the transfected cells or if DI RNA was present in cells infected with VP1.

coronavirus sgmRNA (+)-strand molecules are competent templates for (−)-strand RNA synthesis when transfected into infected cells (16) and that when the transfected sgmRNA contains an intergenic template-switching donor signal (UCUAAAC), sgmRNAs of smaller size are generated from this site in a manner consistent with the Sawicki model. This behavior suggests a mechanism of sgmRNA amplification by a cascading transcription process and not by replication (16). So the question becomes: why are the full-length nascent sgmRNA (−) strands arising from the sgmRNAs following transfection not competent for initiating synthesis of new full-length (+)-strand sgmRNA as are the sgmRNA-length (−)-strand RNAs arising from the full-length genome by discontinuous transcription as proposed in the Sawicki model (4, 6, 10, 11)? In other words, why do the sgmRNAs, the shortest of which is 1.8 kb in length in the mouse and bovine coronaviruses, not replicate following transfection as RNA transcripts into helper virus-infected cells (Fig. 1D, upper panel) (7)?

In contrast to the sgmRNAs, a naturally occurring 2.2-kb defective interfering (DI) RNA from BCoV, which differs from the 1.8-kb sgmRNA7 by only 420 nt that map within the 5′-proximal region of the genome (Fig. 1A), can replicate and be passaged as packaged molecules following transfection of RNA transcripts into helper virus-infected cells (Fig. 1D, upper panel) (7). A similar region of the virus genome is found in all naturally occurring coronavirus DI RNAs described to date, and after cDNA cloning, transcripts of these also replicate following transfection into helper virus-infected cells (2, 17–25). Within the 420-nt 5′-proximal region of the naturally occurring BCoV DI RNA, the 65-nt common leader plus a 3′-ward extension of 9 nt (making 74 nt total) are found in common between the replicating wild-type (WT) DI RNA and the nonreplicating sgmRNA7 (see inset, Fig. 2A). Furthermore, the 5′-proximal 420-nt sequence on the naturally occurring WT BCoV DI RNA can be shortened from its 3′ end to 322 nt without loss of DI RNA-replicating ability (Fig. 1C and D, lower panel, and Fig. 2A to C) (26). This indicates that the 3′-terminal 136 nt of the genomic 5′ untranslated region (UTR) in addition to the 5′-terminal 186 nt of the nsp1-coding region (encoding 62 amino acids, or 25%, of the 246-amino-acid nsp1) are necessary and sufficient for replication competence in the DI RNA (compared to in sgmRNA7) when assayed by transfection of RNA transcripts into helper virus-infected cells (26). (Note that the WT DI RNA encodes 94 amino acids, or 38%, of nsp1).

In previous studies, two kinds of *cis*-replication signals have been associated with the 322-nt region in the BCoV DI RNA: (i) higher-order *cis*-acting RNA structures and (ii) a *cis*-translation requirement of the fused open reading frame (ORF). (i) With regard to the 5′-terminal *cis*-acting RNA structures, stem-loops 1, 2, and 3 map (almost entirely) within the most 5′ 74 nt (7, 27) and may not be components unique to the function of the 322-nt region. Similarly, stem-loop 4 (28), which is nearly identical to its homolog in MHV that has been recently shown not to be required for virus replication (29, 30), also may not be a component unique to the function of the 322-nt region. All of *cis*-acting stem-loops 5 (31, 32), 6 (33), and 7 (26) and possibly a small stem-loop 8, which has been predicted to be but not tested as a *cis*-acting replication signal (26), however, may contribute uniquely to the replication function of the 322-nt region. (Please note that stem-loops 1, 2, and 3 were formerly named stem-loops I and II and stem-loops 4 through 8 were formerly named stem-loops III through VII and are so named in the references noted.) Homologous 5′-proximal *cis*-acting structures in the MHV (30, 32, 34–36) and in the more distantly related severe acute respiratory syndrome coronavirus (SARS-CoV) have been described, although in the SARS-CoV homolog the status for stem-loops downstream of stem-loop 4 is less


A
B
C
Start of nsp1
Start of N ORF in sgmRNA7
Start of N ORF in DI RNA-WT
Start of N ORF in DI RNA-Δ397-498
ΔG for nt 143-364, -80.30 kcal/mol
ΔG for nt 143-364, -71.20 kcal/mol
NdeI

clear (36, 37). More recently, it has been shown from reverse genetic studies with MHV that there is also a long-range RNA-RNA base-paired interaction between a region mapping between stem-loops 4 and 5 within the 5′ UTR and the partial nsp1-coding region in BCoV and MHV that is required for MHV replication (Fig. 2A) (38). Interestingly, the BCoV 5′ UTR and entire nsp1-coding region function together as an integral unit in the MHV genome to produce WT-like MHV, but the two regions are not immediately functional when mismatched, and when they are mismatched, adaptive mutations are found in viable virus progeny (38). The long-range RNA-RNA interaction is also predicted by mfold analyses for other betacoronaviruses, including SARS-CoV, and in alphacoronaviruses (36, 38).

(ii) In regard to a *cis*-translation requirement for the partial nsp1-coding region, one was demonstrated in the context of the WT BCoV DI RNA (39). A similar requirement for translation of the partial nsp1-coding region has been reported for the MHV DI RNA (24, 40), although in these studies it was concluded that it was probably the process of translation and not the product that was required (see Discussion). Therefore, consideration of these two sets of features, the long-range RNA-RNA interaction and the *cis*-translation requirement, brings into sharper focus the question of what properties exist within the 322-nt region that provide replication competence to the BCoV DI RNA (as opposed to the replication incompetence of BCoV sgmRNA7), and we approach that question here.

It should be noted that in addition to the role of the 5′-terminal partial nsp1 structure in RNA replication examined here, the entire nsp1 in coronaviruses has been shown to be a multifunctional protein with RNA binding properties and with features that regulate replication, interferon-dependent signaling, host cell mRNA stability, and pathogenesis (33, 41–52).

Here we investigated the *cis*-acting replication function of the long-range RNA-RNA base-paired structure that maps between the 5′ UTR and partial nsp1-coding region and learned that it, like the 5′-proximal stem-loops 4 through 7, functions as a *cis*-acting replication signal in the (+) strand. We also investigated the *cis*-acting translation requirement of the partial nsp1-coding region and discovered that the presence of a nascent protein product carrying a group A betacoronavirus-conserved octameric amino acid sequence, WAPEFPWM, correlates with BCoV DI RNA replication and that changing the quality of the central four amino acids, PEFP, in various arrangements without changing RNA secondary structure abolished DI RNA replication. Furthermore, reestablishing the WT amino acid sequence with different codons showing the same base-pairing pattern as those in WT restored DI RNA replication. We propose that the protein product of the 5′-proximal partial nsp1, possibly in concert with its associated RNA structure, functions to direct the translating DI RNA genome to a still poorly defined position within the replication compartment where viral enzymes required for RNA replication reside.

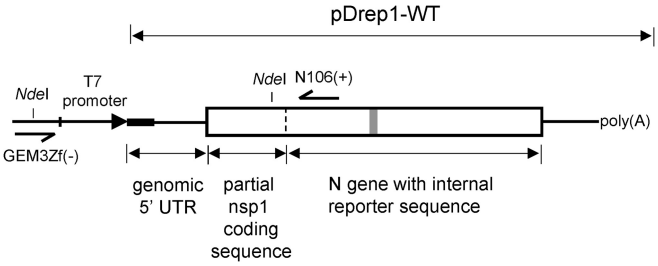


FIG 3 Mutagenesis strategy for the reporter-containing WT DI RNA. Overlap PCR mutagenesis was used to make mutations within the genomic 5′ UTR and partial nsp1-coding sequence of the cloned, reporter-containing DI RNA (WT) named pDrep1-WT. The NdeI sites were used for constructing pDrep1 mutants.

MATERIALS AND METHODS

Cells, virus, and DI RNA. A DI-RNA-free stock of the Mebus strain of BCoV (genome sequence, GenBank accession no. #U00735) at a concentration of 4.5×10^8 PFU/ml was used as a helper virus as described previously (7, 39). The human rectal-tumor cell line HRT-18 (53) was used in all experiments. pDrep1 is a pGEM3Zf(−) (Promega)-based plasmid containing the cDNA clone of a naturally occurring 2.2-kb DI RNA of BCoV modified to carry a 30-nt in-frame reporter (Fig. 1B) (7).

RNA structure predictions. The mfold program of M. Zuker (http://mfold.rna.albany.edu/?q=mfold) (54, 55) was used for RNA structure predictions. The long-range RNA-RNA base-pairing patterns described below were revealed by folding nt 1 to 400 or nt 1 to 500 and from the results of a reverse genetics study with MHV and BCoV chimeric constructs (38).

Construction of mutant DI RNAs and synthesis of RNA transcripts. Modifications of pDrep1 DNA were made by overlap PCR mutagenesis as previously described (56, 57). For this process, the appropriate oligonucleotide primers containing the described mutations and the NdeI restriction endonuclease sites within the pGEM3Zf(−) vector and pDrep1 DNA were used (Fig. 3). Mutations in the final constructs were confirmed by sequencing. The sequence for primer GEM3Zf(−) is 5′-GAGAGTGCAC CATATGCGGTGT-3′, and for primer N106(+), 5′-CTCTTCTACCCC TGGTTTGAAC-3′. The (+) and (−) signs designate the polarity of the RNA to which the primer binds. For synthesis of RNA, 1 μg of MluI-linearized DNA was used with a T7 mMessage mMachine kit (Ambion) according to the manufacturer's protocol to make 5′ m7GpppG-capped RNA. The reaction mix was incubated with 5 U of Turbo DNase (Ambion), and RNA was chromatographed through a Bio-Spin 6 column (Bio-Rad) and quantitated by nanodrop spectrophotometry. *In vitro*-syn-

FIG 2 Predicted higher-order RNA structures in the 5′-proximal 322 nt of the BCoV DI RNA and its (−)-strand counterpart. (A) Higher-order RNA structures. Shown are the mfold-predicted RNA structures at the 5′ end of the BCoV genome (above) and at the 3′ end of the (−)-strand antigenome (below). Structures in the (+) strand are stem-loops 1 through 8. The previously described long-range RNA-RNA interaction between a region within the 5′ UTR (nt 143 through 170) and the 5′-terminal nsp1-coding region (nt 335 through 364) (38) is shown in shaded lettering. Note that the alternate stem-loops 7 and 8 in the (+) strand would not coexist with the long-range RNA-RNA interaction as drawn. The boxed amino acid sequence, WAPEFPWM, is described in the text. The 322-nt region differentiating the minimalized replication-competent BCoV DI RNA from the replication-incompetent sgmRNA7 is comprised of nt 75 through 396. The mfold-predicted ΔG for the long-range higher-order RNA structure (nt 143 through 364) in the (+) strand is −80.30 kcal/mol, and in the (−) strand, −71.20 kcal/mol. (Inset) 5′ UTR of sgmRNA7. Note that the first 74 nt of the genome and of sgmRNA7 are identical. (B) Nucleotides 211 through 396 encoding the 62 aa in the partial nsp1 in the minimalized replication-competent transcripts of pDrep1-Δ397-498. The boxed amino acid sequence, WAPEFPWM, is described in the text. (C) The 102-nt sequence (397 through 498) removed from the 3′ end of the partial nsp1-coding sequence in pDrep1-WT to form the minimalized pDrep1-Δ397-498 (26). Note that the partial nsp1 fusion site is between A494 in the nsp1-coding sequence and A495, the fourth nucleotide upstream of the N start codon in the genome. This fusion formed a codon for glutamic acid (E, underlined). The NdeI endonuclease restriction enzyme site used for *in vitro* mutagenesis in pDrep1-WT is shown.

thesized RNAs were used for transfection in the replication assays and for *in vitro* translation assays. RNA preparations were stored at −80°C.

Northern assay for DI RNA replication and packaging. A Northern assay for detecting reporter-containing DI RNAs was performed as described previously (7, 27). Briefly, cells (~1.5×10^6) at ~80% confluence in a 35-mm dish were infected with BCoV at a multiplicity of 10 PFU per cell and transfected 1 h later with 300 ng of capped RNA, using Lipofectin (Invitrogen). At the indicated times postinfection (see figures), total RNA (approximately 10 μg per plate) was extracted with TRIzol (Invitrogen) and stored as an ethanol precipitate. For passage of progeny virus, supernatant fluids were harvested at 48 h postinfection (hpi) and 500 μl was used to infect freshly confluent cells (~2.0×10^6) in a 35-mm dish (27) from which RNA was extracted at 48 hpi. For electrophoretic separation of RNA in a formaldehyde-agarose gel, 2.5 μg per lane was used. Approximately 5 ng of transcript, identified as RNA in the Northern blot figures, was loaded per lane when used as a marker. RNA was transferred to Nytran membranes by vacuum blotting, and the UV-irradiated blots were probed with oligonucleotide TGEV(+), which had been ^{32}P-labeled at the 5′ end to specific activities of 1×10^6 to 4×10^7 cpm/pmol (5). Probed blots were exposed to Kodak XAR-5 film for 1 to 7 days at −80°C for imaging. Image intensity variations within some figures resulted from differing times of RNA sample and probe preparation. Replication was judged positive when there was an increase in DI RNA abundance over time or when progeny DI RNA was present in cells at 48 h following infection with virus passage 1 (VP1) or VP2, along with evidence that there was no sequence reversion in the progeny (39). The probe used for detecting 18S rRNA was 5′-CTGCTGGCACCAGACTTGCCCTCCAA-3′ (39).

RT-PCR and sequence analysis of progeny from transfected WT and mutant DI RNAs. Reverse transcriptase PCR (RT-PCR) and sequence analyses were carried out as previously described (26). Briefly, RNAs extracted from VP1- and VP2-infected cells were used for cDNA synthesis with SuperScript II reverse transcriptase (Invitrogen) and DI-RNA-specific primer TGEV-8(+) (5′-CATGGCACCATCCTTGGCAACCCAGA-3′). PCR was carried out using primers TGEV-8(+) and leader(−) (5′-GAGCGATTTGCGTGCGTGCATCCCGC-3′), and the PCR product was sequenced directly.

***In vitro* translation and Western blotting.** For *in vitro* translation, 100 ng of transcript was translated for 1 h at 30°C in a 25-μl reaction mixture containing 12.5 μl wheat germ extract (Promega) and 60 mM potassium acetate as recommended by the manufacturer. Proteins were resolved by SDS-PAGE in gels of 10% polyacrylamide (58) and electroblotted onto Hybond ECL nitrocellulose membranes (GE Healthcare). The immobilized proteins were probed with rabbit anti-BCoV N (made by Proteintech Group, Inc., from bacteria-expressed purified N protein; product identification [ID] number 90186) as the primary antibody and with horseradish peroxidase-conjugated goat anti-rabbit IgG (Abcom) as the secondary antibody, and the blot was then incubated in SuperSignal West Pico chemiluminescent substrate (Thermo Scientific) for 1 min and exposed to Kodak XAR-5 film for imaging.

RESULTS

Long-range RNA-RNA base-pairing between the 5′ UTR and the nsp1-coding region is a *cis*-acting requirement for DI RNA replication. Inasmuch as the regions of stem-loops 5 (31), 6 (33), and 7 (26) were each shown to contribute a *cis*-acting function for BCoV DI RNA replication (see the introduction), we thought it possible that the specific regions of base pairing within the long-range interacting domain between the 5′ UTR and the partial nsp1-coding region (38) would act separately as a higher-order *cis*-acting feature for DI RNA replication or as a component of a larger structure connecting the stem-loops. The mfold program of Zuker et al. (54, 55) predicts the (+)-strand and separately the (−)-strand RNAs in this region to be folded as depicted in Fig. 2A were they to exist as single-stranded molecules.

To determine whether the long-range RNA-RNA base-paired structure functions as a *cis*-acting element for replication in DI RNA, we used the cDNA-cloned original WT (i.e., nonminimalized) DI RNA with the reporter sequence (WT pDrep1) (7) for mutation analyses, since it contains a convenient natural NdeI endonuclease site for mutagenesis (Fig. 2B and 3) and the extended length 3′-ward increased the number of potential frameshifting options (described below) while possibly retaining functional RNA structure. With the WT pDrep1 construct, sets of translationally silent mutations were made within the long-range stem that map within three regions of the ascending (left) and descending (right) locations and that were designed to disrupt base pairing in the (+) strand or (−) strand as depicted in Fig. 4A. Transcripts of each of these as well as of mutants containing their associated compensatory double mutations were tested for replication by transfection into BCoV-infected cells. Note that the mutant name corresponds to the panel with the same name. The replication and sequence reversion results are shown in Fig. 4B. Replication was judged positive when there was an increase in DI RNA abundance over time and when progeny DI RNA was present in cells at 48 h following infection with VP1 or VP2 (39). Following transfection of uninfected HRT-18 cells, transcripts of WT pDrep1 have a half-life of less than 2 h (39). For constructs that replicated, sequences of the intracellular DI RNA were determined with the use of reporter-specific primers to identify potential reversion to the WT sequence, which might have occurred via recombination with the helper virus genome.

In the upper left panel of Fig. 4A, reading from the top, note that mutations A167G and C165U retain base pairing in the (+) strand (i.e., G-U and U-G, respectively) but diminish base pairing in the (−) strand (i.e., C A and A C, respectively). Yet replication was nearly as robust as for the WT (compare lanes 2 and 1 in Fig. 4B). In the upper right panel of Fig. 4A, note that mutations U339C and G342A would diminish base pairing in the (+) strand (i.e., A C and C A, respectively) but retain base pairing in the (−) strand (i.e., U-G and G-U, respectively). In this mutant, replication was blocked (compare lanes 3 and 1 in Fig. 4B). The compensatory double mutation A167G, C165U, U339C, and G342A, however, reformed base pairing in both the (+) and (−) strands, and replication returned to near WT levels (compare lanes 4 and 1 in Fig. 4B). Overall, the results from the upper panels suggest that base pairing in the upper section of the stem in the (+) strand but not the (−) strand is important for DI RNA replication, and they also indicate that this part of the double-stranded long-range RNA-RNA structure functions as part of a *cis*-acting replication signal. Although these data and the genetics data from Guan et al. (38) support the existence of full-length long-distance RNA base pairing as depicted, we cannot currently rigorously rule out that stem-loop 7 has a regulatory role in DI RNA replication. More study is needed to evaluate the function of stem-loop 7 in this context.

In the middle left panel of Fig. 4A, note that mutations C158U and C155U would retain base pairing in the (+) strand (i.e., U-G and U-G, respectively) but diminish base pairing in the (−) strand (i.e., A C and A, respectively). Replication was nearly as robust as for the WT; however, there may be less efficient packaging (compare lanes 5 and 1 in Fig. 4B). In the middle right panel of Fig. 4A, note that mutations G348A and G351A would diminish base pair-

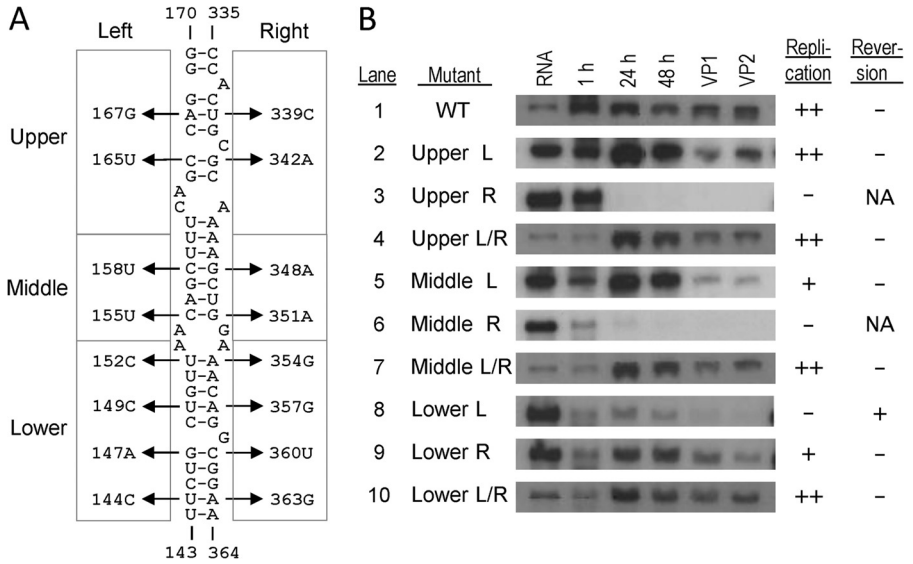


FIG 4 Replication of the DI RNA requires a long-range RNA-RNA interaction in the positive strand. (A) Mutations that alter base pairing in pDrep1-WT were made, and replication of RNA transcripts transfected into infected cells was measured by Northern blotting. (B) Results of Northern blotting. Reversion, reversion to WT sequence as a result of recombination with the helper virus genome; L, left; R, right; NA, not applicable.

ing in the (+) strand (i.e., C A and C A, respectively) but retain base pairing in the (−) strand (i.e., G-U and G-U, respectively). In this case, replication was blocked (compare lanes 6 and 1 in Fig. 4B). The compensatory double mutations C158U and C155U, along with G348A and G351A, however, restored base pairing in both the (+) and (−) strands, and replication was at near WT levels (compare lanes 7 and 1 in Fig. 4B). Overall, the results suggest that, as for the upper section, base pairing in the middle section in the (+) strand is important for DI RNA replication and that this section of the long-range double-stranded structure is part of a *cis*-acting replication signal.

In the lower left panel of Fig. 4A, note that mutations U152C, U149C, G147A, and U144C would diminish base pairing in the (+) strand (i.e., C A, C A, A C, and C A, respectively) but retain base pairing in the (−) strand (i.e., G-U, G-U, U-G, and G-U, respectively). This change appeared to allow weak replication (compare lanes 8 and 1 in Fig. 4B). However, sequencing of the small amount of VP1 progeny revealed that these molecules had reverted to the WT sequence. Therefore, we conclude that this mutant was blocked in replication. In the lower right panel of Fig. 4A, note that mutations A354G, A357G, C360U, and A363G would retain base pairing in the (+) strand (i.e., U-G, U-G, G-U, and U-G, respectively) but diminish base-pairing in the (−) strand (i.e., A C, A C, C A, and A C, respectively). In this case, replication was only slightly less than for the WT (compare lanes 9 and 1 in Fig. 4B). In the lower section, compensatory double mutations caused base pairing in both the (+) strand, (i.e., C-G, C-G, A-U, and C-G) and the (−) strand (G-C, G-C, U-A, and G-C), and replication was as robust as for the WT (compare lanes 10 and 1 in Fig. 4B). Taken together, the results with the lower panels suggest that the base pairing in the (+) strand is important.

Thus, overall, the results suggest that replication is the most robust when there is base pairing in both the (+) and (−) strands of all three sections of the long-range RNA-RNA base-paired structure but that base pairing is required in the (+) strand.

As illustrated, there is potential for a stem-loop 8 at the base of the lower panel that would not coexist with the long-range RNA-RNA interaction as shown. Currently, we have no experimental evidence for this stem-loop, but we note that it may be playing a role in replication.

Evidence for a *cis*-acting replication signal associated with the NH_2-proximal WAPEFPWM amino acid domain within the partial nsp1. It was previously reported that translation was a *cis* requirement for BCoV DI RNA replication (39), and in that study it was shown that the N protein encoded within the 3′-proximal region of the genome was required in *cis*, presumably to form a component of the replication complex similar to what has been described in other (+)-strand RNA viruses (59, 60). This feature is consistent with the association of N with the replication complex (61, 62). However, the 5′-terminal partial nsp1 region was not examined at that time for a *cis*-acting protein function. Precedents for a *cis*-acting protein in the replication of (+)-strand viral RNA genomes have been described (see Discussion) and led us to examine this possibility for BCoV DI RNA despite a remarkable amino acid sequence divergence between BCoV and MHV in this region (63). To determine whether there is a *cis*-acting protein function, we took three mutagenesis approaches: (i) frame shifting mutations designed to change the amino acid content of regions while maintaining predicted native RNA structure as much as possible, (ii) truncating the NH_2 terminal of the expressed protein within the nsp1 ORF to map a putative short 5′-proximal *cis*-acting region of nsp1 learned from the frameshifting experiments, and (iii) using point mutations to test the requirement for a phylogenetically conserved NH_2-proximal WAPEFPWM amino acid sequence that corresponds to the required region identified by the frameshift and truncation experiments. In each approach, replication of mutant DI RNAs was assayed by Northern blotting following transfection into helper virus-infected cells and Western blots of *in vitro* translation products were analyzed for the presence of the previously demonstrated *cis*-acting fused N-containing protein. A summary of all mutants used in these assays and their associated mutations is given in Table 1.

TABLE 1 pDrep1 mutants used in the study

Wild type or mutant	Mutation(s)[a]	Comments
WT		WT is background for M1, M2, M3, M27, M28, M30, M32, and M34 through 41
M1	ΔA217/+U222	Frameshift mutation in WT background
M2	ΔA227/+A234	Frameshift mutation in WT background
M3	A324U/G384U/A438U/A457U	Mutations knock out stop codons at positions 323, 383, 437, and 455 in the +2 reading frame of WT to form M3; M3 is background for M4 through M26
M4	+[CU]222/ΔU229/ΔA230	Frameshift mutation in M3 background
M5	+[CU]222/ΔA392/Δ393	Frameshift mutation in M3 background
M6	A224G/A227C/U229A	Changes amino acids NKY to STN in M3 background
M7	ΔA217	Frameshift mutation in M3 background
M8	ΔA217/+U222	Frameshift mutation in M3 background
M9	ΔA217/+C251	Frameshift mutation in M3 background
M10	ΔA217/+C305	Frameshift mutation in M3 background
M11	ΔA217/+U379	Frameshift mutation in M3 background
M12	ΔA217/+A464	Frameshift mutation in M3 background
M13	ΔA227	Frameshift mutation in M3 background
M14	ΔA227/+A234	Frameshift mutation in M3 background
M15	ΔA227/+C251	Frameshift mutation in M3 background
M16	ΔA227/+C305	Frameshift mutation in M3 background
M17	ΔA227/+U379	Frameshift mutation in M3 background
M18	ΔG233/+C251	Frameshift mutation in M3 background
M19	ΔG233/+C305	Frameshift mutation in M3 background
M20	ΔG233/+U379	Frameshift mutation in M3 background
M21	ΔG359/+A464	Frameshift mutation in M3 background
M22	ΔU176/ΔA224/+C251	Frameshift mutation in M3 background
M23	ΔU176/ΔA224/+C305	Frameshift mutation in M3 background
M24	ΔU176/ΔA224/+U379	Frameshift mutation in M3 background
M25	U273G/+C273/G276A/+C276/+A464	Frameshift mutation in M3 background
M26	+G321/+C331/+A464	Frameshift mutation in M3 background
M27	A189U/A211U/U212A	Changes 211AUG to 211UAG and produces U189 to base pair with A in 211UAG in WT background; M27 is background for M29, M31, and M33
M28	U177A/G178C/C222G/A223U	Changes 220AUC to 220AUG and produces U177A to base pair with A223U and G178C to base pair with C222G in WT background
M29	U177A/G178C/C222G/A223U	Changes 220AUC to 220AUG and produces U177A to base pair with A223U and G178C to base pair with C222G in M27 background
M30	U271A/G274A/A275U/A278U	Changes 274GAG to 274AUG; U271A and A278U strengthen Kozak context in WT background
M31	U271A/G274A/A275U/A278U	Changes 274GAG to 274AUG; U271A and A278U strengthen Kozak context in M27 background
M32	A290U	Changes 289AAG to 289AUG in WT background
M33	A290U	Changes 289AAG to 289AUG in M27 background
M34	G256A/U260C	Changes PEFP to PKSP in WT background
M35	C254U/G256A/U260C/C263U	Changes PEFP to LKSL in WT background
M36	C253U/G256A/U260C/C262U	Changes PEFP to SKSS in WT background
M37	C254U/C263U	Changes PEFP to LEFL in WT background
M38	C253U/C262U	Changes PEFP to SEFS in WT background
M39	A255G	Keeps PEFP but changes codon for P to CCG in WT background
M40	C253U	Changes PEFP to SEFP in WT background
M41	A255G/A258G/A264G	Keeps PEFP but changes codons for underlined amino acids to CCG, GAG, and CCG, respectively, in WT background

[a] Δ, deletion; +, insertion immediately upstream of the numbered position in the background construct.

The design of the frameshift experiments is shown in Fig. 5A and B, and the results are shown in Fig. 5B and C. As is evident from the Northern analyses (Fig. 5B), all frameshifted mutants except for those with changed NH_2-terminal amino acids 3 through 7 showed no evidence of replication. The block in replication could mean (i) that *cis*-acting RNA replication signals were disrupted by mutagenesis or (ii) that the mutated protein product was nonfunctional. The fact that amino acids 3 through 7 could be changed without killing replication (note results with mutants M1, M6, and M8) indicates that their WT character is not necessary for DI RNA replication. Note that mutant M4, made by frame shifting, did not replicate, whereas mutant M6, made by site-specific mutagenesis, did replicate, suggesting that the lesion preventing M4 replication was RNA structure mediated. Lethal results with four other mutants with altered amino acids at positions 8 through 14, however, could mean that WT amino acids within this window are important. These mutants are M9 with altered residues 3 through 14, M15 with altered residues 6 through 14, M18 with altered residues 8 through 14, and M22 with altered residues 5 through 14. Note that for all mutants except M7 and M13, in which the N ORF is out of frame with the upstream ORF, the N fusion protein with an altered partial nsp1 composition was made

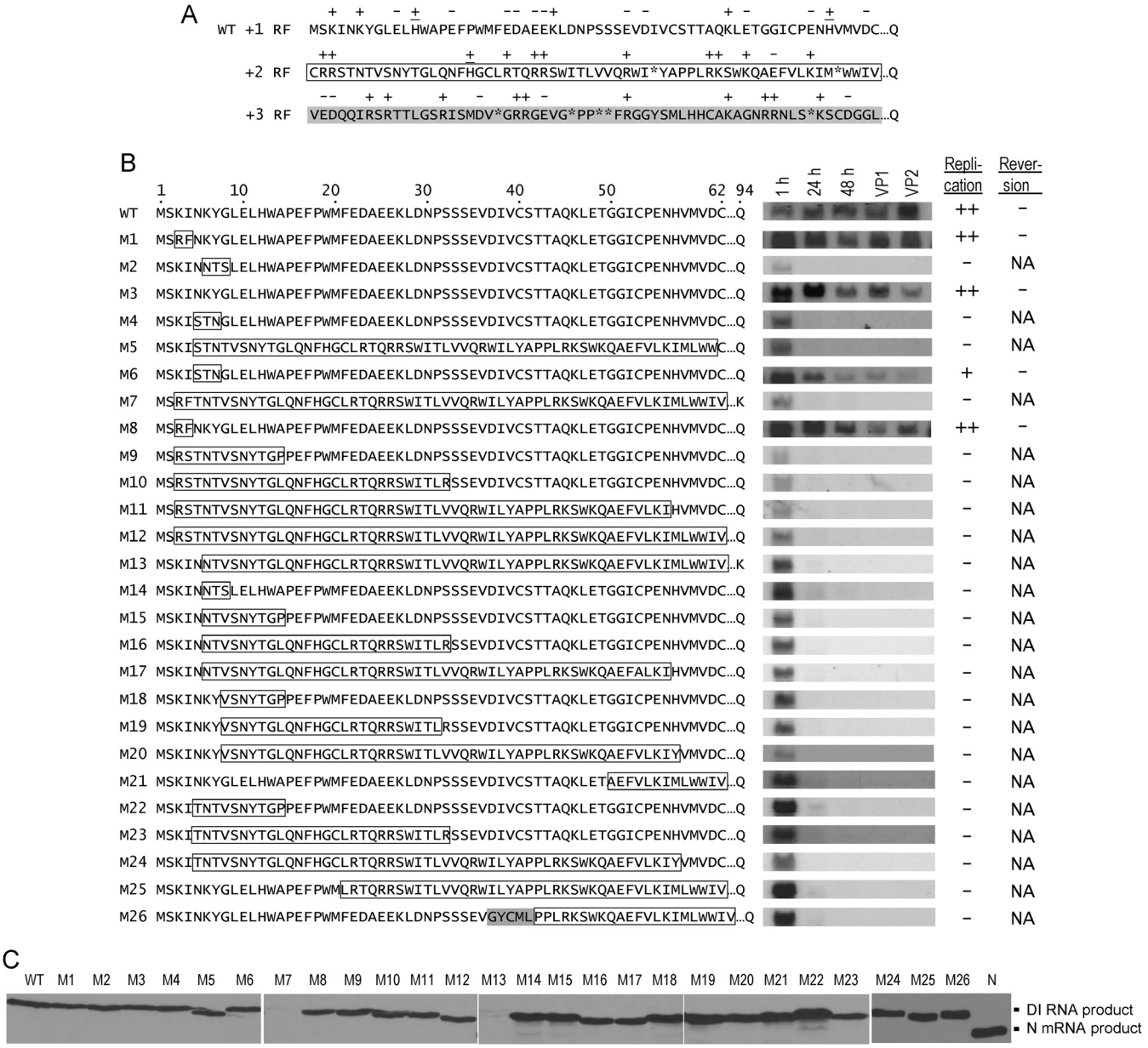


FIG 5 Frameshift mutation analyses reveal that nsp1 amino acids 3 through 7 can be changed from WT without blocking DI RNA replication. (A) Amino acid sequences of the WT (+1 reading frame) partial nsp1 sequence (unmarked), +2 reading frame (boxed lettering), and +3 reading frame (gray shading) for the first 62 amino acids of the partial nsp1 are shown. (B) Summary of the frameshifted sequences and replication results as determined by Northern blotting analyses and sequence reversion to WT as determined by RT-PCR sequencing. (C) Products from *in vitro* translation of each mutant transcript were analyzed by Western blotting with N protein-specific antibody. N, protein translated from transcripts of pNrep2. Both the fusion protein synthesized from DI RNA and that from N mRNA were identified by Western blotting with N-specific antibody. Because of an altered reading frame downstream in M7 and M13, the N ORF was not expressed. Note that mutations that blocked replication did not block translation.

by *in vitro* translation as evidenced by Western blotting with an N-specific antibody (Fig. 5C).

Since the lethal results with frameshifted mutations could have been caused by altered *cis*-acting RNA structures rather than by altered amino acids *per se*, a second mutational approach that was less likely to alter RNA structure was used to test for the importance of amino acids 8 through 14. This entailed replacing the nsp1 AUG start codon at nt 211 with a UAG stop codon (plus a silent A189U to maintain double strandedness at this site) to form M27 and testing for replication with an AUG start codon inserted at positions 4, 20 (site of a natural AUG codon), 22, and 27 in mutants M29, M27, M31, and M33, respectively, made in the M27 background (Fig. 6A, bottom panel). As controls, WT nsp1 constructs were made with the same mutations at these sites and with the natural AUG start codon at nt 211 left in place, and comparisons were made with WT, M28, M30, and M32 (Fig. 6A, top panel). Note that whereas the inserted AUG codon in M29 at amino acid position 4 led to replication, the AUG codon inserted at amino acid position 20 in M27, position 22 in M31, and position 27 in M33 did not (Fig. 6A, bottom panel). *In vitro* translation

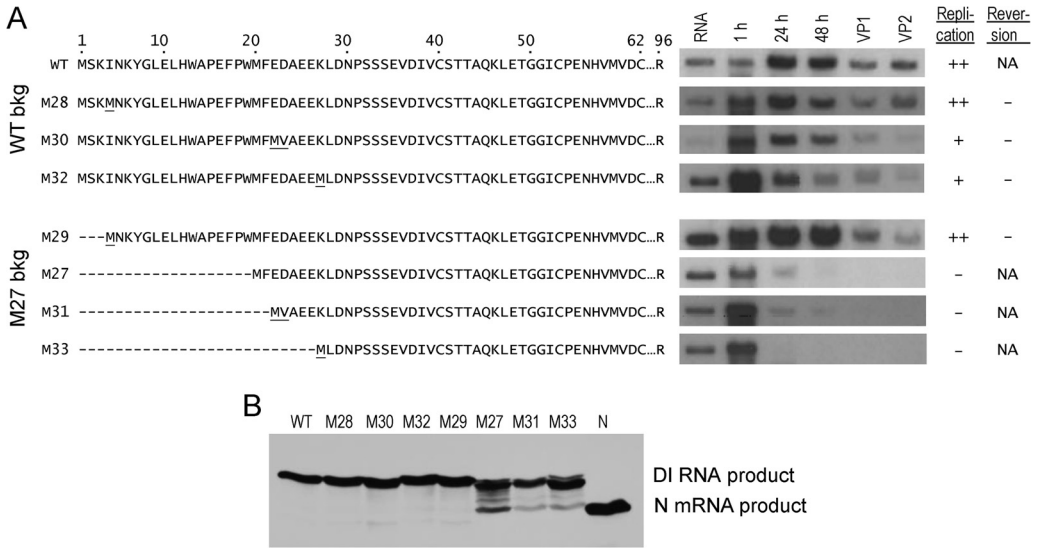


FIG 6 NH_2-terminal truncation of nsp1 amino acids 1 through 19 blocks replication but not translation of the DI RNA. (A) Amino acids synthesized in the WT and mutant constructs. (Upper) WT sequence and sequences for M28, M30, and M32 in which codons for amino acids 4, 22, and 27, respectively, were converted to AUG. Underlining identifies amino acids changed from those of the WT. Northern blotting results for replication and sequencing results for sequence reversion are shown at the right. (Lower) In mutant M27, the AUG codon at amino acid position 1 of WT was converted to UAG, and M27 was used as background for M29, M31, and M33, in which the codons for amino acids 4, 22, and 27, respectively, were converted to AUG. Underlining identifies amino acids changed from those of the WT. Northern blotting results for replication and sequencing results for sequence reversion are shown at the right. Note that no replication was observed when amino acids 1 through 19 were not expressed. NA, not applicable. (B) Western blotting results using N-specific antibody when transcripts used for Northern blot analysis were translated *in vitro*.

and Western blot analysis of these transcripts with an N-specific antibody, however (Fig. 6B), indicated that translation was not blocked and that failed translation *in vivo* is not a likely explanation for blockage in replication. All control constructs, M28, M30, and M32, replicated, although less robustly than the WT, while showing no evidence of reversion to a WT sequence (Fig. 6A, top panel), and they translated well (Fig. 6B). These results together suggest that WT amino acids at positions 5 through 20 are required for DI RNA replication.

A direct comparison of the 62 amino acids in the partial nsp1 between BCoV and MHV shows little sequence conservation; however, an 8-amino-acid stretch, from amino acid 13 through 20, WAPEFPWM, is evident (Fig. 7A). Upstream of amino acid 13, 6 of 12 amino acids (50%) differ, and downstream of amino acid 20, 24 of 42 amino acids (57%) differ. Interestingly, this 8-amino-acid conserved region appeared in an earlier study of eight group A lineage betacoronaviruses that were documented to function as helper viruses for the replication of BCoV pDrep1-WT (viruses in Fig. 7A without an asterisk) (63). The amino acid comparisons are shown here in an updated figure which includes the recently characterized canine respiratory coronavirus (CrCoV) (64) and the rabbit betacoronavirus (RbCoV) (65) (Fig. 7A). Of the 62 amino acids in the partial nsp1 region, those at positions 10 through 33 are coded by the sequence that forms the *cis*-acting stem-loop 6, and the octameric WAPEFPWM sequence is encoded by codons 13 through 20 within the ascending leg of this stem-loop (Fig. 7C). A comparison of stem-loop 6 among the 10 viruses listed in Fig. 7A shows the structures to be quite similar but not identical (Fig. 7C). Likewise, the codons encoding the eight amino acids differ among the viruses (Fig. 7B). This conservation of product from differing codons would suggest that there is an evolutionary pressure to keep the WAPEFPWM sequence. This, along with tolerance for adjacent sequence variations among the helper viruses supporting pDrep1-WT replication (63), also suggests that the WAPEFPWM sequence is important.

It should be noted that the International Committee on Taxonomy of Viruses has recommended that betacoronaviruses of the group A lineage now be organized into three species (of seven total species in the newly characterized betacoronavirus genus): species 1 (the BCoV-like canine respiratory coronaviruses CrCoV, human respiratory coronavirus strain OC43 [HCoV-OC43], human enteric coronavirus [HECoV], porcine hemagglutinating encephalomyelitis virus [HEV], equine coronavirus [ECoV], and rabbit coronavirus [RbCoV]), the MHV species, and the human coronavirus HKU1 (HKU1) species (http://ictvonline.org/virusTaxonomy.asp). With this in mind, we note that, whereas the betacoronavirus species 1 viruses and the MHV species share the entire 8-amino-acid sequence (WAPEFPWM), the sequence in the more distantly related HKU1 virus (66) (Fig. 7D) differs at amino acid positions 18 and 20 (WAPEFRWL) (Fig. 7D).

To test whether the WAPEFPWM is a necessary sequence for replication of DI RNA, we altered the central four amino acids by changing codons predicted to retain the same base-pairing pattern in the RNA secondary structure (Fig. 7E). For this process, two mutants that changed all four amino acids (M35 and M36), three mutants that changed two amino acids (M34, M37, and M38), one mutant that changed one amino acid (M40), and two mutants that retained the WT WAPEFPWM sequence but changed the codons for amino acids at position 15 (M39) and at positions 15, 16, and 18 (M41) were tested. For all mutants in which one or more of the four amino acids was changed, there was no replication, and for both mutants in which the WT amino acid sequence was retained but the codons changed, there was replication at WT levels (Fig. 7E). A Northern blotting assay for 18S rRNA demonstrated even

loading of cellular RNA among the samples (Fig. 7E). For the mutants that replicated (M39 and M41), sequencing of the product indicated there had been no reversion to the WT nucleotide sequence (Fig. 7E and data not shown). *In vitro* translation and a Western blotting assay to detect the N fusion protein, furthermore, demonstrated that translation of all mutants was complete (Fig. 7F), and therefore incomplete translation *in vivo* was not likely to be the cause of replication failure in mutants M34 through M38 and M40. These results therefore indicate that at least the four central amino acids, PEFP, within the 8-amino-acid sequence are important for the *cis*-acting translation function for replication. Whether or not the other specific amino acids from 20 to 62 are important for DI RNA replication was not tested; however, other sites of conserved identity might be important. The wide variation in amino acid composition would suggest that they are not all critical. Thus, we conclude that the WAPEFPWM amino acid sequence is a conserved peptide sequence within nsp1 of which at least the central four amino acids are critical in *cis* for DI RNA replication. To our knowledge, this is the first description of a *cis*-acting translation product in the 5′-proximal region of coronavirus nsp1.

DISCUSSION

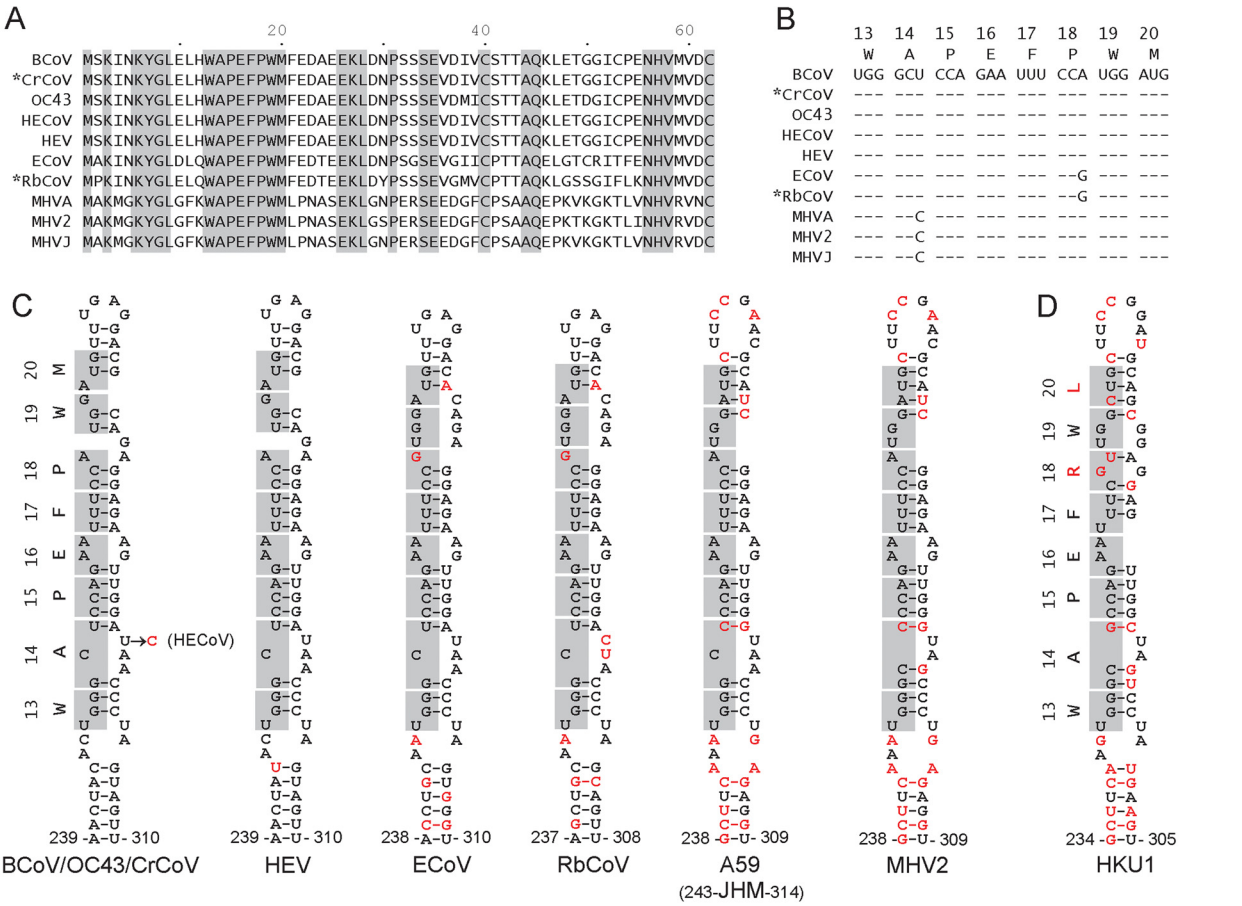
From previous work (38) and from work described here, it has become clear that for BCoV and MHV, different species in the group A lineage betacoronaviruses, the genomic 5′ untranslated region and the region encoding the NH_2-terminal 62 amino acids of nsp1 (identical between the virus genome and DI RNA) are structurally linked in a way that suggests a functional connection. Although the details of how each feature functions remain to be fully explored, the findings in this study indicate that (i) the 5′-proximal long-range higher-order RNA structure probably plays a direct role in genomic and DI RNA replication and maybe also in packaging, and (ii) translation of the NH_2-terminal region of nsp1 fulfills a *cis*-acting requirement for BCoV DI RNA replication. Conceptually each feature contributes by a different mechanism and will be discussed separately.

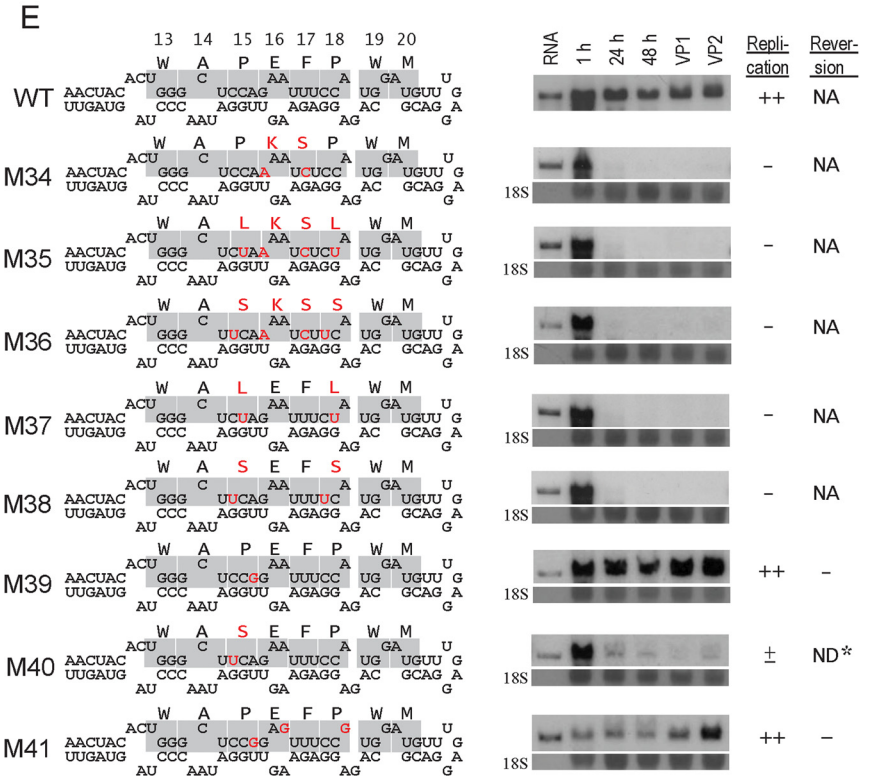
The *cis*-acting long-range RNA-RNA base-paired element. Our analysis of this structure in transfected DI RNA takes place when helper virus replication is well under way (1 h postinfection) (27), which means the replicase-coding region of the helper virus genome (ORF1) has been translated and the RNA-synthesizing machinery within a membrane-protected replication compartment (67–72) is fully active. The data show that DI RNA replication correlates with the 30-nt core of the long-range higher-order RNA structure in the (+) strand (nt 143 through 170 base-paired with nt 335 through 364, shaded in the upper part of Fig. 2A) but not in the (−) strand (shaded in the lower part of Fig. 2A) (Fig. 4). Since the entire long-range higher-order RNA structure (defined here as the 322-nt sequence, nt 75 through nt 396 [Fig. 2A]) is found in molecules that replicate (the genome and DI RNA) and not in molecules that fail to replicate (the sgmRNAs) and since both kinds of molecules can function as templates for (−)-strand synthesis (16, 73), a simple view is that this structure contributes key steps for the initiation of new (+)-strand RNA from the 3′ end of the (−)-strand template. How the higher-order structure facilitates this task is not clear, but we view it or some variation of it as an elaborated 5′-end RNA promoter in the nature of that described by Vogt and Andino for poliovirus and proposed for other (+)-strand RNA viruses (74). That is, the 5′ end functions as a promoter in *trans* locally for initiation of new (+)-strand synthesis (74). However, several elements within the higher-order RNA structure and possibly its extension to the 5′ end of the genome (i.e., nt 1 through 396) are also mechanistically associated with the RdRp template switching that occurs during discontinuous transcription (15, 75–77), and since discontinuous transcription is associated with the initiation of sgmRNA (+)-strand synthesis (11), we suggest an integrated view of the function of this higher-order RNA structure.

In conceptualizing what features the structured "promoter" might have, we envision three. (i) It could facilitate the initiation of (+)-strand synthesis at the 3′ end of the completed genomic (−)-strand template. In this sense, it would mimic aspects of the (+)-strand 5′-end promoter described for poliovirus (74) in which RNA structures engage different components of the polymerase complex. (ii) It could facilitate the RdRp template switching at the 5′ end of the genome by functioning as the acceptor template (UCUAAAAC) for the switch from intergenic donor signals in the genome (11) or in sgmRNAs (16). Functioning as an acceptor site would require that the higher-order RNA structure include much of the very 5′ end of the genome which harbors the leader (nt 1 through 65) (7), the UCUAAAC template-switching signal (nt 64 through 70) (27), the UUUAUAAA template-switching hot spot (nt 71 through 78) (78), and the 65-nt-wide template-switching window (nt 33 through 97) (75). Following the template switch and completion of (−)-strand synthesis, initiation of new (+)-stand synthesis would be facilitated as described above. The involvement of a higher-ordered RNA structure for template switching and for initiation of new (+)-strand synthesis would explain why the sgmRNAs that are missing in this structure fail to make new (+) strands and hence fail to replicate. (iii) It could facilitate the RdRp template switching at the 5′ end of the genome by functioning as the donor template (UCUAAAAC) during use of an alternate pathway for genome replication (75, 78). Although template switching at this site was described in earlier studies (78), the model used to explain the phenomenon was different; i.e., it suggested RdRp template switching takes place during (+)-strand synthesis. It is envisioned that the structures at the 5′ end enabling it to function as an acceptor template (described in section ii above) are the same as those that would enable it to function as a donor template. The higher-order structure might also be involved in an experimentally induced positive-to-negative-strand template switch of the RdRp that takes place in this region of the genome (76).

The long-range higher-order RNA structure might also be a packaging signal for the DI RNA, although the packaging signal described to date for the betacoronaviruses of group A lineage maps to a site within the downstream ORF1b region of the genome (79–82), a sequence that is missing in the BCoV DI RNA. Interestingly, a packaging signal nearly equivalent in position to the 5′-terminal 396-nt region studied here has been described for porcine transmissible gastroenteritis virus, an alphacoronavirus (83). At no time during the current study was replication observed in the absence of packaging, which might be expected if the long-range higher-order RNA structure functioned only as a replication signal. Further studies are needed to characterize the packaging signal for the BCoV DI RNA.

The *cis*-acting function of a nascent partial nsp1 protein in BCoV DI RNA replication. A novel finding in the current study is that translation of the NH_2-terminal portion of a partial nsp1


A
BCoV MSKINKYGLELHWAPEFPWMFEDAEEKLDNPSSSEVDIVCSTTAQKLETGGICPENHVMVDC
*CrCoV MSKINKYGLELHWAPEFPWMFEDAEEKLDNPSSSEVDIVCSTTAQKLETGGICPENHVMVDC
OC43 MSKINKYGLELHWAPEFPWMFEDAEEKLDNPSSSEVDMICSTTAQKLETDGICPENHVMVDC
HECoV MSKINKYGLELHWAPEFPWMFEDAEEKLDNPSSSEVDIVCSTTAQKLETGGICPENHVMVDC
HEV MSKINKYGLELHWAPEFPWMFEDAEEKLDNPSSSEVDIVCSTTAQKLETGGICPENHVMVDC
ECoV MAKINKYGLDLQWAPEFPWMFEDTEEKLDNPSGSEVGIICPTTAQELGTCRITFENHVMVDC
*RbCoV MPKINKYGLELQWAPEFPWMFEDTEEKLDYPSSSEVGMVCPTTAQKLGSSGIFLKNHVMVDC
MHVA MAKMGKYGLGFKWAPEFPWMLPNASEKLGNPERSEEDGFCPSAAQEPKVKGKTLVNHVRVNC
MHV2 MAKMGKYGLGFKWAPEFPWMLPNASEKLGSPERSEEDGFCPSAAQEPKTKGKTLINHVRVDC
MHVJ MAKMGKYGLGFKWAPEFPWMLPNASEKLGNPERSEEDGFCPSAAQEPKVKGKTLVNHVRVDC
B
13 14 15 16 17 18 19 20
W A P E F P W M
BCoV UGG GCU CCA GAA UUU CCA UGG AUG
*CrCoV --- --- --- --- --- --- --- ---
OC43 --- --- --- --- --- --- --- ---
HECoV --- --- --- --- --- --- --- ---
HEV --- --- --- --- --- --- --- ---
ECoV --- --- --- --- --- --G --- ---
*RbCoV --- --- --- --- --- --G --- ---
MHVA --- --C --- --- --- --- --- ---
MHV2 --- --C --- --- --- --- --- ---
MHVJ --- --C --- --- --- --- --- ---
C
U→C (HECoV)
239 -A-U- 310
BCoV/OC43/CrCoV
239 -A-U- 310
HEV
238 -A-U- 310
ECoV
237 -A-U- 308
RbCoV
238 -G-U- 309
A59
(243-JHM-314)
238 -G-U- 309
MHV2
D
234 -G-U- 305
HKU1


E
13 14 15 16 17 18 19 20
W A P E F P W M
WT AACUAC GGG UCCAG UUUCC UG UGUU G UUGAUG CCC AGGUU AGAGG AC GCAG A AU AAU GA AG G
M34 W A P K S P W M AACUAC GGG UCCAA UCUCC UG UGUU G
M35 W A L K S L W M AACUAC GGG UCUAA UCUCU UG UGUU G
M36 W A S K S S W M AACUAC GGG UUCAA UCUUC UG UGUU G
M37 W A L E F L W M AACUAC GGG UCUAG UUUCU UG UGUU G
M38 W A S E F S W M AACUAC GGG UUCAG UUUUC UG UGUU G
M39 W A P E F P W M AACUAC GGG UCCGG UUUCC UG UGUU G
M40 W A S E F P W M AACUAC GGG UUCAG UUUCC UG UGUU G
M41 W A P E F P W M AACUAC GGG UCCGG UUUCC UG UGUU G
RNA
1 h
24 h
48 h
VP1
VP2
Repli-cation
Rever-sion
18S
WT ++ NA
M34 - NA
M35 - NA
M36 - NA
M37 - NA
M38 - NA
M39 ++ -
M40 ± ND*
M41 ++ -
F
WT M34 M35 M36 M37 M38 M39 M40 M41

ORF, and possibly synthesis of only a few amino acids within this stretch, is required in *cis* for replication of the BCoV DI RNA in virus-infected cells. This conclusion differs from that in two previous reports on MHV DI RNA replication in which it was judged that a product of translation in *cis* was not required for replication (24, 40). How might the two sets of results be reconciled? In the first study (24), a naturally occurring DI RNA of MHV-A59 with a long ORF consisting of an in-frame fusion of a 5′-terminal portion of ORF 1a (3,680 nt), a partial ORF 1b, and a partial N ORF replicated even after frameshifting point mutations had severely truncated the fused ORF or when the spike gene, not a normal part of a naturally occurring DI RNA, had been substituted in-frame for parts of the 1a and 1b regions in the original ORF. All mutants replicated, but in each the 5′-terminal WAPEFPWM domain was intact and therefore potentially able to function for replication as observed in the current study. In the second (40), a naturally occurring DI RNA of MHV-JHM in which in-frame fusions of a 5′-terminal partial ORF1a (634 nt), a partial ORF1b, and a partial N ORF formed a long ORF. Three ORF-truncating mutants of this construct were made by inserting an in-frame 12-nt oligonucleotide that ensured a UAG amber (stop) codon within each reading frame. All three mutants replicated, and in each the 5′-terminal WAPEFPWM domain was intact and therefore could have enabled replication as observed in the current study. In the fourth and fifth MHV-JHM DI RNA mutants, however, the AUG codons at the beginning of ORF1a in a truncated construct were changed to non-start codons, and the result was that translation of ORF1a was blocked but replication proceeded unimpaired. We speculate that perhaps genomic RNA structures such as parts of ORF1b retained within the severely truncated MHV DI RNAs but not found in the BCoV DI RNA (39) facilitate an alternate access to the replication compartment.

Although discovered in BCoV DI RNA, the requirement for the WAPEFPWM domain in *cis* for DI RNA replication may have implications for genome replication, although this should be viewed with the caveat that *cis*-acting replication elements for DI RNA are not always the same for the viral genome. An example is the 5′-proximal stem-loop 4 in MHV (29, 30). If the WAPEFPWM domain does function in the genome, it would likely not interact with a replicase protein as postulated for the DI RNA (see below). During virus genome translation immediately following virus entry, autoproteolytic processing of viral polyproteins 1a and 1ab takes place and the hydrophobic membrane-anchoring domains within nsp3, nsp4, and nsp6 of the nsp1-nsp10 polyprotein 1a precursor function to assemble the endoplasmic reticulum membrane-associated replication/transcription complex on the cytoplasmic side of the endoplasmic reticulum membrane (84, 85). Presumably the partial nsp1 product described here for the DI RNA would be functional on the NH_2 terminus of the full-length nsp1 product as well, but what it would interact with at this time is not apparent. Perhaps it interacts with a replicase protein prior to its final assembly in the replicase complex.

With regard to BCoV DI RNA replication, there are no (obvious) hydrophobic domains in the translated product of DI RNA, so other mechanism(s) for engaging the replication apparatus must be involved (84, 85). At the time of DI RNA transfection, viral replication is under way (27) and the replication compartment is formed and all *trans*-acting factors are in place. The transfected BCoV DI RNA, therefore, is totally dependent on the *trans*-acting factors in the infected cell and on its own RNA structures and on the newly synthesized protein to engage the viral replication machinery. It is probable that *cis*-acting RNA structures in the DI RNA are recruited to the replication complexes through interaction with the viral replicase proteins (73). However, one possibility derived from the current study is that the *cis*-acting WAPEFPWM element is a signal to engage some component of the membrane-associated viral replicase. Characterization of this association will require an identification of the molecular partner(s) of the WAPEFPWM moiety. The NH_2-terminal region of the coronavirus nsp1 contained within the first 62 amino acids predicts only a hydrophilic region with charged amino acids, which suggests there may be a protein-protein interaction (Fig. 5A).

Although not identical in detail, precedents for *cis*-acting proteins in other viral systems include some whose function it is to engage a replicase partner in the replication compartment. One of the more characterized is the protein 1a-2a interaction found in two members of the brome mosaic virus family, which may serve as a paradigm for the *cis*-acting behavior of the partial nsp1 protein described here (86, 87). (i) Protein 2a (a polymerase) harbors a *cis*-acting signal for interaction with its 1a partner (a multidomain RNA replication protein) for RNA replication (86). (ii) Protein 2a, while in the process of translating RNA2 and by way of a protein 1a-2a interaction, recruits RNA2 to the replication complex, which results in its replication (86). Precedents for such associations have also been described for cellular mRNAs that are translated within vesicles (reviewed in reference 88). From these examples, we envision a similar picture for the *cis*-acting function of the partial nsp1 (Fig. 8). For the localization of the BCoV DI RNA within the replication compartment, we propose that the

FIG 7 Point mutations that change amino acids within the WAPEFPWM domain block DI RNA replication, whereas mutations that change the codons but keep the WT amino acids enable replication. (A) Alignment of the first 62 amino acids in nsp1 among 10 different group A lineage betacoronaviruses reveals a conserved WAPEFPWM domain. All eight viruses from this group that were tested (those without an asterisk) supported replication of the BCoV WT DI RNA (63). Viruses and GenBank accession numbers: BCoV-Mebus, bovine coronavirus strain Mebus (accession number NC_U00735); CrCoV-K37, canine respiratory coronavirus strain K37 (accession number JX860640); HCoV-OC43, human respiratory coronavirus strain OC43 (accession number NC_005147); HECoV, human enteric coronavirus strain 4408 (accession number AF523844); PHEV-VW572, porcine hemagglutinating encephalomyelitis virus strain VW572 (accession number NC_007732); ECoV-NC99, equine coronavirus strain NC99 (accession number NC_010327); RbCoV-HKU14, rabbit coronavirus strain HKU14 (accession number NC_017083); MHV-A59, mouse hepatitis virus strain A59 (accession number NC_001846); MHV-2, mouse hepatitis virus strain 2 (accession number AF201929); MHV-JHM, mouse hepatitis virus strain JHM (accession number NC_006852). (B) Codon variations among the group A lineage betacoronaviruses that encode the WAPEFPWM domain. (C) Predicted structural differences within stem-loop 6, which carries the nucleotides encoding the WAPEFPWM domain. Red letters identify nucleotides or amino acids that differ from those in WT BCoV. (D) Human coronavirus strain HKU1 (HCoV-HKU1 [accession number NC_006577]), a proposed separate betacoronavirus species, encodes a similar domain but with an R and L at amino acid positions 18 and 20, respectively. (E) Mutations that disrupt the WAPEFPWM domain block DI RNA replication, whereas synonymous codons encoding WT amino acids enable replication. Northern blot analyses show DI RNA abundances at various times posttransfection; analyses for 18S rRNA abundances serve as loading controls. NA, not applicable; ND*, not determined, since in two separate transfection experiments not enough RT-PCR material was obtained for sequencing. (F) Western blot analyses show the synthesis of N within the fusion protein of *in vitro* translation products for mutants M34 through M41.

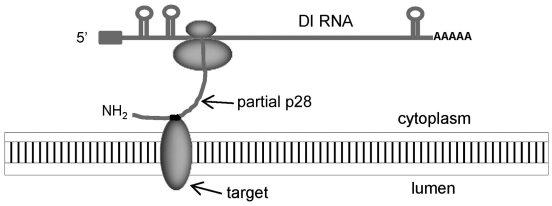


FIG 8 Model for *cis* function of partial nsp1. One model showing how a nascent peptide during translation could direct the mRNA template to a specific membrane-associated RNA replication site. In this model, the WAPEFPWM amino acid sequence interacts with an endoplasmic reticulum membrane-anchored viral (or possibly cellular) protein.

nascent partial nsp1 protein encoded by the DI RNA as it emerges from the 80S ribosomal subunit finds its target cotranslationally. Following this step, replication of the DI RNA would ensue. The target for the WAPEFPWM domain is unknown at this time. Whether the viral genomic RNA employs some part of this scheme remains to be determined.

A cursory search for a sequence analogous to the WAPEFPWM domain in other coronavirus groups has not revealed such a site. However, it is notable that the NH_2-terminal regions of the alphacoronaviruses and SARS-CoV show sites of amino acid similarity (89, 90). It could be that *cis*-acting protein sites are present, but in a more dispersed pattern than found here.

ACKNOWLEDGMENTS

This work was supported by grant AI 14376 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.

REFERENCES

1. **Baric RS, Stohlman SA, Lai MM.** 1983. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. J. Virol. **48:**633–640.
2. **Brian DA, Baric RS.** 2005. Coronavirus genome structure and replication. Curr. Top. Microbiol. Immunol. **287:**1–30. http://dx.doi.org/10.1007/3-540-26765-4_1.
3. **Masters PS.** 2006. The molecular biology of coronaviruses. Adv. Virus Res. **66:**193–292. http://dx.doi.org/10.1016/S0065-3527(06)66005-3.
4. **Sawicki D, Wang T, Sawicki S.** 2001. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. J. Gen. Virol. **82:**385–396. http://vir.sgmjournals.org/content/82/2/385.long.
5. **Sethna PB, Hung SL, Brian DA.** 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. Proc. Natl. Acad. Sci. U. S. A. **86:**5626–5630. http://dx.doi.org/10.1073/pnas.86.14.5626.
6. **Sawicki SG, Sawicki DL.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J. Virol. **64:**1050–1056.
7. **Chang RY, Hofmann MA, Sethna PB, Brian DA.** 1994. A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. J. Virol. **68:**8223–8231.
8. **Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ.** 2006. Nidovirales: evolving the largest RNA virus genome. Virus Res. **117:**17–37. http://dx.doi.org/10.1016/j.virusres.2006.01.017.
9. **Bredenbeek PJ, Pachuk CJ, Noten AF, Charite J, Luytjes W, Weiss SR, Spaan WJ.** 1990. The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. Nucleic Acids Res. **18:**1825–1832. http://dx.doi.org/10.1093/nar/18.7.1825.
10. **Sawicki SG, Sawicki DL.** 2005. Coronavirus transcription: a perspective. Curr. Top. Microbiol. Immunol. **287:**31–55. http://dx.doi.org/10.1007/3-540-26765-4_2.
11. **Sawicki SG, Sawicki DL, Siddell SG.** 2007. A contemporary view of coronavirus transcription. J. Virol. **81:**20–29. http://dx.doi.org/10.1128/JVI.01358-06.
12. **Hofmann MA, Sethna PB, Brian DA.** 1990. Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. J. Virol. **64:**4108–4114.
13. **Sethna PB, Hofmann MA, Brian DA.** 1991. Minus-strand copies of replicating coronavirus mRNAs contain antileaders. J. Virol. **65:**320–325.
14. **van Marle G, Dobbe JC, Gultyaev AP, Luytjes W, Spaan WJ, Snijder EJ.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. Proc. Natl. Acad. Sci. U. S. A. **96:**12056–12061. http://dx.doi.org/10.1073/pnas.96.21.12056.
15. **Zuniga S, Sola I, Alonso S, Enjuanes L.** 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J. Virol. **78:**980–994. http://dx.doi.org/10.1128/JVI.78.2.980-994.2004.
16. **Wu HY, Brian DA.** 2010. Subgenomic messenger RNA amplification in coronaviruses. Proc. Natl. Acad. Sci. U. S. A. **107:**12257–12262. http://dx.doi.org/10.1073/pnas.1000378107.
17. **Brian DA, Spaan WJM.** 1997. Recombination and coronavirus defective interfering RNAs. Semin. Virol. **8:**101–111. http://dx.doi.org/10.1006/smvy.1997.0109.
18. **Makino S, Shieh CK, Soe LH, Baker SC, Lai MM.** 1988. Primary structure and translation of a defective interfering RNA of murine coronavirus. Virology **166:**550–560. http://dx.doi.org/10.1016/0042-6822(88)90526-0.
19. **van der Most RG, Bredenbeek PJ, Spaan WJ.** 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.
20. **Masters PS, Koetzner CA, Kerr CA, Heo Y.** 1994. Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis virus. J. Virol. **68:**328–337.
21. **Mendez A, Smerdou C, Izeta A, Gebauer F, Enjuanes L.** 1996. Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. Virology **217:**495–507. http://dx.doi.org/10.1006/viro.1996.0144.
22. **Penzes Z, Wroe C, Brown TD, Britton P, Cavanagh D.** 1996. Replication and packaging of coronavirus infectious bronchitis virus defective RNAs lacking a long open reading frame. J. Virol. **70:**8660–8668.
23. **Makino S, Fujioka N, Fujiwara K.** 1985. Structure of the intracellular defective viral RNAs of defective interfering particles of mouse hepatitis virus. J. Virol. **54:**329–336.
24. **van der Most RG, Luytjes W, Rutjes S, Spaan WJ.** 1995. Translation but not the encoded sequence is essential for the efficient propagation of the defective interfering RNAs of the coronavirus mouse hepatitis virus. J. Virol. **69:**3744–3751.
25. **Kim YN, Jeong YS, Makino S.** 1993. Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication. Virology **197:**53–63. http://dx.doi.org/10.1006/viro.1993.1566.
26. **Brown CG, Nixon KS, Senanayake SD, Brian DA.** 2007. An RNA stem-loop within the bovine coronavirus nsp1 coding region is a *cis*-acting element in defective interfering RNA replication. J. Virol. **81:**7716–7724. http://dx.doi.org/10.1128/JVI.00549-07.
27. **Chang RY, Krishnan R, Brian DA.** 1996. The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. J. Virol. **70:**2720–2729.
28. **Raman S, Bouma P, Williams GD, Brian DA.** 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. http://dx.doi.org/10.1128/JVI.77.12.6720-6730.2003.
29. **Wu HY, Guan BJ, Su YP, Fan YH, Brian DA.** 2014. Reselection of a genomic upstream open reading frame in mouse hepatitis coronavirus 5′-untranslated-region mutants. J. Virol. **88:**846–858. http://dx.doi.org/10.1128/JVI.02831-13.
30. **Yang D, Liu P, Giedroc DP, Leibowitz J.** 2011. Mouse hepatitis virus stem-loop 4 functions as a spacer element required to drive subgenomic RNA synthesis. J. Virol. **85:**9199–9209. http://dx.doi.org/10.1128/JVI.05092-11.
31. **Raman S, Brian DA.** 2005. Stem-loop IV in the 5′ untranslated region is a *cis*-acting element in bovine coronavirus defective interfering RNA replication. J. Virol. **79:**12434–12446. http://dx.doi.org/10.1128/JVI.79.19.12434-12446.2005.
32. **Guan BJ, Wu HY, Brian DA.** 2011. An optimal *cis*-replication stem-loop IV in the 5′ untranslated region of the mouse coronavirus genome extends

16 nucleotides into open reading frame 1. J. Virol. **85:**5593–5605. http://dx.doi.org/10.1128/JVI.00263-11.
33. **Gustin KM, Guan BJ, Dziduszko A, Brian DA.** 2009. Bovine coronavirus nonstructural protein 1 (p28) is an RNA binding protein that binds terminal genomic *cis*-replication elements. J. Virol. **83:**6087–6097. http://dx.doi.org/10.1128/JVI.00160-09.
34. **Liu P, Li L, Millership JJ, Kang H, Leibowitz JL, Giedroc DP.** 2007. A U-turn motif-containing stem-loop in the coronavirus 5′ untranslated region plays a functional role in replication. RNA **13:**763–780. http://dx.doi.org/10.1261/rna.261807.
35. **Li L, Kang H, Liu P, Makkinje N, Williamson ST, Leibowitz JL, Giedroc DP.** 2008. Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. J. Mol. Biol. **377:**790–803. http://dx.doi.org/10.1016/j.jmb.2008.01.068.
36. **Chen SC, Olsthoorn RC.** 2010. Group-specific structural features of the 5′-proximal sequences of coronavirus genomic RNAs. Virology **401:**29–41. http://dx.doi.org/10.1016/j.virol.2010.02.007.
37. **Kang H, Feng M, Schroeder ME, Giedroc DP, Leibowitz JL.** 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. http://dx.doi.org/10.1128/JVI.00455-06.
38. **Guan BJ, Su YP, Wu HY, Brian DA.** 2012. Genetic evidence of a long-range RNA-RNA interaction between the genomic 5′ untranslated region and the nonstructural protein 1 coding region in murine and bovine coronaviruses. J. Virol. **86:**4631–4643. http://dx.doi.org/10.1128/JVI.06265-11.
39. **Chang RY, Brian DA.** 1996. *cis* Requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. J. Virol. **70:**2201–2207.
40. **Liao CL, Lai MM.** 1995. A cis-acting viral protein is not required for the replication of a coronavirus defective-interfering RNA. Virology **209:**428–436. http://dx.doi.org/10.1006/viro.1995.1275.
41. **Brockway SM, Denison MR.** 2005. Mutagenesis of the murine hepatitis virus nsp1-coding region identifies residues important for protein processing, viral RNA synthesis, and viral replication. Virology **340:**209–223. http://dx.doi.org/10.1016/j.virol.2005.06.035.
42. **Wathelet MG, Orr M, Frieman MB, Baric RS.** 2007. Severe acute respiratory syndrome coronavirus evades antiviral signaling: role of nsp1 and rational design of an attenuated strain. J. Virol. **81:**11620–11633. http://dx.doi.org/10.1128/JVI.00702-07.
43. **Tanaka T, Kamitani W, DeDiego ML, Enjuanes L, Matsuura Y.** 2012. Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. J. Virol. **86:**11128–11137. http://dx.doi.org/10.1128/JVI.01700-12.
44. **Züst R, Cervantes-Barragan L, Kuri T, Blakqori G, Weber F, Ludewig B, Thiel V.** 2007. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. PLoS Pathog. **3:**e109. http://dx.doi.org/10.1371/journal.ppat.0030109.
45. **Kamitani W, Narayanan K, Huang C, Lokugamage K, Ikegami T, Ito N, Kubo H, Makino S.** 2006. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proc. Natl. Acad. Sci. U. S. A. **103:**12885–12890. http://dx.doi.org/10.1073/pnas.0603144103.
46. **Narayanan K, Huang C, Lokugamage K, Kamitani W, Ikegami T, Tseng CT, Makino S.** 2008. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. J. Virol. **82:**4471–4479. http://dx.doi.org/10.1128/JVI.02472-07.
47. **Kamitani W, Huang C, Narayanan K, Lokugamage KG, Makino S.** 2009. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. Nat. Struct. Mol. Biol. **16:**1134–1140. http://dx.doi.org/10.1038/nsmb.1680.
48. **Tohya Y, Narayanan K, Kamitani W, Huang C, Lokugamage K, Makino S.** 2009. Suppression of host gene expression by nsp1 proteins of group 2 bat coronaviruses. J. Virol. **83:**5282–5288. http://dx.doi.org/10.1128/JVI.02485-08.
49. **Huang C, Lokugamage KG, Rozovics JM, Narayanan K, Semler BL, Makino S.** 2011. Alphacoronavirus transmissible gastroenteritis virus nsp1 protein suppresses protein translation in mammalian cells and in cell-free HeLa cell extracts but not in rabbit reticulocyte lysate. J. Virol. **85:**638–643. http://dx.doi.org/10.1128/JVI.01806-10.
50. **Huang C, Lokugamage KG, Rozovics JM, Narayanan K, Semler BL, Makino S.** 2011. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. PLoS Pathog. **7:**e1002433. http://dx.doi.org/10.1371/journal.ppat.1002433.
51. **Lokugamage KG, Narayanan K, Huang C, Makino S.** 2012. Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. J. Virol. **86:**13598–13608. http://dx.doi.org/10.1128/JVI.01958-12.
52. **Lei L, Ying S, Baojun L, Yi Y, Xiang H, Wenli S, Zounan S, Deyin G, Qingyu Z, Jingmei L, Guohui C.** 2013. Attenuation of mouse hepatitis virus by deletion of the LLRKxGxKG region of Nsp1. PLoS One **8:**e61166. http://dx.doi.org/10.1371/journal.pone.0061166.
53. **Tompkins WA, Watrach AM, Schmale JD, Schultz RM, Harris JA.** 1974. Cultural and antigenic properties of newly established cell strains derived from adenocarcinomas of the human colon and rectum. J. Natl. Cancer Inst. **52:**1101–1110.
54. **Zuker M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415. http://dx.doi.org/10.1093/nar/gkg595.
55. **Mathews DH, Sabina J, Zuker M, Turner DH.** 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. **288:**911–940. http://dx.doi.org/10.1006/jmbi.1999.2700.
56. **Horton RM, Cai ZL, Ho SN, Pease LR.** 1990. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques **8:**528–535.
57. **Senanayake SD, Brian DA.** 1995. Precise large deletions by the PCR-based overlap extension method. Mol. Biotechnol. **4:**13–15. http://dx.doi.org/10.1007/BF02907467.
58. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685. http://dx.doi.org/10.1038/227680a0.
59. **Novak JE, Kirkegaard K.** 1994. Coupling between genome translation and replication in an RNA virus. Genes Dev. **8:**1726–1737. http://dx.doi.org/10.1101/gad.8.14.1726.
60. **Weiland JJ, Dreher TW.** 1993. Cis-preferential replication of the turnip yellow mosaic virus RNA genome. Proc. Natl. Acad. Sci. U. S. A. **90:**6095–6099. http://dx.doi.org/10.1073/pnas.90.13.6095.
61. **Verheije MH, Hagemeijer MC, Ulasli M, Reggiori F, Rottier PJ, Masters PS, de Haan CA.** 2010. The coronavirus nucleocapsid protein is dynamically associated with the replication-transcription complexes. J. Virol. **84:**11575–11579. http://dx.doi.org/10.1128/JVI.00569-10.
62. **Hurst KR, Ye R, Goebel SJ, Jayaraman P, Masters PS.** 2010. An interaction between the nucleocapsid protein and a component of the replicase-transcriptase complex is crucial for the infectivity of coronavirus genomic RNA. J. Virol. **84:**10276–10288. http://dx.doi.org/10.1128/JVI.01287-10.
63. **Wu HY, Guy JS, Yoo D, Vlasak R, Urbach E, Brian DA.** 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. http://dx.doi.org/10.1016/S0042-6822(03)00511-7.
64. **Lim SI, Choi S, Lim JA, Jeoung HY, Song JY, Dela Pena RC, An DJ.** 2013. Complete genome analysis of canine respiratory coronavirus. Genome Announc. **1:**e00093–12. http://dx.doi.org/10.1128/genomeA.00093-12.
65. **Lau SK, Woo PC, Yip CC, Fan RY, Huang Y, Wang M, Guo R, Lam CS, Tsang AK, Lai KK, Chan KH, Che XY, Zheng BJ, Yuen KY.** 2012. Isolation and characterization of a novel *Betacoronavirus* subgroup A coronavirus, rabbit coronavirus HKU14, from domestic rabbits. J. Virol. **86:**5481–5496. http://dx.doi.org/10.1128/JVI.06927-11.
66. **Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, Wong BH, Poon RW, Cai JJ, Luk WK, Poon LL, Wong SS, Guan Y, Peiris JS, Yuen KY.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. **79:**884–895. http://dx.doi.org/10.1128/JVI.79.2.884-895.2005.
67. **den Boon JA, Diaz A, Ahlquist P.** 2010. Cytoplasmic viral replication complexes. Cell Host Microbe **8:**77–85. http://dx.doi.org/10.1016/j.chom.2010.06.010.
68. **den Boon JA, Ahlquist P.** 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. Annu. Rev. Microbiol. **64:**241–256. http://dx.doi.org/10.1146/annurev.micro.112408.134012.
69. **van Hemert MJ, van den Worm SH, Knoops K, Mommaas AM, Gorbalenya AE, Snijder EJ.** 2008. SARS-coronavirus replication/

transcription complexes are membrane-protected and need a host factor for activity in vitro. PLoS Pathog. **4:**e1000054. http://dx.doi.org/10.1371/journal.ppat.1000054.
70. **Sethna PB, Brian DA.** 1997. Coronavirus genomic and subgenomic minus-strand RNAs copartition in membrane-protected replication complexes. J. Virol. **71:**7744–7749.
71. **Ulasli M, Verheije MH, de Haan CA, Reggiori F.** 2010. Qualitative and quantitative ultrastructural analysis of the membrane rearrangements induced by coronavirus. Cell. Microbiol. **12:**844–861. http://dx.doi.org/10.1111/j.1462-5822.2010.01437.x.
72. **Hagemeijer MC, Verheije MH, Ulasli M, Shaltiel IA, de Vries LA, Reggiori F, Rottier PJ, de Haan CA.** 2010. Dynamics of coronavirus replication-transcription complexes. J. Virol. **84:**2134–2149. http://dx.doi.org/10.1128/JVI.01716-09.
73. **Züst R, Miller TB, Goebel SJ, Thiel V, Masters PS.** 2008. Genetic interactions between an essential 3′ *cis*-acting RNA pseudoknot, replicase gene products, and the extreme 3′ end of the mouse coronavirus genome. J. Virol. **82:**1214–1228. http://dx.doi.org/10.1128/JVI.01690-07.
74. **Vogt DA, Andino R.** 2010. An RNA element at the 5′-end of the poliovirus genome functions as a general promoter for RNA synthesis. PLoS Pathog. **6:**e1000936. http://dx.doi.org/10.1371/journal.ppat.1000936.
75. **Wu HY, Ozdarendeli A, Brian DA.** 2006. Bovine coronavirus 5′-proximal genomic acceptor hotspot for discontinuous transcription is 65 nucleotides wide. J. Virol. **80:**2183–2193. http://dx.doi.org/10.1128/JVI.80.5.2183-2193.2006.
76. **Wu HY, Brian DA.** 2007. 5′-proximal hot spot for an inducible positive-to-negative-strand template switch by coronavirus RNA-dependent RNA polymerase. J. Virol. **81:**3206–3215. http://dx.doi.org/10.1128/JVI.01817-06.
77. **Ozdarendeli A, Ku S, Rochat S, Williams GD, Senanayake SD, Brian DA.** 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. http://dx.doi.org/10.1128/JVI.75.16.7362-7374.2001.
78. **Zhang X, Lai MM.** 1996. A 5′-proximal RNA sequence of murine coronavirus as a potential initiation site for genomic-length mRNA transcription. J. Virol. **70:**705–711.
79. **Chen SC, van den Born E, van den Worm SH, Pleij CW, Snijder EJ, Olsthoorn RC.** 2007. New structure model for the packaging signal in the genome of group IIa coronaviruses. J. Virol. **81:**6771–6774. http://dx.doi.org/10.1128/JVI.02231-06.
80. **Kuo L, Masters PS.** 2013. Functional analysis of the murine coronavirus genomic RNA packaging signal. J. Virol. **87:**5182–5192. http://dx.doi.org/10.1128/JVI.00100-13.
81. **Molenkamp R, Spaan WJ.** 1997. Identification of a specific interaction between the coronavirus mouse hepatitis virus A59 nucleocapsid protein and packaging signal. Virology **239:**78–86. http://dx.doi.org/10.1006/viro.1997.8867.
82. **Cologna R, Hogue BG.** 2000. Identification of a bovine coronavirus packaging signal. J. Virol. **74:**580–583. http://dx.doi.org/10.1128/JVI.74.1.580-583.2000.
83. **Morales L, Mateos-Gomez PA, Capiscol C, del Palacio L, Enjuanes L, Sola I.** 2013. Transmissible gastroenteritis coronavirus genome packaging signal is located at the 5′ end of the genome and promotes viral RNA incorporation into virions in a replication-independent process. J. Virol. **87:**11579–11590. http://dx.doi.org/10.1128/JVI.01836-13.
84. **Oostra M, Hagemeijer MC, van Gent M, Bekker CP, te Lintelo EG, Rottier PJ, de Haan CA.** 2008. Topology and membrane anchoring of the coronavirus replication complex: not all hydrophobic domains of nsp3 and nsp6 are membrane spanning. J. Virol. **82:**12392–12405. http://dx.doi.org/10.1128/JVI.01219-08.
85. **Kanjanahaluethai A, Chen Z, Jukneliene D, Baker SC.** 2007. Membrane topology of murine coronavirus replicase nonstructural protein 3. Virology **361:**391–401. http://dx.doi.org/10.1016/j.virol.2006.12.009.
86. **Chen J, Noueiry A, Ahlquist P.** 2003. An alternate pathway for recruiting template RNA to the brome mosaic virus RNA replication complex. J. Virol. **77:**2568–2577. http://dx.doi.org/10.1128/JVI.77.4.2568-2577.2003.
87. **Vlot AC, Laros SM, Bol JF.** 2003. Coordinate replication of alfalfa mosaic virus RNAs 1 and 2 involves *cis*- and *trans*-acting functions of the encoded helicase-like and polymerase-like domains. J. Virol. **77:**10790–10798. http://dx.doi.org/10.1128/JVI.77.20.10790-10798.2003.
88. **Weis BL, Schleiff E, Zerges W.** 2013. Protein targeting to subcellular organelles via MRNA localization. Biochim. Biophys. Acta **1833:**260–273. http://dx.doi.org/10.1016/j.bbamcr.2012.04.004.
89. **Wang Y, Shi H, Rigolet P, Wu N, Zhu L, Xi XG, Vabret A, Wang X, Wang T.** 2010. Nsp1 proteins of group I and SARS coronaviruses share structural and functional similarities. Infect. Genet. Evol. **10:**919–924. http://dx.doi.org/10.1016/j.meegid.2010.05.014.
90. **Jansson AM.** 2013. Structure of alphacoronavirus transmissible gastroenteritis virus nsp1 has implications for coronavirus nsp1 function and evolution. J. Virol. **87:**2949–2955. http://dx.doi.org/10.1128/JVI.03163-12.