Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Virology 517 (2018) 44–55

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Structural and functional conservation of *cis*-acting RNA elements in coronavirus 5'-terminal genome regions

Ramakanth Madhugiri[a], Nadja Karl[a], Daniel Petersen[a], Kevin Lamkiewicz[b,d], Markus Fricke[b,d], Ulrike Wend[a], Robina Scheuer[a], Manja Marz[b,c,d], John Ziebuhr[a,d,*]

[a] Institute of Medical Virology, Justus Liebig University, Giessen, Germany
[b] Faculty of Mathematics and Computer Science, Friedrich Schiller University, Jena, Germany
[c] FLI Leibniz Institute for Age Research, Jena, Germany
[d] European Virus Bioinformatics Center, Jena, Germany

ARTICLE INFO

Keywords:
Coronavirus
Replication
cis-acting RNA element
Stem-loop
RNA structure
Coronavirus phylogeny

ABSTRACT

Structure predictions suggest a partial conservation of RNA structure elements in coronavirus terminal genome regions. Here, we determined the structures of stem-loops (SL) 1 and 2 of two alphacoronaviruses, human coronavirus (HCoV) 229E and NL63, by RNA structure probing and studied the functional relevance of these putative *cis*-acting elements. HCoV-229E SL1 and SL2 mutants generated by reverse genetics were used to study the effects on viral replication of single-nucleotide substitutions predicted to destabilize the SL1 and SL2 structures. The data provide conclusive evidence for the critical role of SL1 and SL2 in HCoV-229E replication and, in some cases, revealed parallels with previously characterized betacoronavirus SL1 and SL2 elements. Also, we were able to rescue viable HCoV-229E mutants carrying replacements of SL2 with equivalent betacoronavirus structural elements. The data obtained in this study reveal a remarkable degree of structural and functional conservation of 5′-terminal RNA structural elements across coronavirus genus boundaries.

1. Introduction

Cis-acting RNA elements play important roles in the life cycle of plus-strand (+) RNA viruses, including RNA replication, viral gene expression and genome packaging (Barton et al., 2001; Liu et al., 2009b; Firth and Brierley, 2012; Goto et al., 2013; Kuo and Masters, 2013; Morales et al., 2013; Nicholson and White, 2014; Keane et al., 2015). Compared to many other +RNA viruses, information on *cis*-acting RNA elements of coronaviruses, including their specific functions, structures and interactions, remains limited. Particularly, this applies to viruses from genera outside the genus *Betacoronavirus* (for reviews, see Brian and Baric, 2005; Masters, 2007; Liu and Leibowitz, 2010; Madhugiri et al., 2014; Yang and Leibowitz, 2015; Madhugiri et al., 2016). Historically, RNA structures and sequences required for (beta)coronavirus RNA synthesis were characterized using defective interfering (DI) RNA-based systems (Chang et al., 1994, 1996; Raman et al., 2003; Raman and Brian, 2005; Brown et al., 2007; Gustin et al., 2009). Thus, for example, RNA structure probing studies of mouse hepatitis virus (MHV) and bovine coronavirus (BCoV)-derived RNAs led to the identification of up to four stem-loops within the 5′-terminal 215 nt of the genome (for recent reviews, see Liu and Leibowitz, 2010; Madhugiri et al., 2014, 2016; Yang and Leibowitz, 2015). In many cases, potential functional roles of RNA structural elements present in the 5'-terminal genome region could be confirmed by mutational analyses. More recently, genus- and subfamily-wide RNA structure-based alignments using all currently approved coronavirus species in the respective genera of the *Coronavirinae* were performed for this highly divergent genome region. The studies led to a model of three highly conserved stem-loop structures, called SL1, SL2, and SL4, in the 5′-terminal, ~150-nt genome region (Kang et al., 2006; Liu et al., 2007; Chen and Olsthoorn, 2010; Madhugiri et al., 2014). Furthermore, nuclear magnetic resonance (NMR) spectroscopy provided structural support for SL1 and SL2 in three betacoronaviruses, MHV, BCoV, and HCoV-OC43 (Liu et al., 2007, 2009a; Li et al., 2008). Also, a selective 2′-hydroxyl acylation and primer extension (SHAPE) analysis *in virio* and *ex virio* confirmed the predicted SL1, SL2, and SL4 structures for MHV-A59 (Yang et al., 2015).

Possible biological functions of betacoronavirus 5′-terminal SL1 and SL2 structures in viral replication could be substantiated by reverse genetics studies (Kang et al., 2006; Liu et al., 2007, 2009a; Li et al., 2008). For example, an MHV study revealed that destabilization of the upper part of SL1 produces viruses with replication defects, while

* Correspondence to: Institute of Medical Virology, Biomedical Research Center, Justus Liebig University, Schubertstr. 81, 35392 Giessen, Germany.
E-mail address: john.ziebuhr@viro.med.uni-giessen.de (J. Ziebuhr).

https://doi.org/10.1016/j.virol.2017.11.025
Received 1 October 2017; Received in revised form 28 November 2017; Accepted 29 November 2017
Available online 06 December 2017
0042-6822/

compensatory mutations restoring these base-pairing interactions led to viruses with near-wildtype growth kinetics (Li et al., 2008). In contrast, disruption of the basal part of SL1 was largely tolerated, while compensatory mutations that restored these base-pairing interactions proved to be lethal, suggesting a critical role for the RNA sequence (rather than structure) in this lower part of SL1. Based on these and other data, SL1 was suggested to require an optimal stability suitable to establish transient long-range (RNA- and/or protein-mediated) interactions between the 5′- and 3′-UTRs that may be required for genome replication and subgenomic (sg) mRNA synthesis. Other reverse genetics studies confirmed that the 5′-terminal SL2 is also required for MHV RNA synthesis (Liu et al., 2007, 2009a). Based on phylogenetic analyses, the SL2 was proposed to be the most conserved RNA secondary structure in coronaviruses (Kang et al., 2006; Liu et al., 2007; Chen and Olsthoorn, 2010). It is composed of a 5-bp stem and a conserved loop sequence, 5′-CUUGY-3′, that was shown to adopt a 5′-uCUYG(U)a-3'- or a 5'-uYNMG(U)a-3′-like tetraloop structure (Liu et al., 2009a).

To extend these studies and corroborate predictions on alphacoronavirus-associated 5'-terminal RNA structural elements, we used a combination of bioinformatics, biochemical and reverse genetics approaches, focusing on structures and functions of the 5′-terminal SL1 and SL2 structures in the HCoV-229E genome (genus *Alphacoronavirus*). The data obtained in this study provide evidence for the existence of two SL structures (SL1 and SL2) in the ~80-nt, 5'-terminal HCoV-229E and HCoV-NL63 genome regions. The structures were found to be required for viral replication and appear to be (largely) conserved between alpha- and betacoronaviruses. Thus, for example, we were able to show that the HCoV-229E SL2 structure can be replaced with that of the betacoronaviruses BCoV and SARS-CoV, respectively, providing experimental support for our previous hypothesis that (some) RNA structural elements in coronavirus untranslated genome regions may be more conserved than previously thought, even across genus boundaries (Madhugiri et al., 2014).

2. Material and methods

2.1. Cells and viruses

Wildtype HCoV-229E and HCoV-229E mutants were propagated in Huh-7 cells. HCoV-229E titers were determined by plaque assay using Huh-7 cells. Recombinant vaccinia viruses were propagated in CV-1 and BHK-21 cells, and plaque purifications of single virus clones were performed using CV-1 and D980R cells as described previously (Isaacs et al., 1990; Thiel et al., 2001).

2.2. Mutagenesis of the HCoV-229E full-length cDNA clone

HCoV-229E mutants (HCoV-229E_C11G, _C16G, _G45C, _C47G, _C11G-G34C, _C16G-G29C, _G45C-C55G, and _C47-G53C) were generated using the recombinant vaccinia virus vHCoV-inf-1, which contains a full-length HCoV-229E cDNA (GenBank accession number NC_002645). Site-directed mutagenesis of the HCoV-229E cDNA insert in vHCoV-inf-1 was done using previously described methods (Thiel et al., 2001). To construct vHCoV-inf-1 derivatives containing nucleotide substitutions in the HCoV-229E 5′-UTR, we used the plasmid pBS-5′GPT for recombination with vaccinia virus vHCoV-inf-1. This pBluescriptII-derived plasmid was constructed to contain the *E. coli gpt* gene flanked by (i) a 500-bp fragment representing the vaccinia DNA sequence located upstream of the HCoV-229E cDNA insert in vHCoV-inf-1 and (ii) a 500-bp fragment representing the cDNA sequence of nts 1001–1500 of the HCoV-229E genome RNA. Next, a *gpt*-positive vHCoV-inf-1 derivative, called vRec-5′GPT, was selected from CV-1 cells infected with vHCoV-inf-1 and transfected with pBS-5′ plasmid DNA. In a second selection step, D980R cells were infected with vRec-5′GPT and transfected with an appropriate pBS-5′UTR-mut plasmid DNA. Using appropriate selection conditions (Hertzig et al., 2004), *gpt*-negative vHCoV-inf-1 derivatives (called vHCoV_5'UTR-mut) that contained the desired mutation(s) in the 5' UTR cDNA sequence were isolated. The pBS-5′UTR-mut plasmid constructs used to produce the recombinant vHCoV_5'UTR-mut vaccinia viruses contained the 500-bp vaccinia virus sequence described above followed by a cDNA copy of HCoV-229E nts 1–1500, with appropriate mutations being introduced by PCR-based mutagenesis. Sequences of vHCoV_5'UTR-mut vaccinia virus constructs were verified by Southern blotting and sequence analysis as described (Thiel et al., 2001). The vHCoV-inf-1 derivatives generated in this study were called vHCoV_5′UTR-C11G, vHCoV_5′UTR-C16G, vHCoV_5′UTR-G45C, vHCoV_5′UTR-C47G, vHCoV_5′UTR-C11G + G34C, vHCoV_5′UTR-C16G + G29C, vHCoV_5′UTR-G45C + C55G, and vHCoV_5′UTR-C47G + G53C. Genome-length HCoV-229E RNAs were prepared by T7-based *in vitro* transcription (RiboMAX Large Scale RNA Production System, Promega) using purified genomic DNA from vHCoV-inf-1 and its mutant derivatives, respectively. 1.25 μg of *in vitro*-transcribed genome-length RNAs and 0.75 μg of *in vitro*-transcribed HCoV-229E nucleocapsid (N) protein mRNA (Schelle et al., 2005; Almazan et al., 2006) were used to transfect 1×10^6 Huh-7 cells using the TransIT® mRNA transfection kit according to the manufacturer's instructions (Mirus Bio LLC). At 72 h posttransfection (p.t.), cell culture supernatants were collected to determine viral titers, and total RNA was isolated for subsequent Northern blot and genome sequence analyses.

2.3. RNA extraction and Northern blot analysis

At 72 h p.t., intracellular RNA was extracted using TRizol reagent (Invitrogen) according to the manufacturer's instructions. To analyze viral RNAs by Northern blot hybridization, 10 μg total RNA was denatured for 10 min at 65 °C in loading buffer (50% deionized formamide, 18% formaldehyde, 1x MOPS) and separated in a 1% (w/v) agarose and 2.2 M formaldehyde-containing, 1x MOPS-buffered gel at 16 V for 16–17 h. The gel was soaked in buffer A (50 mM NaOH, 150 mM NaCl) for 30 min and then in buffer B (100 mM Tris-HCl / pH 7.5, 150 mM NaCl) for 30 min. Next, the RNA was transferred onto a positively charged nylon membrane by vacuum blotting. The RNA was cross-linked to the membrane and hybridized with an [α-^{32}P]dCTP-labeled DNA probe specific for HCoV-229E nucleotides 26857–27277 and the negative-strand complement of this sequence (TaKaRa Bio Inc). Following hybridization, membranes were rinsed 2 times with 2x SSC/0.01% (w/v) SDS at room temperature and 2 times with 0.2x SSC/0.01% (w/v) SDS at 55 °C for 30 mins. Hybridization signals were visualized by autoradiography using a Typhoon 9200 imager (GE Healthcare).

2.4. Genome sequence analysis of virus progeny

At 72 h p.t., cell culture supernatants were collected (passage zero [p0]) and used to determine virus titers and plaque sizes (see below). From the cell pellet, total RNA was extracted using TRIzol reagent (Invitrogen). Following reverse transcription (RT)-PCR amplification, the 5' and 3'-terminal HCoV-229E genome regions (nts 25–750 and nts 25323–27317, respectively) were sequenced. The following primer pairs were used to produce two amplicons for subsequent sequence analyses: (1) HCoV-229E-25up (5′-ACTTAAGTACCTTATCTATCTACAG-3′) and HCoV-229E-750dn (5′-GAAATTATCATCAATGGTCATACTTAC-3′) and (2) HCoV-229E-25323up (5′-CATGGAATCCTGAGGTTAATGCAATC-3′) and HCoV-229E-oligo(dT) (5′-TTTTTTTTTTGTGTATCCATATCG-3′). To determine the 5′-terminal nts 1–25 of progeny virus genomes, the FirstChoice™ RLM-RACE kit was used according to the manufacturer's instructions (Invitrogen). PCR products used for sequence analyses were gel purified (innuPREP Gel Extraction Kit, Analytik Jena) and subjected to automated Sanger sequencing (LGC Genomics). The 5′-UTR and 3'-UTR amplicons were sequenced using oligonucleotides HCoV-229E-750dn and HCoV-229E-27317dn,

respectively.

2.5. Virus titration

Virus titers in the supernatants of cells transfected with the appropriate full-length HCoV-229E RNA (see above) were determined as follows. Nearly confluent monolayers of Huh-7 cells that were grown in 96-well plates using DMEM (supplemented with 10% fetal bovine serum [FBS] and antibiotics) were inoculated with 100 µl per well of a serial logarithmic dilution of cell culture supernatants obtained from transfected cells. Following incubation for 5–6 d at 33 °C, titers of infectious virus progeny (given as 50% tissue culture infectious dose [$TCID_{50}$] per ml) were determined using the method described by Read and Muench (Reed and Muench, 1938).

Virus plaque assays were performed using confluent Huh-7 cells that were grown in 6-well plates. Cell monolayers were inoculated with a 10-fold serial dilution of virus-containing culture supernatants. At 1 h p.i., the inoculum was removed. Cells were washed with PBS and overlaid with 2 ml MEM containing 10% FBS, 1.25% Avicell (Sigma), and antibiotics. At 4 days p.i., the medium was removed and cell monolayers were stained with 0.1% crystal violet solution to visualize virus plaques.

3. *In vitro* RNA structure probing

3.1. In vitro transcription using T7 RNA polymerase

Typically, 1 µg of PCR product representing the 5'-terminal ~100 nts of the HCoV-229E and HCoV-NL63 genome, respectively, was used as template in *in vitro* transcription reactions. The reactions were performed using the T7 RiboMAX™ express large scale RNA production system (Promega) according to the manufacturer's instructions. DNA templates were digested using 1U of RNase-free RQ1 DNase (Promega). Free nucleotides were removed using G25 microspin columns (GE healthcare). The RNA was purified by phenol-chloroform-isoamyl alcohol (Roth) extraction and precipitated.

3.2. RNA structure probing

RNA structure probing experiments were done as described previously (Ehresmann et al., 1987; Luo et al., 1998) with minor modifications. Typically, 0.6 µg of *in vitro*-synthesized RNA was heat-denatured at 90 °C for 1 min and then cooled on ice for 5 min. The RNA was renatured in AN buffer (50 mM sodium cacodylate, pH 7.5, 5 mM $MgCl_2$, 60 mM KCl) for 20 min at room temperature. Next, the samples (total volume of 8 µl) were mixed with 1 µl of yeast tRNA (2 mg/ml, Ambion) and 1 µl of dimethyl sulfate (DMS, Aldrich, #D186309) solution (diluted to 1/2, 1/5, 1/10 and 1/20, respectively, in 20% ethanol). Control reactions were done under equal conditions in the absence of DMS. Following incubation for 5 min at room temperature, the reactions were terminated by ethanol precipitation in the presence of 1/10 vol 3 M sodium acetate (pH 5.2). DMS modifications of specific nucleotides were determined by primer extension analysis.

3.3. Primer extension assay

To analyze DMS modifications, reverse transcription reactions were performed using one of the following oligonucleotides: 229E-1 (5′-CGACTCCAGCATCAAAGATGC-3′; complementary to nts 93–113 of the HCoV-229E 5'-UTR) and NL63-1 (5′-CGAAATTTCAATTACACTAG GAC-3′; complementary to nts 107–129 of the HCoV-NL63 5'-UTR). Aliquots of chemically modified RNAs (3 pmol) were hybridized with 1–3 pmol of 5′-end labeled primer (1–2 × 10^5 dpm). Following a brief heating step (90 °C, 2 min), the reaction was cooled slowly (5 min at 75 °C, 10 min at 50 °C, 5 min at 37 °C, 10 min at room temperature). Next, the primer annealing mixture was used to set up a 20-µl reverse transcription reaction in 1 × SuperScript® III RTase reaction buffer supplemented with 170 units of SuperScript® III RTase (Invitrogen), 20 units RNaseOUT (Invitrogen), and 1 mM of each dNTP. The reaction was performed at 42 °C for 50 min and then at 55 °C for 60 min. Reactions were terminated by the addition of 1/10 vol of 3 M sodium acetate, pH 5.2, and 10 volumes of ice-cold ethanol. Following centrifugation, the pellets were washed with 70% ethanol. The dried pellets were resuspended in water and treated with DNase-free RNase A for 20 min at 37 °C (0.2 mg/ml, Invitrogen). Next, PCR-grade Proteinase K (Invitrogen) was added to a final concentration of 1 mg/ml and the reaction was incubated for another 15 min at 55 °C. Reactions were stopped by adding Fu-mix (6 M urea, 80% deionized formamide, 1x TBE, 0.1% (w/v) Bromophenol blue, and 0.1% (w/v) Xylene cyanol). Reaction products were separated in TBE-buffered 8% polyacrylamide gels containing 7 M urea. Signals were visualized using a Typhoon 9200 imager (GE Healthcare) and analyzed using Quantity One software (BioRad).

3.4. Bioinformatic analyses

RNA secondary structures were calculated using RNAfold, version 2.4.1 (Lorenz et al., 2011). To calculate base-pairing probabilities, parameters were set to –noLP and -p. RNA secondary structures were visualized using VARNA (version 3.93) (Darty et al., 2009). The color codes used in Figs. 1, 2, 5B, and 8 indicate base-pairing probabilities derived from dot plots generated by RNAfold (see Suppl. Figures 1 and 2). Structure-based alignments were calculated with LocARNA, version 1.8.11 (Will et al., 2012). Consensus secondary structures were calculated with RNAalifold –noLP –color -r -p (version 2.4.1) (Lorenz et al., 2011). Herein, the color code represents the numbers of different base-pairing types and numbers of incompatible bases, respectively. Sequence conversation was visualized using WebLogo 3.5.0 (Fig. 5C) (Crooks et al., 2004).

4. Results

4.1. RNA structure probing analysis of alphacoronavirus 5'-terminal genome regions

To provide experimental support for our RNA structure model of alphacoronavirus 5'-terminal genome regions (Madhugiri et al., 2014) (Suppl. Figure 1), we performed a series of RNA structure probing experiments. RNA transcripts representing the 5′-terminal ~100 nt of the HCoV-229E and HCoV-NL63 genome, respectively, were produced *in vitro* and treated with the methylating agent DMS (Ehresmann et al., 1987). N1 methylation of unpaired adenosines and N3 methylation of unpaired cytidines was identified by primer extension analysis using reverse transcriptase. 5'-[^{32}P]-labeled products obtained in these reactions were separated in denaturing polyacrylamide gels and visualized by phosphorimaging (see Materials and Methods). RNA that was not treated with DMS was included as a control to detect potential non-specific termination products of the reverse transcriptase reaction. Autoradiograms that are representative of an extensive set of DMS structure probing experiments and a summary of the structure probing information obtained in these experiments are shown in Fig. 1. Our RNA secondary structure predictions (Madhugiri et al., 2014) (Suppl. Figure 1) and the *in vitro* DMS structure probing data obtained in this study (Fig. 1) lead us to propose a model in which the 5′-terminal 80-nt regions of the HCoV-229E and NL63 genome RNAs fold into two conserved stem-loops (SL), called SL1 and SL2, while the adjacent 3' region containing the leader-associated transcription regulatory sequence (TRS-L) (Zuniga et al., 2004) does not adopt a stable structure. The SL1 structures of HCoV-229E and NL63 appear to be fairly stable as none of the principal nucleotides forming the predicted SL1 stem structure were accessible to DMS modification, while several nucleotides predicted to be part of bulge or loop regions were modified by DMS (Fig. 1B and D).

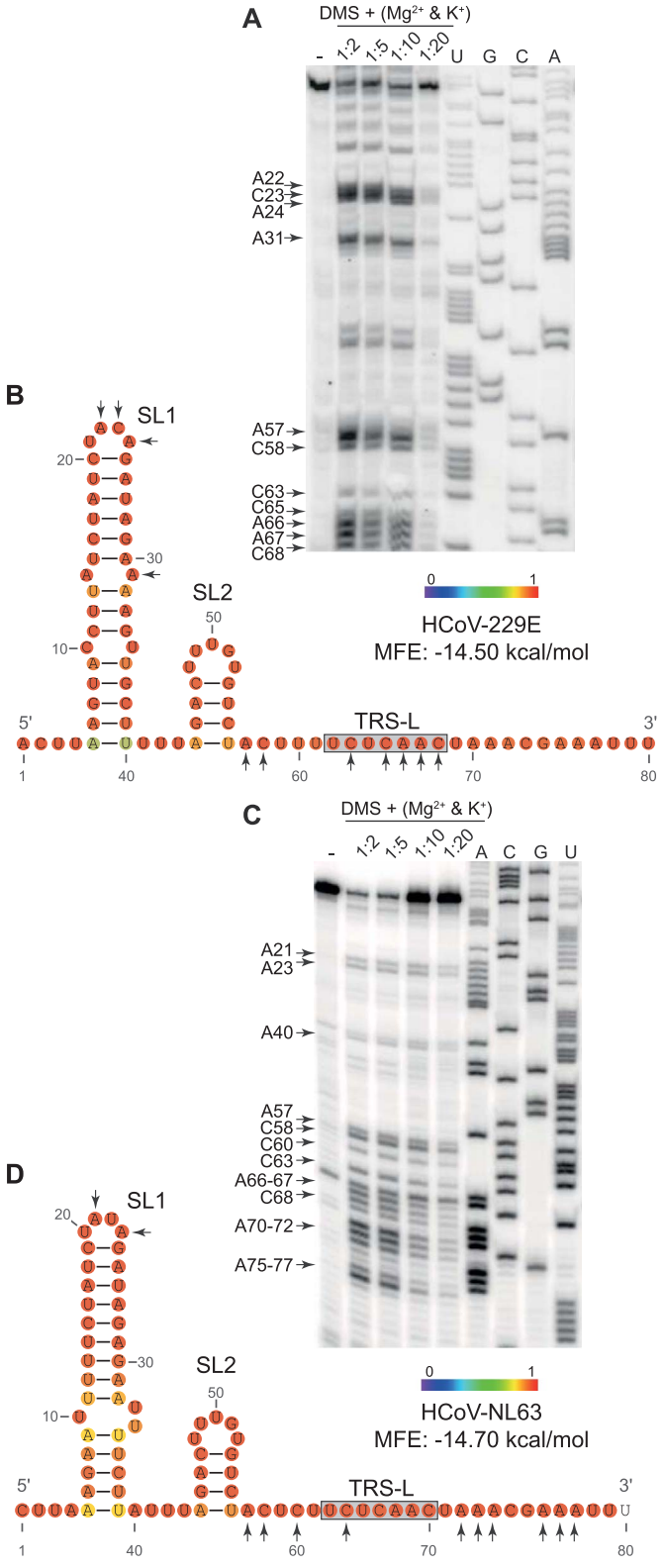

Fig. 1. RNA structure probing analysis of 5′-terminal 80-nt genome regions of HCoV-229E and HCoV-NL63. A and C, Chemical probing of *in vitro*-transcribed RNAs representing the 5′-terminal 80 nts of the genome RNAs of HCoV-229E and HCoV-NL63. RNAs were modified with DMS in the presence of Mg^{2+} and K^{+} ions (for details, see Material and Methods). DMS modifications of unpaired adenosines and cytidines were identified by primer extension analysis. Shown are the autoradiograms of representative denaturing polyacrylamide gels. Lanes: –, reaction performed in the absence of DMS; 1:2, 1:5, 1:10 and 1:20, DMS dilutions used in the respective reactions; T, G, C, A, sequencing reactions using the indicated dideoxynucleotide. B and D, RNA secondary structure models of the 5′-terminal 80-nt genome regions of HCoV-229E and HCoV-NL63. SL1 and SL2 structures are indicated and the TRS-L sequence is highlighted as a gray box. Black arrowheads indicate positions of DMS modifications that were reproducible in repeated experiments. RNA structures were predicted by RNAfold. The color code indicates base-pairing probabilities calculated with RNAfold.

Previous analyses suggested that (beta)coronavirus 5'-SL2 elements are made up of a 5-bp stem and a pentaloop (Kang et al., 2006). The DMS structure probing data presented in this study suggest that the SL2 elements of both HCoV-229E and HCoV-NL63 are composed of a 4-bp (rather than 5-bp) stem and a pentaloop sequence (Fig. 1). Thus, for both viruses, A57 (positioned at the base of SL2) was regularly found to be modified by DMS while none of the other adenosine and cytidine residues predicted to be part of the SL2 structures were found to be accessible to DMS modification. Base-pairing probabilities calculated with RNAfold further indicate that the basal part of HCoV-229E SL2 may have a certain degree of flexibility (Figs. 1 and 2).

Furthermore, the DMS structure probing data obtained for HCoV-NL63 (and, to a slightly lesser extent, HCoV-229E) suggest that the TRS-L element located downstream of SL2 is part of an unstructured region. As shown in Fig. 1, nucleotides of the TRS-L core sequence and nucleotides adjacent to the TRS region were accessible to DMS modification, confirming that they are part of single-stranded regions. In conclusion, our structure model is consistent with previous betacoronavirus studies (Kang et al., 2006; Liu et al., 2007; Yang et al., 2015) and supports the idea that, in most coronaviruses, TRS-L is part of an unstructured region rather than a stable SL structure (Van Den Born et al., 2004; Dufour et al., 2011) (see also Discussion).

4.2. Disruption of SL1 and SL2 by single-nucleotide substitutions causes major defects in HCoV-229E RNA synthesis and virus reproduction

Having established the existence of two conserved RNA structural elements in the 5′-leader regions of HCoV-229E and HCoV-NL63 (this study and Madhugiri et al., 2014), we sought to investigate the functional significance of these elements in alphacoronavirus RNA synthesis using a reverse genetics system developed for HCoV-229E (Thiel et al., 2001). To this end, HCoV-229E genome-length RNAs containing appropriate nucleotide substitutions in SL1 and SL2, respectively, were generated and used to investigate possible effects on viral replication in cell culture (for details, see Material and Methods). Using RNAfold (Lorenz et al., 2011), we predicted the most probable structures for RNAs containing specific nucleotide substitutions in the HCoV-229E 5'-terminal genome region and, based on these predictions, designed a set of mutations to be introduced in the HCoV-229E genome RNA for subsequent cell culture studies. The first set of mutants (HCoV-229E_C11G, _C16G, _G45C, and _C47G) contained single-nucleotide substitutions predicted to disrupt specific base-pair interactions in SL1 or SL2, resulting in a destabilization (or restructuring) of the respective secondary structures. Another set of mutants (HCoV-229E_C11G + G34C, _C16G + G29C, _G45C + C55G, and _C47G + G53C) contained a second, compensatory mutation that restored the respective base-pairing interaction and, thus, preserved the stability of the stem structure (Fig. 2, panels C, E, G, and I).

For the C11G mutation, computer-assisted RNA structure analyses predicted a partial destabilization, resulting in a larger bulge in the middle of SL1 and, accordingly, a reduction of the calculated minimal free energy (Fig. 2A and B). This structural change is also reflected by reduced base-pair probabilities in SL1 (compare color codes in Fig. 2A and B, Suppl. Figure 2B). Restoration of the base-pair interaction between nt 11 and 34 in the HCoV-229E_C11G + G34C mutant was predicted to preserve the wildtype RNA structure (Fig. 2A and C). For the C16G mutation, the calculation of base-pair probabilities (Fig. 2D and Suppl. Figure 2) predicted a profound destabilization of the upper

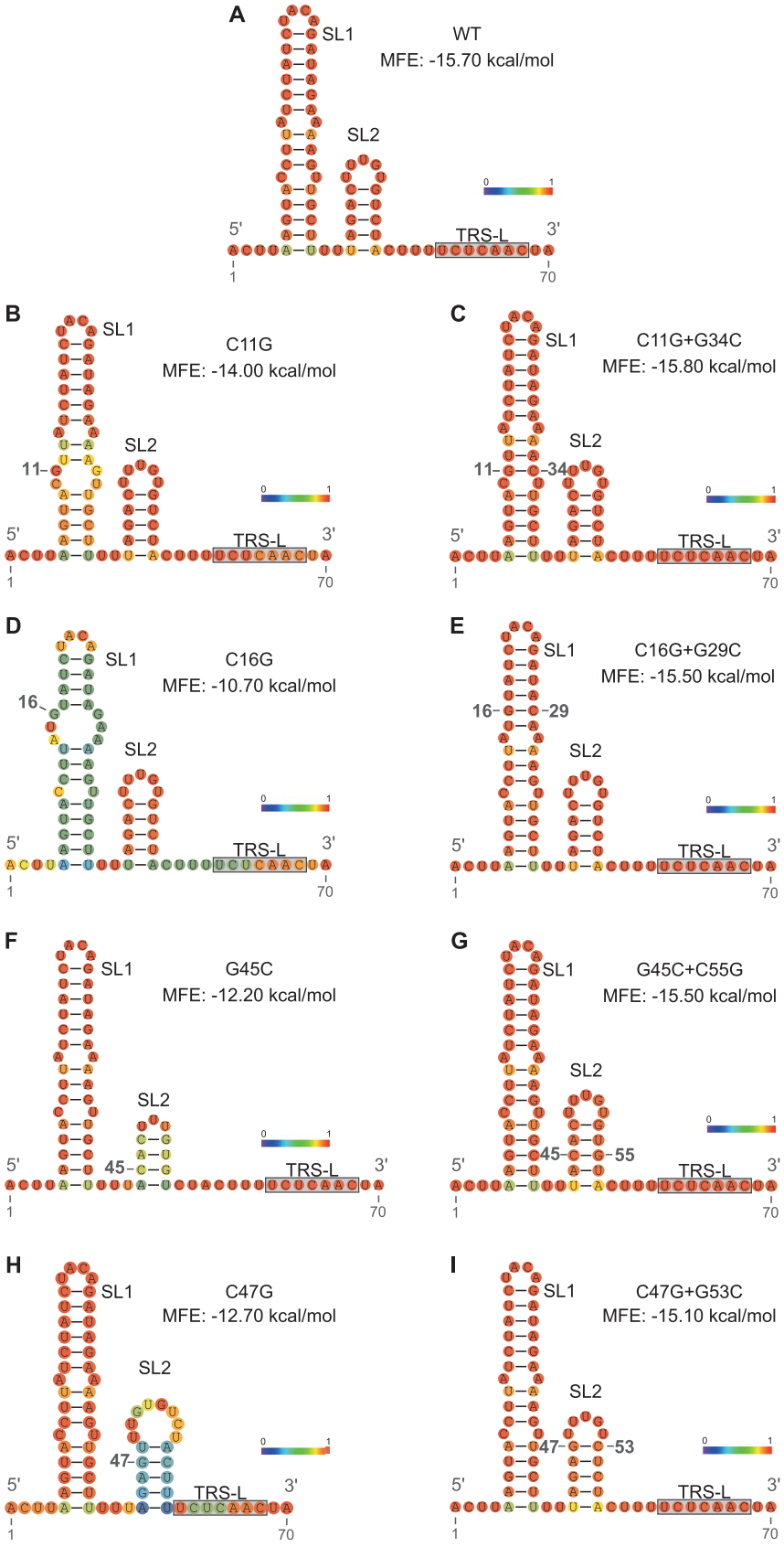

Fig. 2. RNA secondary structure predictions for HCoV-229E 5′-terminal genome sequences containing nucleotide substitutions. For each mutant, positions of substituted nucleotide(s) in the HCoV-229E genome are indicated (with numbers in boldface). SL1 and SL2 structures are indicated and the TRS-L sequence is highlighted as a gray box. Nucleotides are numbered according to the wildtype sequence, starting from the first nucleotide at the 5′-end of the genome. RNA secondary structures were predicted by RNAfold. The color code represents the base-pairing probability of the individual bases predicted by RNAfold. Dot plots for each structure are provided in the supplement.

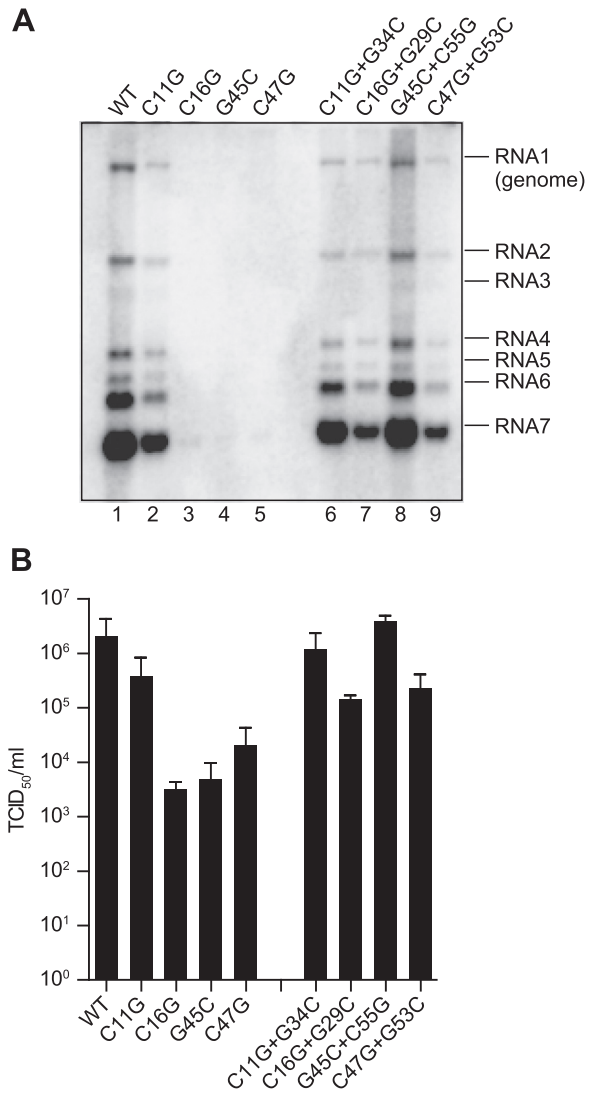

Fig. 3. Analysis of viral RNA accumulation and virus titers of HCoV-229E SL1 and SL2 mutants. Recombinant HCoV-229E (WT) and HCoV-229E SL1 and SL2 mutants were generated by co-transfecting Huh-7 cells with the appropriate (mutant or wildtype) *in vitro*-transcribed genome-length HCoV-229E RNA and HCoV-229E N mRNA as described in Material and Methods. At 72 h p.t, the virus titer in the cell culture supernatant was determined and viral RNA was analyzed by Northern blotting. A, Northern blot analysis of HCoV-229E-specific RNAs produced in cells transfected with the indicated HCoV-229E full-length RNAs. Virus-specific RNAs were detected using a [^{32}P]-labeled DNA probe specific for the HCoV-229E 3′-UTR (nts 26857–27277). B, Virus titers of HCoV-229E wildtype (WT) and mutants were determined by end-point dilution using Huh-7 cells. Virus titers (means ± SEM) are represented as $TCID_{50}$/ml and were determined from three independent transfection experiments.

segment of SL1 but also changes in adjacent regions (Fig. 2D and Suppl. Figure 2). Again, introduction of an additional compensatory mutation (G29C) was predicted to preserve the wildtype structure (Fig. 2E). For the G45C mutation, drastic structural changes were predicted for SL2, including a significantly less stable (4-bp) stem structure and a smaller loop size (3 instead of 5 nts) (Fig. 2F). Even more profound effects were predicted for the C47G replacement, resulting in a complete destabilization of the SL2 stem structure (Fig. 2H, Suppl. Figure 2). Taken together, these RNA structure predictions suggested that the single-nucleotide and two-nucleotide substitutions introduced in the HCoV-229E genome RNA were suitable to study possible roles of the HCoV-229E SL1 and SL2 structures in viral replication. The predictions also confirmed that the second-site (compensatory) substitutions were suitable to preserve (near-) wildtype structures for both SL1 and SL2.

To study the effects of the structural changes caused by the

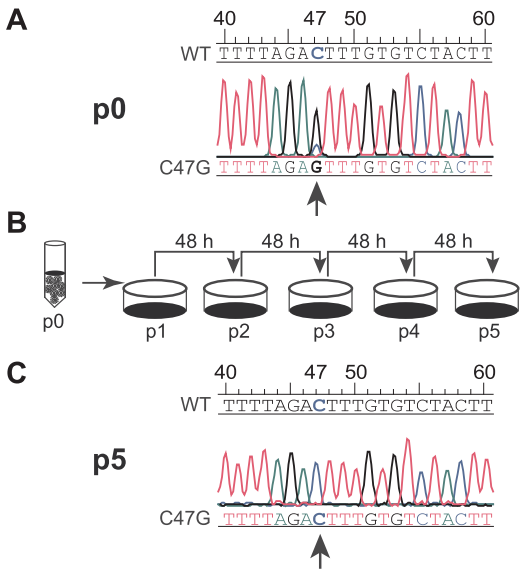

Fig. 4. Sequence analysis of the SL2_C47G mutant. A and C, Sequence analysis of RT-PCR products obtained from Huh-7 cells transfected with full-length HCoV-229E_C47G RNA at passage 0 (at 72 h p.t) and after serial passaging of the recombinant virus (passage 5). The position of the nucleotide substitution and reversion is indicated by an arrow. B, Serial passaging of the HCoV-229E_C47G mutant.

nucleotide substitutions, *in vitro* transcribed full-length HCoV-229E RNAs (wildtype and mutants, respectively) and N mRNA were co-transfected into 90% confluent Huh-7 cells. At 72 h p.t., cell culture supernatants were collected to determine virus titers and intracellular RNA was isolated for Northern blot analysis of viral RNA replication. Using a [^{32}P]-labeled probe specific for the 3′ end of the genome, we analyzed the full set of 3'-coterminal genomic and subgenomic HCoV-229E RNAs. We found that, except for mutant C11G (see below), the SL1 and SL2 single-nucleotide mutants displayed severe defects in viral RNA accumulation, suggesting that the structural integrity of SL1 and SL2 is essential for viral replication. In the case of C11G, only minor defects in viral RNA accumulation were observed, suggesting that the stability of the basal part of SL1 is less critical, with some structural flexibility being tolerated in this case. However, our observation that the double mutant (C11G + G34C) with a fully preserved SL1 structure replicates more efficiently than the C11G mutant shows that an intact basal part of SL1 is beneficial (though not essential) for virus replication (Fig. 2C, Fig. 3A, lane 6). Interestingly, similar observations were also reported for murine hepatitis virus (MHV). If the lower part of the (presumably equivalent) MHV SL1 structure was disrupted, infectious virus progeny could still be recovered, while disruption of the upper part proved to be lethal (Li et al., 2008). For the HCoV-229E_C16G mutant, a major replication defect was observed which could be reversed (albeit not completely) by restoring the base-pair interaction between nts 16 and 29 in SL1 (C16G + G29C) (Fig. 3A, lane 7). For the G45C and C47G mutations in SL2, we found that both mutations cause major defects in RNA replication (Fig. 3A, lanes 4 and 5). The corresponding double mutant HCoV-229E_G45C + C55G replicated with near-wildtype efficiency, while viral RNA accumulation was not fully restored in the double mutant C47G + G53C, suggesting additional constraints. Taken together, the mutagenesis study shows that the structural integrity of SL2 is essential for efficient HCoV-229E replication. Similar observations were previously made for the betacoronavirus MHV, where nucleotide substitutions that destabilized the SL2 stem region resulted in a drastic reduction of viral RNA synthesis and production of infectious virus progeny (Liu et al., 2007). As mentioned above, all double mutants replicated more efficiently than their single-mutation counterparts (Fig. 3A, lanes 2–5 and 6–9), providing strong evidence for the existence and functional relevance of the HCoV-229E

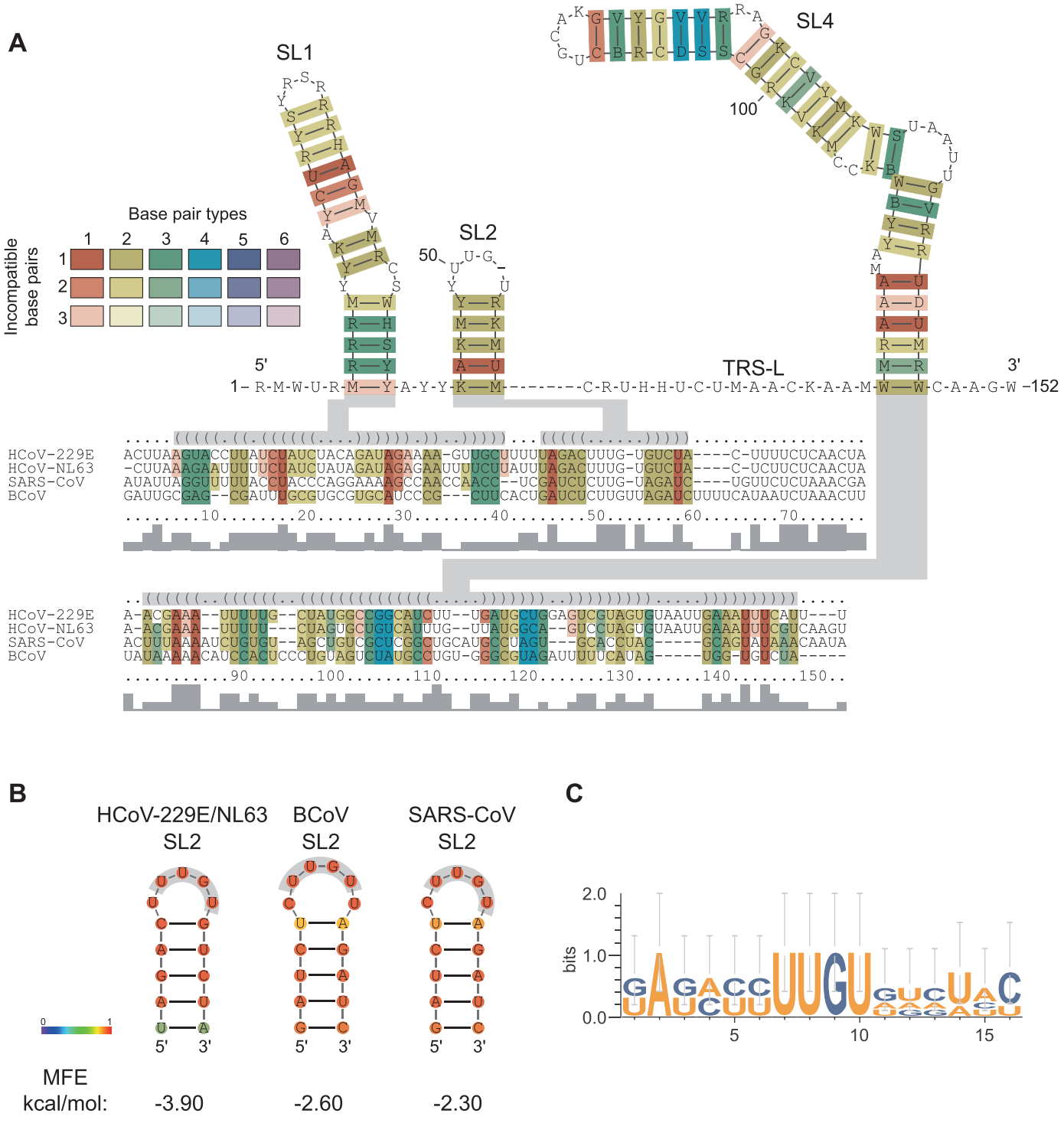

Fig. 5. Conservation of 5'-terminal SL1 and SL2 structures in alpha- and betacoronaviruses. A, Secondary structure prediction of the 5′-terminal genome regions of 4 coronaviruses representing the genera *Alphacoronavirus* (HCoV-229E, HCoV-NL63) and *Betacoronavirus* (SARS-CoV, BCoV). The alignment was calculated by LocARNA and the structure by RNAalifold. The consensus sequence is represented using the IUPAC code. Colors are used to indicate conserved base pairs: from red (conservation of only one base pair type) to purple (all six base pair types are found); from dark (all sequences contain this base pair) to light colors (1 or 2 sequences are unable to form this base pair). The gray bars below the alignment indicate the extent of sequence conservation at a given position. Gray shadows are used to link RNA structures with the corresponding dot-bracket notations above the alignment. To refine the alignment, an anchor at the highly conserved SL2 was used. B, Individual SL2 structures were predicted by RNAfold. Nucleotides in gray boxes indicate the conserved loop sequence. Colors represent base-pairing probabilities calculated by RNAfold. C, WebLogo representation of the conserved loop sequence of SL2 (Crooks et al., 2006). For the structure-based alignment and WebLogo representation, sequences from human coronavirus (HCoV) 229E (NC_002645), HCoV-NL63 (isolate Amsterdam 1, NC_005831), bovine coronavirus (isolate BCoV-ENT, NC_003045), and severe acute respiratory syndrome coronavirus (SARS-CoV, strain Tor2, NC_004718) were used.

SL1 and SL2 structures.

Previous betacoronavirus studies suggested that, with few exceptions, preservation of the SL1 and SL2 secondary structures is more important for viral replication than preservation of a specific nucleotide sequence (Liu et al., 2007; Li et al., 2008). By and large, our functional analysis of the HCoV-229E SL1 and SL2 elements supports these earlier proposals of the Leibowitz and Giedroc laboratories for the SL1 and SL2 equivalents in MHV. For the HCoV-220E_C11G + G34C and

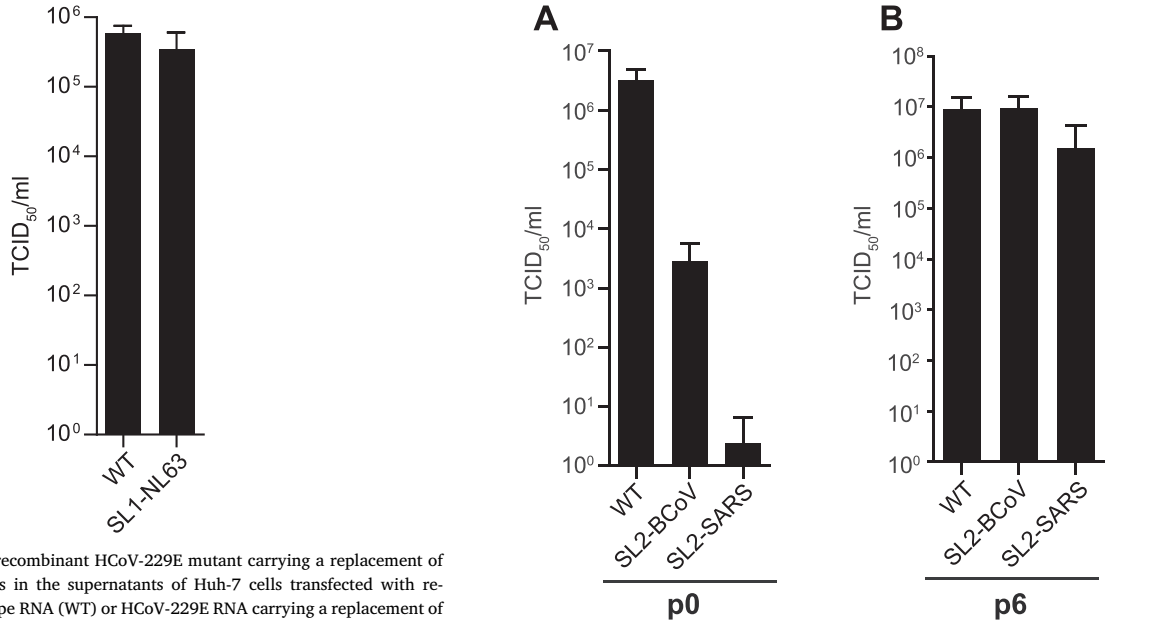

Fig. 6. Characterization of a recombinant HCoV-229E mutant carrying a replacement of the SL1 structure. Virus titers in the supernatants of Huh-7 cells transfected with recombinant HCoV-229E wildtype RNA (WT) or HCoV-229E RNA carrying a replacement of the cognate HCoV-229E SL1 element with the structural counterpart from HCoV-NL63 (SL1-NL63) were determined as described in Material and Methods. Virus titers (means ± SEM) are represented as $TCID_{50}$/ml and were determined from three independent transfection experiments.

G45C + C55G mutants, we were able to show that RNA synthesis returned to wildtype levels if the base-pairing potential was restored by introducing appropriate compensatory mutations (Fig. 3A, lane 6 and 8). In contrast, genome replication and sg mRNA synthesis and virus progeny production (see below) did not revert to wildtype levels in the case of C16G + G29C and C47G + G53C even though the stem structures were restored in these mutants (Fig. 3A, lanes 7 and 9). Our results strongly suggest that not only the SL1 and SL2 structures but also the nucleotide sequence plays an important role in RNA synthesis and the production of infectious virus progeny (see below).

Along with the analysis of viral RNA accumulation, we measured virus titers in the supernatants of transfected cells (Fig. 3B). Titers are given as mean values and standard error of the mean (± SEM) and were determined from three independent transfection experiments. In all SL1 and SL2 mutants, transfection of full-length (wildtype or mutant) genome RNA gave rise to infectious virus progeny, but virus titers varied greatly among the different mutants (Fig. 3B). Substitution of C11 with G resulted in an approximately 10-fold reduced titer compared to the wildtype virus, consistent with the moderate reduction in viral RNA synthesis observed for this mutant (Fig. 3A, lane 2). In contrast, the C16G, G45C, and C47G substitutions caused a drastic 100–1000-fold) reduction of virus titers, suggesting severe defects in viral replication. Upon restoration of the base pairing in the SL1_C11G + G34C and SL2_G45C + C55G SL2 mutants, the production of infectious virus progeny returned to (near) wildtype levels. Consistent with the Northern blot data presented above, restoration of base-pairing interactions in the C16G + G29C and C47G + G53C mutants failed to restore the full replication potential. Overall, the titers obtained for the HCoV-229E SL1 and SL2 mutants correlate very well with the RNA replication data (Fig. 3A), suggesting that the introduced mutations primarily affect viral RNA synthesis rather than a late step in the viral life cycle.

4.3. Rapid reversion of the C47G substitution to the wildtype sequence

As illustrated in Fig. 2, single-nucleotide substitutions in the stem of SL2 were predicted to cause major structural rearrangements and, consistent with the presumed *cis*-acting function of this element, resulted in severe defects in viral RNA synthesis and reproduction (Fig. 3). To test if the introduced mutations were stable enough to allow their phenotypes to be analyzed in passage 0 (p0), we subjected viral RNAs isolated from p0 virus stocks of the SL1 and SL2 mutants to

(caption on next page)

Fig. 7. Characterization of recombinant HCoV-229E mutants carrying replacements of the SL2 structure. A, Virus titers of HCoV-229E wildtype (WT) and HCoV-229E mutants carrying SL2 elements from BCoV (SL2-BCoV) and SARS-CoV (SL2-SARS). Cell culture supernatants were collected at 72 h p.t. and titers were determined as described in Material and Methods. B, Recombinant HCoV-229E (WT) and HCoV-229E mutants carrying SL2 structures from BCoV or SARS-CoV (SL2-BCoV, SL2-SARS-CoV) were serially passaged and virus titers were determined for passage 6 (p6) virus stocks using cell culture supernatants collected at 48 h p.i.. Virus titers (means ± SEM) are represented as $TCID_{50}$/ml and were determined from three independent transfection (p0) and infection (p6) experiments, respectively. C, Plaques sizes of the indicated recombinant viruses. D, Sequence analysis of the genome region containing the indicated SL2 replacement in the HCoV-229E genome using RNA isolated at passage 0 (72 h p.t.). Positions of wildtype and betacoronavirus-derived SL2 sequences are indicated below the chromatograms.

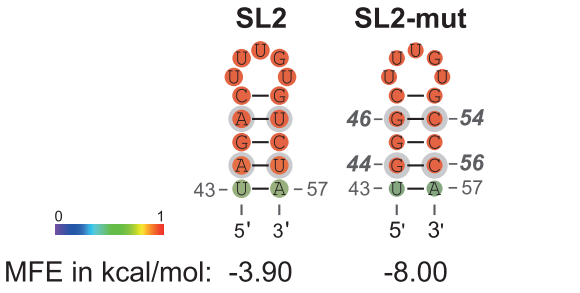

Fig. 8. Predicted RNA secondary structure of an HCoV-229E SL2 element carrying two A/U to G/C replacements. Positions of nucleotide substitutions are indicated with numbers (in boldface) in the SL2-mut structure. RNA secondary structures were predicted by RNAfold. Colors represent the base-pairing probability calculated by RNAfold.

partial genome sequence analyses covering the entire 5'-UTR and 3'-UTR regions. In all cases, the introduced mutations were found to be retained, suggesting that the virus titration and Northern blot data shown Fig. 3 reflect 'true' phenotypes of the SL1 and SL2 mutants generated in this study. Only in one case, SL2_C47G, we obtained evidence for a rapid reversion back to the wildtype sequence. As shown in Fig. 4A, we observed an additional C peak at position 47. To corroborate this observation, the C47G mutant was subjected to 5 serial passages in Huh-7 cells (Fig. 4B) and RNA isolated from virus-infected cells was used for sequence analysis. The sequence data confirmed a complete replacement of the C47-to-G mutation with the wildtype nucleotide C47, thereby restoring the Watson-Crick base pairing of nucleotides 47 and 53 in this revertant (Fig. 4C). The data lead us to suggest a critical role for the C47-G53 base pair in supporting specific SL2 structure-function relationships required for coronavirus RNA synthesis. The data also suggest that the (low) titer determined for the C47G p0 virus stock may represent an overestimate of the real replication efficiency of the C47G mutant because, even at this early time point p.t., a significant proportion of virus genomes had reverted to the wildtype sequence. We did not observe reversions or second-site substitutions in the SL2-C47G + G53C double mutant, where Watson-Crick base pairing was restored (Fig. 2), suggesting that RNA synthesis and viral reproduction, although being reduced compared to the wildtype, were sufficient for virus growth in cell culture. Nevertheless, the partial growth defect of this double mutant indicates that not only the helical stem structure but also the sequence might play a role in viral RNA synthesis (Fig. 3).

4.4. HCoV-229E SL2 is functionally exchangeable with betacoronavirus SL2 elements

Genus-wide consensus secondary structural models indicated that the 5′-terminal ~150 nt of alpha- and betacoronaviruses folds into three highly conserved structures that are generally referred to as SL1, SL2, and SL4 (Madhugiri et al., 2014, 2016; Yang and Leibowitz, 2015). This conservation is illustrated in Fig. 5A for 4 coronaviruses representing the genera *Alpha-* and *Betacoronavirus*. Previous studies suggested that SL2 represents the most conserved RNA structural element in coronavirus genomes (Kang et al., 2006; Chen and Olsthoorn, 2010). As illustrated for 4 representative coronaviruses in Fig. 5B, SL2 is always comprised of a 5-base-pair helical stem and, in most cases, a pentaloop structure. The loop sequence (5′-(C/U)UUG(U/C)−3′) is highly conserved while the stem sequence is variable (Fig. 5C). Previous studies in betacoronavirus systems revealed that intra-genus replacements of 5′-terminal RNA structural elements may result in viable viruses (Kang et al., 2006), supporting the high degree of both structural and functional conservation among betacoronavirus *cis*-acting elements in the 5' genome region. These earlier and our own studies led us to suggest that several RNA structural elements in the 5' and 3' UTRs are not only conserved among alphacoronaviruses but also across different genera of the *Coronavirinae* (Madhugiri et al., 2014). To test this hypothesis, we constructed a mutant in which the HCoV-229E SL1 was replaced with the equivalent structure of HCoV-NL63. Transfection of *in vitro*-transcribed HCoV-229E genome RNA containing this intra-genus replacement gave rise to infectious virus progeny that replicated to near-wildtype titers (Fig. 6), confirming the functional conservation of this structure among alphacoronaviruses. As the SL2 sequences of HCoV-229E and HCoV-NL63 are identical, a replacement of the SL2 structures was dispensable. Instead, we decided to study the extent of SL2 conservation across genus boundaries by extending our studies to the SL2 elements of betacoronaviruses. To our knowledge, inter-genus exchanges of *cis*-acting elements in coronavirus 5'-proximal genome regions have not been performed previously. Using the vaccinia virus-based reverse genetics system, we replaced the HCoV-229E SL2 structure in the full-length HCoV-229E cDNA sequence with the structural counterpart from BCoV and SARS-CoV, respectively, representing different lineages of the genus *Betacoronavirus*. *In vitro*-transcribed full-length 'chimeric' and wildtype HCoV-229E genome RNAs were transfected into Huh-7 cells. At 72 h p.t., virus titers in cell culture supernatants collected from transfected cells were determined from three independent transfection experiments. In all cases, we were able to recover infectious virus progeny (Fig. 7). Exchange of the HCoV-229E SL2 with that of BCoV resulted in a viral titer of $> 10^3$ $TCID_{50}$/ml (Fig. 7A), while a replacement with the SL2 of SARS-CoV resulted in significantly lower titers (Fig. 7A) and slightly smaller plaques sizes compared to the parental HCoV-229E virus (Fig. 7C). Sequence analyses of cDNA obtained from the chimeric virus progeny confirmed that the exchanges introduced into the HCoV-229E genome were retained (Fig. 7D) and no additional second-site mutations were identified in the 5′- and 3'-UTRs in the recovered viruses. Furthermore, we passaged the chimeric viruses 'blindly' for six times. Titration of these serially passaged viruses revealed near-wildtype growth for the virus carrying the BCoV SL2, while the virus carrying the SARS-CoV structure replicated to slightly lower titers (Fig. 7D). The increase in titers might be due to the accumulation of compensatory mutations. To address this possibility, viral cDNA produced from the p6 virus stocks of the BCoV-SL2 and SARS-SL2 mutants, respectively, were subjected to sequence analysis covering the entire 5'- and 3'-UTRs and the replicase gene sequences encoding nsp7 to 12. Analysis of the consensus sequence of the virus stocks confirmed that the introduced SL2 replacements were retained after six passages (data not shown) and no compensatory mutations were detected in this partial genome sequence analysis. To understand how these chimeric viruses evolved to almost wild-type like titers (while retaining the engineered substitutions), complete sequence analyses of single plaque-purified (high-passage) chimeric viruses remain to be performed in future studies. Taken together, these inter-genus exchange data provide experimental proof for our hypothesis that several coronavirus *cis*-acting RNA elements are conserved, both structurally and functionally, among different coronavirus genera.

4.5. Mutations predicted to increase the stability of SL2 are not tolerated

As shown in Fig. 3, nucleotide substitutions predicted to destabilize SL2 cause major defects in HCoV-229E RNA synthesis and virus

reproduction, demonstrating the functional relevance of this SL structure. Interestingly, the helical SL2 stem of most coronaviruses is composed of three A-U and two G-C base pairs, respectively (Fig. 5A and B, Fig. 8A). We therefore asked the question of whether this pattern of base-pair interactions reflects a finely balanced stability of this structure. To address this question, we stabilized the SL2 structure by introducing 4 mutations. The mutations replaced two A/U base pairs at positions 44 + 56 and 46 + 54 with G/C base pairs and were predicted to decrease the minimal free energy of this secondary structure (Fig. 8). An *in vitro*-transcribed full-length genome RNA containing this set of mutations and a wildtype control RNA, respectively, were transfected into Huh-7 cells and virus titers in cell culture supernatants collected at 72 h p.t. were determined. In repeated experiments, we failed to recover viable virus containing these 'stabilizing' mutations in the stem region of SL2, while the wild-type virus was readily recovered with high titers. The data demonstrate specific sequence requirements for the SL2 stem region. It remains to be studied in further experiments if these sequence constraints reflect a requirement for an 'optimal stability' of SL2 or rather the presence of specific nucleotides at specific positions.

5. Discussion

In this study, we used a combined bioinformatics, biochemical and reverse genetics approach to characterize the structures and functions of the putative *cis*-acting SL1 and SL2 elements of the alphacoronavirus HCoV-229E (Madhugiri et al., 2014, 2016; Yang and Leibowitz, 2015). Based on RNA structure probing information obtained for *in vitro*-transcribed RNAs representing appropriate genome sequences of HCoV-229E and a second alphacoronavirus (HCoV-NL63), combined with bioinformatics studies, we present a robust RNA secondary structure model for the 5′-terminal 80 nts of the HCoV-229E genome that we think to be representative for other alphacoronaviruses. We also provide evidence that (i) the SL1 and SL2 RNA structural elements are required for viral replication and (ii) the structures and functions of these elements are conserved among alphacoronaviruses (and, probably, betacoronaviruses). The study revealed a number of interesting parallels to the SL1 and SL2 elements of betacoronaviruses which have been characterized extensively in previous studies and shown to be required for BCoV and MHV genome replication and sg mRNA synthesis (Kang et al., 2006; Liu et al., 2007, 2009a; Li et al., 2008).

The RNA structure probing data presented in this study (Fig. 1) are consistent with models developed previously for a range of beta- and, to a lesser extent, alphacoronaviruses (Kang et al., 2006; Liu et al., 2007, 2009a; Li et al., 2008; Madhugiri et al., 2014; Yang et al., 2015). The probing information provides experimental support for the presence of stable SL1 and SL2 structures in the 5'-terminal genome region and suggests that the TRS-L, together with flanking sequences, is part of an unstructured region. The latter conclusion is consistent with previous studies in which the TRS-L was proposed to be located in an unstructured region in the majority of coronavirus genomes (Liu et al., 2007; Madhugiri et al., 2014; Yang et al., 2015). Furthermore, most structure prediction applications place (or can be forced to place) the coronavirus TRS-L in SL structures that would only be supported by two conserved base pairs, arguing against a major role of such a structure (Raman et al., 2003; Liu et al., 2007; Chen and Olsthoorn, 2010; Yang et al., 2015; Madhugiri et al., 2016). In this context, it should be noted that there is also evidence that, in a subset of alphacoronaviruses (TGEV), betacoronaviruses (including BCoV) and gammacoronaviruses, an additional structural element (SL3, also called SL-II in several BCoV studies) may exist. For example, studies on the TRS-L element of TGEV using NMR spectroscopy, UV thermal denaturation experiments and a reverse genetics approach suggested the existence of a defined hairpin structure in this genome region (Dufour et al., 2011). Most of the TGEV TRS-L core sequence was proposed to be located in a heptaloop region of a hairpin structure with moderate (possibly optimized) thermal stability. Both the structure and stability of this TRS-L hairpin structure was shown to play a role in TGEV replication and transcription where it was proposed to act as a landing platform for the nascent minus-strand RNA, similar to the similarity-assisted RNA recombination model proposed earlier by Nagy et al. (Nagy and Simon, 1997). Taken together, the available information suggests that the TRS-L region may be structurally flexible and adopt alternative structures to regulate specific steps of viral RNA synthesis.

The HCoV-229E mutagenesis data obtained in our reverse genetics study of SL1 and SL2 mutants provide experimental support for the functional relevance of these elements in viral replication. Based on computer-assisted structure predictions for SL1 and SL2 variants containing specific mutations in stem regions, a set of mutants was designed in which the respective structures were destabilized or disrupted. Possible effects of the mutations on viral replication were subsequently studied in cell culture. The mutagenesis data obtained for the SL1 mutants confirm a critical role for SL1 in HCoV-229E replication. The data also revealed that destabilizing mutations in the upper and lower parts of the SL1 structure have quite different effects on viral replication. Furthermore, the incomplete restoration of the *in vitro* growth characteristics of the C16G + G29E mutant suggests additional (sequence) constraints which remain to be investigated in further studies. The less critical role observed for C11 (which acts to stabilize the lower part of SL1) in viral replication and the observation that the double mutant C11G + G34C replicated with near-wildtype characteristics suggest that the lower part of SL1 tolerates some structural changes, possibly indicating flexibility in this part of the structure. These observations are reminiscent of data reported by the Giedroc laboratory for MHV (Li et al., 2008). In this case, the upper part of the SL1 stem was found to be required for efficient MHV replication while a less stable structure of the lower part was largely tolerated. Interestingly, the MHV study also detected second-site suppressor mutations in the 5′- and 3′-UTRs in some of the SL1 mutants. Based on these second-site mutations, a “dynamic SL1” model was proposed in which the lower part of SL1 is required to have an optimized flexibility to mediate physical interactions between the 5′- and 3′-UTRs that, for example, may stimulate sg mRNA synthesis. To date, we failed to detect any second-site suppressor mutations in the 5′- and 3′-UTRs of HCoV-229E in our SL1 mutants. Possible reasons for this discrepancy from the MHV data remain to be studied but may relate to the more drastic (deletion) mutations introduced in the MHV SL1 structure, which may have forced the development and fixation of compensatory mutations in the MHV mutants (Li et al., 2008) while, in our own study, single-nucleotide substitutions were introduced in the HCoV-229E SL1.

As mentioned above, SL2 represents the most conserved *cis*-acting RNA element in coronaviruses (Kang et al., 2006; Chen and Olsthoorn, 2010), suggesting an important function in coronavirus replication. Our HCoV-229E SL2 mutagenesis data strongly support this hypothesis. Thus, any disruption of G-C base pair interactions predicted to destabilize the HCoV-229E SL2 stem structure (Fig. 2F and H) caused major defects in viral replication (Fig. 3) while restoration of the helical stem in the double mutant G45C + C55G resulted in a wildtype phenotype (Fig. 3). Surprisingly, the other double mutant (C47G + G53C) that was predicted to preserve the SL2 stem structure (Fig. 2, panel I) was found to have partial defects in RNA replication and production of infectious virus progeny (Fig. 3). Furthermore, the rapid reversion of the C47G mutant to the wildtype sequence (Fig. 4) indicates a strong selection pressure to maintain this particular base pair interaction. While our mutagenesis data, combined with extensive MHV SL2 mutagenesis studies (Kang et al., 2006; Liu et al., 2007), establish an essential role for SL2 in alpha- and betacoronavirus replication, more studies will be required to investigate the precise roles of residues in the HCoV-229E SL2 stem and loop regions, including base pair interactions within the loop or at its base (C47), and, possibly, also unravel the special role of the C47-G53 pair in the function of SL2.

Based on our own and other studies (see below), it was tempting to suggest that the (structurally) conserved 5′-proximal *cis*-acting RNA

elements, including SL1 and SL2, may also be *functionally* conserved across coronavirus genera. To test this idea, we constructed HCoV-229E mutants in which the cognate SL2 element was replaced with that of BCoV and SARS-CoV, respectively. Our study was guided by earlier bioinformatic analyses (Madhugiri et al., 2014) that suggested conservation of RNA secondary structures in the UTRs of viruses from the same genus but also other coronavirus genera. Previously, such a structural and functional conservation had only been confirmed for members of the same genus (Goebel et al., 2004; Kang et al., 2006). In the present study, we were able to extend these previous conclusions to alphacoronaviruses by showing that a replacement of the HCoV-229E SL1 with the equivalent structure from HCoV-NL63 (Fig. 6) was tolerated very well, with titers of the chimeric virus approaching that of the wildtype virus. More importantly, we were able to show that the SL2s of BCoV and SARS-CoV, respectively, can act, at least in part, as functional substitutes for the cognate SL2 structure in the HCoV-229E genome. To our knowledge, this is the first experimental proof that the 5′-terminal SL2 is both structurally and functionally conserved among alpha- and betacoronaviruses. In line with previous studies (Kang et al., 2006), the four coronaviruses included in the present study (representing the genera *Alpha-* and *Betacoronavirus*) share a short 4–5-bp) helical stem and a highly conserved pentaloop sequence, 5′-(U/C)UUGU-3′ (Fig. 5). Although the sequence of the predicted helical stems is not conserved (Fig. 5C), viable viruses carrying the BCoV or SARS-CoV SL2 counterparts could be recovered as confirmed by virus titration and sequence analysis (Fig. 7A and D). We did not observe second-site mutations in the introduced SL2 structures and the entire 5′- and 3′-UTR regions in virus stocks collected at 72 h p.t. (p0) and after serial passaging (p6), respectively (Fig. 7). Complete genome analyses of plaque-purified mutants remain to be performed in future studies to exclude compensatory mutations in other genome regions. Although several features are conserved among alpha- and betacoronavirus SL2 structures, which may explain the functionality of the betacoronavirus SL2 structure in an alphacoronavirus context, more studies are required to fully understand the critical parameters required for the SL2 function(s) in virus replication. One of these parameters might be an optimal stability of the 4–5-bp stem of SL2. In this context, the observed lethal phenotype of an HCoV-229E mutant carrying an SL2 with two additional G-C base pairs (Fig. 8) provides preliminary evidence to suggest that both the nucleotide sequence and the stability of the stem may play a more important role than previously thought and, thus, deserve further studies.

Taken together, our functional characterization of alphacoronavirus SL1 and SL2 elements strongly supports the idea that, despite very limited sequence conservation, a number of *cis*-acting RNA elements including SL1 and SL2 are structurally conserved and have similar functions in coronavirus replication.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 1021, A01, to J.Z.).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2017.11.025.

References

Almazan, F., Dediego, M.L., Galan, C., Escors, D., Alvarez, E., Ortego, J., Sola, I., Zuniga, S., Alonso, S., Moreno, J.L., Nogales, A., Capiscol, C., Enjuanes, L., 2006. Construction of a severe acute respiratory syndrome coronavirus infectious cDNA clone and a replicon to study coronavirus RNA synthesis. J. Virol. 80, 10900–10906.

Barton, D.J., O'Donnell, B.J., Flanegan, J.B., 2001. 5' cloverleaf in poliovirus RNA is a *cis*-acting replication element required for negative-strand synthesis. EMBO J. 20, 1439–1448.

Brian, D.A., Baric, R.S., 2005. Coronavirus genome structure and replication. Curr. Top. Microbiol. Immunol. 287, 1–30.

Brown, C.G., Nixon, K.S., Senanayake, S.D., Brian, D.A., 2007. An RNA stem-loop within the bovine coronavirus nsp1 coding region is a cis-acting element in defective interfering RNA replication. J. Virol. 81, 7716–7724.

Chang, R.Y., Hofmann, M.A., Sethna, P.B., Brian, D.A., 1994. A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. J. Virol. 68, 8223–8231.

Chang, R.Y., Krishnan, R., Brian, D.A., 1996. The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. J. Virol. 70, 2720–2729.

Chen, S.C., Olsthoorn, R.C., 2010. Group-specific structural features of the 5'-proximal sequences of coronavirus genomic RNAs. Virology 401, 29–41.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res 14, 1188–1190.

Darty, K., Denise, A., Ponty, Y., 2009. VARNA: interactive drawing and editing of the RNA secondary structure. Bioinformatics 25, 1974–1975.

Dufour, D., Mateos-Gomez, P.A., Enjuanes, L., Gallego, J., Sola, I., 2011. Structure and functional relevance of a transcription-regulating sequence involved in coronavirus discontinuous RNA synthesis. J. Virol. 85, 4963–4973.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.P., Ehresmann, B., 1987. Probing the structure of RNAs in solution. Nucleic Acids Res 15, 9109–9128.

Firth, A.E., Brierley, I., 2012. Non-canonical translation in RNA viruses. J. Gen. Virol. 93, 1385–1409.

Goebel, S.J., Taylor, J., Masters, P.S., 2004. The 3' *cis*-acting genomic replication element of the severe acute respiratory syndrome coronavirus can function in the murine coronavirus genome. J. Virol. 78, 7846–7851.

Goto, H., Muramoto, Y., Noda, T., Kawaoka, Y., 2013. The genome-packaging signal of the influenza A virus genome comprises a genome incorporation signal and a genome-bundling signal. J. Virol. 87, 11316–11322.

Gustin, K.M., Guan, B.J., Dziduszko, A., Brian, D.A., 2009. Bovine coronavirus nonstructural protein 1 (p28) is an RNA binding protein that binds terminal genomic cis-replication elements. J. Virol. 83, 6087–6097.

Hertzig, T., Scandella, E., Schelle, B., Ziebuhr, J., Siddell, S.G., Ludewig, B., Thiel, V., 2004. Rapid identification of coronavirus replicase inhibitors using a selectable replicon RNA. J. Gen. Virol. 85, 1717–1725.

Isaacs, S.N., Kotwal, G.J., Moss, B., 1990. Reverse guanine phosphoribosyltransferase selection of recombinant vaccinia viruses. Virology 178, 626–630.

Kang, H., Feng, M., Schroeder, M.E., Giedroc, D.P., Leibowitz, J.L., 2006. Putative cis-acting stem-loops in the 5' untranslated region of the severe acute respiratory syndrome coronavirus can substitute for their mouse hepatitis virus counterparts. J. Virol. 80, 10600–10614.

Keane, S.C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., Barton, S., Hosic, A., Florwick, A., Santos, J., Bolden, N.C., McCowin, S., Case, D.A., Johnson, B.A., Salemi, M., Telesnitsky, A., Summers, M.F., 2015. RNA structure. Structure of the HIV-1 RNA packaging signal. Science 348, 917–921.

Kuo, L., Masters, P.S., 2013. Functional analysis of the murine coronavirus genomic RNA packaging signal. J. Virol. 87, 5182–5192.

Li, L., Kang, H., Liu, P., Makkinje, N., Williamson, S.T., Leibowitz, J.L., Giedroc, D.P., 2008. Structural lability in stem-loop 1 drives a 5' UTR-3' UTR interaction in coronavirus replication. J. Mol. Biol. 377, 790–803.

Liu, P., Leibowitz, J., 2010. RNA higher-order structures within the coronavirus 5' and 3' untranslated regions and their roles in viral replication. In: Lal, S.K. (Ed.), Molecular Biology of the SARS-Coronavirus. Springer-Verlag, Berlin Heidelberg, Germany, pp. 47–61.

Liu, P., Li, L., Keane, S.C., Yang, D., Leibowitz, J.L., Giedroc, D.P., 2009a. Mouse hepatitis virus stem-loop 2 adopts a uYNMG(U)a-like tetraloop structure that is highly functionally tolerant of base substitutions. J. Virol. 83, 12084–12093.

Liu, P., Li, L., Millership, J.J., Kang, H., Leibowitz, J.L., Giedroc, D.P., 2007. A U-turn motif-containing stem-loop in the coronavirus 5' untranslated region plays a functional role in replication. RNA 13, 763–780.

Liu, Y., Wimmer, E., Paul, A.V., 2009b. *Cis*-acting RNA elements in human and animal plus-strand RNA viruses. Biochim Biophys. Acta 1789, 495–517.

Lorenz, R., Bernhart, S.H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., Hofacker, I.L., 2011. ViennaRNA Package 2.0. Algorithms Mol. Biol. 6, 26.

Luo, D., Condon, C., Grunberg-Manago, M., Putzer, H., 1998. In vitro and in vivo secondary structure probing of the thrS leader in Bacillus subtilis. Nucleic Acids Res 26, 5379–5387.

Madhugiri, R., Fricke, M., Marz, M., Ziebuhr, J., 2014. RNA structure analysis of alphacoronavirus terminal genome regions. Virus Res 194, 76–89.

Madhugiri, R., Fricke, M., Marz, M., Ziebuhr, J., 2016. Coronavirus cis-acting RNA elements. Adv. Virus Res. 96, 127–163.

Masters, P.S., 2007. Genomic *cis*-acting elements in coronavirus RNA replication. In: Thiel, V. (Ed.), Coronaviruses - Molecular and Cellular Biology. Caister Academic Press, Norfolk, United Kingdom, pp. 65–80.

Morales, L., Mateos-Gomez, P.A., Capiscol, C., del Palacio, L., Enjuanes, L., Sola, I., 2013. Transmissible gastroenteritis coronavirus genome packaging signal is located at the 5' end of the genome and promotes viral RNA incorporation into virions in a replication-independent process. J. Virol. 87, 11579–11590.

Nagy, P.D., Simon, A.E., 1997. New insights into the mechanisms of RNA recombination. Virology 235, 1–9.

Nicholson, B.L., White, K.A., 2014. Functional long-range RNA-RNA interactions in positive-strand RNA viruses. Nat. Rev. Microbiol. 12, 493–504.

Raman, S., Bouma, P., Williams, G.D., Brian, D.A., 2003. Stem-loop III in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J. Virol. 77, 6720–6730.

Raman, S., Brian, D.A., 2005. Stem-loop IV in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J. Virol. 79,

12434–12446.
Reed, L.J., Muench, H., 1938. A simple method of estimating fifty percent endpoints. Am. J. Epidemiol. 27, 493–497.
Schelle, B., Karl, N., Ludewig, B., Siddell, S.G., Thiel, V., 2005. Selective replication of coronavirus genomes that express nucleocapsid protein. J. Virol. 79, 6620–6630.
Thiel, V., Herold, J., Schelle, B., Siddell, S.G., 2001. Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. J. Gen. Virol. 82, 1273–1281.
Van Den Born, E., Gultyaev, A.P., Snijder, E.J., 2004. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. RNA 10, 424–437.
Will, S., Joshi, T., Hofacker, I.L., Stadler, P.F., Backofen, R., 2012. LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. RNA 18, 900–914.
Yang, D., Leibowitz, J.L., 2015. The structure and functions of coronavirus genomic 3' and 5' ends. Virus Res 206, 120–133.
Yang, D., Liu, P., Wudeck, E.V., Giedroc, D.P., Leibowitz, J.L., 2015. SHAPE analysis of the RNA secondary structure of the Mouse Hepatitis Virus 5' untranslated region and N-terminal nsp1 coding sequences. Virology 475, 15–27.
Zuniga, S., Sola, I., Alonso, S., Enjuanes, L., 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J. Virol. 78, 980–994.