

Subgenomic messenger RNA amplification in coronaviruses Author(s): Hung-Yi Wu, David A. Brian and Peter Palese Source: Proceedings of the National Academy of Sciences of the United States of America, Vol. 107, No. 27 (July 6, 2010), pp. 12257-12262 Published by: National Academy of Sciences Stable URL: <u>http://www.jstor.org/stable/20724237</u> Accessed: 05-02-2016 12:44 UTC

REFERENCES

Linked references are available on JSTOR for this article: http://www.jstor.org/stable/20724237?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <u>http://www.jstor.org/page/info/about/policies/terms.jsp</u>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



National Academy of Sciences is collaborating with JSTOR to digitize, preserve and extend access to Proceedings of the National Academy of Sciences of the United States of America.

Subgenomic messenger RNA amplification in coronaviruses

Hung-Yi Wu¹ and David A. Brian²

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

Edited by Peter Palese, Mount Sinai School of Medicine, New York, NY, and approved May 17, 2010 (received for review January 14, 2010)

Coronaviruses possess the largest known RNA genome, a 27- to 32-kb (+)-strand molecule that replicates in the cytoplasm. During virus replication, a 3' coterminal nested set of five to eight subgenomic (sg) mRNAs are made that are also 5' coterminal with the genome, because they carry the genomic leader as the result of discontinuous transcription at intergenic donor signals during (-)-strand synthesis when templates for sgmRNA synthesis are made. An unanswered question is whether the sgmRNAs, which appear rapidly and abundantly, undergo posttranscriptional amplification. Here, using RT-PCR and sequence analyses of head-totail-ligated (-) strands, we show that after transfection of an in vitro-generated marked sgmRNA into virus-infected cells, the sgmRNA, like the genome, can function as a template for (-)-strand synthesis. Furthermore, when the transfected sgmRNA contains an internally placed RNA-dependent RNA polymerase templateswitching donor signal, discontinuous transcription occurs at this site, and a shorter, 3' terminally nested leader-containing sgmRNA is made, as evidenced by its leader-body junction and by the expression of a GFP gene. Thus, in principle, the longer-nested sgmRNAs in a natural infection, all of which contain potential internal template-switching donor signals, can function to increase the number of the shorter 3'-nested sgmRNAs. One predicted advantage of this behavior for coronavirus survivability is an increased chance of maintaining genome fitness in the 3' one-third of the genome via a homologous recombination between the (now independently abundant) WT sgmRNA and a defective genome.

bovine coronavirus | discontinuous transcription | negative-strand RNA ligation | negative-strand RNA synthesis | severe acute respiratory syndrome

By virtue of an RNA-dependent RNA polymerase (RdRp) template switch, coronaviruses, which include the severe acute respiratory syndrome (SARS) virus, generate a 3'-coterminal nested set of subgenomic mRNAs (sgmRNAs) that also contain the leader of the genome (1). The leader is 65–90 nt long, depending on the species of coronavirus, and the template switch likely occurs during synthesis of (–)-strand templates for sgmRNA synthesis (2, 3) (Fig. 1, *Upper*), although some switching during (+)-strand synthesis (4) may occur as well (5). With few exceptions, sgmRNAs generated from 3'-proximal template- switching donor sites on the genome are progressively more abundant [up to 70-fold more at the peak time of RNA synthesis, at 6–8 h postinfection (6)] than those generated from 5'-proximal sites. The sgmRNAs are translated to virion structural proteins or to nonstructural accessory proteins, both of which may function as virulence factors or as inhibitors of host immune responses (7, 8).

The function of the common leader on sgmRNAs is not known, but we postulated that it [in its (-)-strand form, called the antileader] serves as a promoter for sgmRNA replication, providing a means for sgmRNA amplification (6, 9). The replication model was deemed feasible because the antileader is longer than the reported <20-nt replication promoters in togaviruses and orthomyxoviruses (10, 11). The sgmRNA replication hypothesis led to a search for (-)-strand copies of coronavirus sgmRNAs in porcine transmissible gastroenteritis virus (TGEV) and bovine coronavirus (BCoV) (6, 9). The presence of sgmRNA (-) strands and their properties (6, 9) seemed to be consistent

with a sgmRNA replication model, because (i) the rates of appearance of sgmRNA (+) and (-) strands were inversely related to their length (6, 9); (ii) sgmRNA-length dsRNA replicative intermediates [also called transcriptive intermediates (12)] were found that were (presumably reiteratively) active in (+)-strand sgmRNA synthesis in TGEV and mouse hepatitis coronavirus (MHV) (9, 12–14); and (iii) the 3'-terminal 55 nt and poly(A) tail together were found to be sufficient for (-)-strand synthesis from a defective interfering (DI) RNA in MHV (15).

Experiments that directly tested whether a reporter-containing sgmRNA underwent replication after transfection into helper virusinfected cells did not support a sgmRNA replication model for-BCoV (16) or MHV (17), however. For BCoV, no accumulation of the reporter-containing 1.7-kb sgmRNA (shown as sgmRNA⁷-1 in Fig. 2) was found when compared with that for a cotransfected helper virus-dependent replicon, a 2.2-kb DI RNA containing the same reporter (shown as DI RNA-1 in Fig. 2) (16).

In the present work, using BCoV as a model, we reexamined the potential for sgmRNA amplification by a mechanism other than replication and postulated that if sgmRNAs function as templates for (-)-strand synthesis as does the genome, then perhaps internal template-switching donor signals on the longer sgmRNAs during (-)-strand synthesis would lead to production of shorter internally encoded sgmRNAs (Fig. 1, Lower). Using RT-PCR and sequencing of head-to-tail-ligated (-)-strand molecules, we demonstrated that a marked synthetic sgmRNA can function as a template for (-)-strand RNA synthesis, and using RT-PCR sequencing of leader-body junctions and expression of a reporter GFP from a sgmRNA, we demonstrated that a template-switching signal within the sgmRNA can lead to the generation of a shorter internally nested sgmRNA. We found no evidence to indicate that the transfected sgmRNAs within the 24- to 72-h duration of the experiments, in the absence of experimentally applied selection pressures, had recombined with the helper virus genome in such a way as to generate the shorter reporter-containing sgmRNA from the genome. Our findings indicate that in principle, coronaviral sgmRNAs, once made from the genome, can be amplified by acting as templates for a cascading transcription process. We speculate about the biomedical implications of this behavior, including how it might aid in the survival of the large coronavirus genome.

Author contributions: H-Y.W. and D.A.B. designed research, H-Y.W. performed research, H-Y.W. and D.A.B. analyzed data, and H-Y.W. and D.A.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the Gen-Bank database (accession nos. U00735 for the Mebus strain of BCoV and NC_001846 for the A59 strain of MHV).

¹Present address: Graduate Institute of Veterinary Pathobiology, College of Veterinary Medicine, National Chung Hsing University, Taichung 40227, Taiwan.

²To whom correspondence should be addressed. E-mail: dbrian@utk.edu

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10. 1073/pnas.1000378107/-/DCSupplemental.



Fig. 1. Model for coronavirus subgenomic mRNA amplification. (Upper) Model for generating sgmRNA (-) strands (dashed gray line) from the viral genome (solid black line) during (-)-strand synthesis. In this model, the coronaviral RdRp pauses at intergenic template-switching donor signals (vertical black bars) and is transferred by a copy-choice mechanism to a highly similar acceptor site near the 5' end of the genome to copy the 5'-terminal leader. In this way, a complete (-)-strand template for sgmRN A synthesis is made. Synthesis of a full-length (-) strand (the antigenome, not depicted) would be used for genome replication. (Lower) The model for subgenomic mRNA amplification tested in this study. This model proposes that the RdRp sees the sgmRNA (solid black line) as a small genome, initiates (-)-strand synthesis on it, and switches template at an internal donor signal to make a (-)-strand template (dashed gray line) to synthesize a shorter internally nested sgmRNA. The leader is indicated by a filled box. An open box indicates the (-)-strand copy of the leader (called an antileader). Note that in this figure, sgmRNA 6 also will come from template switching on the viral genome (not depicted), as well as from sgmRNA 3 (depicted).



Fig. 2. Positive-strand RNA molecules used as templates for testing (–)-strand synthesis and sgmRNA production in helper virus-infected cells. At the top is a truncated diagram of the BCoV genome indicating its parts that are present in a naturally occurring BCoV defective interfering RNA (WT DI RNA). RNA molecules to be transfected were generated in vitro with T7 RNA polymerase from Mlu I-linearized plasmid DNAs. Reporters are T, a 30-nt in-frame sequence derived from the transmissible gastroenteritis coronavirus N gene; MHV 3'UTR, the 301-nt 3' UTR from mouse hepatitis virus-A59; EGFP, the 720-nt GFP gene in the -1 reading frame with respect to the upstream DI RNA and N ORFs; and gD, a 92-nt sequence containing the heptameric template-switching signal UCUAAAC. The 65-nt BCoV leader is identified by a solid black rectangle. The position of the 421 nt differentiating the DI RNA from sgmRNA 7 is shown.

Results

Transfected sgmRNAs Function as Templates for (-)-Strand Synthesis. Northern blot analyses have proven feasible for detecting BCoV sg(-)-strand RNAs during natural infection, where the numbers of (+)-strand sgmRNAs range from ~20 molecules/cell for the viral genome to ~5,000 molecules/cell for sgmRNA 7 [the nucleocapsid (N) protein-encoding, 3'-terminal, and most abundant of eight sgmRNAs] at peak abundance (6-8 h postinfection) and the (-)-strand counterparts range from 20- to 50-fold less (6). These have not proven feasible for detecting the low numbers of sgmRNA (-) strands generated from DI RNA parents transfected into helper virus-infected cells where sgmRNA production for a given template-switching signal is far less robust; for example, for the sgmRNA 7 template-switching signal, it is ~13 molecules per cell from DI RNA, versus ~2,000 from the viral genome at 24 h postinfection (18). Therefore, identification of sgmRNA (-) strands made from transfected (+)-strand templates must rely on a more sensitive method, such as RT-PCR analyses, which demonstrate an increase in numbers over time, followed by sequence confirmation of the RT-PCR product.

To determine whether a (+)-strand sgmRNA transfected into BCoV helper virus-infected cells can function as a template for sgmRNA (-)-strand synthesis by viral RdRp, we used RT-PCR with cDNA-cloned BCoV sgmRNA 7 marked with the MHV 3' UTR (sgmRNA⁷-2 in Fig. 2) to distinguish it from BCoV helper virus sgmRNAs. A 3'-proximal segment within the 301-nt MHV 3' UTR, nt number 46–156 from the 3' end [which differs by ~60% in sequence from the comparable region in the 288-nt BCoV 3' UTR (19, 20), and which is identified as a hypervariable region among



Fig. 3. Strategy for detecting viral RdRp-synthesized (-)-strand products from marked transfected DI RNA-2 and sgmRNA⁷-2. (A) The MHV-specific PCR primer (primer 2) amplifies sequences from the MHV 3' UTR, but not the BCoV 3' UTR. (Upper) Diagram of the MHV 3' UTR. The poly(A)-detecting oligo(dT₁₈) primer is used for reverse-transcription and PCR, and primer 2, which binds the MHV (--) strand complementary to nt 99-122 from the poly (A) tail in the (+) strand, is used for PCR. (Lower) A ~140-nt RT-PCR band (arrowhead) is produced with MHV RNA, but not with BCoV RNA. The conditions for the RT-PCR are the same as those described for Fig. 4, except that the number of cycles is 25 instead of 34. M, dsDNA molecular size markers; MHV, RNA from MHV-infected cells used for RT-PCR; BCoV, RNA from BCoV-infected cells used for RT-PCR. (B) Strategy for specifically detecting the head-to-tail-ligated RdRp-generated marked (-)-strand RNAs. The primers used do not generate an RT-PCR product from input copy-back (-)-strand RNA synthesized by the T7 RNA polymerase, but they do generate a product from head-to-tail-ligated (--)-strand RNA synthesized by the viral RdRp (Fig. 4). The vertical bar depicts the ligation site between the (-)-strand 5' monophosphorylated and 3' hydroxylated ends. The MHV 3' UTR-specific primer 2 shown in A is the same as the RT primer shown in B and also in Fig. 4. The (+)-strand–binding primers are depicted above the sequence, and the (-)-strand-binding primers are depicted below the sequence.

Wu and Brian

^{12258 |} www.pnas.org/cgi/doi/10.1073/pnas.1000378107

group 2 coronaviruses (19)], was exploited for specific RT-PCR differentiation between BCoV and MHV 3' UTRs (Fig. 3A). Despite this difference, BCoV DI RNA with the MHV 3' UTR (DI RNA-2 in Fig. 2) replicated well (Fig. S1), as anticipated, because the BCoV 3' UTR supports replication in the MHV genome (21); thus, we used it as a positive control for testing RdRp-dependent (-)-strand synthesis from transfected (+)-strand sgmRNA⁷-2 (described below).

Because T7 RNA polymerase generates transcripts in vitro that can be up to a 1% copy-back sequence (22) (Fig. 3B), direct RT-PCR sequencing of (-) strands obtained from cells cannot differentiate between (-) strands made by T7 RNA polymerase



Fig. 4. Transfected sgmRNA⁷-2 marked with the MHV 3' UTR functions as a template for (-)-strand synthesis. (A) Predicted results with transfected DI RNA-2 (control). (B) RT-PCR with primers described in Fig. 3B reveals a ~210nt product of ligated (-)-strands from BCoV-infected cells (Center, arrowhead), but not from uninfected cells (Left) or from an in vitro ligation reaction with RNA from infected cells mixed with 1 ng of DI RNA-2 (Right). (C) Sequence of the cDNA-cloned ligated junction from 12 hpt [lane 11 of B. shown in the (+) strand), in which the MHV 3' UTR (ending in ... ATCAC3') with the poly(A) tail is shown to be joined to the 5-base-truncated BCoV leader 5' end (5'TGAGC...) (vertical bar). (D) Predicted results with transfected sgmRNA7-2. (E) RT-PCR reveals a ~210-nt product from ligated (-)-strands from BCoV-infected cells (Center, arrowhead), but not from uninfected cells (Left) or from an in vitro ligation reaction with RNA from infected cells mixed with 1 ng of sgmRNA7-2 (Right). (F) Sequence of the cDNA-cloned ligated junction from 12 hpt (lane 11 of E, showing the same features as the sequence depicted in C). M, ds DNA size markers in nt pairs. (before transfection) and (-) strands made by viral RdRp (after transfection). Thus, we carried out head-to-tail ligation of RNA from cells and developed an RT-PCR primer set that uses the MHV 3' UTR (-)-strand-specific primer (Fig. 3A) and a BCoV leader (+)-strand-specific primer (Fig. 3B) to detect the headto-tail-ligated RdRp-generated sg(-) strands (Fig. 3B). These primers do not yield a product from copy-back RNA (Fig. 3B and control data described in Fig. 4). These experiments presumed that some fraction of the RdRp-generated sgmRNA(-)strands would be 5' monophosphorylated and thus ligatable with T4 RNA ligase 1 (23). Although both circular [the expected predominant form (24)] and tandemly ligated molecules are possible after ligation, the results from both would lead to the same conclusion because of the specificity of the RT primer for the MHV 3' UTR. A head-to-tail ligation method has been similarly applied to identify terminal features on influenza virus RNAs (24).

In BCoV-infected cells transfected with DI RNA-2 (the positive control) (Fig. 4A), a ~210-nt RT-PCR product from putative head-to-tail-ligated RdRp-generated (-)-strand RNA was found (Fig. 4B, lanes 9-14, arrowhead) that was not found in transfected mock-infected cells (Fig. 4B, lanes 2-7) or in control reactions with in vitro-ligated mixed components containing infected cell RNA and input DI RNA-2 (Fig. 4B, lanes 16-21). The sequence of the cDNA-cloned ~210-nt ligated product obtained at 12 h posttranscription (hpt) in BCoV-infected cells (Fig. 4B, lane 11) confirmed a head-to-tail molecule and demonstrated a 5' poly(U) tail of ~30 nt (Fig. 4C). Interestingly, the 3' terminus of the BCoV leader (-) strand was missing five 3'terminal bases, consistent with an incompletely understood sequence hypervariablity reported previously for the (+)-strand BCoV leader (25). In BCoV-infected cells transfected with sgmRNA⁷-2 RNA (Fig. 4D), a ~210-nt product from head-totail-ligated, putative RdRp-generated (-)-strand RNA was found (Fig. 4E, lanes 9-14, arrowhead) that was not found in transfected mock-infected cells (Fig. 4E, lanes 2-7) or in control reactions with in vitro-ligated mixed components containing infected cell RNA and input sgmRNA-2 (Fig. 4E, lanes 16-21). RT-PCR sequencing of the cDNA-cloned ~210-nt product obtained at 12 hpt (Fig. 4E, lane 11) confirmed a ligated head-totail (-) strand and also demonstrated a 5'-poly(U) tail of ~ 30 nt and a 3'-terminal 5-nt deletion (Fig. 4F). Strikingly, in contrast to the pattern seen for DI RNA, the sgmRNA (-) strands were expressed for a shorter period (6-18 hpt, with peak abundance at 12 hpt), suggesting that replication of the sgmRNA⁷-2 molecule might not occur or that if it does occur, it is short-lived.

These findings also suggest that the initiation of (-) strand synthesis is likely derived from a cleaved primer within the 5' poly(U) tail, because the (-)-strand RNAs were ligatable with T4 RNA ligase 1 and thus necessarily 5'-monophosphorylated (23) (Discussion).

Because a high rate of homologous RNA recombination between coronavirus molecules is well documented (25% over the entire genome) (26-28), and, along with experimentally applied selection methods, served as the basis for the first reverse-genetics system developed in coronaviruses (1, 29, 30), we conducted concurrent analyses of the helper virus genome to search for evidence of a recombinant genome that might have given rise to the observed marked ligated (-)-strand molecules. For this, the 6- to 24-h samples from DI RNA-2 and sgmRNA⁷-2 that demonstrated head-to-tail-ligated (-) strands (Fig. 4 *B* and *E*) were tested by RT-PCR for a 1,639-nt 3'-proximal helper virus genome-integrated reporter-containing sequence between the membrane (\overline{M}) protein gene and the 3' UTR. This was done under PCR conditions (with a 90-s extension time vs. 30 s as used in Fig. 4 B and E) that can readily identify a 1,637-nt fragment from a helper virus genome (Fig. S2B, Top and Bottom, lane 15). None was found (Fig. S2A, lanes 2-9), leading us to conclude that

Wu and Brian

PNA5 | July 6, 2010 | vol. 107 | no. 27 | 12259



Fig. 5. A short sgmRNA is produced from transfected sgmRNA⁷-3, which carries a transcription signal and EGFP gene. (A) Predicted results with DI RNA-3 (control). (B) RT-PCR reveals a 150-nt leader-body product from transfected BCoV-infected cells (lanes 17-24, arrowhead), but not from uninfected cells (lanes 2-11) or from infected cells alone (RT-PCR control a), from a mixture of RNA from infected cells at 13 h postinfection with 1 ng of DI RNA-3 (RT-PCR control b), or from an RT-PCR mix with primers alone (RT-PCR control c). The 150-nt RT-PCR product is not observed in cells infected with progeny helper virus passages 1 and 2 (lanes 25 and 26) or in blind passages from uninfected cells (lanes 12 and 13). A 1,757-nt RT-PCR product from input DI RNA-3 is identified with an asterisk. (C) Sequence of the cDNAcloned 150-nt RT-PCR product from lane 21 in B showing the leader-body junction (vertical bar), the heptameric template-switching signal (under lined), and the AUG translation start codon (overlined) for EGFP. (D) Predicted results with sgmRNA7-3. (E) RT-PCR reveals a 150-nt leader-body product from transfected BCoV-infected cells (lanes 17-24, arrowhead), but not from uninfected cells (lanes 2-11), from infected cells alone (RT-PCR control a), from a mixture of RNA from infected cells at 13 h postinfection with 1 ng of sgmRNA⁷-3 (RT-PCR control b), or from an RT-PCR mix with primers alone (RT-PCR control c). The 150-nt RT-PCR product likewise is not observed in cells infected with progeny helper virus passages 1 and 2 (lanes

the sgmRNA⁷⁻² (-) strands detected in the experiment did not come from a recombinant helper virus genome.

An Intergenic Transcription-Associated (Template-Switching) Signal on the Transfected sgmRNA (+) Strand Leads to Generation of a Smaller Internally Nested Reporter-Containing sgmRNA. To test the hypothesis that sgmRNAs containing an internal template-switching signal can generate a shorter sgmRNA, we compared sgmRNA⁷ 3 with DI RNA-3 (the positive control) (Figs. 2 and 5A). For sister constructs of DI RNA-3, sgmRNA generation from internal donor template-switching signals is well documented (18, 31, 32). From BCoV-infected cells transfected with DI RNA-3, a 150-nt nested RT-PCR product was obtained that first appeared at 1 hpt and was present throughout the 72-h experiment, but was most abundant at 6-48 hpt (Fig. 5B, lanes 17-24, arrowhead). A similar product was not detectable in mock-infected cells (Fig. 5B, lanes 2-11) or in in vitro mixtures of reaction components (Fig. 5B, lanes 27-29), indicating that the 150-nt product arose from an in vivo RdRp-generated internally nested short sgmRNA. The sequence of the cDNA-cloned 150-nt molecule obtained at 12 hpt (Fig. 5B, lane 21) confirmed the expected sgmRNA leader-body junction (Fig. 5C). A 1,757-nt nested RT-PCR product from primers binding within the transfected parent DI RNA was obtained with RNA from both infected and uninfected cells (Fig. 5B, asterisk). From BCoV-infected cells transfected with sgmRNA'-3 RNA (Figs. 2 and 5D), a 150-nt RT-PCR product from the leaderbody junction of the putative short reporter-containing sgmRNA was found that first appeared at 1 h and was present throughout the 72-h experiment, but was most abundant at 12 hpt (Fig. 5E, lanes 17-24, arrowhead). The 150-nt product was not found in RNA from transfected uninfected cells (Fig. 5E, lanes 2-11) or in in vitro mixtures of reaction components (Fig. 5E, lanes 27-29). A 1,336-nt nested RT-PCR product from primers binding within the transfected parent sgmRNA was obtained with RNA from both infected and uninfected cells (Fig. 5E, asterisk). The sequence of the cDNA-cloned 150-nt product at 12 hpt (Fig. 5E, lane 21) confirmed its origin at the template-switching signal in sgmRNA⁷-3 (Fig. 5F). Interestingly, the relatively short-lived peak abundance of the 150-nt leader-body product from sgmRNA'-3 (maximal at 12 hpt) compared with DI RNA-3 (maximal at 6-48 hpt) resembled that for the (-)-strand RNA synthesized from sgmRNA[']-2 (Fig. 4E, lanes 10-12), suggesting that the transfected sgmRNA⁷-3 molecule might not be replicated, or if it is replicated, it is short-lived.

In a second experimental approach, fluorescing cells were sought as evidence of EGFP expression from the internally encoded small sgmRNA. Because the EGFP ORF in both the DI RNA-3 and sgmRNA⁷-3 templates is in the -1 reading frame with respect to the upstream ORFs, fluorescence of EGFP as a fusion protein was not expected, and was not observed, in transfected uninfected cells (Fig. 5G, panels 1 and 3). It was found in ~0.125% of cells for DI RNA-3 and in ~0.00015% of cells for sgmRNA⁷-3 in the presence of helper virus (Fig. 5G, panels 2 and 4). The fluorescence in ~1,000-fold fewer cells for sgmRNA⁷-3 likely reflects the transient transcription observed from this construct.

Two experimental approaches indicated that the small reportercontaining sgmRNA did not come from a recombinant helper vi-

^{12260 |} www.pnas.org/cgi/doi/10.1073/pnas.1000378107

²⁵ and 26) or in blind passages from uninfected cells (lanes 12 and 13). A 1,336-nt RT-PCR product from input sgmRNA⁷-3 is identified with an asterisk. (*F*) Sequence of the cDNA-cloned 150-nt RT-PCR product from lane 21 in *E* showing the same features as described in C. (G) Fluorescence evidence of EGFP expression from the sgmRNA⁷-3-encoded short sgmRNA. Panels 1 and 3 are mock-infected cells transfected with DI RNA-3 and sgmRNA⁷-3, respectively; panels 2 and 4 are BCoV-infected cells transfected with DI RNA-3 and sgmRNA⁷-3, respectively. In all cases, cells were examined for fluorescence at 16 h posttransfection.

rus genome. In the first approach, RT-PCR tests on RNA extracted at 12 hpt from cells infected with progeny helper virus from transfected cells (identified as virus passages 1 and 2) demonstrated no evidence of the small reporter-containing sgmRNA (Fig. 5 *B* and *E*, lanes 25 and 26). In the second approach, RT-PCR tests on RNA from the transfected cells described in Figs. 5 *B* and *E*, lanes 15–24, in which an upstream M gene–specific primer replaced the leader-specific primer 2 in three separate PCR primer combinations, demonstrated no evidence of a recombinant helper virus genome (Fig. S2*B*). Together, these results indicate that a longer sgmRNA with an internal RdRp template-switching donor signal can serve as a template for generating shorter internally encoded sgmRNAs.

Discussion

We have demonstrated two previously undescribed features of coronavirus sgmRNA synthesis: (i) sgmRNA (+) strands can function as templates for sgmRNA (-)-strand RNA synthesis, and (ii) template-switching signals on sgmRNA (+) strands can trigger synthesis of internally nested (+)-strand sgmRNAs. The first finding is consistent with two previous reports describing putative requirements for initiation of (-)-strand synthesis in coronaviruses. Lin et al. (15) demonstrated by RNase protection assays that the 3'-terminal 55 nt along with the poly(A) tail in coronavirus DI RNA (and by implication, sgmRNA) are sufficient for (-)-strand synthesis in vivo. Züst et al. (33) demonstrated, through genetic and phylogenetic evidence, that an intramolecular pairing between nine bases at the genome 3' terminus and bases within loop 1 of the 3' UTR pseudoknot mapping 220 nt upstream, even without the nt 30-170 sequence from the 3' end, is required for MHV genome replication. The model of Züst et al. (33) proposes that the nine-base interaction is necessary for assembly of the replication complex and initiation of (-)-strand synthesis. These authors further suggested that this required base pairing might have occurred in trans in the earlier study of Lin et al. (15). Because the polyadenylated 3' UTR on sgmRNAs is identical to that on the genome, the 3'-proximal structural requirements for (-)-strand synthesis as defined previously (15, 33) would be expected to be met. The present study extends the observations on initiation of (-)-strand synthesis by showing that poly(U) is part of the nascent (-) strand. The identity of the primer used by the primer-dependent viral RdRp (nsp 12) remains unresolved by our experiments (34, 35). We presume that the 5' monophosphorylated U enabling head-to-tail ligation by T4 RNA ligase 1 shown in Fig. 4 is the product of an as-yet-unidentified endonuclease, because initiation of coronavirus (-)-strand synthesis likely uses a 5' triphosphorylated single nucleotide primer [as would be the case for de novo initiation (36)] or an oligonucleotide primer [as would be made by the viral primase (nsp 8) (37)]. The coronaviral endonuclease (EndoU) (nsp 15) is an unlikely candidate for this cleavage, because it leaves a 2'-3' cyclic phosphate (38, 39) that is not ligatable with T4 RNA ligase 1 (23). The importance of the cleavage in the nascent (-) strand as observed here, if any, remains to be determined.

We did not extend our analyses to explore how the genomespecific 5'-proximal (+)-strand 421 nt differentiating DI RNA from sgmRNA 7 (Fig. 2) (16) influences the initiation of (–)-strand synthesis. However, it clearly is not required in *cis* for initiation of sgmRNA (–)-strand synthesis, as has been postulated (3), although it might function to enhance initiation. Certainly the 5'-proximal 421-nt sequence contributes to long-term accumulation of DI RNA compared with sgmRNA⁷ (Figs. 4 and 5) (16), suggesting that some feature within it might be required for the initiation of new (+)-strand synthesis. The (–)-strand complement of the 421 nt, for example, might function to anchor the (–)-strand template within the membranous replication/transcription complex for reiterated (+)-strand synthesis, as has been suggested for the (–) strand of tomato bushy stunt virus (40, 41). In the absence of an anchor, sgmRNA (-) strands might be ushered from the replication/transcription compartment and thus provide only a short-term template function. To the best of our knowledge, the only report of posttranscriptional sgmRNA replication is that for flock house virus (42), whose features might prove instructive for further analysis of coronavirus sgmRNA.

Our second finding indicates that the larger sgmRNAs transcribed from the genome via the template-switching process could themselves serve as donor templates for discontinuous transcription and the generation of internally nested sgmRNAs, as depicted in Fig. 1, *Lower*. This mechanism likely would contribute to the progressively greater abundance of the 3'-proximal sgmRNAs. It also suggests that RdRp template-switching rates at signaling sites on the sgmRNA would be influenced by the flanking and other more distant enhancer elements known to affect template switching on the genome (31, 43–47).

We envision the biomedical implications of sgmRNA amplification in coronaviruses to include the following:

- A cascading mechanism for sgmRNA production would relieve pressure on the large genome as the sole template for sgmRNA synthesis and thus might facilitate genomelength (-)-strand synthesis.
- (ii) A rapid appearance of large numbers of sgmRNAs encoding virulence factors, whether the resulting proteins act directly to cause lesions (8) or indirectly to inhibit host immune responses (7), could significantly enhance early disease development.
- (iii) Although copy-choice recombination between the transfected marked sgmRNAs and helper virus genome was not observed in the experiments of short duration performed here, much experimental evidence documenting coronavirus recombination has unequivocally established that it would occur over time and under selective pressures (27-30, 48). Furthermore, recombination between coronavirus genome and sgmRNAs occurring during natural infections has been offered as an explanation for the skewing of higher recombination rates toward the 3' end of the MHV genome (28, 48). Thus, we postulate that a pool of sgmRNAs having been amplified posttranscriptionally from a fit genome could recombine with a posttranscriptionally (mutated) debilitated genome to restore fitness. The demonstration that coreplicating subgenome-length defective picornavirus RNA molecules can give rise to a full-length fit standard virus genome through recombination (49) lends support to this idea. Thus, a sgmRNA-assisted, recombination-based fitness restoration mechanism as postulated here could, along with a proofreading function for the coronavirus-encoded $3' \rightarrow 5'$ exonuclease (NendoU; nsp 14) (50), contribute significantly to the survival of the largest known viral RNA genome.

Materials and Methods

Plasmid Construction, in Vitro T7 RNA Polymerase Synthesis of Marked DI RNAs and sgmRNAs, Transfection, and Virus Propagation. Construction of pDI RNA-1 and psgmRNA7-1 (formerly called pDrep1 and pNrep2, respectively) (Fig. 2) have been described previously (16). pDI RNA-2 and psgmRNA⁷-2 (Fig. 2) were made by replacing the 288-nt 3' UTR of BCoV-Mebus (GenBank accession no. U00735) in pDI RNA-1 and psgmRNA⁷-1 with the 301-nt 3' UTR of MHV-A59 (GenBank accession no. NC_001846) from fragment G DNA (51). pDI RNA-3 and psgmRNA7-3 (Fig. 2) are identical to pDI RNA-1 and psgmRNA⁷-1, respectively, except that inserted at the Nsil site is an 836-nt sequence containing, in order, an 18-nt intergenic sequence harboring the underlined heptameric template-switching signal (5'AATATCTAAACTT-TAAGG3'), the 720-nt EGFP gene including a stop codon [from pIRES1-EGFP DNA (Clontech)], a 6-nt KpnI recognition sequence, and a 92-nt HSV 1-gD epitope-encoding sequence (31). In both pDI RNA-3 and psgmRNA⁷-3, the EGFP gene is in the -1 reading frame with respect to the upstream DI RNA and N gene coding regions. For each, uncapped T7 RNA polymerase transcripts were made from plasmid linearized with Mlul (Fig. 2), treated with RNase-free DNase (Promega), and chromatographed through a Biospin 6 column (Bio-Rad) before use in transfection (18).

BCoV was propagated on human rectal tumor (HRT)-18 cells, and MHV was propagated on delayed brain tumor (DBT) cells (18, 20). For transfection, HRT in 35-mm dishes at ~80% confluency (~8 × 10⁵ cells/dish) was infected with BCoV at a multiplicity of infection of 5 PFU per cell and transfected 1 h later with 600 ng of transcript RNA using Lipofectin (Invitrogen) (16).

Head-to-Tail Ligation of Viral (–)-Strand RNA Products and Sequence Analysis of the Junction. RNA was extracted with TRIzol (Invitrogen) (18), and one fourth of the volume from one plate (~10 μ g total per plate) in 25 μ L of water was heat-denatured at 95 °C for 5 min and then quick-cooled. Then 3 μ L of 10× ligase buffer and 2 U (in 2 μ L) of T4 RNA ligase 1 (New England Biolabs) were added, and the mix was incubated for 16 h at 16 °C. Phenol-chloroform–extracted ligated RNA in 20 μ L of reaction buffer was used for the RT reaction with SuperScript II reverse-transcriptase (Invitrogen), primer MHV3'UTR99-122(–) (Table S1), and RT primer in Fig. 4, which binds 99–122 nt from the polyU tail in the (–)-strand of the MHV-A59 3' UTR. Of this, 5 μ L was used in a 50- μ L PCR with AccuPrime Taq DNA polymerase (Invitrogen), primer BCVL29-54(+) (Table S1), primer 2 in Fig. 4, which binds t29-54 in the (+)-stand leader of BCoV, and primer MHV3'UTR99-122(–). The resulting mixture was heated to 94 °C for 2 min, then subjected to 34 cycles

- 1. Masters PS (2006) The molecular biology of coronaviruses. Adv Virus Res 66:193–292. 2. Zúñiga 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. 3. Sawicki SG, Sawicki DL, Siddell SG (2007) A contemporary view of coronavirus
- transcription. J Virol 81:20–29. 4. Lai MM, Patton CD, Baric RS, Stohlman SA (1983) Presence of leader sequences in the
- mRNA of mouse hepatitis virus. J Virol 46:1027–1033.
- 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.
- Hofmann MA, Sethna PB, Brian DA (1990) Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. J Virol 64:4108–4114.
- Kopecky-Bromberg SA, Martínez-Sobrido L, Frieman M, Baric RA, Palese P (2007) Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. J Virol 81:548–557.
- Cowley TJ, Long SY, Weiss SR (2010) The murine coronavirus nucleocapsid gene is a determinant of virulence. J Virol 84:1752–1763.
- 9. Sethna PB, Hung SL, Brian DA (1989) Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc Natl Acad Sci USA* 86:5626–5630.
- 10. Li X, Palese P (1992) Mutational analysis of the promoter required for influenza virus virion RNA synthesis. J Virol 66:4331–4338.
- Levis R, Weiss BG, Tsiang M, Huang H, Schlesinger S (1986) Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. Cell 44:137–145.
- 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.
- Baric RS, Yount B (2000) Subgenomic negative-strand RNA function during mouse hepatitis virus infection. J Virol 74:4039–4046.
- 14. Sawicki SG, Sawicki DL (1990) Coronavirus transcription: Subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J Virol* 64:1050–1056.
- Lin YJ, Liao CL, Lai MM (1994) Identification of the cis-acting signal for minus-strand RNA synthesis of a murine coronavirus: Implications for the role of minus-strand RNA in RNA replication and transcription. J Virol 68:8131–8140.
- 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.
- Makino S, Joo M, Makino JK (1991) A system for study of coronavirus mRNA synthesis: A regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. J Virol 65:6031–6041.
- 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.
- Goebel SJ, Miller TB, Bennett CJ, Bernard KA, Masters PS (2007) A hypervariable region within the 3' cis-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis. J Virol 81:1274–1287.
- Wu H-Y, et al. (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.
- Peng D, Koetzner CA, McMahon T, Zhu Y, Masters PS (1995) Construction of murine coronavirus mutants containing interspecies chimeric nucleocapsid proteins. J Virol 69: 5475–5484.
- Schenborn ET, Mierendorf RC, Jr (1985) A novel transcription property of SP6 and T7 RNA polymerases: Dependence on template structure. *Nucleic Acids Res* 13:6223–6236.
- Nichols NM, Tabor S, McReynolds LA (2008) RNA ligases. Curr Protoc Mol Biol unit 3:15.
 Szymkowiak C, et al. (2003) Rapid method for the characterization of 3' and 5' UTRs
- of influenza viruses. J Virol Methods 107:15–20. 25. Hofmann MA, Senanayake SD, Brian DA (1993) A translation-attenuating intraleader
- open reading frame is selected on coronavirus mRNAs during persistent infection. Proc Natl Acad Sci USA 90:11733–11737.
- 26. Lai MM (1992) Genetic recombination in RNA viruses. *Curr Top Microbiol Immunol* 176:21–32.

12262 | www.pnas.org/cgi/doi/10.1073/pnas.1000378107

of 30 s at 94 °C, 30 s at 55 °C, and 30 s at 72 °C. The ~210-bp PCR product resolved by a nondenaturing agarose (2.5%) gel electrophoresis was sequenced either directly or after cloning into TOPO XL PCR (Invitrogen).

Leader-Junction Sequence Analyses of (+)-Strand sgmRNAs Derived from Transfected DI RNA and sgmRNA. RNA was extracted with TRIzol, and RT-PCR reactions were carried out with oligonucleotide 5'GD(+) (Table S1) [the RT primer in Fig. 5 which binds within the (+)-strand of the partial HSV 1 gD gene (31)] for RT, oligonucleotide leader20(-) (Table S1) [primer 2 in Fig. 5 which binds the 3'-terminal 20 nt within the (-)-strand of the BCoV leader], and oligonucleotide EGFP2(+) (Table S1) [primer 3 in Fig. 5 which binds the (+)-strand of the EGFP gene] for PCR. RT-PCR was carried out as described for head-to-tail ligation assays, except that 5 μ L of the RT reaction mixture was used in a 50- μ L PCR mix that was heated to 94 °C for 2 min and then subjected to 29 cycles of 30 s at 94 °C, 30 s at 55 °C, and 90 s at 72 °C. The agarose gel-purified ~150-bp PCR products were sequenced either directly or after cloning into TOPO XL PCR (Invitrogen).

ACKNOWLEDGMENTS. This work was funded by US Public Health Service Grant Al14367 and the University of Tennessee, College of Veterinary Medicine Center of Excellence in Livestock Diseases and Human Health.

- Baric RS, Fu K, Schaad MC, Stohlman SA (1990) Establishing a genetic recombination map for murine coronavirus strain A59 complementation groups. *Virology* 177:646–656.
 Fu K, Baric RS (1994) Map locations of mouse hepatitis virus temperature-sensitive
- mutants: Confirmation of variable rates of recombination. J Virol 88:7458–7466.
- Koetzner CA, Parker MM, Ricard CS, Sturman LS, Masters PS (1992) Repair and mutagenesis of the genome of a deletion mutant of the coronavirus mouse hepatitis virus by targeted RNA recombination. J Virol 66:1841–1848.
- 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.
- Ozdarendeli A, et al. (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.
- Krishnan R, Chang RY, Brian DA (1996) Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. *Virology* 218:400–405.
- 33. 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.
- Cheng A, et al. (2005) Expression, purification, and characterization of SARS coronavirus RNA polymerase. Virology 335:165–176.
- te Velthuis AJ, Arnold JJ, Cameron CE, van den Worm SH, Snijder EJ (2010) The RNA polymerase activity of SARS-coronavirus nsp12 is primer-dependent. Nucleic Acids Res 38:203–214.
- Kao CC, Singh P, Ecker DJ (2001) De novo initiation of viral RNA-dependent RNA synthesis. Virology 287:251–260.
- Imbert I, et al. (2006) A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. EMBO J 25:4933–4942.
- Ivanov KA, et al. (2004) Major genetic marker of nidoviruses encodes a replicative endoribonuclease. Proc Natl Acad Sci USA 101:12694–12699.
- Bhardwaj K, Sun J, Holzenburg A, Guarino LA, Kao CC (2006) RNA recognition and cleavage by the SARS coronavirus endoribonuclease. J Mol Biol 361:243–256.
- Wang RY, Nagy PD (2008) Tomato bushy stunt virus co-opts the RNA-binding function of a host metabolic enzyme for viral genomic RNA synthesis. *Cell Host Microbe* 3:178–187.
- 41. Wang X, Ahlquist P (2008) Filling a GAP(DH) in asymmetric viral RNA synthesis. Cell Host Microbe 3:124–125.
- Eckerle LD, Albariño CG, Ball LA (2003) Flock House virus subgenomic RNA3 is replicated and its replication correlates with transactivation of RNA2. *Virology* 317:95–108.
- Fischer F, Stegen CF, Koetzner CA, Masters PS (1997) Analysis of a recombinant mouse hepatitis virus expressing a foreign gene reveals a novel aspect of coronavirus transcription. J Virol 71:5148–5160.
- Moreno JL, Zúňiga S, Enjuanes L, Sola I (2008) Identification of a coronavirus transcription enhancer. J Virol 82:3882–3893.
- Sola I, Moreno JL, Zúñiga S, Alonso S, Enjuanes L (2005) Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. J Virol 79:2506–2516.
- Hsue B, Masters PS (1999) Insertion of a new transcriptional unit into the genome of mouse hepatitis virus. J Virol 73:6128–6135.
- 47. An S, Makino S (1998) Characterizations of coronavirus *cis*-acting RNA elements and the transcription step affecting its transcription efficiency. *Virology* 243:198–207.
- Fu K, Baric RS (1992) Evidence for variable rates of recombination in the MHV genome. Virology 189:88–102.
- García-Arriaza J, Ojosnegros S, Dávila M, Domingo E, Escarmís C (2006) Dynamics of mutation and recombination in a replicating population of complementing, defective viral genomes. J Mol Biol 360:558–572.
- Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR (2007) High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J Virol* 81:12135–12144.
 Yount B, Denison MR, Weiss SR, Baric RS (2002) Systematic assembly of a full-length
- infectious cDNA of mouse hepatitis virus strain A59. J Virol 76:11065–11078.