JOURNAL OF VIROLOGY, Jan. 2003, p. 1175–1183 Vol. 77, No. 2
0022-538X/03/$08.00+0 DOI: 10.1128/JVI.77.2.1175–1183.2003

The Stability of the Duplex between Sense and Antisense Transcription-Regulating Sequences Is a Crucial Factor in Arterivirus Subgenomic mRNA Synthesis

Alexander O. Pasternak, Erwin van den Born, Willy J. M. Spaan, and Eric J. Snijder*

Molecular Virology Laboratory, Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

Received 5 July 2002/Accepted 7 October 2002

Subgenomic mRNAs of nidoviruses (arteriviruses and coronaviruses) are composed of a common leader sequence and a "body" part of variable size, which are derived from the 5′- and 3′-proximal part of the genome, respectively. Leader-to-body joining has been proposed to occur during minus-strand RNA synthesis and to involve transfer of the nascent RNA strand from one site in the template to another. This discontinuous step in subgenomic RNA synthesis is guided by short transcription-regulating sequences (TRSs) that are present at both these template sites (leader TRS and body TRS). Sense-antisense base pairing between the leader TRS in the plus strand and the body TRS complement in the minus strand is crucial for strand transfer. Here we show that extending the leader TRS-body TRS duplex beyond its wild-type length dramatically enhanced the subgenomic mRNA synthesis of the arterivirus *Equine arteritis virus* (EAV). Generally, the relative amount of a subgenomic mRNA correlated with the calculated stability of the corresponding leader TRS-body TRS duplex. In addition, various leader TRS mutations induced the generation of minor subgenomic RNA species that were not detected upon infection with wild-type EAV. The synthesis of these RNA species involved leader-body junction events at sites that bear only limited resemblance to the canonical TRS. However, with the mutant leader TRS, but not with the wild-type leader TRS, these sequences could form a duplex that was stable enough to direct subgenomic RNA synthesis, again demonstrating that the stability of the leader TRS-body TRS duplex is a crucial factor in arterivirus subgenomic mRNA synthesis.

Discontinuous subgenomic (sg) mRNA synthesis is a hallmark of the replication cycle of arteriviruses and coronaviruses, members of the positive-strand RNA virus order *Nidovirales* (reviewed in references 4, 18, and 29). This process involves fusion of sequences that are noncontiguous in the viral RNA: the leader, which is derived from the 5′ end of the genome and is common to all sg mRNA species, is joined to different body segments, which are derived from the 3′ end of the genome. The sg mRNAs generated in this way are both 5′- and 3′-coterminal with the genome. The joining of leader and body occurs cotranscriptionally, most likely during minus-strand RNA synthesis (see below). Central to this process are short transcription-regulating sequences (TRSs). They are present in the genome (Fig. 1A) both at the 3′ end of the leader (leader TRS) and at different locations (body TRSs) throughout the 3′-proximal part of the genome which contains the structural protein genes. Usually, a body TRS precedes a structural protein-coding region and gives rise to an sg mRNA species that is used for structural protein expression. In arteriviruses, structural protein genes are frequently preceded by multiple body TRSs, which can be used with quite different efficiency, leading to the synthesis of major and minor sg transcripts from which the same protein can be translated (24; for a review, see reference 29).

Using the arterivirus *Equine arteritis virus* (EAV) as a model system, we have shown that the leader-body junction event involves a sense-antisense base pairing interaction between the leader TRS (5′-UCAACU-3′ in the plus sense) and the complement of the body TRS (3′-AGUUGA-5′ in the minus sense) (32). However, more than mere leader TRS-body TRS complementarity is required for a leader-body junction event to occur, as the EAV genomic RNA contains a number of sequences that match the leader TRS precisely but nevertheless are not used for sg mRNA synthesis (6, 24). Our recent findings argue that EAV sg mRNA synthesis depends on both sequence identity between leader TRS and body TRS and sequence-specific functions of the body TRS or, more likely, the primary and/or secondary structure of the sequences that flank the body TRS (25). This is consistent with the discontinuous minus-strand synthesis model (Fig. 1B), in which the body TRS is considered to be the attenuation signal for the replicase synthesizing the minus strand (26). The replicase would then transfer the nascent strand, with the body TRS complement at its ultimate 3′ end, to base pair with the leader TRS, which is probably presented by a stem-loop structure in the plus-strand template (Fig. 1C) (32). After that, nascent strand synthesis would be resumed to copy the leader sequence and complete the full-length sg minus strand, from which an sg mRNA would be transcribed. Recently, using genetic and biochemical approaches to analyze arterivirus and coronavirus sg mRNA production, strong evidence in favor of this model has been ob-

* Corresponding author. Mailing address: Molecular Virology Laboratory, Department of Medical Microbiology, Leiden University Medical Center, LUMC P4-26, P.O. Box 9600, 2300 RC Leiden, The Netherlands. Phone: 31 71 5261657. Fax: 31 71 5266761. E-mail: e.j.snijder@lumc.nl.

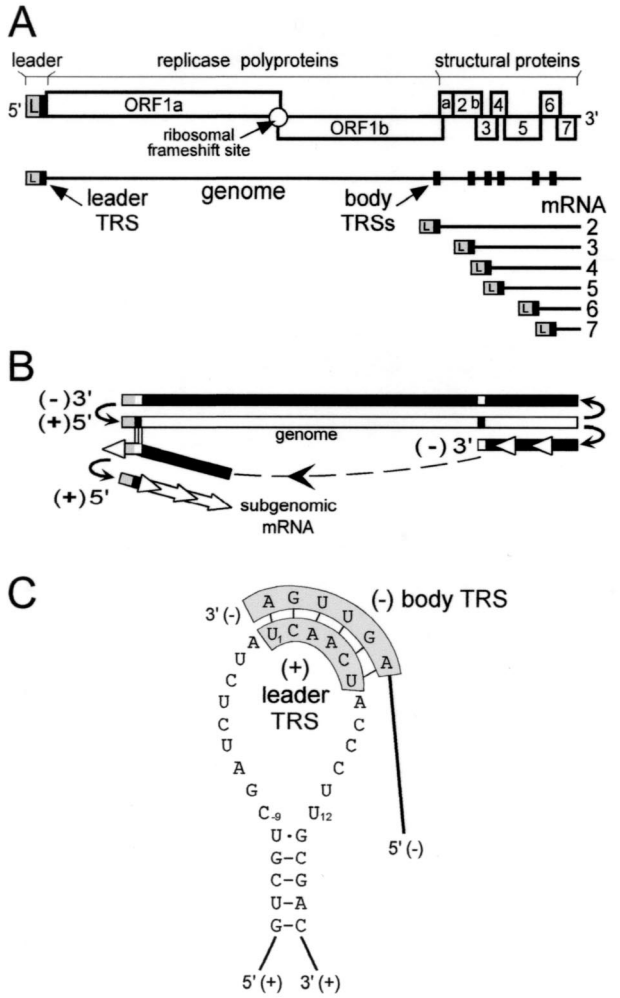

FIG. 1. (A) Schematic diagram of the genome organization and expression of EAV, the arterivirus prototype. The regions of the genome specifying the leader (L) sequence, the replicase gene (open reading frames 1a and 1b), and the structural protein genes are indicated. The nested set of EAV mRNAs (genome and sg mRNAs 2 to 7) is depicted below. The black boxes in the genomic RNA indicate the positions of leader and major body TRSs. (B) Discontinuous minus-strand synthesis model as proposed by Sawicki and Sawicki (26). This model proposes that the joining of the leader and body of sg mRNAs occurs by discontinuous minus-strand synthesis, producing the sg minus-strand template that is subsequently used to transcribe the sg mRNA. For details, see text. (C) RNA structure prediction for the EAV leader TRS region, which has been proposed to reside in a stable hairpin structure (32); only the loop and the top of the stem of the predicted hairpin structure are depicted here. The core of the leader TRS is indicated, with nt 1 of the TRS being labeled as U_1. Also depicted is the proposed base pairing interaction between the sense leader TRS and the antisense body TRS in the 3′ end of the nascent minus strand, as proposed in the model in panel B.

tained (3, 25, 27, 32), and consequently we will use it as our working model in this study.

To prove the role of the leader TRS-body TRS base pairing interaction in sg mRNA synthesis, we have previously employed site-directed mutagenesis of an EAV full-length cDNA clone to disrupt the leader TRS-body TRS duplex. We observed that, irrespective of the position of the mismatch in the TRS hexanucleotide, sg mRNA synthesis was significantly impaired in these mutants (25, 32). This suggested that TRS-TRS duplex stability is one of the factors determining the rate of sg mRNA synthesis. Here, we have investigated the role of TRS-TRS duplex stability in more detail. Instead of disrupting the duplex, we have mutagenized nucleotides directly downstream of the leader TRS to extend the possible duplex with several body TRSs that direct the synthesis of major EAV sg mRNAs (major body TRSs). At the same time, the possible duplex with other major body TRSs was reduced. We observed dramatic changes (in two directions) in the levels at which these mutants produced the various sg mRNAs. In most of the cases, the relative amounts of sg mRNAs directly correlated with the calculated stability of the corresponding leader TRS-body TRS duplex. We also observed that various leader TRS mutants produced novel minor sg RNA species. Synthesis of these RNA species involved leader-body junction events at sites that bear little resemblance to the canonical TRS. However, with a specific mutant leader TRS, but not with the wild-type (wt) leader TRS, these sequences were able to form a duplex that is stable enough to direct sg RNA synthesis. This argues that leader TRS-body TRS duplex stability is an important factor for the regulation of the synthesis of sg mRNAs and, consequently, structural proteins in arteriviruses.

MATERIALS AND METHODS

Site-directed mutagenesis and RNA transfections. Site-directed mutagenesis of EAV leader TRS and body TRSs was carried out as described by van Marle et al. (32), and all mutant constructs were checked by sequence analysis. Following in vitro transcription from infectious cDNA clones, full-length EAV RNA was introduced into BHK-21 cells by electroporation as described by van Dinten et al. (31).

RNA isolation and analysis. For RNA analyses, cells were lysed at 14 h posttransfection. Intracellular RNA isolation was performed by using the acidic phenol method as described by Pasternak et al. (24). Total intracellular RNA was resolved in denaturing agarose-formaldehyde gels. Hybridization of dried gels with the radioactively labeled oligonucleotide probe E154 (5′-TTGGTTCCTG GGTGGCTAATAACTACTT-3′), which is complementary to the 3′ end of the EAV genome and recognizes all viral mRNA molecules (genomic and subgenomic), and phosphorimager quantitation of individual bands were performed as described by Pasternak et al. (24). To determine the leader-body junction sequence of sg mRNAs, sg mRNA-specific reverse transcription (RT)-PCRs were done as described by van Marle et al. (34) using antisense (RT and PCR) primers from the RNA7, RNA6, or RNA2 body regions and a sense PCR primer matching a part of the leader sequence. RT-PCR products were sequenced directly as described by Pasternak et al. (24) using the leader-derived primer, an ABI PRISM sequencing kit (Perkin-Elmer), and an ABI PRISM 310 Genetic Analyzer (Perkin-Elmer).

Calculation of TRS duplex stability. The stability of leader TRS-body TRS duplexes was calculated in the form of the free energy of the duplex using the free-energy tables (version 3.0 at http://www.bioinfo.rpi.edu/~zukerm/rna/energy/) for RNA folding at 37°C (12, 20). Briefly, the free energy of a leader TRS-body TRS duplex was calculated as the sum of stacking energies of all pairs of adjacent base paired nucleotides in the duplex, energies of bulges and internal loops (where appropriate), and the single base stacking energies of the unpaired nucleotides at both ends of the duplex. The inclusion of additional (nonmatching) nucleotides up- or downstream of the TRS-TRS duplex did not influence the outcome of the analysis. For each TRS-TRS duplex, the structure with the lowest free energy was used (Tables 1 and 2).

RESULTS

Increasing the size of the leader TRS–body TRS duplex enhances sg mRNA synthesis. To study the influence of leader TRS-body TRS duplex stability on EAV sg mRNA synthesis,

TABLE 1. Leader TRS-body TRS duplexes[a]

Body TRS	WT	LU_6C	LA_7U	LC_8G	L_{678}
RNA2	+L 5′-UCAACUAC-3′ -B 3′-AGUUGAAG-5′ -11.0 kcal/mol[b]	+L 5′-UCAACCAC-3′ -B 3′-AGUUGAAG-5′ -8.9 kcal/mol	+L 5′-UCAACUUCC-3′ -B 3′-AGUUGAAGA-5′ -14.3 kcal/mol	+L 5′-UCAACUAG-3′ -B 3′-AGUUGAAG-5′ - 11.0 kcal/mol	+L 5′-UCAACCUG-3′ -B 3′-AGUUGAAG-5′ -8.9 kcal/mol
RNA3.1	+L 5′-UCAACUACCCU-3′ -B 3′-AGUU-AUGGGU-5′ -14.3 kcal/mol	+L 5′-UCAACCACCCU-3′ -B 3′-AGUU-AUGGGU-5′ -12.8 kcal/mol	+L 5′-UCAACUUCCCU-3′ -B 3′-AGUU-AUGGGU-5′ -12.9 kcal/mol	+L 5′-UCAACUAGCCU-3′ -B 3′-AGUU-AUGGGU-5′ - 9.8 kcal/mol	+L 5′-UCAACCUGCCU-3′ -B 3′-AGUU--AUGGG-5′ -12.1 kcal/mol
RNA3.2	+L 5′-UCAACUAC-3′ -B 3′-AGUUGACA-5′ -11.0 kcal/mol	+L 5′-UCAACCAC-3′ -B 3′-AGUUGACA-5′ -8.9 kcal/mol	+L 5′-UCAACUUC-3′ -B 3′-AGUUGACA-5′ -10.4 kcal/mol	+L 5′-UCAACUAG-3′ -B 3′-AGUUGACA-5′ - 11.0 kcal/mol	+L 5′-UCAACCUGC-3′ -B 3′-AGUUG-ACA-5′ -9.2 kcal/mol
RNA4.2	+L 5′-UCAACUAC-3′ -B 3′-AGUUGAAA-5′ -11.0 kcal/mol	+L 5′-UCAACCAC-3′ -B 3′-AGUUGAAA-5′ -8.9 kcal/mol	+L 5′-UCAACUUC-3′ -B 3′-AGUUGAAA-5′ -11.3 kcal/mol	+L 5′-UCAACUAG-3′ -B 3′-AGUUGAAA-5′ - 11.0 kcal/mol	+L 5′-UCAACCUG-3′ -B 3′-AGUUGAAA-5′ -8.9 kcal/mol
RNA5.1	+L 5′-UCAACUAC-3′ -B 3′-AGUUGAAC-5′ -11.0 kcal/mol	+L 5′-UCAACCAC-3′ -B 3′-AGUUGAAC-5′ -8.9 kcal/mol	+L 5′-UCAACUUC-3′ -B 3′-AGUUGAAC-5′ -11.3 kcal/mol	+L 5′-UCAACUAG-3′ -B 3′-AGUUGAAC-5′ - 11.0 kcal/mol	+L 5′-UCAACCUGC-3′ -B 3′-AGUUGAACC-5′ -9.6 kcal/mol
RNA6	+L 5′-UCAACUAC-3′ -B 3′-AGUUGGAA-5′ -11.0 kcal/mol	+L 5′-UCAACCAC-3′ -B 3′-AGUUGGAA-5′ -13.1 kcal/mol	+L 5′-UCAACUUC-3′ -B 3′-AGUUGGAA-5′ -11.0 kcal/mol	+L 5′-UCAACUAG-3′ -B 3′-AGUUGGAA-5′ - 11.0 kcal/mol	+L 5′-UCAACCUG-3′ -B 3′-AGUUGGAA-5′ -14.3 kcal/mol
RNA7	+L 5′-UCAACUACC-3′ -B 3′-AGUUGAUGA-5′ -14.5 kcal/mol	+L 5′-UCAACCACC-3′ -B 3′-AGUUGAUGA-5′ -10.0 kcal/mol	+L 5′-UCAACUUCC-3′ -B 3′-AGUUGAUGA-5′ -10.3 kcal/mol	+L 5′-UCAACUAG-3′ -B 3′-AGUUGAUG-5′ - 12.3 kcal/mol	+L 5′-UCAACCUGCC-3′ -B 3′-AGUUG-AUGA-5′ -9.5 kcal/mol

[a] For each duplex, the structure with the lowest free energy is shown. No base pairing is shown in cases where this is not energetically favorable. Mutated nucleotides are underlined.
[b] The calculated free energy is shown for every leader TRS-body TRS duplex, as described in Materials and Methods.

TABLE 2. Novel sg RNA species produced by leader TRS mutants[a]

Leader TRS mutation	Genome position of leader-body junction	Duplex between body TRS and mutant leader TRS	Potential duplex between body TRS and wild-type leader TRS[b]
LU_1C	8,736	+L 5′-**cua**C**CAAC**Uaccc-3′ -B 3′-uua**gguuguuuua**-5′ -11.0 kcal/mol[c]	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-uuagguuguuuua-5′* *-8.7 kcal/mol*
LU_1A	8,941	+L 5′-**cua**A**CAA-**Cuaccc-3′ -B 3′-acc**uguuagacaaa**-5′ -7.1 kcal/mol	*+L 5′-cuaUCAA-Cuaccc-3′* *-B 3′-accuguuagacaaa-5′* *-5.2 kcal/mol*
LU_1G	9,416	+L 5′-**cua**G**CA**ACUaccc-3′ -B 3′-**gaucgucgacaca**-5′ -13.6 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-gaucgucgacaca-5′* *-7.0 kcal/mol*
LC_5G	12,299	+L 5′-**cua**UCAAGUaccc-3′ -B 3′-**cauuguucgaaca**-5′ -9.5 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-cauuguucgaaca-5′* *-4.5 kcal/mol*
LU_1A[d]	12,299	+L 5′-**cua**A**CAA**CUaccc-3′ -B 3′-**cauuguuugagca**-5′ -8.6 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-cauuguucgaaca-5′* *-4.5 kcal/mol*
LU_1A	12,373	+L 5′-**cucua**A**CAA**-CUacc-3′ -B 3′-cg**gau-guucgaugu**-5′ -11.7 kcal/mol	+L 5′-**cucua**UCAA-CUacc-3′ -B 3′-cg**gau-guucgaugu**-5′ -11.7 *kcal/mol*
LU_1G	12,377	+L 5′-cuaGCAACUaccc-3′ -B 3′-**guucgauguuacu**-5′ -8.1 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-guucgauguuacu-5′* *-4.4 kcal/mol*
LA_4C	12,386	+L 5′-**cua**UCA**C**CUaccc-3′ -B 3′-gu**uacuggaugac**-5′ -12.5 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-guuacuggaugac-5′* *-8.4 kcal/mol*
LU_1A	12,522	+L 5′-**cua**A**CAAC**Uaccc-3′ -B 3′-gca**uguugccaag**-5′ -9.9 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-gcauguugccaag-5′* *-7.4 kcal/mol*
LA_4C	12,580	+L 5′-**cua**UCA**C**CUaccc-3′ -B 3′-**cau-guggcaguc**-5′ -9.5 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-gcauguggcaguc-5′* *-3.5 kcal/mol*
LA_3C	12,610	+L 5′-cuaUC**C**ACUaccc-3′ -B 3′-**uuuaggucgcguu**-5′ -9.8 kcal/mol	*+L 5′-cuaUCAACUaccc-3′* *-B 3′-uuuaggucgcguu-5′* *-5.5 kcal/mol*

[a] Leader TRS nucleotides are shown in capital letters; mutated nucleotides are underlined; nucleotides present in the sg RNA leader-body junction are shown in bold type.

[b] Leader-body TRS duplexes, which do not give rise to detectable amounts of sg RNAs, are shown in italics.

[c] The calculated free energy is shown for every leader TRS-body TRS duplex.

[d] Two additional nucleotide substitutions were introduced into the LU_1A construct for cloning purposes: G→A (12305) and U→C (12308).

we have chosen to mutagenize nucleotides directly downstream of the leader TRS. Thus, it was possible to differentially modulate the stability of the potential duplexes between the leader TRS and different body TRSs, because the nucleotides directly downstream of the TRS vary between different body TRSs (Fig. 2) (6). For example, a uridine residue at TRS position 6 is conserved in all EAV major body TRSs, except for the sg RNA6 body TRS. This TRS contains a C instead of U (5′ UCAACC 3′) (6), leading to a potential U-G base pair in the duplex between sense leader TRS and antisense body TRS. If disruption of the duplex by site-directed mutagenesis downregulates sg mRNA synthesis (25, 32), then transcription might be upregulated in TRS mutants in which the duplex is stabilized and/or expanded. This prompted us to check whether mRNA6 synthesis was enhanced in the LU_6C mutant (25). In this mutant, nucleotide (nt) 6 of the leader TRS has been changed from U to C, and therefore a stronger C-G base pair is present at position 6 of the leader TRS–RNA6 body TRS duplex (Table 1). At the same time, we expected the level of other sg mRNAs to drop, because the LU_6C substitution should destabilize all other TRS-TRS duplexes (Table 1). Figure 3 shows that the sg mRNA6 level was indeed 1.75-fold increased in mutant LU_6C, whereas the levels of all other sg mRNAs were 2.5- to 4-fold decreased. This implied that sg mRNA synthesis was quite sensitive to stabilizing or destabilizing the TRS-TRS duplex. Moreover, it was possible to in-

TRS		1	2	3	4	5	6	7	8	
Leader		5'	U	C	A	A	C	U	A	C	3'
RNA2	body	5'	U	C	A	A	C	U	U	C	3'
RNA3.1	body	5'	U	C	A	A	–	U	A	C	3'
RNA3.2	body	5'	U	C	A	A	C	U	G	U	3'
RNA4.2	body	5'	U	C	A	A	C	U	U	U	3'
RNA5.1	body	5'	U	C	A	A	C	U	U	G	3'
RNA6	body	5'	U	C	A	A	C	C	U	U	3'
RNA7	body	5'	U	C	A	A	C	U	A	C	3'

FIG. 2. Alignment of the EAV leader TRS and major body TRSs (plus sense). The variable positions downstream of the 5′ UCAAC 3′ core TRS, which were targeted in this study, are indicated with different shades of gray.

crease the synthesis of an sg mRNA above the level of the wt virus by mutating the leader TRS.

The fact that stabilizing the TRS-TRS duplex could enhance sg mRNA synthesis incited us to extend the duplex further downstream (in the plus strand). A uridine residue is present at position 7 of the body TRSs of mRNAs 2, 4.2, 5.1, and 6 (Fig. 2). The RNA 4.2 and 5.1 body TRSs are the major body TRSs for sg mRNAs 4 and 5, respectively, directing 97.5% or more of their synthesis (24). If the leader TRS contained a U at position 7, the formation of a 7-bp duplex between the leader TRS and four major body TRSs would be possible. To assess whether such an expansion of the duplex influenced the corresponding sg mRNA levels, the LA_7U mutant was constructed and tested. Indeed, sg mRNA levels, with the exception of that of mRNA7, were upregulated in this mutant (Fig. 3). At this point, we cannot explain the moderate (compared with the other sg mRNAs) upregulation of the sg mRNA3 level in the LA_7U mutant. On the other hand, the dramatic (3.5-fold) increase of the mRNA2 level can be explained by the fact that the nucleotide at position 8 of the RNA2 body TRS also matches with the leader TRS (Fig. 2), thus allowing the formation of an eight-base pair duplex in the LA_7U mutant (Table 1). Concurrently, the mRNA7 level was decreased in this mutant. This was expected, because the leader-body duplex for mRNA7, which consists of eight base pairs in the wt situation, had gained a mismatch and therefore was reduced to six consecutive base pairs. Again, these results demonstrate that expanding the normal 6-bp leader TRS-body TRS duplex beyond its wt size can enhance sg mRNA synthesis, whereas destabi-

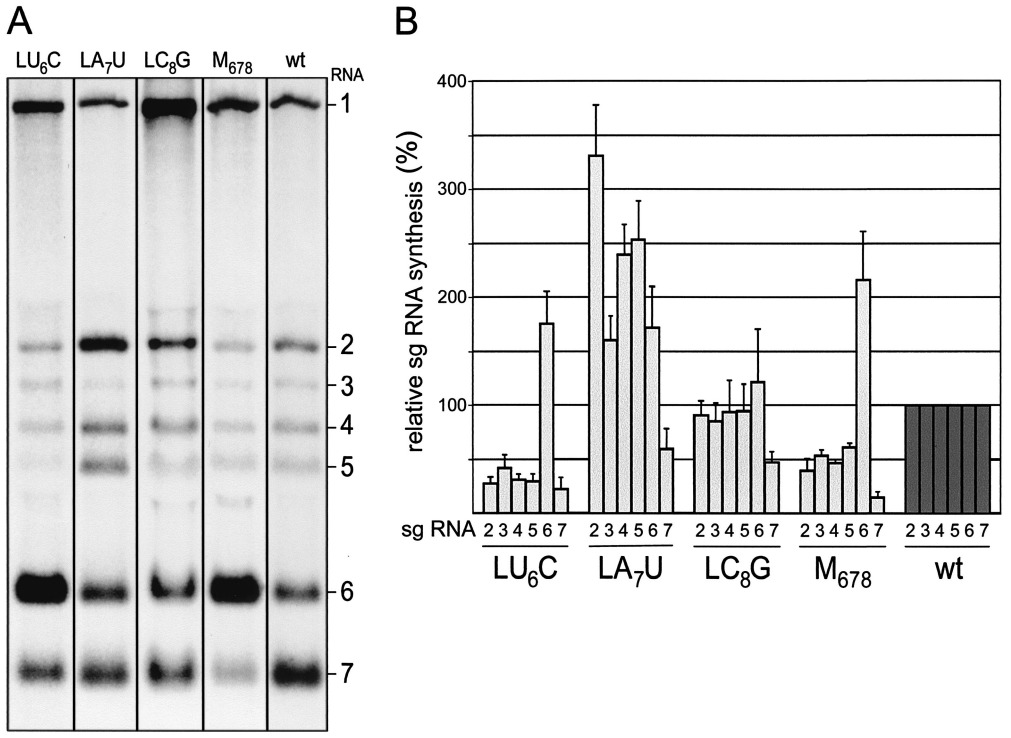

FIG. 3. (A) Northern blot analysis of EAV-specific RNA isolated from cells transfected with in vitro transcripts derived from constructs with a mutant leader TRS or with the wt EAV control. (B) Relative amount of sg mRNAs produced in cells transfected with leader TRS mutants or the wt control. For all samples, the level of sg mRNA synthesis was corrected for the level of genomic RNA, which was used as an internal standard as described previously (25). Subsequently, sg mRNA synthesis was related to the level of the same sg mRNA in the wt control in the same experiment, which was set at 100%. For each mutant construct, the experiment was repeated at least three times, and average sg mRNA levels and standard deviations are shown. The bands that appear to be present above sg mRNA2 and below sg mRNA5 in some of the lanes are artifacts of electrophoresis and hybridization, caused by the presence of larger amount of rRNA in these samples.

lization of a relatively large 8-bp duplex can suppress sg mRNA production.

In contrast to the two mutations described above, the C-to-G nucleotide substitution at position 8 of the leader TRS was not predicted to influence the size or stability of the leader TRS-body TRS duplex for all major body TRSs, with the exception of the RNA3.1 and RNA7 TRSs (Table 1). To check the effects of this substitution on sg mRNA synthesis, we constructed and tested the LC_8G mutant. As expected, no significant effects on the synthesis of sg mRNAs 2 to 6 were observed. The relatively weak effect on sg mRNA3 synthesis can be explained by the fact that the mutation affected the TRS-TRS duplex stability of only one of the mRNA3 subspecies, mRNA3.1, whereas the mRNA3.2 TRS-TRS duplex remained unchanged (Table 1). On the other hand, as in the case of the two mutants described above, the amount of mRNA7 produced by the LC_8G mutant was reduced to ~50% of the wt level (Fig. 3), underlining that base pairing at all eight positions of the leader TRS-RNA7 body TRS duplex is necessary for the optimal production of mRNA7.

Cumulative effect of TRS mutations on SG mRNA synthesis in a tri-nucleotide mutant. To assess how the simultaneous substitution of 3 nt downstream of the leader TRS can influence sg mRNA levels, we constructed the triple mutant L_{678}, in which the tri-nucleotide UAC at positions 6 to 8 of the leader TRS was changed to CUG. For each sg mRNA, the effect of this mutant was likely to reflect the combined effects of the corresponding single substitution mutants, described in the previous chapter. Indeed, the levels of SG mRNAs 2 to 5 were ~50% of the wt level (Fig. 3), slightly higher than those of LU_6C, but much lower than those of LA_7U, suggesting that the suppressing effect of the position 6 mutation was stronger than the enhancing effect of the position 7 mutation. This was expected because, in spite of the stabilizing effect of the acquired base pairing possibility at position 7, the duplex was still reduced to five consecutive base pairs due to the mismatch at position 6 (Table 1). The level of mRNA7 was downregulated stronger than in the LU_6C and LA_7U single-nucleotide mutants, because for this TRS, the duplex was reduced from eight to five consecutive base pairs. On the other hand, the mRNA6 amount was increased to a higher level than in the LU_6C or LA_7U mutants, because here a leader-body duplex of 7 bp was formed, with a stronger C-G base pair at position 6 (Table 1). Taken together, these data firmly established a role in transcriptional regulation for the 2 nt (TRS positions 7 and 8) immediately downstream of the conserved, 6-nt core (5′-UCA ACU-3′) of the EAV TRS.

Leader TRS mutations can activate cryptic body TRSs. In the course of our RT-PCR experiments to establish the sequence of the leader-body junction regions of sg mRNAs produced by various TRS mutants (25), we came across a remarkable phenomenon: the sg mRNA-specific RT-PCR products of some of the leader TRS mutants contained additional DNA fragments. Sequencing of these fragments revealed that they were derived from minor sg RNA molecules, which were in turn derived from leader-body junction events that involved various sequences in the open reading frame 1b region (upstream of the RNA2 body TRS) and the 3′-proximal region of the genome (downstream of the RNA7 body TRS). Although they bear little resemblance to the canonical EAV TRS, these cryptic "body TRS-like sequences" were apparently able to establish duplexes with some of the mutant leader TRSs. Our sequence analyses indicated that these duplexes probably also involved nucleotides up- and downstream of the leader TRS (Table 2). In all but one of these cases, we could not detect the corresponding sg RNA species in cells infected with wt EAV (reference 24 and data not shown). This could consistently be explained by the fact that the duplex formed between these cryptic body TRS-like sequences and the mutant leader TRSs was calculated to be considerably more stable than the duplex that could be formed between these sequences and the wt leader TRS (Table 2). In one case (LU_1A mutant and the junction site at position 12,373), the mutated nucleotide at position 1 of the leader TRS was predicted to be unpaired and therefore not to contribute to the duplex stability (Table 2). In agreement with that, the corresponding minor sg RNA species could also be detected in cells infected with wt EAV (Table 2) (24). However, neither the LU_1C nor the LU_1G mutant produced this sg RNA species (data not shown). Remarkably, the sg RNA species (RNA7.1) generated by the novel leader-body junction event in the LA_4C mutant (the site at nt 12,386) was abundant enough to be detected by Northern analysis (Fig. 4). This correlated well with the fact that the corresponding duplex between mutant leader TRS and body TRS was calculated to be more stable than any other duplex involving cryptic TRS-like sequences located in the 3′-proximal region of the genome, except for one junction site at position 9,416 from which a cryptic sg RNA could be produced by the LU_1G mutant (Table 2). However, the latter sg RNA could not be detected by Northern analysis (data not shown), probably because the site at nt 9,416 is upstream of the major body TRSs in the genome, whereas the site at nt 12,386 is

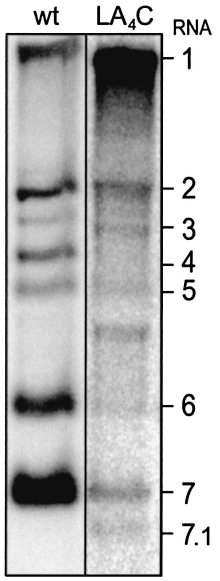

FIG. 4. Northern blot analysis of EAV-specific RNA isolated from cells transfected with the LA_4C leader TRS mutant or the wt EAV control. This analysis revealed the synthesis of a novel sg transcript (RNA7.1) by mutant LA_4C is indicated, due to the activation of a cryptic body TRS in the genomic 3′ end that matches the mutated leader TRS. For details, see text.

downstream. As predicted by the discontinuous minus strand synthesis model, much less minus strand-synthesizing RNA-dependent RNA polymerase complexes would reach the nt 9,416 site than the 12,386 site, and consequently less RNA molecules would be available for transfer to the leader TRS.

Origin of leader-body junctions in novel sg mRNAs. In accordance with our previous studies (25), sequence analysis of the leader-body junction region of the novel minor sg RNAs revealed that these were mostly body TRS-derived (Table 2), as predicted by the discontinuous minus-strand extension model (Fig. 1B and C). For most of the sg RNAs, it was not possible to map the crossover site precisely because of the nucleotide complementarity in the 5′ part of the duplex. Nevertheless, we were able to map the exact crossover site in the LA_3C mutant. The leader-body junction region of this mutant was completely body TRS derived. Interestingly, in three mutants (LA_4C, LA_3C, LU_1G), crossovers occurred upstream of the duplex (Table 2), resembling the pattern found with certain sg mRNAs of two other arteriviruses, *Lelystad virus* (an isolate of *Porcine reproductive and respiratory syndrome virus*) and *Simian hemorrhagic fever virus*. There, the leader-to-body fusion actually occurs 2 nt upstream of the conserved TRS (9, 21).

For the LU_1A mutant and the leader-body junction at position 12,373, a mixed population of sg mRNAs was found, with one part containing the leader TRS-derived mutant adenosine at TRS position 1 and the other lacking it. This implied that for the latter part leader-to-body joining occurred upstream of the leader TRS. Interestingly, an analogous sg mRNA species, which was isolated from cells infected with the wt EAV and was derived from a leader-body junction event involving the same cryptic body TRS-like sequence and the wt leader TRS, was homogeneous with respect to the origin of the leader-body junction sequence, and did not contain leader TRS-derived nucleotides (Table 2). This resembled the situation with the origin of sg mRNA7 leader-body junction sequences produced by various leader TRS or RNA7 body TRS position 1 mutants (25). There, analogous mixed populations of sg mRNAs were also found, again depending on the nature of the nucleotide at TRS position 1.

DISCUSSION

The data presented above confirm and extend our previous findings (24, 25, 32) in which an important role in EAV sg mRNA synthesis was attributed to the stability of the leader TRS-body TRS duplex. Previously, we analyzed the effects on sg mRNA synthesis of nucleotide substitutions in the conserved core sequence (5′-UCAACU-3′) of the EAV leader TRS and body TRS. These substitutions consistently decreased sg mRNA synthesis compared to the levels observed for wt virus. We have established that this effect was due both to disruption of the leader TRS-body TRS duplex and to an additional body TRS-specific function that is yet unknown but distinct from the TRS-TRS base pairing interaction (25). The leader TRS was postulated not to have any other function than formation of the duplex with the body TRS. Thus, in order to study exclusively the role of leader TRS-body TRS duplex stability in sg mRNA synthesis, we had to focus on leader TRS mutants only. In the present study, instead of mutagenizing the 6-nt core TRS (5′-UCAACU-3′) that is conserved between the

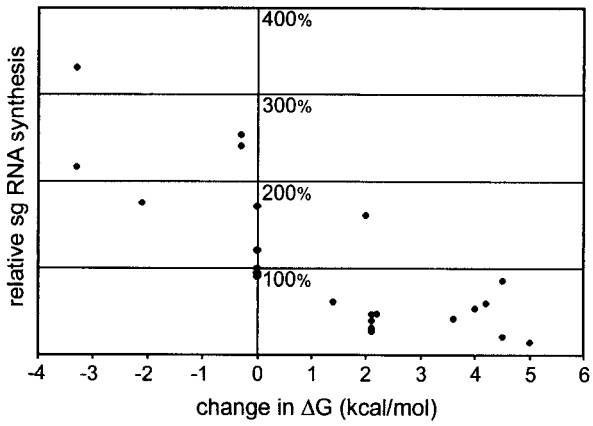

FIG. 5. Scatter plot showing the general correlation between relative sg mRNA synthesis and the change in free energy of the leader TRS-body TRS duplex. For each of the four leader TRS mutants in Table 1, the change of free energy of the TRS-TRS duplex (compared to the free energy of the same duplex in the wt virus) was plotted against the change in the synthesis of the corresponding sg mRNA, compared to that of the wt virus.

EAV leader TRS and all major body TRSs (with the exception of nt 6 of the RNA6 TRS), we substituted the nucleotides directly downstream of the leader TRS. Because the corresponding nucleotides downstream of the various body TRSs are not identical (Fig. 2), we expected that each of these mutant leader TRSs would differentially affect the synthesis of the various sg mRNAs, depending on whether the TRS-TRS duplex had been stabilized or destabilized.

We observed that extending the EAV leader TRS-body TRS duplex beyond its wt length (6 bp for all major body TRSs, except for the sg mRNA7 TRS) to 7 or 8 bp clearly enhanced the synthesis of the corresponding sg mRNA. At the same time (and in the same mutants), reducing the size of the leader TRS-body TRS duplex from 6 to 5 bp (or, in the case of sg mRNA7, from 8 bp to 5 or 6 bp) downregulated the level of the corresponding sg mRNA. Generally, relative sg mRNA levels correlated with the predicted relative stability of the leader TRS-body TRS base pairing interaction, which was calculated in the form of the duplex free energy (see Materials and Methods) for all mutants and duplexes studied (Fig. 5). However, in some cases, sg mRNA amounts correlated with the number of potential base pairs in the duplex, rather than with duplex stability. For example, the amount of sg mRNA6 in the LA_7U mutant was increased almost twice compared to that of the wt virus (Fig. 3). This probably reflected the gain of an additional U-A base pair in the leader TRS–RNA6 body TRS base pairing interaction, but not a change in calculated duplex stability, which was constant despite the increased duplex size. On the other hand, in the LU_6C mutant, the size of the duplex was not extended, but the duplex was stabilized (ΔG was reduced by 2.1 kcal/mol) due to the substitution of the U-G base pair at position 6 with a more stable C-G base pair. This, again, increased sg mRNA6 synthesis almost twice, indicating that both size and stability of the duplex are likely to play a role in sg mRNA synthesis.

We have previously observed that the synthesis of sg mRNA7 was often almost completely abolished when nucleo-

tides of the conserved core of the leader TRS were changed (25). If the size and/or stability of the leader TRS-body TRS duplex were the only factor regulating sg mRNA synthesis, the impact of these mutations would probably not be so large. For example, in the LU_1A mutant, the duplex size was reduced from 8 to 7 bp (whereas most body TRSs can form only a 6-bp duplex with the leader TRS), and its stability was reduced by 1.9 kcal/mol only (data not shown). Nevertheless, this mutant produced only 0.5% of the wt sg mRNA7 amount (25). The LA_7U mutant, on the other hand, in which the duplex size was reduced to 6 bp and its stability decreased by 4.2 kcal/mol, retained 60% of wt mRNA7 synthesis (Table 1 and Fig. 3). It is therefore tempting to speculate that the transcription mechanism requires base pairing at the core TRS nucleotides to properly position the nascent strand on the template before elongation. Additional nucleotide substitutions downstream of the core TRS, which extend and/or thermodynamically stabilize the duplex beyond its wt size and/or stability, could then be expected to enhance sg mRNA synthesis in a quantitative manner.

For the coronavirus *Mouse hepatitis virus* (MHV), which also belongs to the order of *Nidovirales*, the extent of potential base pairing between the leader TRS and the six major body TRSs (which are commonly referred to as intergenic sequences in coronaviruses) ranges from 7 to 18 bp (5, 16, 28). In MHV-infected cells (and also in EAV-infected cells), individual sg mRNAs are produced in constant but different amounts. For MHV, sg mRNA abundance correlates with the size of the leader TRS-body TRS duplex, which led Shieh et al. to postulate that duplex size determines sg mRNA levels (28). However, such a direct correlation does not exist for a number of coronaviruses (1, 10, 16). Although the EAV RNA7 body TRS, which directs synthesis of the most abundant sg mRNA, also forms the most-stable duplex with the leader TRS, several lines of evidence from the EAV system also argue against the hypothesis of Shieh et al. (28). EAV sg mRNAs are produced in molar ratios that generally are inversely proportional to their length and do not correlate with the size or calculated stability of the TRS-TRS duplex (Table 1). For example, mRNA3 is the least-abundant major sg mRNA species of EAV, despite the fact that there are two mRNA3 subspecies, mRNA3.1 and mRNA3.2, each of which originates from a body TRS that can form a quite long and stable duplex with the leader TRS (Table 1) (6). We have recently observed that the amount of sg mRNA7 is also not determined solely by TRS-TRS duplex stability, since many leader TRS/RNA7 body TRS double mutants in which the duplex was stabilized compared to the wt virus (for example, by the substitution of a U-A base pair with a C-G base pair as in the DU_1C mutant) (25) still showed dramatically reduced sg mRNA7 synthesis.

Having studied the effects of leader TRS-body TRS duplex size on sg RNA synthesis in a defective interfering (DI) RNA system, van der Most et al. (30) formally refuted the hypothesis of Shieh et al. (28) for the coronavirus MHV. It was found that a body TRS that allowed for the formation of a 17-bp duplex with the leader TRS did not yield higher sg RNA levels than a TRS with a 10-nt match. Similar results were recently obtained for the coronavirus *Transmissible gastroenteritis virus* by Alonso et al. (1). It was argued that, once the TRS-TRS duplex has achieved a certain stability, extended base pairing will not further increase sg RNA levels. Our present investigation using EAV yielded different results in that we have not observed this "saturation" of sg RNA synthesis, although it is possible that further extension of the duplex would eventually show this effect. This notion is supported by secondary RNA structure predictions of the leader TRS region in the EAV plus strand genomic template. The leader TRS is predicted to reside in the loop of a hairpin structure (32; E. van den Born and E. J. Snijder, unpublished observations) and also nucleotides C_{-9} to U_{12}, counting from the first nucleotide of the TRS, are predicted to be located in this leader TRS loop (Fig. 1C). One could argue that extending the duplex possibilities further upstream or downstream of this loop region would not further increase sg mRNA levels, because nucleotides upstream of C_{-9} and downstream of U_{12} are base paired and therefore cannot easily participate in duplex formation with the body TRS. If this hypothesis were correct, then sg mRNA synthesis would reach saturation with a 12-bp duplex. Unfortunately, it was not possible to study the effects of leader TRS mutations in the leader of coronavirus DI RNAs (30), because of the practical limitations of this experimental system (high-frequency leader recombination between DI RNA and helper virus). Consequently, in TRS mutagenesis studies with coronavirus DI RNA systems, only body TRSs could be targeted (1, 13, 14, 19, 30). Our data obtained with the EAV full-length cDNA clone (25) indicate that such body TRS mutations, in addition to disrupting the leader TRS-body TRS duplex, probably have additional consequences for the regulation of sg RNA synthesis. On the other hand, given the relatively large evolutionary distance between the two virus groups, differences may exist between the arterivirus and coronavirus transcription mechanisms. In this light, it should be noted that there is a conspicuous difference between the (average) size of the arterivirus TRS (~6 nt) and that of its coronavirus counterpart (10 to 12 nt).

We established that leader-body junction events can occur at a dozen noncanonical sites throughout the 3′-proximal part of the EAV genome (Table 2), provided that the complement of these sequences can form a duplex of sufficient stability with the (mutant) leader TRS. This could imply not only that formation of a stable leader TRS-body TRS duplex is necessary for sg RNA synthesis, but also that it can be sufficient for a junction event to occur. This idea is also supported by the finding that a number of minor sg RNA species originated from a foreign sequence (the green fluorescent protein gene) inserted into the genome of either MHV (8) or EAV (7). On the other hand, the fact that there are a number of sequences in the EAV genome that match the leader TRS precisely, but are not used for leader-body junction (6, 24), argues against this view. In addition, much less sg RNA was generated from the cryptic site at nt 12,373 (see above and Table 2) than would be expected on the basis of a comparison of the calculated stability of this TRS-TRS duplex and values obtained for various major TRSs in the wt virus (Table 1). Moreover, for many leader TRS mutants that produced minor sg RNA species from cryptic body TRS-like sequences (Table 2), additional potential TRS-like sequences were present in the 3′ end of the genome that could have generated TRS-TRS duplexes of comparable stability with the mutant leader TRS. Still, these sites

did not promote the synthesis of sg RNA molecules that could be detected with a sensitive RT-PCR assay (data not shown).

In a number of studies on coronavirus and arterivirus sg mRNA synthesis, a major role in the regulation of sg mRNA synthesis has been assigned to the primary sequences or RNA secondary structures that flank the body TRS (1, 2, 11, 22, 23, 24). Recently, we obtained data strongly suggesting that the EAV body TRS, in addition to being crucial for leader TRS-body TRS duplex formation, has a distinct function in sg mRNA synthesis, which may be either primary or secondary RNA structure specific (25). Thus, sequence context and the stability of the duplex that a given TRS-like sequence in the genome can form with the leader TRS may synergistically influence sg RNA synthesis. In addition, the relative amounts of sg mRNAs may be determined by the relative order of corresponding body TRSs in the viral genome (11, 15, 16, 17, 33; A. O. Pasternak and E. J. Snijder, unpublished observations).

Finally, the data presented in this study show that for all major sg mRNAs of EAV, with the possible exception of sg mRNA7, the stability of the leader TRS-body TRS duplex and, consequently, the amount of sg mRNA produced have not been maximized. Though we were able to experimentally enhance the transcription of several major sg mRNAs, it probably will not be possible to design a mutant leader TRS that would upregulate the synthesis of all major sg mRNAs. These data and earlier work (24) indicate that the leader and body TRSs of nidoviruses have evolved to optimize the transcription rate of the various sg mRNAs in order to coordinate the expression of the various structural proteins.

ACKNOWLEDGMENTS

We are grateful to Alexander Gultyaev for help with free energy calculations, Jessika Dobbe for technical assistance, and Marieke Tijms for helpful discussions.

A.O.P. was supported by grant 700-31-020 from the Council for Chemical Sciences of The Netherlands Organization for Scientific Research.

REFERENCES

1. **Alonso, S., A. Izeta, I. Sola, and L. Enjuanes.** 2002. Transcription regulatory sequences and mRNA expression levels in the coronavirus transmissible gastroenteritis virus. J. Virol. **76:**1293–1308.
2. **An, S., and S. Makino.** 1998. Characterizations of coronavirus cis-acting RNA elements and the transcription step affecting its transcription efficiency. Virology **243:**198–207.
3. **Baric, R. S., and B. Yount.** 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. J. Virol. **74:**4039–4046.
4. **Brian, D. A., and W. J. M. Spaan.** 1997. Recombination and coronavirus defective interfering RNAs. Semin. Virol. **8:**101–111.
5. **Budzilowicz, C. J., S. P. Wilczynski, and S. R. Weiss.** 1985. Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. J. Virol. **53:**834–840.
6. **den Boon, J. A., M. F. Kleijnen, W. J. M. Spaan, and E. J. Snijder.** 1996. Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. J. Virol. **70:**4291–4298.
7. **de Vries, A. A. F., A. L. Glaser, M. J. B. Raamsman, and P. J. M. Rottier.** 2001. Recombinant equine arteritis virus as an expression vector. Virology **284:**259–276.
8. **Fischer, F., C. F. Stegen, C. A. Koetzner, and P. S. Masters.** 1997. Analysis of a recombinant mouse hepatitis virus expressing a foreign gene reveals a novel aspect of coronavirus transcription. J. Virol. **71:**5148–5160.
9. **Godeny, E. K., A. A. F. de Vries, X. C. Wang, S. L. Smith, and R. J. de Groot.** 1998. Identification of the leader-body junctions for the viral subgenomic mRNAs and organization of the simian hemorrhagic fever virus genome: evidence for gene duplication during arterivirus evolution. J. Virol. **72:**862–867.
10. **Hofmann, M. A., R. Y. Chang, S. Ku, and D. A. Brian.** 1993. Leader-mRNA junction sequences are unique for each subgenomic mRNA species in the bovine coronavirus and remain so throughout persistent infection. Virology **196:**163–171.
11. **Hsue, B., and P. S. Masters.** 1999. Insertion of a new transcriptional unit into the genome of mouse hepatitis virus. J. Virol. **73:**6128–6135.
12. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
13. **Jeong, Y. S., J. F. Repass, Y. N. Kim, S. M. Hwang, and S. Makino.** 1996. Coronavirus transcription mediated by sequences flanking the transcription consensus sequence. Virology **217:**311–322.
14. **Joo, M., and S. Makino.** 1992. Mutagenic analysis of the coronavirus intergenic consensus sequence. J. Virol. **66:**6330–6337.
15. **Joo, M., and S. Makino.** 1995. The effect of two closely inserted transcription consensus sequences on coronavirus transcription. J. Virol. **69:**272–280.
16. **Konings, D. A., P. J. Bredenbeek, J. F. Noten, P. Hogeweg, and W. J. M. Spaan.** 1988. Differential premature termination of transcription as a proposed mechanism for the regulation of coronavirus gene expression. Nucleic Acids Res. **16:**10849–10860.
17. **Krishnan, R., R. Y. Chang, and D. A. Brian.** 1996. Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. Virology **218:**400–405.
18. **Lai, M. M. C., and D. Cavanagh.** 1997. The molecular biology of coronaviruses. Adv. Virus Res. **48:**1–100.
19. **Makino, S., and M. Joo.** 1993. Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. J. Virol. **67:**3304–3311.
20. **Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner.** 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. **288:**911–940.
21. **Meulenberg, J. J. M., E. J. de Meijer, and R. J. M. Moormann.** 1993. Subgenomic RNAs of Lelystad virus contain a conserved leader-body junction sequence. J. Gen. Virol. **74:**1697–1701.
22. **Nelsen, C. J., M. P. Murtaugh, and K. S. Faaberg.** 1999. Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. J. Virol. **73:**270–280.
23. **Ozdarendeli, A., S. Ku, S. Rochat, G. D. Williams, S. D. Senanayake, and D. A. Brian.** 2001. Downstream sequences influence the choice between a naturally occurring noncanonical and closely positioned upstream canonical heptameric fusion motif during bovine coronavirus subgenomic mRNA synthesis. J. Virol. **75:**7362–7374.
24. **Pasternak, A. O., A. P. Gultyaev, W. J. M. Spaan, and E. J. Snijder.** 2000. Genetic manipulation of arterivirus alternative mRNA leader-body junction sites reveals tight regulation of structural protein expression. J. Virol. **74:**11642–11653.
25. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J. **20:**7220–7228.
26. **Sawicki, S. G., and D. L. Sawicki.** 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. Adv. Exp. Med. Biol. **380:**499–506.
27. **Sawicki, D. L., Wang, T., and S. G. Sawicki.** 2001. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. J. Gen. Virol. **82:**385–396.
28. **Shieh, C. K., L. H. Soe, S. Makino, M. F. Chang, S. A. Stohlman, and M. M. C. Lai.** 1987. The 5′-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. Virology **156:**321–330.
29. **Snijder, E. J., and J. J. M. Meulenberg.** 1998. The molecular biology of arteriviruses. J. Gen. Virol. **79:**961–979.
30. **van der Most, R. G., R. J. de Groot, and W. J. M. 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.
31. **van Dinten, L. C., J. A. den Boon, A. L. M. Wassenaar, W. J. M. Spaan, and E. J. Snijder.** 1997. An infectious arterivirus cDNA clone: identification of a replicase point mutation that abolishes discontinuous mRNA transcription. Proc. Natl. Acad. Sci. USA **94:**991–996.
32. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. Proc. Natl. Acad. Sci. USA **96:**12056–12061.
33. **van Marle, G., W. Luytjes, R. G. van der Most, T. van der Straaten, and W. J. M. Spaan.** 1995. Regulation of coronavirus mRNA transcription. J. Virol. **69:**7851–7856.
34. **van Marle, G., L. C. van Dinten, W. J. M. Spaan, W. Luytjes, and E. J. Snijder.** 1999. Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. J. Virol. **73:**5274–5281.