JOURNAL OF VIROLOGY, May 2011, p. 4963–4973 Vol. 85, No. 10
0022-538X/11/$12.00 doi:10.1128/JVI.02317-10
Copyright © 2011, American Society for Microbiology. All Rights Reserved.

# Structure and Functional Relevance of a Transcription-Regulating Sequence Involved in Coronavirus Discontinuous RNA Synthesis▽

David Dufour,[1,2]† Pedro A. Mateos-Gomez,[3]† Luis Enjuanes,[3]* José Gallego,[1,2]* and Isabel Sola[3]

*Centro de Investigación Príncipe Felipe, Avda. Autopista del Saler 16, 46012 Valencia, Spain[1]; Instituto de Investigación Viña Giner, Universidad Católica de Valencia, Quevedo 2, 46001 Valencia, Spain[2]; and Centro Nacional de Biotecnología, CSIC, Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain[3]*

Received 5 November 2010/Accepted 26 February 2011

**Transmissible gastroenteritis coronavirus (TGEV) genomic RNA transcription generates 5′- and 3′-coterminal subgenomic mRNAs. This process involves a discontinuous step during the synthesis of minus-sense RNA that is modulated by transcription-regulating sequences located at the 3′ end of the leader (TRS-L) and also preceding each viral gene (TRS-Bs). TRSs include a highly conserved core sequence (CS) (5′-CUAAAC-3′) and variable flanking sequences. It has been previously proposed that TRS-Bs act as attenuation or stop signals during the synthesis of minus-sense RNAs. The nascent minus-stranded RNA would then be transferred by a template switch process to the TRS-L, which acts as the acceptor RNA. To study whether the TRS-L is structured and to determine whether this structure has a functional impact on genomic and subgenomic viral RNA synthesis, we have used a combination of nuclear magnetic resonance (NMR) spectroscopy and UV thermal denaturation approaches together with site-directed mutagenesis and *in vivo* transcriptional analyses. The results indicated that a 36-nucleotide oligomer encompassing the wild-type TRS-L forms a structured hairpin closed by an apical AACUAAA heptaloop. This loop contains most of the CS and is isolated from a nearby internal loop by a short Watson-Crick base-paired stem. TRS-L mutations altering the structure and the stability of the TRS-L hairpin affected replication and transcription, indicating the requirement of a functional RNA hairpin structure in these processes.**

Transmissible gastroenteritis virus (TGEV) is an enveloped virus of the *Coronaviridae* family containing a single-stranded positive-sense RNA genome of 28.5 kb. The first two-thirds of the genome encodes two replicase polyproteins, while the last third comprises the genes encoding structural and nonstructural viral proteins (Fig. 1A) (10, 18).

Coronaviruses (CoVs) share with other plus-stranded RNA viruses the synthesis of subgenomic mRNAs (sgmRNAs) as a mechanism to regulate the expression of the proteins encoded at the 3′ end of the genome (9, 10, 18, 22, 27). CoV sgmRNAs have a terminal leader sequence of around 60 to 93 nucleotides (nt) derived from the 5′ end of the genome fused to distant RNA coding sequences. This implies a discontinuous step during the synthesis of sgRNAs (22, 29, 35). This step is guided by transcription-regulating sequences (TRSs) found at the 3′ end of the leader (TRS-L) and also preceding each gene (TRS-Bs). TRSs comprise a common core sequence (CS) (5′-CUAAAC-3′), together with 5′- and 3′-flanking nucleotide sequences that vary depending on the gene but that also are essential for sgmRNA production (29, 35).

According to the current model of discontinuous transcription in CoVs (26, 29, 35) (Fig. 1B) and related arteriviruses (23, 24, 32), the TRS-Bs would act as slow-down and detachment signals for the transcription complex during the synthesis of minus-stranded RNA, possibly by forming a precomplex with the TRS-L and proteins (29). Then, base pairing between the complement of TRS-B (cTRS-B) in the nascent minus-stranded RNA and the TRS-L, located at the 5′ untranslated region (5′-UTR) of the genomic RNA, would allow a template switch. Resumption of minus-stranded RNA synthesis would add the complementary sequence of the 93-nt leader sequence to the minus-stranded sgRNA (Fig. 1B). Discontinuous transcription in CoVs would require at least three steps: (i) physical proximity of the distal TRS-B and TRS-L sequences, (ii) detachment of nascent minus-stranded RNA and transcription complex from the TRS-B, and (iii) base pairing between the TRS-L and the cTRS-B in the nascent RNA. This process resembles a high-frequency similarity-assisted RNA recombination in which the TRS-L would act as an acceptor for the cTRS-B donor sequence. This similarity-assisted RNA recombination usually requires hairpin structures present in the acceptor RNA (3, 7, 21, 23).

The sequence requirements for replication in TGEV and other CoVs have been determined (4, 11, 14, 17). Both 5′ and 3′ ends of viral genome contain *cis*-acting signals that are essential for replication. Modifications of the TRS-L sequence, located within an essential region of viral 5′ end, could affect genome replication, and the discontinuous RNA synthesis could also be affected by the availability of genome molecules as templates for transcription.

To assess the possible impact of the TRS-L structure on CoV genome replication and on template switching, we have used TGEV, a member of CoV genus α, as a model. First, the structure and stability of the TGEV TRS-L were determined

* Corresponding author. Mailing address for Luis Enjuanes: Department of Molecular and Cell Biology, Centro Nacional de Biotecnología, CSIC, Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain. Phone: 34-915854555. Fax: 34-915854915. E-mail: L.Enjuanes@cnb.csic.es. Mailing address for José Gallego: Instituto de Investigación Viña Giner, Universidad Católica de Valencia, Quevedo 2, 46001 Valencia, Spain. Phone: 34-993289680. Fax: 34-963289701. E-mail: jose.gallego@ucv.es.

† These authors contributed equally to the work.

▽ Published ahead of print on 9 March 2011.



Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

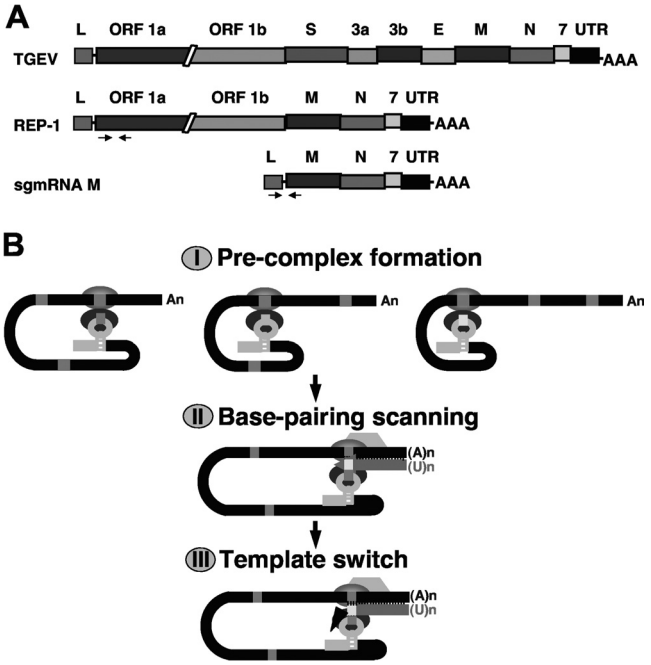


FIG. 1. TGEV genome organization and discontinuous transcription model. (A) Genetic structure of TGEV (upper line), the TGEV-derived replicon used in this study (middle line), and the sgmRNA M (lower line). TGEV genes are represented by letters above the boxes. The 93-nt TGEV leader is represented by an "L" above the rectangle. The arrows below the REP-1 and sgmRNA M schemes represent the binding sites of qRT-PCR primers. (B) Scheme showing the three-step working model of coronavirus discontinuous transcription. The leader sequence and TRS-L hairpin are represented in gray, and the CS sequences are depicted as light rectangles. (I) Formation of complexes between the TRS-L and the TRSs-B found along the coronavirus genome, mediated by sequence identity and RNA-protein interactions. These complexes would act as slow-down signals for the transcription complex. (II) Evaluation of base-pairing complementarity between the TRS-L hairpin and the complement of TRS-B in the nascent minus-stranded RNA. (III) Template switch of the nascent minus RNA to the TRS-L to resume the synthesis of minus sgRNA. This would occur only when TRS-L and cTRS-B form a stable duplex.

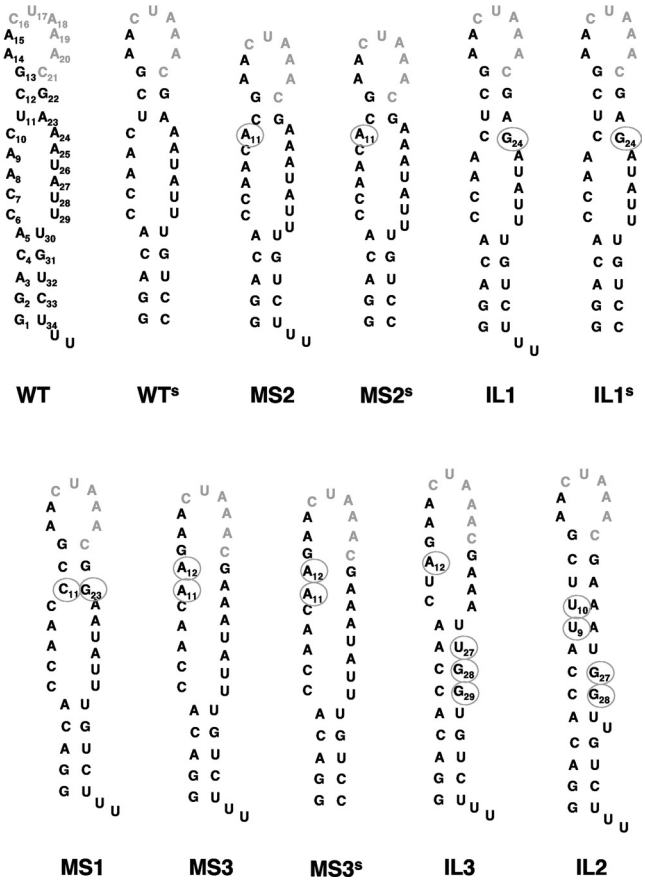


FIG. 2. Sequences and secondary structures of wild-type and mutant TRS-Ls of TGEV. The 36-nucleotide RNA oligomers comprise the 6-nucleotide CS (indicated in gray, from C-16 to C-21) flanked by 15 nucleotides on each side. Relative to the wild-type sequence (WT), the TRS-L mutant sequences change the stability of the upper stem (MS1), disrupt the upper stem (MS2, MS3, and IL-3), or close the adjacent internal loop (IL-1, IL-2, and IL-3) to different degrees. To reduce the tendency of some of these RNA hairpins to dimerize, several sequences were also synthesized with the terminal G1 · U34 pair replaced with G1 · C34 and with no 3′-UU overhang; these are designated with the same name followed by an "S" superscript. Circled nucleotides indicate the mutated positions.

by nuclear magnetic resonance (NMR) spectroscopy and UV thermal denaturing experiments. Then, specific domains of this structure were modified by site-directed mutagenesis and their effects on virus RNA replication and transcription were studied. The results indicated that the TRS-L of TGEV forms a well-defined hairpin structure closed by an apical 5′-AACUA AA-3′ heptaloop. This loop contains most of the leader CS (CS-L) and is isolated from a nearby internal loop by a short stem of three Watson-Crick pairs. Mutations that disrupt the structure of the TRS-L hairpin or modify its thermal stability had a significant effect on viral replication and transcription.

## MATERIALS AND METHODS

**Large-scale RNA transcription.** Unlabeled RNA oligonucleotides encompassing the TGEV TRS-L wild-type (WT) and mutant sequences (Fig. 2) were prepared by T7 polymerase *in vitro* transcription from oligonucleotide DNA templates. For the wild-type TRS-L sequence, unlabeled 36-nucleotide (WT) as well as unlabeled, $^{15}$N-labeled, and $^{13}$C/$^{15}$N-labeled 34-nucleotide (WT$^S$) variants were synthesized. $^{15}$N-labeled and $^{13}$C/$^{15}$N-labeled nucleoside triphosphates (NTPs) were prepared from cellular RNA as described previously (25). After a purification procedure based on gel electrophoresis and dialysis, all samples were microdialyzed in aqueous solutions containing 10 mM sodium phosphate (pH 6.0) and 0.1 mM EDTA. The final RNA concentration in the NMR samples ranged between 0.1 and 0.3 mM. The wild-type sample and several mutants were also analyzed in the same buffer supplemented with 100 mM NaCl or 5 mM $MgCl_2$ to evaluate the effects of ionic strength and divalent cations on the structures.

**NMR spectroscopy.** The NMR spectroscopy technique used provides inter-proton distances (through the nuclear Overhauser effect [NOE]) and hydrogen bonding information, which are useful for studying the three-dimensional (3D) structure of RNA molecules in solution (for a methodological review, see reference 13 and references therein). NMR spectra were acquired on 600-MHz (cryoprobe-equipped) and 500-MHz Bruker Avance III spectrometers, processed with Topspin 1.3 (Bruker Biospin), and analyzed using Sparky 3.110 (T. D. Goddard and D. G. Kneller, University of California, San Francisco). For the unlabeled samples, spectra recorded in $D_2O$ were acquired at several temperatures (typically 27°C and 34°C) and included series of 2D NOE spectroscopy (NOESY) (mixing times 80, 120, and 250 ms), double-quantum filtered correlation spectroscopy (DQF-COSY), total correlation spectroscopy (TOCSY) (60 ms), and rotating-frame NOESY (ROESY) experiments. NOESY spectra were also recorded in $H_2O$, generally at 8°C with a 150-ms mixing time. The relaxation delay was 2 s for all experiments with unlabeled samples. For the $^{15}$N-labeled

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

TABLE 1. Oligonucleotides used for PCR-directed mutagenesis

| Mutant | Oligonucleotide | Sequence (5′ → 3′)[a] |
|---|---|---|
| IL-1 | IL-1 VS | GTGCTAGATTTTGTCTTCGGACACCAACTCGAACTAAACGAGATATTTGTCTTTCTATG |
| MS1 | MS1 VS | GTGCTAGATTTTGTCTTCGGACACCAACCCGAACTAAACGGAATATTTGTCTTTCTATG |
| L1 | L1 VS | GTGCTAGATTTTGTCTTCGGACACCAACTCGTCCTAAACGAAATATTTGTC |
| IL-2 | IL-2 VS | GTGCTAGATTTTGTCTTCGGACACCATTTCGAACTAAACGAAATGGTTGTCTTTCTATG |
| IL-3 | IL-3 VS | GTGCTAGATTTTGTCTTCGGACACCAACTAGAACTAAACGAAATTGGTGTC TTTCTATG |
| MS2 | MS2 VS | GTGCTAGATTTTGTCTTCGGACACCAACACGAACTAAACGAAATATTTGTC |
| MS3 | MS3 VS | GTGCTAGATTTTGTCTTCGGACACCAACAAGAACTAAACGAAATATTTGTC |
| Common | Oli 5′I | CGC**GAATT**CGATGATAAGCTGTCAAAC |
| | Oligo RS | GTTGGTGTCCGAAGACAAAATCTAGCAC |
| | Oli 3′D | CGC**GAATT**CCTCTACTACTTTCCAAGCGT |

[a] The mutated nucleotides are underlined. Restriction sites are in bold (EcoRI, GAATTC).

$WT^S$ samples, 2D $^1H$-$^{15}N$ heteronuclear single-quantum correlation (HSQC) was recorded in $H_2O$, as well as HNN-COSY experiments allowing detection of hydrogen bonds between bases via two-bond N-N couplings (6). For the $^{13}C$/$^{15}N$-labeled $WT^S$ sample, 2D $^1H$-$^{13}C$ HSQC, 3D $^{13}C$-edited NOESY-heteronuclear multiple-quantum correlation (HMQC) (100- and 200-ms mixing times), 3D HCCH exclusive correlation spectroscopy (E-COSY), 2D HCCH-TOCSY, 3D HCP, and 3D $^{13}C$-edited $^1H$-$^{31}P$ heteronuclear correlation (HETCOR) spectra were also acquired at 27°C in $D_2O$.

**UV thermal denaturation experiments.** The thermal stability of the RNA structures was monitored by measuring the UV absorbance at 260 nm as a function of temperature in a Varian Cary 300 UV/visible spectrophotometer. The temperature was raised from 5 to 95°C at a gradient of 1°C $min^{-1}$ and subsequently decreased at the same rate to evaluate the reversibility of the process. The experiments were carried out using RNA at 0.1 to 0.6 optical density units (ODU)/ml (0.5 to 1.6 μM), and the thermal denaturing profiles were examined in three different aqueous solutions: one with 0.1 mM EDTA and 10 mM sodium cacodylate (pH 6.0) and the other two additionally containing either 100 mM NaCl or 5 mM $MgCl_2$.

**Cells and transfection.** Baby hamster kidney (BHK) cells stably transformed with the porcine aminopeptidase N gene (5) and with the Sindbis virus replicon pSINrep1 (12), expressing TGEV N protein (BHK-pAPN-N cells), were grown in Dulbecco's modified medium supplemented with 5% fetal calf serum (FCS) and G418 (1.5 mg/ml) as a selection agent. BHK-pAPN-N cells were grown to 95% confluence on 35-mm-diameter plates and transfected with 4 μg of each TGEV replicon, representing on average 100 molecules per cell, by using 12 μg of Lipofectamine 2000 (Invitrogen) according to the manufacturer's specifications. To minimize transfection variability, each pair of data used for comparison (mutant gRNA-WT gRNA and mutant sgmRNA-WT sgmRNA) came from the same transfection and quantitative reverse transcription-PCR (qRT-PCR) experiment. Furthermore, the conditions in transfection experiments were strictly controlled: (i) the same number of cells per well was seeded ($5 \times 10^5$ cells/well); (ii) the same amount of cDNA was always transfected (100 molecules per cell); and (iii) cDNA was purified using a Large-Construct kit (Qiagen) including an exonuclease treatment to remove bacterial DNA contamination and damaged plasmids, thus providing ultrapure DNA plasmid for transfection.

**Plasmid constructs.** cDNAs of TGEV-derived replicons were generated by PCR-directed mutagenesis. All mutants were generated from pBAC-TGEV plasmid (2), containing the TGEV genome (GenBank accession no. AJ271965) as a template, and specific oligonucleotides. Two overlapping fragments were obtained by PCR. The first one, common to all mutants, was generated with Oli 5′I and Oligo RS oligonucleotides (Table 1). The second fragment, specific for each mutant, was generated with Oli 3′D and the specific forward oligonucleotide (Table 1). After the overlapping PCR, the resulting fragments with SfiI and ApaLI restriction sites on the 5′ and 3′ ends, respectively, were introduced into the same sites of a pBAC plasmid containing only the first 4,423 nt of the TGEV genome. From this intermediate plasmid, the SfiI-ClaI fragments were obtained and then introduced into the same sites of pBAC-TGEV-REP1ΔCla (1) lacking the ClaI-ClaI fragment. The ClaI-ClaI fragment was subsequently introduced into all bacterial artificial chromosomes (BACs) to obtain complete pBAC-TGEV-REP1 including TRS-L mutant sequences.

**RNA analysis by real-time RT-PCR.** Total intracellular RNA was extracted at 24 h posttransfection (hpt) from transfected BHK-pAPN-N cells and purified with an RNeasy minikit (Qiagen) according to the manufacturer's specifications. To remove transfected DNA from samples for quantitative RT-PCR analysis, 7 μg of each RNA in 100 μl was treated with 20 U of DNase I (Roche) for 30 min at 37°C. DNA-free RNAs were repurified using the RNeasy minikit (Qiagen). cDNAs were synthesized using 60 ng RNA with the MultiScribe reverse transcriptase (high-capacity cDNA reverse transcription kit; Applied Biosystems). Quantitative analysis of RNAs from TGEV-derived replicons was performed by real-time RT-PCR. Oligonucleotides used for PCRs were designed with Primer Express software (Applied Biosystems) (Table 2). SYBR green PCR master mix (Applied Biosystems) was used in the PCR step, according to the manufacturer's specifications. Detection was performed with an ABI PRISM 7000 sequence detection system (Applied Biosystems). Data were analyzed with ABI PRISM 7000 SDS version 1.2.3 software.

**RNA hybridization assays.** RNA hybridization assays were performed between wild-type or mutant TRS-L sequences and the complement of TRS-7 (cTRS-7). Synthetic unlabeled 34-nt RNAs (Microsynth), containing wild-type or mutant TRS-L sequences similar to those analyzed by NMR, were incubated with a 16-nt biotinylated RNA with the cTRS-7 sequence. The hybridization reactions were performed with each pair of RNAs (0.025 μM each RNA) in a total volume of 10 μl of annealing buffer (40 mM Tris-HCl [pH 7.5], 0.1 M NaCl, 0.02 mM EDTA, 2% glycerol). RNAs alone were previously heated to 95°C for 2 min and then slowly cooled down to 37°C. Then annealing buffer was added, and reaction mixtures were incubated for 10 min at 37°C. Reactions were stopped by adding 2 μl of stop solution (50 mM EDTA, 2.5% SDS, 25% glycerol, 0.01% xylene cyanol, 0.01% bromophenol blue). RNAs were separated in 20% nondenaturing polyacrylamide gels and transferred to nylon membranes (BrightStar Plus; Ambion). Biotinylated RNAs were detected using BrightStar BioDetect kit (Am-

TABLE 2. Oligonucleotides used for real-time RT-PCR analysis

| Amplicon | Forward primer | | Reverse primer | |
|---|---|---|---|---|
| | Name | Sequence (5′ → 3′) | Name | Sequence (5′ → 3′) |
| gRNA | RT-REP-VS | TTCTTTTGACAAAACATACGGTGAA | RT-REP-RS | CTAGGCAACTGGTTTGTAACATCTTT |
| sgmRNA-7 | Ldrt-VS | CGTGGCTATATCTCTTCTTTTACTTT AACTAG | 7(38)-RS | AAAACTGTAATAAATACAGCATGGA GGAA |
| sgmRNA-N | Ldrt-VS | CGTGGCTATATCTCTTCTTTTACTTT AACTAG | N(82)-RS | TCTTCCGACCACGGGAATT |
| sgmRNA-M | Ldrt-VS | CGTGGCTATATCTCTTCTTTTACTTT AACTAG | mRNAM-RS | GCATGCAATCACACACGCTAA |

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

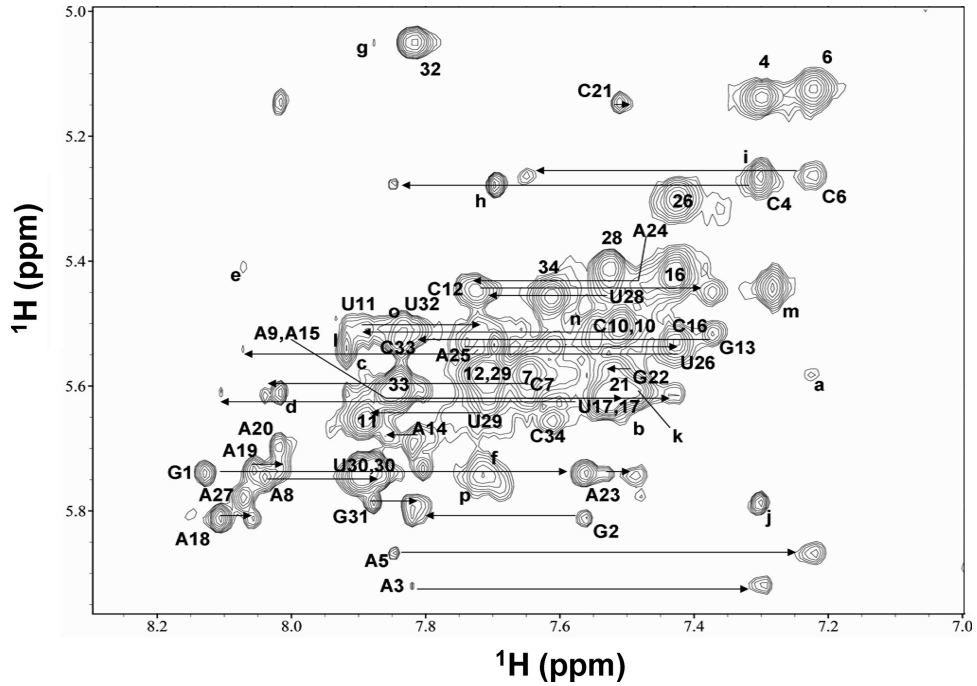


FIG. 3. Two-dimensional NOESY spectrum of wild-type TGEV TRS-L. The assignments of the H2/H6/H8-H1′/H5 WT$^S$ NOESY region (250-ms mixing time, 27°C) are indicated. Intraresidue H1′-H6/H8 cross peaks are labeled with residue name and number, intraresidue H5-H6 cross peaks are labeled with residue number, and sequential NOE connectivities are indicated with horizontal arrows. Cross peaks a to p are assigned as follows: a, C6 H6-C7 H5; b, C10 H6-U11 H5; c, U11 H6-C12 H5; d, A20 H8-C21 H5; e, A27 H8-U28 H5; f, U29 H6-U30 H5; g, G31 H8-U32 H5; h, A3 H2-C4 H1′; i, A5 H2-C6 H1′; j, A5 H2-G31 H1′; k, A8 H2-A9 H1′; l, A9 H2-U26 H1′; m, A23 H2-C12 H1′ and A23 H2-A24 H1′(overlapped); n, A24 H2-U11 H1′; o, A25 H2-C10 H1′; and p, A27 H2-A8 H1′. All of the assignments are supported by analyses of two- and three-dimensional NOESY, HCCH-TOCSY, and HCCH-COSY data obtained from a $^{15}$N/$^{13}$C-labeled WT$^S$ sample.

bion). The hybridization rate was estimated with band densitometry software (Quantity One; Bio-Rad).

***In silico* analysis.** Potential base-paring score calculations for mutant TRS-L sequences were performed as previously described (35). $\Delta G$ calculations were performed using a two-state hybridization server (http://www.bioinfo.rpi.edu/applications/hybrid/twostate.php) (19). Predictions of secondary structure were performed using the Mfold Web server for nucleic acid folding and hybridization prediction (http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi) (33). The analysis of sequences was performed using DNASTAR Lasergene software 7.0.

## RESULTS

**The TRS-L of TGEV forms a hairpin structure in solution.** Using unlabeled as well as $^{15}$N- and $^{15}$N/$^{13}$C-labeled samples, we have examined by NMR spectroscopy the conformation of the TGEV TRS-L, represented by 36-nucleotide and 34-nucleotide oligomers. The 36-nt wild-type oligomer (WT) contained the 6-nt CS-L core sequence and 15 nt flanking the CS-L 5′ and 3′ ends, whereas the 34-nt variant (WT$^S$) was generated by replacement of the terminal G1 · U34 with G · C and removal of the 3′-UU overhang (Fig. 2). Both oligonucleotides encompassed the transcription-regulating sequences previously shown to be relevant for TGEV discontinuous transcription (29, 35), and their lengths were appropriate for NMR analyses of wild-type and mutant TRS-Ls. Furthermore, previous Mfold predictions also supported the formation of a hairpin structure comprising this region of the TGEV leader sequence.

The WT and WT$^S$ oligonucleotides fold back on themselves, forming similar hairpin motifs in solution (Fig. 2), as indicated by the hydrogen-bonding and NOE interactions detected by the NMR analyses (Fig. 3 and 4). These hairpins were closed by an apical 5′-AACUAAA-3′ heptaloop comprising most of the CS. A short stem of three Watson-Crick pairs separated this apical loop from an adjacent 5′-CCAAC · AAUAUU-3′ internal loop, which was closed on the other side by a terminal stem of five base pairs (Fig. 2). The presence of the Watson-Crick pairs forming the upper and lower stems was clearly demonstrated by HNN-COSY experiments allowing detection of hydrogen bonds between bases via two-bond N-N couplings (6) and by NOESY data (Fig. 3), as well as by detection of C3′-endo conformations typical of nucleotides embedded in standard A helices. Detailed analyses of two- and three-dimensional NOESY WT$^S$ data indicated that most of the bases of both the apical and internal loops of the TRS-L hairpin were not opened but rather formed well-defined stacked structures. In the apical 5′-AACUAAA-3′ heptaloop, the C16 and U17 nucleotides made the backbone turn, but sugar-base and base-base NOE interactions were detected for the G13-A15 and A18-C21 segments of the loop (Fig. 3), clearly indicating nucleotide stacking. Likewise, the nucleotides of both strands of the internal loop also formed continuous stacks, as indicated by sugar-base and base-base sequential NOE interactions as well as by interstrand adenine H2-sugar H1′ NOEs (Fig. 3). This was expected, given the similar lengths of both sides of the loop (5 and 6 nucleotides). However, these interactions were weaker than those observed in Watson-Crick stems. The presence of this internal loop also broadened the imino resonances

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

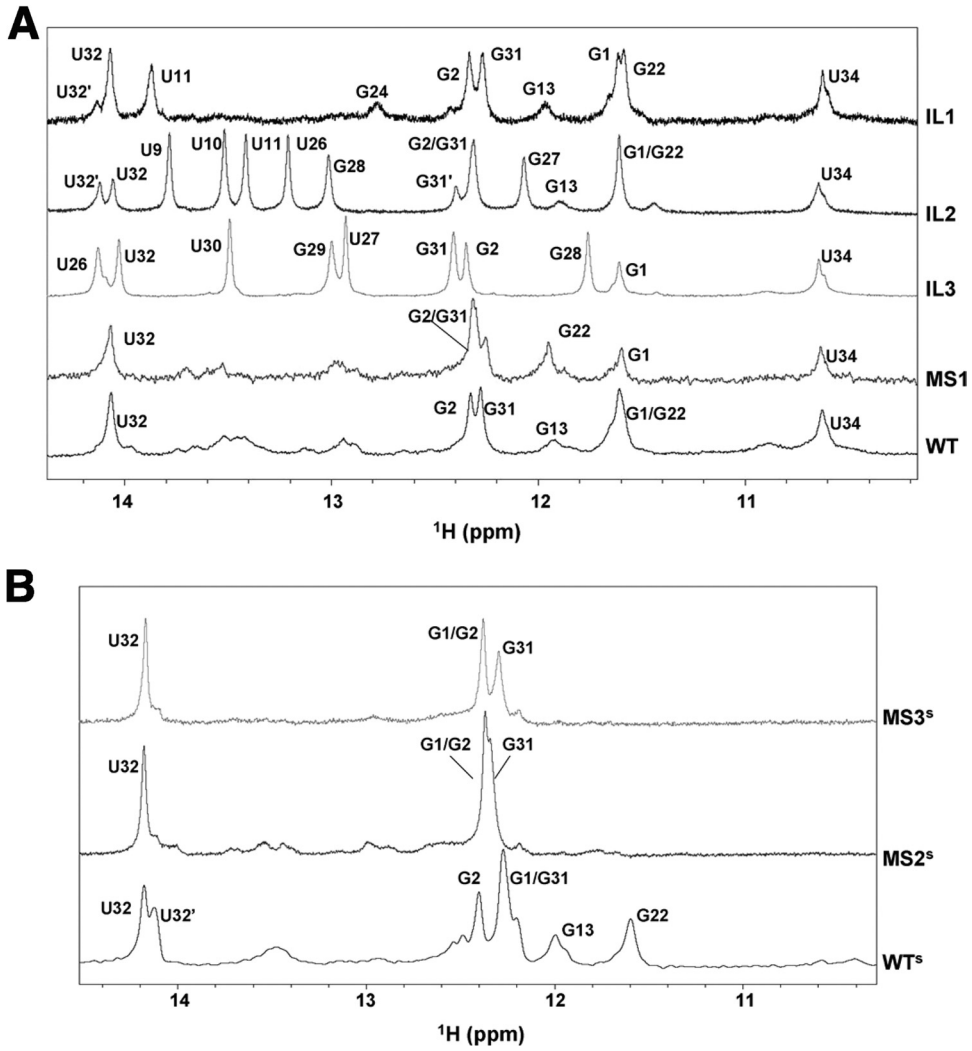


FIG. 4. Imino proton NMR spectra of wild-type and mutant TGEV TRS-Ls. In all cases, the data were recorded at 12°C and the assignments were obtained from NOESY and TOCSY two-dimensional spectra. (A) Spectra of the 36-nucleotide WT and all mutants except MS2 and MS3, which aggregated under all NMR conditions tested. (B) Spectra of the 34-nucleotide $WT^S$ and mutants $MS2^S$ and $MS3^S$. The U32′ and G31′ resonances in IL-1, IL-2, and $WT^S$ represent a secondary conformation of the lower stem related to the presence of the U29/U30 bulge opposite A5.

of the upper stem (Fig. 4) and significantly affected the thermal stability of the WT and $WT^S$ hairpins (see below).

Gel electrophoresis and UV thermal denaturation experiments indicated that a fraction of the WT hairpins dimerized in solution (Fig. 5). This hairpin dimerization process was mostly abolished in the 34-nt RNA oligomer in which 3′-terminal UU nucleotides were deleted ($WT^S$ system), as revealed by the presence of a monomeric band in gel electrophoresis experiments (data not shown) and by monophasic UV denaturing profiles. The $WT^S$ RNA motif also led to better-quality NMR spectra (Fig. 3). Analysis of HNN-COSY spectra confirmed that the base-pairing interactions detected by NMR were established within the monomeric hairpins shown in Fig. 2. This analysis was performed by comparison of HNN-COSY spectra of $WT^S$ obtained with a 100% $^{15}$N-labeled sample and with a mixture of 50% unlabeled and 50% $^{15}$N-labeled RNA (8).

**The TRS-L mutants modulate the structure and stability of the hairpin.** Six mutants were designed to modulate the structural and thermodynamic properties of the TRS-L hairpin (Fig. 2) and were analyzed by one- and two-dimensional NMR spectroscopy (Fig. 4) and UV thermal denaturation experiments (Table 3). Internal loop (IL) mutants IL-1 and IL-2 were designed to close this loop to different extents (Fig. 2). In mutant IL-2, the internal loop was almost completely replaced by a Watson-Crick base-paired stem, as demonstrated by its exchangeable and nonexchangeable proton NMR spectra (Fig. 4A). This mutant exhibited the strongest thermal stabilization with respect to the wild type (Table 3). In the single mutant IL-1, the internal loop was partially closed from the apical side by adding one additional C10 · G24 base pair to the upper stem. This was confirmed by the NMR data and is indicated in Fig. 4A by the presence of a G24 Watson-Crick imino and by

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

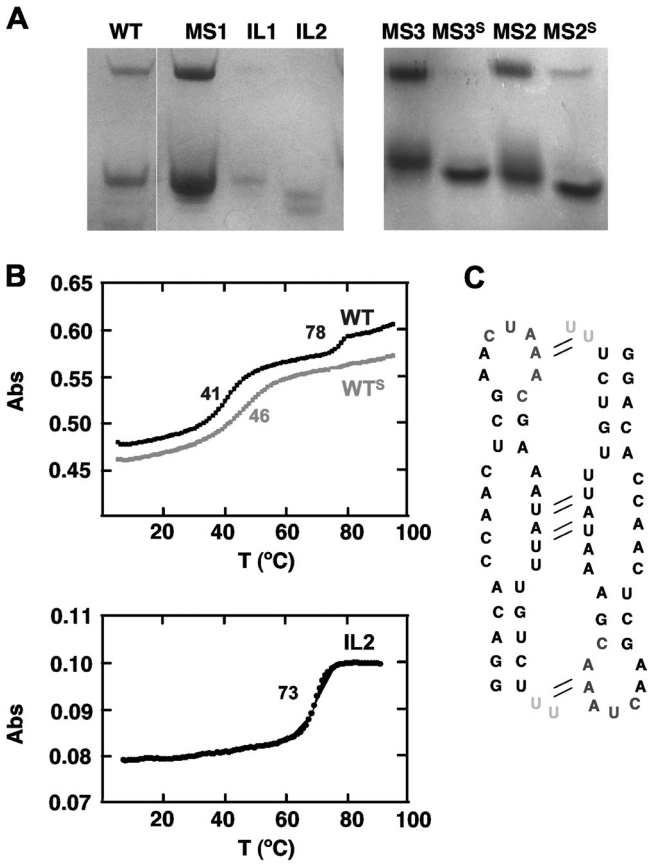


FIG. 5. Gel electrophoresis and UV thermal denaturation profiles of wild-type and mutant TRS-Ls. (A) Native gels comparing the electrophoretic mobilities of the indicated RNA oligonucleotides. Experimental conditions were as follows: 4°C, 20% polyacrylamide, 10 μM RNA in 10 mM sodium cacodylate (pH 6.0), and 0.1 mM EDTA. The upper band exhibited by the WT, MS1, IL-1, MS2, and MS3 RNAs corresponds to a dimeric species that is weakly populated or absent in the sequences where the internal loop is closed (IL-2) or the 3′-UU tail is eliminated ($MS2^S$ and $MS3^S$). (B) UV-monitored thermal denaturation curves of WT, $WT^S$, and IL-2 oligonucleotides. Melting temperatures are indicated on the curves. Sequences giving rise to biphasic curves have a sequence-dependent lower transition and an essentially sequence-independent upper transition that likely corresponds to the melting of the dimer (Table 3). WT RNA exhibits a biphasic thermal curve with two transitions, whereas $WT^S$ (lacking a 3′-UU tail) and IL-2 (without an internal loop) show a monophasic profile with only one melting temperature. (C) Model of an antiparallel dimer of hairpins likely responsible for the upper electrophoretic bands and the biphasic melting curves exhibited by some TRS-L sequences. Dimer formation is facilitated by interactions between the self-complementary AAUAUU sequence present in the internal loop and by head-to-tail interactions between the AACUAAA apical loop and the 3′-UU tail (indicated by gray lines). HNN-COSY experiments conducted with $WT^S$ confirmed that all base-pairing interactions detected by NMR are established within intramolecular hairpins.

the stabilization of the nearby U11 imino. The thermal stability of this mutant hairpin was also increased in relation to the wild type but to a lesser extent relative to IL-2 (Table 3). The double mutant MS1 (Fig. 2) was designed to increase the stability of the TRS-L hairpin without altering its structure by replacing the terminal U11 · A23 pair of the upper microstem (MS) with a C11 · G23 pair. The MS1 structure was confirmed by one- and two-dimensional NMR experiments (Fig. 4A). The effect of these mutations on thermal stability was weaker than that observed in the IL-1 and IL-2 structures, as expected (Table 3).

In contrast to the IL-1, IL-2, and MS1 mutants, which were designed to close the internal loop or stabilize the upper stem, mutants IL-3, MS2, and MS3 were intended to destabilize the upper stem located below the CS-L to different degrees. In the IL-3 mutant the internal loop was partially closed at the terminal side, while the Watson-Crick pairs of the upper stem were completely disrupted (Fig. 2). This was confirmed by two-dimensional spectroscopy data. The absence of the upper-stem U11, G13, and G22 imino resonances and the appearance of new lower-stem U26, U27, G28, and G29 Watson-Crick iminos can be observed in Fig. 4A. The thermal stability of the IL-3 mutant was higher than that of the wild-type sequence (Table 3). Mutants MS2 and MS3 were designed to eliminate one or two Watson-Crick pairs from the upper stem, respectively (Fig. 2). These two adenine-rich mutants exhibited lower thermal stability than the wild-type hairpin, as expected (Table 3). Since the MS2 and MS3 mutants aggregated under all NMR conditions tested, NMR analyses focused instead on 34-nt $MS2^S$ and $MS3^S$ variants lacking the 3′-UU overhang (Fig. 2), which could be studied by one- and two-dimensional spectroscopy experiments (Fig. 4B). The exchangeable G13 and G22 iminos, already broadened in the WT and $WT^S$ spectra, were not detectable in either mutant (Fig. 4B). However, examination of two-dimensional NOESY and TOCY data at 15°C revealed that the nonexchangeable proton chemical shift and cross peak patterns in $MS2^S$ were similar to those observed in the $WT^S$ hairpin (data not shown). In contrast, these patterns were different in $MS3^S$, indicating that the apical region of this mutant hairpin did not retain the structure adopted by the wild-type sequence.

TABLE 3. Thermal stabilities and hybridization energies of wild-type and mutant TRS-Ls of TGEV

| Sequence | $T_m$ (°C)[a] | $\Delta G_{hyb}$[b] (kcal mol$^{-1}$) | Activity[c] (% of WT) |
|---|---|---|---|
| WT | 39 ± 1 | −14.1 | 100 |
| MS2 | 34 ± 1 | −14.1 | 103 |
| IL-1 | 51 ± 1 | −13.8 | 50 |
| MS1 | 45 ± 2 | −13.6 | 40 |
| MS3 | 35 ± 2 | −14.0 | 35 |
| IL-3 | 67 ± 2 | −14.0 | 1 |
| IL-2 | 72 ± 2 | −13.8 | 0.2 |

[a] Melting temperatures obtained from UV-monitored thermal denaturation profiles. The averages and standard deviations were obtained from two to four experiments carried out in aqueous solutions containing 10 mM sodium cacodylate (pH 6.0), 100 mM NaCl, and 0.1 mM EDTA. Most of the intramolecular hairpins containing a 3′-UU tail (Fig. 2) underwent a hairpin dimerization process characterized by a biphasic UV thermal transition and an additional electrophoretic band. The reported $T_m$ values correspond to the first transition, which reflects the melting of the intramolecular hairpins. The second transition is essentially sequence independent (76 ± 4°C for all sequences except IL-2 and IL-3) and likely corresponds to the separation of the dimer (Fig. 5).

[b] Base-pairing energies between the 5′-**CGAA**CUAAAC**GAAA**-3′ TRS-L (+) sequence and the 5′-UUUGUUUAGUUCG-3′ complement of TRS-M, obtained with the DINAMELT software (19). The core sequences are underlined, and the nucleotides that vary in the different TRS-L (+) mutants are indicated in bold.

[c] The values correspond to sgmRNA synthesis levels of replicon mutants.

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

TABLE 4. Hybridization energies of wild-type and mutant TRS-Ls of TGEV with the cTRS-Bs of N and 7 genes

| TRS-L sequence | mRNA N | | mRNA 7 | |
|---|---|---|---|---|
| | $\Delta G_{hyb}$ (kcal mol$^{-1}$)[a] | Activity (% of WT)[b] | $\Delta G_{hyb}$ (kcal mol$^{-1}$)[c] | Activity (% of WT)[b] |
| WT | −8.6 | 100 | −15.1 | 100 |
| MS2 | −8.6 | 107 ± 12 | −15.1 | 110 ± 27 |
| IL-1 | −-8.6 | 61 ± 13 | −18.2 | 49 ± 12 |
| MS1 | −8.6 | 35 ± 12 | −14.3 | 36 ± 11 |
| MS3 | −8.6 | 32 ± 6 | −15.0 | 39 ± 15 |
| IL-3 | −8.6 | 1.3 ± 0.9 | −15.0 | 1.1 ± 0.6 |
| IL-2 | −8.6 | 0.4 ± 0.4 | −15.1 | 0.3 ± 0.2 |

[a] Base-pairing energies between the wild-type (5′-CGAACUAAACG**AA**A-3′) or mutant TRS-L (+) sequences and the complement (−) of TRS-N (5′-AGA AGUUUAGUUAU-3′) were obtained with the DINAMELT software (19). The complementary nucleotides between TRS-L and (−) TRS-N sequences are underlined, and the positions that vary in the different TRS-L mutants are indicated in bold. Mutations introduced in mutant TRS-L sequences are outside the complementary region and do not affect $\Delta G$ values.

[b] The values correspond to the sgmRNA synthesis levels in each replicon mutant compared to those in the wild-type replicon.

[c] Base-pairing energies between the wild-type (5′-CGAACUAAACG**AA**A-3′) or mutant TRS-L (+) sequences and the complement (−) of TRS-7 (5′-UCU CGUUUAGUUCG-3′). The complementary nucleotides between TRS-L and (−) TRS-7 sequences are underlined, and the positions that vary in the different TRS-L mutants are indicated in bold.

**The structure and stability of the TRS-L hairpin affect TGEV replication and transcription levels.** The TRS-L is embedded within the 5′-UTR, and this region is involved in CoV replication (9, 14). Therefore, TRS-L mutations might have an effect on replication, which would in turn influence the total amount of mRNAs accumulated. As a consequence, it was decided to analyze the replication of the TRS-L mutants to evaluate whether the effects observed in transcription were also due to changes in replication leading to reduced levels of template gRNA. Transfected replicon cDNA is first transcribed in the nucleus by cellular RNA polymerase II (pol II) to yield replicon gRNA. These replicon gRNA molecules are then transported to the cytoplasm, where they are translated to produce the viral replicase, which amplifies replicon gRNA and also transcribes sgmRNAs. The amount of gRNA synthesized by RNA pol II in transfected BHK-pAPN-N cells was quantified using a mutant replicon with an inactive replicase. This amount represented 1% of gRNA levels in a replication-competent replicon (data not shown). Therefore, the amount of replicon gRNA quantified in transfection experiments was mainly (99%) produced by the replicon self-amplification activity.

The effect of the TRS-L mutations on replication and transcription activities was analyzed using a reverse genetics system (2). The mutations were introduced into the cDNA of a TGEV replicon, and the replication and transcriptional activities of mutant replicons transfected into BHK-pAPN-N cells (1) were evaluated by quantitative RT-PCR. N protein is essential for viral transcription (34); therefore, as in some mutants N gene expression might be affected, N protein was provided in *trans* by using BHK-pAPN-N cells in all the experiments.

Mutations in CS-L flanking sequences affect the complementarity between the TRS-L and the cTRS-B. The extent of base pairing is the main determinant for transcriptional activity (29). To assess the potential influence of complementarity between the TRS-L mutant sequences and the complement of genes M, N, and 7 TRS-Bs on transcription, base-pairing energies were calculated (19). No significant differences were observed relative to the wild-type TRS-L sequence (Tables 3 and 4), indicating that the observed variations in transcriptional activity were most likely not due to changes in base-pairing extent between the TRS-L and the cTRS-Bs.

The self-replication activity of the IL-1, MS1, and MS3 mutant replicons was between 55 to 70% of the wild-type reference levels (Fig. 6), whereas sgmRNA levels in the same mutants were reduced to a greater extent (37 to 50% relative to the wild type). In the IL-2 and IL-3 mutants, gRNA levels were reduced to around 25% of wild-type RNA levels, whereas transcription levels were 1% or less than those of the wild type (Fig. 6). In contrast, mutant MS2 maintained both wild-type replication and transcription levels. These results indicated that the modification of TRS-L sequence and structure led to a reduction of both the replication and the transcription levels and that TRS-L mutations reduced transcription levels to a greater extent than expected from the reduction in replication rates. For mutants IL-1 and MS3, the differences between replication and transcription levels were at the limit of statistical significance (IL-1, $P = 0.064$; MS3, $P = 0.065$), while for mutants MS1, IL-2, and IL-3, reduction of transcription was significantly higher than that of replication (MS1, $P = 0.005$; IL-2, $P = 0.000001$; IL-3, $P = 0.0001$). In all mutants, the reduction in sgmRNA levels was similar for the M (Fig. 6), N, and 7 genes (Table 4), indicating that the TRS-L mutations varying the structure or thermal stability affected the transcription of all genes to the same extent, as could be expected.

Mutant MS1 stabilized the microstem below the CS-L, whereas mutant IL-1 extended this stem by one base pair (Fig. 2 and 4A). Both mutants moderately increased the thermal stability of the TRS-L hairpin (Table 3) and showed a significant reduction in replication and in viral sgmRNA levels (Fig.

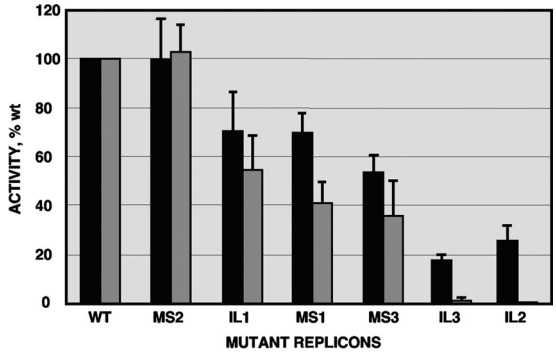


FIG. 6. Transcriptional activity and replication of TGEV TRS-L mutant replicons. Quantification of replication (black bars) and transcription (gray bars) in TRS-L mutants was by real-time RT-PCR. Transcription and replication levels of each mutant replicon are expressed in relation to those of the wild-type replicon. The data are the averages from four independent transfection experiments. Quantitative RT-PCR analysis was performed in duplicate in each case. Error bars represent the standard deviations. In each transfection experiment, qRT-PCR data were processed by comparing the levels of gRNA and sgmRNAs of the TRS-L mutants with those of gRNA and sgmRNAs in the wild-type replicon. This relative quantification was performed using the $2^{-\Delta\Delta CT}$ method, which compares cycle threshold values (16).

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

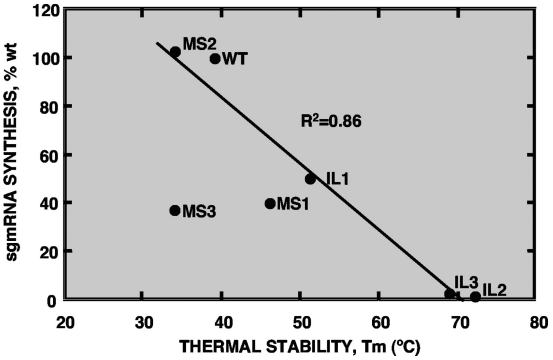


FIG. 7. Correlation between thermal stability and sgmRNA synthesis in TGEV TRS-L mutants. sgmRNA synthesis in mutant replicons was expressed in relation to that of the wild type, and thermal stability was determined from UV-monitored thermal denaturation experiments. The MS3 and IL-3 mutants were not included in the calculation of the correlation coefficient.

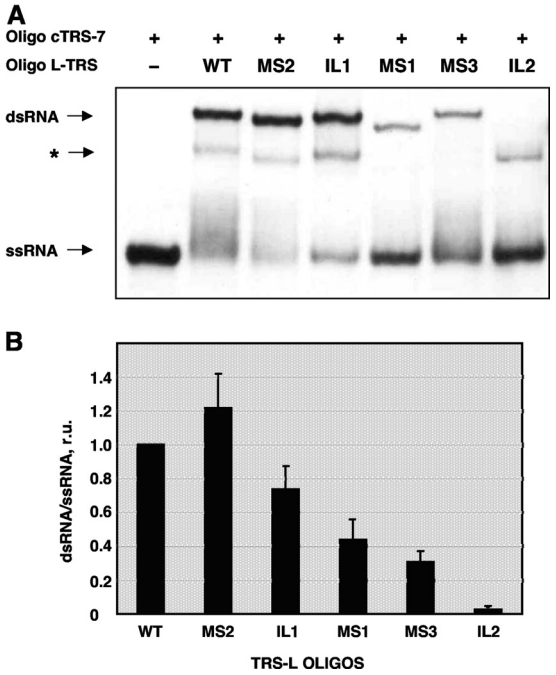


FIG. 8. *In vitro* hybridization assay. (A) Hybridization assay with a biotinylated cTRS-7 oligonucleotide and different unlabeled TRS-L oligonucleotides. Double-stranded RNA (dsRNA), RNA duplex; single-stranded RNA (ssRNA), cTRS-7. The band marked with the asterisk might correspond to a structure formed by the TRS-L sequence partially hybridized with cTRS-7 oligonucleotide. (B) The efficiency of hybridization between cTRS-7 and TRS-L was estimated as the ratio of dsRNA and ssRNA and expressed as a percentage of wild-type TRS-L efficiency. The dsRNA and ssRNA bands used for quantification are indicated by arrows in panel A. The amounts of dsRNA and ssRNA were determined by densitometry. The data are the averages from four independent hybridization experiments. Quantification by densitometry was performed in duplicate in each case. Error bars represent the standard deviations.

6). Mutant IL-2 almost completely replaced the internal loop located below the upper stem with a Watson-Crick base-paired stem (Fig. 2 and 4A) and exhibited the highest thermal stabilization relative to the wild type (Table 3). In this case replication was significantly reduced, and the synthesis of sgmRNA was almost completely abolished, as it was reduced to 0.2% of wild-type levels (Fig. 6).

The IL-3 mutant was designed to study the effect of disrupting the upper stem in the context of a structure with a partially closed internal loop. This disruption would place the CS-L and its flanking sequences in a more opened conformation, in principle favoring the base pairing between the TRS-L and the cTRS-B. The NMR data indicated that in IL-3 the Watson-Crick pairs of the upper stem were completely disrupted, while the internal loop was partially closed from the terminal side (Fig. 2 and 4A). The resulting thermal stability was significantly higher than that of the wild type (Table 3). IL-3 showed significantly reduced replication levels and M gene sgmRNA synthesis of 1% relative to the wild-type level (Fig. 6). Thus, disruption of the base pairs of the upper stem, leading to an opened conformation of CS-L and its flanking sequences, did not favor transcription in the context of a stable structure with a partially closed internal loop. This result indicated that a high thermal stability of the hairpin associated with an alteration of the TRS-L upper stem almost abolished transcription.

Mutants MS2 and MS3 were designed to assess the effect on sgmRNA synthesis of a destabilization of the upper stem below the CS loop without simultaneously closing the internal loop (Fig. 2). In principle, this would promote the CS-L and its flanking sequences to adopt a conformation more accessible for base paring with the cTRS-B in the nascent RNA (Fig. 2). The destabilization of the upper stem was confirmed by the NMR data for both mutants (Fig. 4B) and led to a reduction in thermal stability in relation to that of the wild-type hairpin (Table 3). Two-dimensional NMR data confirmed that the $MS2^S$ mutant hairpin retained a structure similar to that of the wild type. In contrast, the conformation of the apical region of the $MS3^S$ hairpin was different from that of the wild type. In agreement with these observations, RNA synthesis by mutant MS2 was similar to that of the wild-type replicon, while the transcriptional activity of the MS3 mutant was reduced to 35% of wild-type levels (Fig. 6), indicating that the TRS-L conformation in the MS3 replicon was less active in replication and transcription.

A good correlation between sgmRNA synthesis levels and thermal stability was observed for those TRS-L mutants that preserved the structure of the wild-type hairpin apical region (MS1, MS2, IL-1, and IL-2) ($r^2 = 0.86$) (Fig. 7). The mutations within IL-3 led to both structural changes and increased thermal stability. The MS3 mutant had a different apical region conformation, while exhibiting a thermal stability similar to that of MS2. The fact that the MS3 replicon had low transcriptional activity and fell out of the observed correlation (Fig. 7) indicated that in addition to optimal thermal stability, proper hairpin structure was required for efficient replication and transcription levels. A similar correlation between thermal stability and gRNA synthesis levels was confirmed (data not shown).

**TRS-L structure affects hybridization between TRS-L and cTRS-B.** The extent of the hybridization between a 16-nucleotide 5′-biotinylated RNA oligonucleotide complementary to the gene 7 TRS-B (cTRS-7) (5′-AUCUCGUUUAGUUCGU-3′) and the 34-nucleotide sequences representing wild-type

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

TRS-L and mutants IL-1, MS1, IL-2, MS2 and MS3 was determined (Fig. 8). The extent of hybridization was calculated as the ratio between the single-stranded (cTRS-7) and double-stranded RNA electrophoretic bands determined by densitometry. Mutant MS2, which forms a hairpin retaining the wild-type conformation, showed a hybridization efficiency similar to that of the wild type (Fig. 8A and B), whereas mutants MS1, IL-1, and IL-2, forming hairpins with thermal stabilities higher than that of the wild-type hairpin (Table 3), showed hybridization rates lower than that shown by the wild-type sequence (Fig. 8A and B). IL-2, forming the most stable hairpin (Table 3), showed the lowest hybridization rate, which probably was responsible for the reduced sgmRNA synthesis observed with this mutant (Fig. 6). Altogether, these results strongly suggested that the increase in stability of the TRS-L hairpin negatively affected its ability to hybridize with the complement of the TRS-Bs. As explained above, the MS3 mutant, with a thermal stability close to that of the wild type, adopted a TRS-L RNA conformation different from that of the wild type due to the disruption of the Watson-Crick pairs below the CS-L. This suboptimal structure resulted in a reduced hybridization with cTRS-7 relative to that of the wild type, leading to the observed decrease in sgmRNA levels (Fig. 6).

## DISCUSSION

To determine the relevance of TRS-L structure to TGEV replication and discontinuous transcription, structural and functional studies were performed on wild-type and mutant TRS-Ls. The NMR spectroscopy analyses indicated that the TRS-L of TGEV forms a hairpin structure capped by an apical 5′-AACUAAA-3′ heptaloop that comprises most of the CS-L. This heptaloop was separated from an adjacent CCAAC · AAUAUU internal loop by a short upper stem consisting of three Watson-Crick pairs (Fig. 2 and 3). Both the apical heptaloop and the internal loop formed well-defined stacked structures. Nucleotides C16 and U17 formed the backbone turn in the apical loop, whereas the rest of the bases of both loops were stacked with their neighbors (Fig. 3). Thermal denaturation experiments (Table 3) and broadening of imino resonances in the adjacent Watson-Crick stem (Fig. 4) showed that the presence of the internal loop significantly affected the thermal stability of the TRS-L hairpin.

Comparison of the structures, stabilities, and replication and transcription levels of wild-type and mutant TRS-Ls indicated that viral RNA synthesis was affected by the structure and stability of the TRS-L hairpin. For mutants that preserved the apical heptaloop structure (MS1, MS2, IL-1, and IL-2) (Fig. 2), a good correlation between thermal stability and the levels of sgmRNA synthesis was observed (Table 3 and Fig. 7). Mutants MS1 and IL-1, stabilizing or extending by one base pair the upper stem below the apical CS-L loop, moderately increased the thermal stability of the TRS-L hairpin without producing significant changes in structure (Fig. 2 and 4A). The sgmRNA levels of the MS1 and IL-1 mutants were reduced by around 50% compared to that of the wild type (Fig. 6 and 7), indicating that TRS-L thermal stability was critical for optimal RNA synthesis. In the IL-2 mutant, the internal loop below the CS-L was mostly replaced with a Watson-Crick paired stem (Fig. 2 and 4A). IL-2 gave rise to the highest thermal stabilization (Table 3) and almost completely abolished TGEV sgmRNA synthesis (Fig. 6 and 7). For this mutant, however, we cannot formally exclude the possibility that the lack of sgmRNA synthesis was also related to the elimination of the internal loop.

In mutant replicon IL-3, the internal loop was partially closed and the upper stem was completely disrupted (Fig. 2 and 4A). In addition, the thermal stability of the IL-3 TRS-L hairpin was increased (Table 3). In this mutant, replication was severely affected and transcription almost abolished.

The MS2 and MS3 constructs (Fig. 2) were designed to assess the effect of a destabilization of the upper stem below the CS-L loop on sgmRNA synthesis. A total or partial disruption of the upper stem would make the CS-L flanking sequences more accessible to the hybridization with the cTRS-B in the nascent RNA, favoring, in principle, the process of the template switch. The destabilization of the upper stem was confirmed by the NMR data for both mutants (Fig. 4B) and led to a reduction in thermal stability in relation to that of the wild-type hairpin (Table 3). The MS2 mutant, where only a terminal U11 · A23 pair in the upper stem was disrupted (Fig. 2), retained the structure of the wild-type hairpin and showed RNA synthesis levels similar to those of the wild type. In contrast, the apical region of the $MS3^S$ mutant hairpin adopted an alternative structure, different from that of wild type, and a significant reduction of RNA synthesis was observed for this mutant (Fig. 6). The results obtained with the MS3 mutant showed the relevance of the TRS-L hairpin structure, indicating that proper thermal stability was not sufficient for optimal RNA synthesis.

Altogether, the analysis of wild-type and mutant sequences indicated that the TRS-L needed to form a well-defined hairpin structure, with a specific optimal thermal stability, for efficient CoV replication and transcription. The upper stem and the internal loop probably contribute to provide a functional conformation to the apical CS-L heptaloop. This statement was supported by the reduction of sgmRNA levels observed when the upper stem was stabilized (MS1, IL-1, and IL-2 mutants) or disrupted (MS3 and IL-3) or when the internal loop was closed (IL-2 and IL-3). Conversely, mutant MS2, which mostly maintained all the structural elements and had a thermal stability similar to that of the wild type, showed unaffected RNA synthesis. In the IL-2 and IL-3 mutants the internal loop or the upper stem was not present and their thermal stability was highly increased, leading to severely reduced replication and transcription. These requirements could be expected according to a similarity-assisted RNA recombination model for the transcription template switch that is assisted by a structural motif (3, 7, 21, 23).

The reduction in sgmRNA synthesis was similar for sgmRNAs M, N, and 7 in all mutants studied (Tables 3 and 4), indicating that changes in TRS-L affected synthesis of all sgmRNAs to the same extent. This result is consistent with observations of arteriviruses, where mutations in the leader TRS significantly influenced the synthesis of all sgmRNAs (31, 32).

A hairpin RNA secondary structure was described for the leader transcription-regulating sequence of the equine arteritis virus (EAV) arterivirus (30). This leader hairpin included a loop, containing the CS-L, significantly larger than that of TGEV and without an internal loop as that demonstrated for

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

TGEV. Similar hairpin structures for leader TRSs were predicted for other arterivirus genomes, although no experimental support was provided (30). RNA secondary structure predictions for TRS-L regions in different coronaviruses also suggested the formation of hairpin structures similar to that demonstrated for TGEV in this report, but these were not validated with experimental data (30). For both arteriviruses and coronaviruses, the CS-L is exposed, at least in part, in the apical loop according to the predicted hairpin structure. Thus, it seems that the TRS-L hairpin structure is important for leader TRS function in coronaviruses and arteriviruses, which share similar replication and discontinuous transcription mechanisms.

The 5′-UTR is involved in replication (9, 14); therefore, it was expected that TRS-L mutations would also have an effect on TGEV replication, reducing the amount of gRNA available as a template for transcription and indirectly affecting sgmRNA levels. The results obtained supported this hypothesis. In line with our observations on the relevance of the TRS-L of TGEV in RNA synthesis, the lability of a stem-loop contained in the 5′-UTR of mouse hepatitis coronavirus has also been identified as an essential factor for replication (15, 28). The implication of the TRS-L structure in replication and transcription has also been observed in arteriviruses (30). In EAV, mutants potentially affecting TRS-L secondary structure showed a severely reduced replication competence, although the effects on replication in these studies were not quantified and the mutant TRS structures were not experimentally assessed. Altogether, these results suggested that the involvement of TRS structures in replication and transcription is a conserved feature of the *Nidovirales* order.

The wild-type hairpin, and some mutant hairpins containing a 3′-UU overhang, showed a tendency to form dimers that were most likely antiparallel hairpin dimers. This dimerization process was mostly abolished in the $WT^S$ and mutant hairpins lacking the 3′-UU overhang (Fig. 2 and 5). The hairpin heptaloop is likely involved in this dimerization process via head-to-tail interactions with the U-rich 3′-terminal end of the other hairpin monomer. This observation indicates that the AAC UAAA heptaloop is well suited to function as a landing platform for U-rich sequences. It is tempting to speculate that during discontinuous transcription, the template switch of the GUUUAG tract contained in the cTRS-Bs of the nascent minus RNA chain is facilitated by the specific orientation of adenine bases in the CS-L heptaloop active structure.

According to the current model of TGEV transcription (9, 28, 29), the TRS-L hybridizes with the complement of TRS-B during the process of template switching. In agreement with this hypothesis, band shift experiments indicated that both the structure and the stability of the TRS-L hairpin affected the efficiency of hybridization between TRS-L and a complement of TRS-7 (Fig. 8). On the other hand, the fact that the upper stem and AACUAAA apical heptaloop are required for transcription suggested that the three-dimensional structure of this region could also be involved in the template switch process. These observations provide experimental evidence for the selection of TRS-L during the template switch step, excluding the other genome TRS-Bs, albeit these TRSs also included the CS sequence. Only the CS-L, located in a sequence context leading to the optimal secondary structure and stability for template switching, may be used as a landing site for the newly synthesized minus-stranded RNA. The relevance of the extent of complementarity between TRS-L and the complement of the TRS-B in the nascent RNA in transcription was previously described (29, 35). In addition, in this report the relevance of TRS-L structure and stability in transcription efficiency has been shown. The replication and transcription results obtained with the mutant TRS-Ls in the replicon system would also be expected in the viral context, since the process of RNA synthesis is identical in both systems (20). Nevertheless, we consider that the regulation of transcription in the *Nidovirales* is a multifactorial process (9), in which most likely other factors contribute to the fine-tuning of CoV transcription, including TRS-L interaction with viral and host cell proteins.

In conclusion, CoV transcription requires a discontinuous step during the synthesis of the minus-stranded sgRNAs, and this template switch requires the hybridization of the TRS-L with the cTRS-B in the nascent RNA. This report shows that the TRS-L of TGEV forms a well-defined hairpin structure capped by an apical 5′-AACUAAA-3′ heptaloop that comprises most of the CS-L. In this hairpin, the majority of the bases are stacked within the helix rather than disordered. The structure and the stability of this TRS-L hairpin play an important role in replication and transcription.

## ACKNOWLEDGMENTS

This work was supported by grants from the Ministry of Science and Innovation of Spain (SAF2007-60243 to J.G., BIO2007-60978 to L.E., and PET2008-0310 to I.S. and L.E.), the Communities of Madrid (project S-SAL-0185-2006 to L.E.) and Valencia (project ACOMP 2009-075 to J.G.), the U.S. National Institutes of Health (ARRA-W000151845 to L.E.), and Pfizer Animal Health (to L.E.). The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under the projects "EMPERIE" (EC Grant Agreement 223498) and "PoRRSCon" (EC Grant Agreement 245141) to L.E. I.S., P.A.M.-G., and D.D. received contracts supported by grants from the Ministry of Science and Innovation of Spain (BIO2007-60978 to I.S. and BES-2008-001932 to P.A.M.-G.) and the CIPF (to D.D.).

We gratefully acknowledge Luís González, Lorena Pérez, and Margarita González for technical assistance.

## REFERENCES

1. **Almazan, F., C. Galan, and L. Enjuanes.** 2004. The nucleoprotein is required for efficient coronavirus genome replication. J. Virol. **78:**12683–12688.
2. **Almazan, F., et al.** 2000. Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. Proc. Natl. Acad. Sci. U. S. A. **97:**5516–5521.
3. **Brian, D. A., and W. J. M. Spaan.** 1997. Recombination and coronavirus defective interfering RNAs. Semin. Virol. **8:**101–111.
4. **Dalton, K., et al.** 2001. *cis*-acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. J. Virol. **75:**125–133.
5. **Delmas, B., J. Gelfi, H. Sjöström, O. Noren, and H. Laude.** 1993. Further characterization of aminopeptidase-N as a receptor for coronaviruses. Adv. Exp. Med. Biol. **342:**293–298.
6. **Dingley, A. J., and S. Grzesiek.** 1998. Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide 2JNN couplings. J. Am. Chem. Soc. **120:**8293–8297.
7. **Draghici, H. K., and M. Varrelmann.** 2010. Evidence for similarity-assisted recombination and predicted stem-loop structure determinant in potato virus X RNA recombination. J. Gen. Virol. **91:**552–562.
8. **Duszczyk, M. M., K. Zanier, and M. Sattler.** 2008. A NMR strategy to unambiguously distinguish nucleic acid hairpin and duplex conformations applied to a Xist RNA A-repeat. Nucleic Acids Res. **36:**7068–7077.
9. **Enjuanes, L., F. Almazan, I. Sola, and S. Zuniga.** 2006. Biochemical aspects of coronavirus replication and virus-host interaction. Annu. Rev. Microbiol. **60:**211–230.
10. **Enjuanes, L., et al.** 2008. The Nidovirales, p. 419–430. *In* B. W. J. Mahy, M.

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV

Van Regenmortel, P. Walker, and D. Majumder-Russell (ed.), Encyclopedia of virology, 3rd ed. Elsevier Ltd., Oxford, United Kingdom.
11. **Escors, D., A. Izeta, C. Capiscol, and L. Enjuanes.** 2003. Transmissible gastroenteritis coronavirus packaging signal is located at the 5′ end of the virus genome. J. Virol. **77:**7890–7902.
12. **Frolov, I., et al.** 1996. Alphavirus-based expression vectors: strategies and applications. Proc. Natl. Acad. Sci. U. S. A. **93:**11371–11377.
13. **Furtig, B., C. Richter, J. Wohnert, and H. Schwalbe.** 2003. NMR spectroscopy of RNA. Chembiochem **4:**936–962.
14. **Izeta, A., et al.** 1999. Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. J. Virol. **73:**1535–1545.
15. **Li, L., et al.** 2008. Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. J. Mol. Biol. **377:**790–803.
16. **Livak, K. J., and T. D. Schmittgen.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(−delta delta C(T)) method. Methods **25:**402–408.
17. **Luytjes, W., H. Gerritsma, and W. J. M. Spaan.** 1996. Replication of synthetic defective interfering RNAs derived from coronavirus mouse hepatitis virus-A59. Virology **216:**174–183.
18. **Masters, P. S.** 2006. The molecular biology of coronaviruses. Adv. Virus Res. **66:**193–292.
19. **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.
20. **Moreno, J. L., S. Zuniga, L. Enjuanes, and I. Sola.** 2008. Identification of a coronavirus transcription enhancer. J. Virol. **82:**3882–3893.
21. **Nagy, P. D., and A. E. Simon.** 1997. New insights into the mechanisms of RNA recombination. Virology **235:**1–9.
22. **Pasternak, A. O., W. J. Spaan, and E. J. Snijder.** 2006. Nidovirus transcription: how to make sense? J. Gen. Virol. **87:**1403–1421.
23. **Pasternak, A. O., E. van den Born, W. J. Spaan, and E. J. Snijder.** 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J. **20:**7220–7228.
24. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2003. The stability of the duplex between sense and antisense transcription-regulating sequences is a crucial factor in arterivirus subgenomic mRNA synthesis. J. Virol. **77:**1175–1183.
25. **Price, S. R., C. Oubridge, G. Varani, and K. Nagai.** 1998. Preparation of RNA-protein complexes for X-ray crystallography and NMR, p. 37–74. *In* C. Smith (ed.), RNA-protein interaction: practical approach. Oxford University Press, Oxford, United Kingdom.
26. **Sawicki, D. L., T. Wang, 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:**386–396.
27. **Sawicki, S. G., and D. L. Sawicki.** 1998. A new model for coronavirus transcription. Adv. Exp. Med. Biol. **440:**215–219.
28. **Sola, I., P. A. Mateos-Gomez, F. Almazan, S. Zuñiga, and L. Enjuanes.** 1 March 2011, posting date. RNA-RNA and RNA-protein interactions in coronavirus replication and transcription. RNA Biol. [Epub ahead of print.] doi:10.4161/hv.8.2.14991.
29. **Sola, I., J. L. Moreno, S. Zuñiga, S. Alonso, and L. Enjuanes.** 2005. Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. J. Virol. **79:**2506–2516.
30. **van den Born, E., A. P. Gultyaev, and E. J. Snijder.** 2004. Secondary structure and function of the 5′-proximal region of the equine arteritis virus RNA genome. RNA **10:**424–437.
31. **van den Born, E., C. C. Posthuma, A. P. Gultyaev, and E. J. Snijder.** 2005. Discontinuous subgenomic RNA synthesis in arteriviruses is guided by an RNA hairpin structure located in the genomic leader region. J. Virol. **79:**6312–6324.
32. **van Marle, G., et al.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. Proc. Natl. Acad. Sci. U. S. A. **96:**12056–12061.
33. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415.
34. **Zuñiga, S., et al.** 2010. Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription. J. Virol. **84:**2169–2175.
35. **Zuñiga, S., I. Sola, S. Alonso, and L. Enjuanes.** 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J. Virol. **78:**980–994.

Downloaded from http://jvi.asm.org/ on May 8, 2015 by ST ANDREWS UNIV