CORONAVIRUSES USE DISCONTINUOUS EXTENSION FOR SYNTHESIS OF SUBGENOME-LENGTH NEGATIVE STRANDS

Stanley G. Sawicki and Dorothea L. Sawicki

Department of Microbiology
Medical College of Ohio
P.O. Box 10008
Toledo, Ohio 43699

ABSTRACT

We have developed a new model for coronavirus transcription, which we call discontinuous extension, to explain how subgenome-length negatives stands are derived directly from the genome. The current model called leader-primed transcription, which states that subgenomic mRNA is transcribed directly from genome-length negative-strands, cannot explain many of the recent experimental findings. For instance, subgenomic mRNAs are transcribed directly via transcription intermediates that contain subgenome-length negative-strand templates; however subgenomic mRNA does not appear to be copied directly into negatives strands. In our model the subgenome-length negative strands would be derived using the genome as a template. After the polymerase had copied the 3'-end of the genome, it would detach at any one of the several intergenic sequences and reattach to the sequence immediately downstream of the leader sequence at the 5'-end of genome RNA. Base pairing between the 3'-end of the nascent subgenome-length negative strands, which would be complementary to the intergenic sequence at the end of the leader sequence at the 5'-end of genome, would serve to align the nascent negative strand to the genome and permit the completion of synthesis, i.e., discontinuous extension of the 3'-end of the negative strand. Thus, subgenome-length negative-strands would arise by discontinuous synthesis, but of negative strands, not of positive strands as proposed originally by the leader-primed transcription model.

Coronaviruses contain an unusually long (27-32,000 ribonucleotides) positive-sense RNA genome (1, 2) that is polyadenylated at the 3' end and capped at the 5' end. During the coronavirus replication cycle, genome and subgenomic mRNAs are produced that together comprise a 3' co-terminal nested set. At their 5' ends they also all possess an identical sequence, called the leader RNA, of 60-80 nucleotides. The original proposal (3) that transcription of subgenomic mRNA was discontinuous was based on the published report (4) that only genome-length negative-strand templates existed in infected cells. Splicing of

the genome had been ruled out because the rate of inactivation by ultraviolet light (UV) of the synthesis of genomic and subgenomic mRNA was proportional to the length of the RNA (5, 6); therefore, it appeared that the leader RNA was joined to the body of subgenomic mRNA during positive strand synthesis. The leader-primed transcription model was proposed (7) because only genome-length negative strands were reported to be found in replication intermediates containing nascent subgenomic mRNA that already had acquired their leader RNA. Taken together the experimental evidence suggested RNA synthesis initiated at the 3' end of the negative strand copy of the genome and then proceeded either continuously to produce genomic RNA or discontinuously to produce subgenomic mRNA, in which case the leader RNA together with the viral transcriptase detached or jumped after copying the leader RNA and started again further downstream at specific intergenic sequence (IS) elements. The evidence for leader-primed transcription is indirect (reviewed in 1, 2) and based mostly on the fact that the IS elements possess a consensus sequence, which is found also at the end of the leader RNA, and that the IS elements determine subgenomic mRNA synthesis (8, 9, 10, 11).

Several years after the leader-primed transcription model had been proposed and generally accepted as correct, David Brian's laboratory (12) demonstrated that transmissible gastroenteritis virus (TGEV) infected cells contained subgenome-length negative strands. Shortly afterward we demonstrated (13) that replication intermediates (RIs) containing subgenome-length negative strands were transcriptionally active in mouse hepatitis virus (MHV) infected cells. Finally, Brian's laboratory demonstrated (14) that the subgenome-length negative strands contained at their 3'-end a sequence that was complementary to the leader RNA. Taken together, these observations showed unequivocally that subgenomic-length templates were utilized for the production of subgenomic mRNA. Figure 1 illustrates one of the approaches that we are using currently to study the synthesis of genome-length and subgenome-length negative strands in MHV infected cells.

Oligodeoxyribonucleotides were used as primers to first copy with reverse transcriptase a defined region of either positive or negative strand MHV RNA into cDNA. The cDNA was then amplified with PCR. To detect negative strands, we used during reverse transcription a primer that was identical to a sequence very near the 5' end of the leader RNA and that would anneal to genome-length negative strands as well as subgenome-length negative strands if they both possessed at their 3' ends a complementary copy of the leader RNA. To detect positive strands, we used two primers for reverse transcription: one that was complementary to a unique sequence 550 nucleotides downstream of the 5' end of the genome RNA and a second primer that was complementary to a sequence 370 nucleotides downstream of the 5' end of mRNA-7. Because the positive strands form a 3' nested set, the second primer would recognize also sequences that are present not only in RNA-7 but also in the other subgenomic mRNAs as well as in the genomic RNA. However, because we used only 5 min of incubation for the reverse transcriptase reaction, mostly RNA-7 was copied into cDNA. During the PCR amplification all three primers were present. Figure 1 shows the results when we used RNA isolated from cells seven hours after infection with MHV. The primer pairs for the genome and the negative strand copy of the genome gave a PCR product of the expected length of 550 nucleotides; and the primer pairs for RNA-7 and the negative-strand copy of RNA-7 gave a PCR product of the expected length of 370 nucleotides. Therefore, MHV infected cells possess both subgenome-length and genome-length negative strands that have a complementary copy of the leader RNA at their 3' ends. If we mix the primer pairs appropriately, we can detect in infected cells the genome and subgenomic mRNAs and their negative strand copies. Figure 2 shows that uninfected cells do not produce a PCR product using these primers.

Cells infected with MHV in the cold (4°C) and harvested after removing unadsorbed virus contained only genomic RNA and not the negative strand copy of the genome.
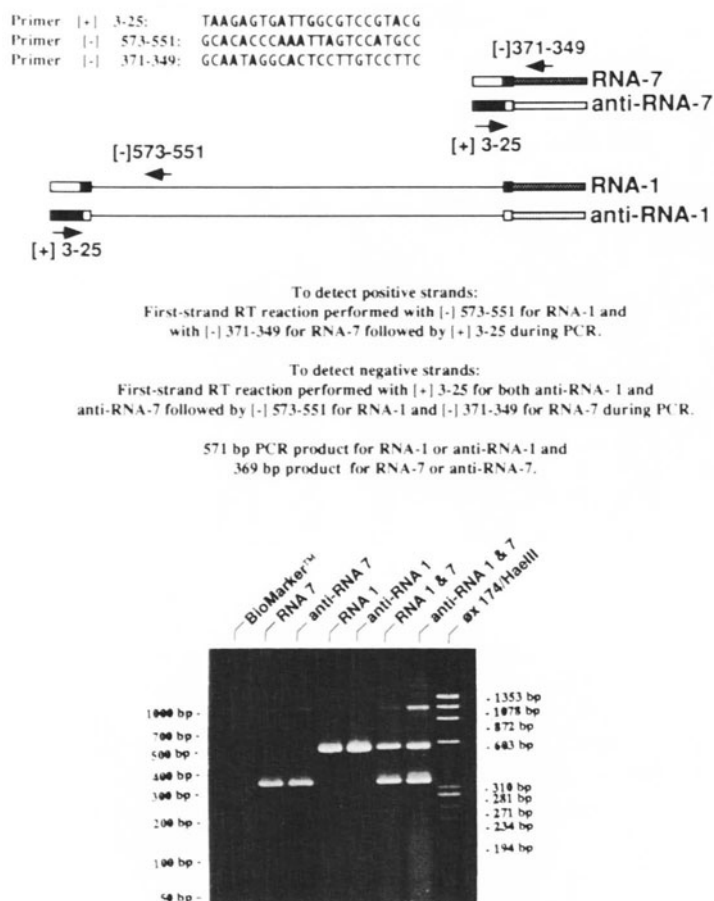
Figure 1. Detection by PCR of genomes and subgenomic mRNA and their negative strand templates. Super-script-II reverse transcriptase (Life Technologies, Bethesda, MD) for first strand synthesis and Taq DNA polymerase (Boehringer-Mannheim, Indianapolis, IN) for PCR were used according to the manufacturer's directions. After denaturation at 95°C for 5 min and annealing with the first strand primer, reverse transcription reactions were carried out for 5 min at 42°C. After heating to 99°C for 2 min and cooling to 4°C, the second strand primer(s) were added with the Taq polymerase and DNA amplification was performed for 35 cycles: 1 min at 65°C for annealing, 5 min at 72°C for extension and 1 min at 95°C for denaturation.

Interestingly, they did not have either RNA-7 or the negative strand copy of RNA-7, although RNA-7, but not its negative strand copy, was present in preparations of purified virions. The virion RNA was obtained from virions that had been purified twice by isopycnic tartrate-gradient centrifugation. Although subgenomic mRNAs were present in particles that have the same density as virions, they were not adsorbed and, therefore, were not in the same particles as the genome. By seven hours post-infection, the infected cell possessed all of the aforementioned RNA molecules. These experiments confirm those from Brian's laboratory (14) and extend their findings to MHV: MHV infected cells contain negative strand copies of subgenomic mRNA including a sequence at their 3′ end that is complementary to the leader RNA. Furthermore, these experiments demonstrate directly that subgenomic mRNA and subgenome-length negative strands arise from genomic RNA because that was the only RNA that was in the virions that adsorbed to the cells.
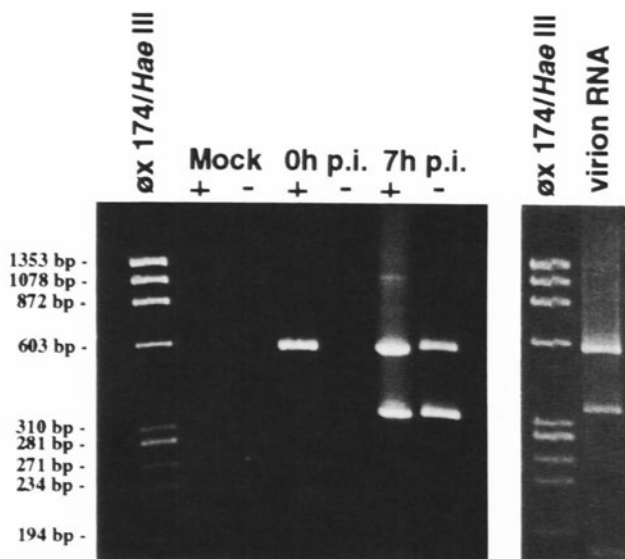
Figure 2. Only virions containing genome RNA adsorb to 17Cl-1 cells and cause the synthesis of genomes and subgenomic mRNA and their negative strand templates. The conditions for reverse transcription and DNA amplification by PCR were as described in Figure 1. Virion RNA was obtained from virus that had been purified twice by isopycnic tartrate-gradient centrifugation (15).

Previously, we published (13) experiments that demonstrated that ^{3}H-uridine accumulated rapidly and at the same rate in the smallest RI, which would produce mRNA-7, as in the largest RI, which would produce genomes. Figure 3 shows the results of this type of experiment.

These experiments demonstrated conclusively that RIs with subgenome-length templates were transcriptionally active in mRNA synthesis. These RIs were synthesizing positive strand RNAs because negative strand synthesis had virtually ceased at the time of labeling
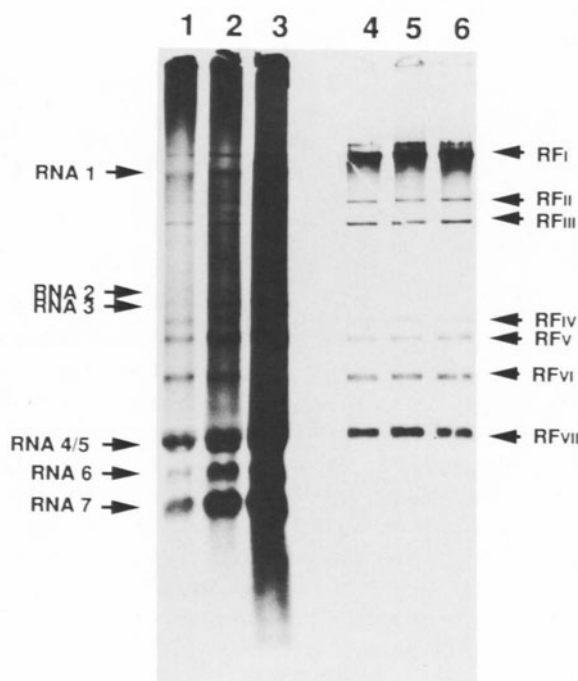

Figure 3. Transcriptionally active MHV subgenome-length RIs. 17Cl-1 cells (3 x 10⁶) in 60 mm petri dishes were infected with MHV at 20 pfu/cell and labeled with 1 mCi of ^{3}H-uridine/ml for 2 min [lanes 1 and 4], with 0.25 mCi/ml for 5 min [lanes 2 and 5], or with 0.1 mCi/ml for 20 min [lanes 3 and 6]. The cells were solubilized, extracted with phenol and chloroform, and ethanol precipitated. Extracts from 5 x 10⁵ cells were treated with DNase and either were not treated [lanes 1-3] or treated [lanes 4-6] with 1μg of RNase A/ml before electrophoresis on an 0.8% agarose gel in TBE buffer. [Reproduced from Sawicki and Sawicki (13) by courtesy of ASM.]

(13, 15). We also noted that the relative amount of ^{3}H-uridine incorporated into each RI did not vary with the length of the labeling period and that the relative amount of label in each species of RI RNA was proportional to the synthesis of genome and subgenomic mRNAs. Therefore, it appeared from the data that all the subgenomic mRNAs were produced from RIs containing subgenome-length templates; and the RIs containing genome-length templates were producing exclusively genomes. This conclusion was consistent with the UV inactivation studies (5, 6) and the observations from Brian's laboratory (12, 14, 16) and our own laboratory (15) that negative strands accumulated during the early phase of the replication cycle when viral RNA synthesis was increasing at an exponential rate and stopped accumulating when the rate of viral RNA synthesis became constant. Thus, contrary to the recent suggestion (17) that subgenome-length negative strands were dead-end products, the experimental data (13) clearly demonstrated they were active as templates for subgenomic mRNA synthesis.

How do subgenome-length templates arise? Sethna, Hofmann and Brian (14) proposed that subgenomic mRNAs would be replicated either after they were produced by leader-primed transcription or directly from subgenomic mRNAs that entered the cells because they were packaged within the virions. They reasoned the coronavirus replicase would be able to recognize the subgenomic mRNAs because they possessed the same 3' and 5' ends as genomic RNA. However, subsequently it was shown that infected cells did not replicate transfected subgenomic mRNAs or subgenomic DI-RNAs (8, 9, 10). Subgenomic mRNAs themselves, therefore, appear unable to participate directly in replication. Based on this observation, it appears that sequences downstream of the leader RNA, which are present in genomes and DI RNAs, are required to form replication complexes. If subgenomic mRNAs are incapable of acting as templates to initiate the formation of replication complexes, it would appear that discontinuous transcription of subgenomic mRNA or leader-primed transcription would not have a role to play in coronavirus transcription as originally proposed.

If subgenome-length negative-strand templates are not created by copying subgenomic mRNAs, then they must arise directly from the genome as we had proposed (13). Based on the fact that subgenome-genome length negative-strand templates are transcriptionally active, we propose a model, called 3'-discontinuous extension of negative-strand templates, to explain the discontinuous synthesis of subgenome-length negative stands. This model is depicted in Figure 4.

The major points of the model are that the viral replicase begins synthesis at the 3' end of the genome and pauses at the IS elements; it then either elongates through the IS elements or discontinues synthesis and switches to the 5' end of the genome and completes synthesis by copying the leader RNA. The result of 3' discontinuous extension during negative strand synthesis is that the subgenome-length negative strands acquire a complementary copy at their 3' end of the leader RNA. The replication complex that synthesized a subgenome-length negative strand would retain it as a template and go on to transcribe continuously, not discontinuously, subgenomic mRNA. The replication complexes of coronaviruses, like alphaviruses, would retain the negative strand template they created by copying the genome and use it as the preferred template (18). If the viral negative strand replicase elongated continuously to the 5' end of the genome, the replication complex would become loaded with a genome-length negative strand template and synthesize genomic RNA. Leader switching (19) would occur during negative strand synthesis if the incomplete or nascent subgenome negative strand switched templates and copied the 5' end of a different positive strand RNA. The extent of leader switching would be governed by the efficiency at which a particular genome was replicated, by the nature and location of the IS elements and by the sequence at the end of the leader RNA, i.e. whether it is a 2-repeat or 3- repeat element or has insertions or deletions. The coronavirus negative strand replicase might copy the
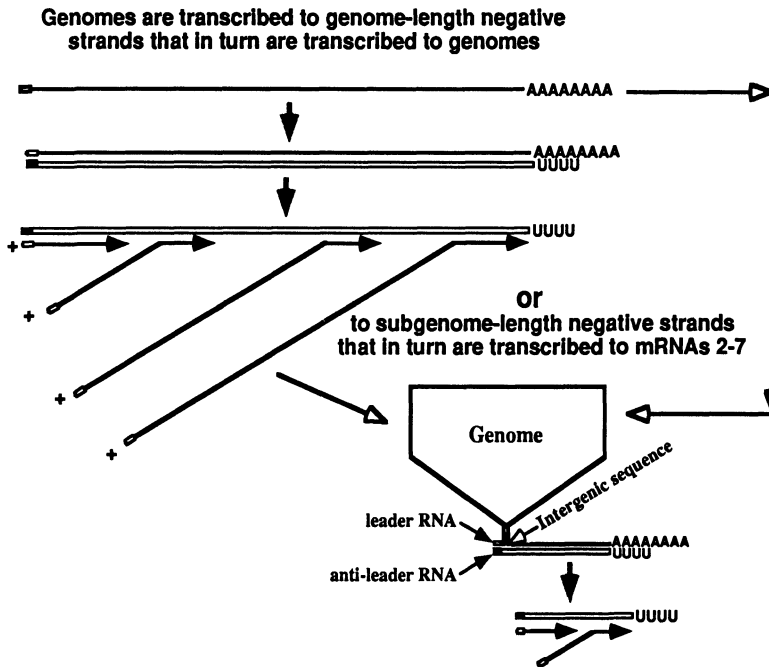
Figure 4. A model for coronavirus negative strand synthesis: 3'-discontinuous extension of negative-strand RNA synthesis.

template RNA in a fashion analogous to DNA-dependent RNA polymerases that are capable of retraction (20) and remain associated with the growing nascent strand rather then with the template (21): If retraction occurred after the viral replicase had copied the IS element, the exposed 3' end of the nascent negative strand would relocate and align precisely to complementary sequences at the 3' side of the leader RNA at the 5' end of the genome and complete synthesis of the negative strand. Therefore, the alignment for joining the body of the mRNA with the leader RNA would occur during the formation of the negative-strand template for the subgenomic mRNAs and not during mRNA synthesis. The interaction of the nascent negative strand with the 5' end of the genome might be mediated by protein:protein interaction between the replicase attached to the growing negative strand and a protein associated with the 5' end of the genome. Only positive strand templates that possess the promoter for negative strand synthesis at their 3' end and the appropriate sequences at the 5' end would be capable of forming active replication complexes. Therefore, positive strands might get copied into negative strands that would be incapable of producing positive strands because they would lack the promoter element at their 3' end.

Discontinuous extension of the 3' end of the negative strands has many attractive features. It would explain for instance why shorter mRNAs are overproduced relative to the longer mRNAs. van der Most et al (11) have shown that if the IS element for mRNA-3 was moved close to the 3' end of the genome, it became as active as the IS element for mRNA-7. The negative-strand template for the shortest mRNA would have the greatest probability of being produced since its IS element is promoter proximal. The sequence of the IS element also influenced the abundance of subgenomic mRNA (11). Therefore, position effect would not be the only factor governing the relative rate of synthesis of a given species of subgenomic mRNA. The open reading frame encoding the viral replicase might be translated simultaneously on the genome serving as a template for subgenome-length negative-strand synthesis. If a viral protein were to bind to the 5' end of the genome and determine the

3'-discontinuous extension of subgenome-length negative strands, its absence during the initial round of replication would guarantee the genome was copied first continuously into a genome-length negative strand template. Because coronavirus replication and transcription requires continuous translation (15), presumably of ORF1, subgenomic mRNA synthesis would be dependent on genome RNA synthesis because a constant supply of viral polymerase proteins would be needed to fuel transcription. Thus, the danger of the subgenome-length templates acting as DI RNA would be mitigated by several factors: Firstly, subgenome-length negative strand templates would arise from the genome, not from subgenomic mRNA; and secondly, the production of subgenomic mRNA would be dependent on the presence of enough genomic mRNA to supply the viral nonstructural proteins that would be required for transcription. To our knowledge, there are no experimental data that disprove such a model. Furthermore, it has many features that make it an attractive and informative model for the elucidation of the strategy that coronaviruses use to express their genetic information and replicate.

ACKNOWLEDGMENTS

REFERENCES

1 Spaan, W, D Cavanagh, and M C Horzinek 1988 Coronaviruses structure and genome expression J Gen Virol 69 2939-2952

2 Lai, M M C 1990 Coronavirus-organization, replication and expression of genome Annu Rev Microbiol 44 303-333

3 Spaan, W, H Delius, M Skinner, J Armstrong, P Rottier, S Smeekens, B A M van der Zeijst, and S G Siddell 1983 Coronavirus mRNA synthesis involves fusion of noncontiguous sequences EMBO J 2 1939

4 Lai, M M C , C D Patton, and S A Stohlman 1982 Replication of mouse hepatitis virus negative-stranded RNA and replicative form RNA are of genome length J Virol 44 487

5 Jacobs, L , W J M Spaan, M C Horzinek, and B A M van der Zeijst 1981 Synthesis of subgenomic mRNAs of mouse hepatitis virus is initiated independently evidence from UV transcription mapping J Virol 39 401-406

6 Stern, D F , and B M Sefton 1982 Synthesis of coronavirus mRNAs kinetics of inactivation of IBV RNA synthesis by UV light J Virol 42 755-759

7 Baric, R S , S A Stohlman, and M M C Lai 1983 Characterization of replicative intermediate RNA of mouse hepatitis virus presence of leader RNA sequences on nascent chains J Virol 48 633

8 Brian, D A , R -Y Chang, M A Hofmann, and P B Sethna 1994 Role of subgenomic minus-strand RNA in coronavirus replication Arch Virol (Suppl) 9 173-180

9 Makino, S , M Joo, and J K Makino 1991 A system for study of coronavirus mRNA synthesis a regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion J Virol 65 6031-6041

10 Masters, P S , C A Koetzner, C A Kerr, and Y Heo 1994 Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis J Virol 68 328-337

11 van der Most, R G , R J de Groot, and W J W Spaan 1994 Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus a study of coronavirus transcription initiation J Virol 68 3656-3666

12 Sethna, P B S-L Hung, and D A Brian 1989 Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons Proc Natl Acad Sci USA 86 5626-5630

13 Sawicki, S G, and D L Sawicki 1990 Coronavirus transcription subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis J Virol 64 1050-1056

14 Sethna, P B M A Hofmann, and D A Brian 1991 Minus-strand copies of replication coronavirus mRNAs contain antileaders J Virol 65 320-325

15 Sawicki, S G , and D L Sawicki 1986 Coronavirus minus-strand RNA synthesis and the effect of cycloheximide on coronavirus RNA synthesis J Virol 25 19-27

16 Hofmann, M A , P B Sethna, and D A Brian 1990 Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture J Virol 64 4108-4114

17 Jeong, Y S , and S Makino 1992 Mechanism of coronavirus transcription duration of primary transcription initiation activity and effects on subgenomic RNA transcription on RNA replication J Virol 66 3339-3346

18 Sawicki, S G, and D L Sawicki 1994 Alphavirus positive and negative strand RNA synthesis and the role of polyproteins in formation of viral replication complexes Arch Virol (Suppl) 9 393-405

19 Jeong, Y S , and S Makino 1994 Evidence for coronavirus discontinuous transcription J Virol 68 2615-2623

20 Kassavetis, G A , and E P Geiduschek 1993 RNA polymerase marching backward Science 259 944-945

21 Johnson, T L , and M J Chamberlin 1994 Complexes of yeast RNA polymerase II and RNA are substrates for TFIIS-induced RNA cleavage Cell 77 217-224