JOURNAL OF VIROLOGY, Jan. 2004, p. 669–682 Vol. 78, No. 2
0022-538X/04/$08.00+0 DOI: 10.1128/JVI.78.2.669–682.2004

Characterization of the RNA Components of a Putative Molecular Switch in the 3′ Untranslated Region of the Murine Coronavirus Genome

Scott J. Goebel,[1] Bilan Hsue,[1,2]† Todd F. Dombrowski,[1] and Paul S. Masters[1,2]*

Wadsworth Center, New York State Department of Health,[1] and Department of Biomedical Sciences, State University of New York,[2] Albany, New York 12201

Received 11 July 2003/Accepted 4 October 2003

RNA virus genomes contain *cis*-acting sequence and structural elements that participate in viral replication. We previously identified a bulged stem-loop secondary structure at the upstream end of the 3′ untranslated region (3′ UTR) of the genome of the coronavirus mouse hepatitis virus (MHV). This element, beginning immediately downstream of the nucleocapsid gene stop codon, was shown to be essential for virus replication. Other investigators discovered an adjacent downstream pseudoknot in the 3′ UTR of the closely related bovine coronavirus (BCoV). This pseudoknot was also shown to be essential for replication, and it has a conserved counterpart in every group 1 and group 2 coronavirus. In MHV and BCoV, the bulged stem-loop and pseudoknot are, in part, mutually exclusive, because of the overlap of the last segment of the stem-loop and stem 1 of the pseudoknot. This led us to hypothesize that they form a molecular switch, possibly regulating a transition occurring during viral RNA synthesis. We have now performed an extensive genetic analysis of the two components of this proposed switch. Our results define essential and nonessential components of these structures and establish the limits to which essential parts of each element can be destabilized prior to loss of function. Most notably, we have confirmed the interrelationship of the two putative switch elements. Additionally, we have identified a pseudoknot loop insertion mutation that appears to point to a genetic interaction between the pseudoknot and a distant region of the genome.

Embedded in the genomes of RNA viruses are *cis*-acting sequence and structural elements that participate in replication, transcription, translation, and packaging. Some of these signals are thought to facilitate the selective interactions of viral RNAs with the machinery of RNA synthesis, while others enable or modulate events that occur during viral protein synthesis or assembly. The coronaviruses and arteriviruses, members of the order *Nidovirales*, are positive-strand RNA viruses with very large genomes (13 to 15 kb for arteriviruses, 27 to 31 kb for coronaviruses). In addition to the production of progeny genome copies, nidovirus replication entails the synthesis of a 3′ nested set of subgenomic (sg) mRNAs (15). Each sgRNA contains a 5′ leader segment connected to a body segment that is identical to the 3′ end of the genome, starting at a given point preceding one of the downstream open reading frames. Accumulating evidence, particularly from landmark studies using a full-length infectious cDNA of equine arterivirus, supports a model in which the discontinuous step in sgRNA formation occurs during negative-strand RNA synthesis (1, 28, 31, 35). Thus, for nidoviruses, the earliest steps of both genome replication and sgRNA transcription initiate at the 3′ end of the genome; therefore, it is reasonable to expect that at least part of the regulation of these processes is implicit in the sequence and structure of the 3′ untranslated region (3′ UTR). Indeed, studies of defective interfering (DI) RNAs, parasitic RNAs with extensive deletions that replicate at the expense of helper virus, have established that the entire 3′ UTR is at least part of the essential *cis*-acting region at the 3′ end of the coronavirus genome (11, 16, 34).

We have studied the 3′ UTR of the group 2 coronavirus mouse hepatitis virus (MHV). This 301-nucleotide (nt) genomic region has greater than 99% sequence conservation among different strains of MHV (27), and for the closest relative of this virus, bovine coronavirus (BCoV), there is 69% sequence identity in the 3′ UTR. Moreover, we have shown that the two regions are functionally homologous, since the MHV 3′ UTR can be replaced in its entirety by the BCoV 3′ UTR. A partially completed picture of the MHV 3′ UTR can be assembled from three separate lines of work (Fig. 1). Our laboratory previously identified an essential bulged stem-loop secondary structure at the upstream end of the MHV 3′ UTR, immediately downstream of the nucleocapsid (N) gene stop codon (6, 7). Shortly thereafter, the Brian laboratory reported the discovery of an essential classical hairpin-type pseudoknot in an adjacent downstream position of the BCoV 3′ UTR (37). The MHV genome harbors a structurally identical counterpart to this pseudoknot, which is >90% conserved in sequence. Finally, the Leibowitz laboratory has presented evidence for a complex RNA secondary structure that accounts for most of the remainder of the MHV 3′ UTR (18) (Fig. 1). Overall, this latter structure is not well conserved with that predicted for the corresponding region of the BCoV 3′ UTR, but mutations made in one stem that is highly conserved between the two viruses were found to be deleterious to the replication of DI RNA. Additionally, the downstream end of this structure contains the minimal signal needed for the initiation of MHV

* Corresponding author. Mailing address: David Axelrod Institute, Wadsworth Center, New York State Department of Health, New Scotland Ave., P.O. Box 22002, Albany, NY 12201-2002. Phone: (518) 474-1283. Fax: (518) 473-1326. E-mail: masters@wadsworth.org.
† Present address: Stratagene, La Jolla, CA 92037.

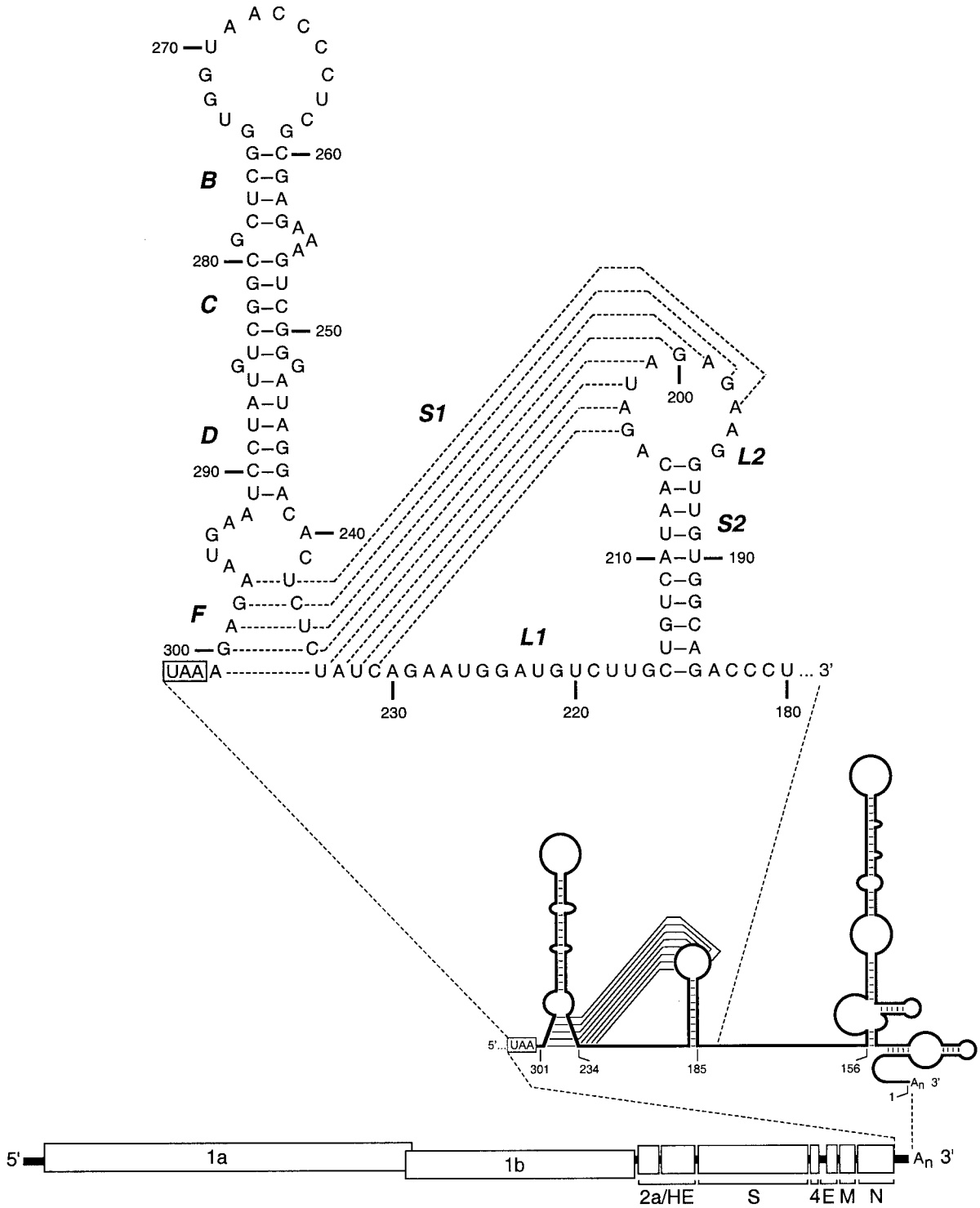

FIG. 1. Landscape of the MHV 3′ UTR. The organization of the 31.3-kb MHV genome is shown at the bottom. Above this is an expanded view of the secondary structure of the 301-nt 3′ UTR, which includes the bulged stem-loop demonstrated previously for nt 234 through 301 (6, 7), a pseudoknot comprising nt 185 through 238 (37), and a complex structure containing multiple stem-loops and bulges proposed for nt 1 through 156 (18). At the top is a detailed view of the upstream end of the 3′ UTR. Stem segments B, C, D, and F are shown as determined by Hsue et al. (6) (reprinted from reference 6). Pseudoknot stems S1 and S2 and loops L1 and L2 are labeled as determined by Williams et al. (37). Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A), and the N-gene stop codon is boxed. Broken lines in the upper diagram indicate alternative base pairings for stem segment F or pseudoknot stem 1.

negative-strand RNA synthesis, which has been mapped by the Lai laboratory to the last 55 bases of the 3′ UTR (17).

An intriguing feature of the two upstream RNA structural elements is that they partially overlap (Fig. 1). Thus, the bulged stem-loop and pseudoknot would not be able to fold up simultaneously. This observation led us to propose that the two elements are alternate states of a molecular switch, in which formation of one structure prevents the formation of the other. In the work reported here, we have extensively mutagenized the bulged stem-loop and pseudoknot, defining both their essential and nonessential components. Most importantly, we have confirmed the previously hypothesized interrelationship of the two structures. In addition, we have identified one mutant that may point to a genetic interaction between the pseudoknot and a distant region of the genome.

MATERIALS AND METHODS

Cells and viruses. Wild-type MHV-A59 and MHV mutants were propagated in mouse 17 clone 1 (17Cl1) cells; plaque assays and plaque purification were performed in mouse L2 cells. The interspecies chimeric viruses MHV (12) and fMHV.v2 (this study) were propagated in feline AK-D fetal lung cells.

Plasmid construction. For the purpose of engineering fMHV.v2, a new version of feline-murine interspecies chimeric coronavirus, we constructed a derivative of the transcription vector pFM1, which contains a cDNA copy of the 3′-most 8.6 kb of the MHV genome with the ectodomain of the feline infectious peritonitis virus (FIPV) spike (S) protein substituted for its MHV counterpart (12). The *Sse*8387I-*Pac*I fragment of pFM1, comprising all genes downstream of the S gene, was removed and replaced with the corresponding fragment from pA122 (3), which contains the same genes in a rearranged order, resulting in plasmid pFM2.

To generate the 45 mutants described in this study, donor RNA transcription vectors were derived from pMH54 (12) by two general strategies. In the first strategy (used for mutants 1 through 18 and 41 through 43), mutations were first constructed in a smaller vector, pBL85. In pBL85, an *Mlu*I site had previously been inserted near the end of the N gene, and an *Eco*RV site had been created by a single base change at position 221 in the 3′ UTR (6). Mutations were incorporated into pBL85 using oligonucleotide cassettes and were then transferred to pMH54 by exchange of the *Nhe*I-*Sac*I fragment that encompasses most of the N gene and 3′ UTR, as described previously (6). In the second strategy (used for mutants 19 through 40, 44, and 45), a coding-silent unique *Bsp*EI site was created in pMH54 near the end of the N gene by changing the sequence of codons 444 through 446 from GTGCCAGAT to GTTCCGGAT (*Bsp*EI site underlined), thus producing plasmid pSG6. Mutations were then created by splicing-overlap extension PCR (5) and cloned into pSG6 between the *Bsp*EI site and the unique *Bcl*I site at bases 21 through 26 of the 3′ UTR. For every plasmid constructed, overall composition was monitored by restriction analysis, and the sequences of all ligation junctions and all regions created by PCR were verified by automated sequencing.

Targeted RNA recombination. A variant of fMHV, designated fMHV.v2, was selected by targeted recombination between recipient virus Alb4 and donor RNA transcribed from pFM2 exactly as described previously for the construction of fMHV (12). To incorporate 3′ UTR mutations into the MHV genome, host range-based selection was then used, with either fMHV or fMHV.v2 as the recipient virus, as described previously (6, 12–14). Briefly, monolayers of feline AK-D cells were infected with fMHV or fMHV.v2 at a multiplicity in the range of 1 PFU per cell for 2.5 to 4 h at 37°C; infected cells were then suspended by treatment with a low concentration of trypsin. Capped donor RNAs were synthesized with a T7 mMessage mMachine transcription kit (Ambion) using *Pac*I-truncated templates. Synthetic donor RNA (roughly 5 to 10 μg) was transfected into 1.25×10^7 to 2.5×10^7 infected AK-D cells by two pulses with a Gene Pulser electroporation apparatus (Bio-Rad) at settings of 975 μF and 0.3 kV. The infected and transfected AK-D cells were either overlaid onto murine 17Cl1 cell monolayers or directly plated into 10-cm^2 wells. Progeny virus released into the supernatant medium was harvested when syncytia or cytopathic effect was apparent in cells, typically at 18 to 24 h postinfection at 37°C. Recombinant candidates were identified and purified through two rounds of plaque titration on murine L2 monolayers at 37°C. For particular mutations that appeared to be lethal, targeted recombination experiments and selections were also performed at 33 and 39°C.

Each purified recombinant candidate was used to infect a 25-cm^2 monolayer of 17Cl1 cells, and total cellular RNA was extracted at 18 to 24 h postinfection using Ultraspec reagent (Biotecx). Reverse transcription of RNA was performed with a random hexanucleotide primer (Boehringer Mannheim) and avian myeloblastosis virus reverse transcriptase (Life Sciences) under standard conditions (30); the relevant region of the genome was then amplified by PCR under standard conditions using AmpliTaq polymerase (Perkin-Elmer). The resulting products were analyzed directly by agarose gel electrophoresis or purified with Quantumprep columns (Bio-Rad) prior to automated sequencing. For most mutants, multiple independent isolates were obtained; the Alb strain designation for one representative isolate of each viable mutant is noted in the figure that shows the 3′ UTR sequence of that mutant.

We had previously found that the *Mlu*I and *Eco*RV sites originating from pBL85 had no phenotypic consequence when they were transduced into the MHV genome (6). Similarly, prior to construction of any mutants with donor RNAs transcribed from descendants of pSG6, we first determined whether the coding-silent *Bsp*EI site in the N gene had any measurable impact on viral phenotype. Recombinants selected from fMHV and donor RNA transcribed from pSG6 were observed to arise at a frequency comparable to that obtained with pMH54 donor RNA, and the growth characteristics of viruses shown to harbor the *Bsp*EI marker were indistinguishable from those of the wild type.

Radiolabeling of viral RNA. Virus-specific RNA was metabolically labeled as described previously (6, 7, 21). Confluent 20-cm^2 monolayers of 17Cl1 cells were infected with mutant or wild-type MHV at a multiplicity of 5 PFU per cell and incubated at 37°C. Infected cells were starved from 2 to 7 h postinfection in Eagle's minimal essential medium containing 5% dialyzed fetal bovine serum and 1/10 of the normal concentration of phosphate. Labeling was then performed from 7 to 9 h postinfection in 1 ml of phosphate-free medium containing 100 μCi of $^{33}P_i$ (ICN) per ml, 5% dialyzed fetal bovine serum, and 20 μg of actinomycin D (Sigma) per ml. Total cytoplasmic RNA was purified, and samples containing equal amounts of radioactivity were analyzed by electrophoresis through 1% agarose gels containing formaldehyde, and visualized by fluorography.

RESULTS

Selection of mutants. To define functionally critical features of *cis*-acting elements at the 3′ end of the MHV genome, we performed a detailed genetic analysis of the bulged stem-loop and pseudoknot that are formed by nt 185 through 301 of the 3′ UTR (Fig. 1). We generated MHV mutants using targeted RNA recombination with fMHV, an interspecies chimeric virus, as described previously (3, 12–14). fMHV is identical to MHV, except that its S glycoprotein contains the ectodomain of the FIPV S protein, by which it has simultaneously gained the ability to grow in feline cells and lost the ability to grow in murine cells (12). To create a given mutant, a donor synthetic RNA containing both the MHV S-gene ectodomain and a 3′ UTR mutation was transfected into fMHV-infected feline cells, and recombinant 3′ UTR mutants were selected on the basis of reacquisition of the ability to grow in murine cells (Fig. 2, top). It was also possible that, at a lower frequency, wild-type MHV recombinants would be selected by this scheme, as the result of a second crossover event downstream of the S gene but upstream of the 3′ UTR (Fig. 2, top). We have previously used this method to obtain extremely defective mutants, and such mutants were easily differentiated from wild-type recombinants on the basis of plaque size (13, 14; L. Kuo, K. R. Hurst, and P. S. Masters, unpublished data).

Midway through the study reported here, an even more stringent selection was developed through construction of a variant of fMHV, designated fMHV.v2, in which the order of genes downstream of the S gene was rearranged (Fig. 2, bottom). In this case, the region in which a second crossover could occur during targeted recombination was made extremely small, thus effectively precluding the formation of wild-type recombinants; that is, second-crossover events occurred at a

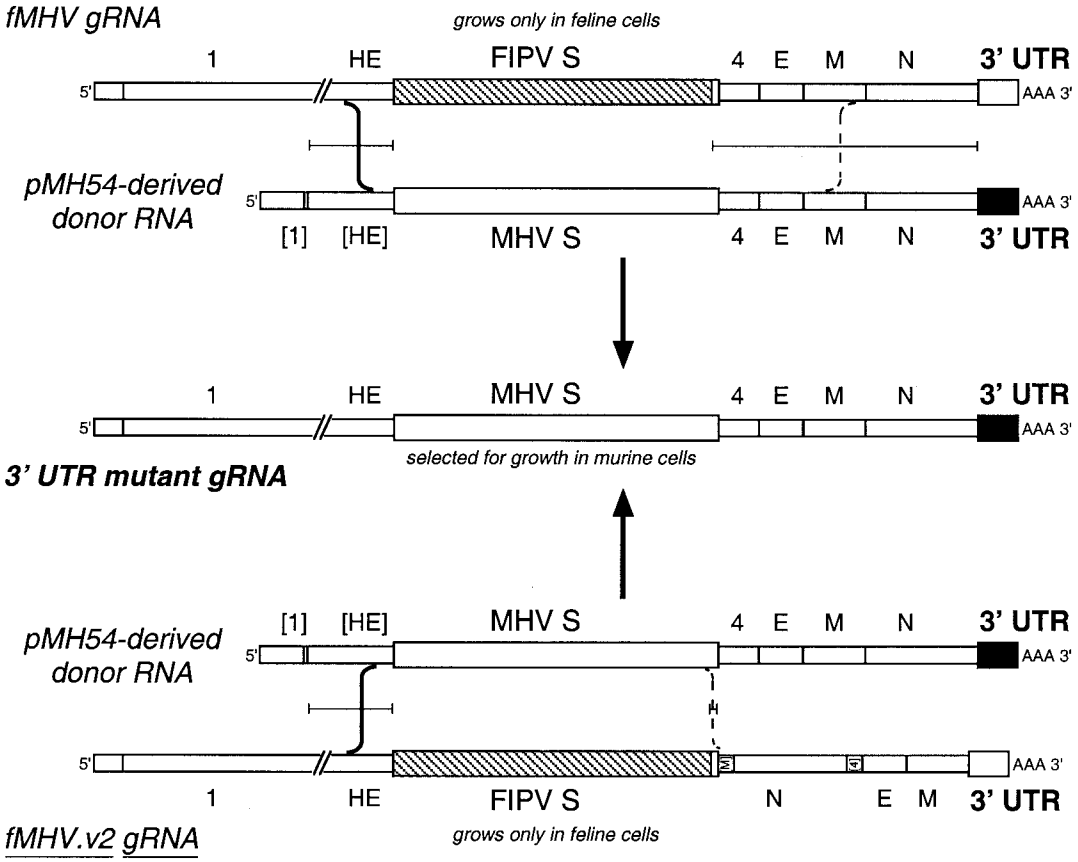

FIG. 2. Selection of MHV 3′ UTR mutants by targeted RNA recombination. Interspecies chimeric viruses fMHV (12) and fMHV.v2 each contain the ectodomain-encoding region of the FIPV S gene (hatched rectangle) and are consequently able to grow in feline cells, but not in murine cells. Mutated 3′ UTRs (black rectangles) can be transduced into either one of these recipient viral genomes by recombination with donor RNA transcribed from derivatives of plasmid pMH54. A single crossover (solid line) within the HE gene should generate a recombinant that has simultaneously reacquired the MHV S ectodomain and the ability to grow in murine cells and has also incorporated the mutated 3′ UTR. A potential second crossover (broken line), downstream of the S-gene ectodomain region, would produce a recombinant retaining the wild-type 3′ UTR. Bracketed horizontal bars indicate the boundaries of the region in which each type of crossover could occur. For pMH54-derived RNA, [1] and [HE] indicate 467- and 1,185-nt fragments of the 5′ end of the MHV genome and HE gene, respectively, as described previously (12). In the fMHV.v2 genome, [M] and [4] indicate 126- and 169-nt fragments of the M gene and gene 4, respectively, as described previously for plasmid pA122 (3).

frequency that was usually below the limits of the scale of the assay. The two types of interspecies chimeric viruses therefore allowed powerful selections for 3′ UTR mutants. Additionally, they provided clear criteria for establishing the lethality of a given mutation. A mutation was judged to be lethal if multiple targeted recombination experiments yielded only (double-crossover) wild-type recombinants with fMHV or a clean background of no recombinants with fMHV.v2, in parallel with positive controls with wild-type donor RNA that yielded a robust frequency of recombinants. A mutant was scored as viable only after it was purified by two rounds of plaque titration, and the presence of its mutation (and the absence of extraneous mutations) was confirmed by direct sequencing of reverse transcription-PCR products amplified from viral RNA.

Functional requirements of the bulged stem-loop. In a previous work, we confirmed and refined details of the structure of the bulged stem-loop, using mutational analysis and chemical and enzymatic probing (6). In particular, we found that the structures, but not the primary sequences, of stem segments C and D (Fig. 1 and 3) were essential for virus replication. By contrast, stem segment B could be disrupted without apparent consequence. Additionally, two other computer-predicted stem segments, designated A and E, were found to actually be part of the terminal loop and an internal loop, respectively (6).

To assess the effect of selective stabilization or destabilization of the essential stem segments, we took advantage of the decreasing gradient of thermodynamic stability in the progression from G-C to A-U to G-U base pairs. Increased stabilization of the bulged stem-loop had no apparent effect on viral replication. Mutants in which all base pairs were changed to G-C base pairs in stem segment C (Fig. 3, mutant 1) or in both stem segments C and D (mutant 2) each had a wild-type phenotype. Conversely, however, it was lethal to change all G-C and A-U base pairs to G-U base pairs (mutant 3) or to simply change all G-C base pairs to G-U base pairs (mutants 4 and 5). This indicated that there was a minimal essential requirement for stability of stem segments C and D.

To further examine this requirement, we first truncated stem

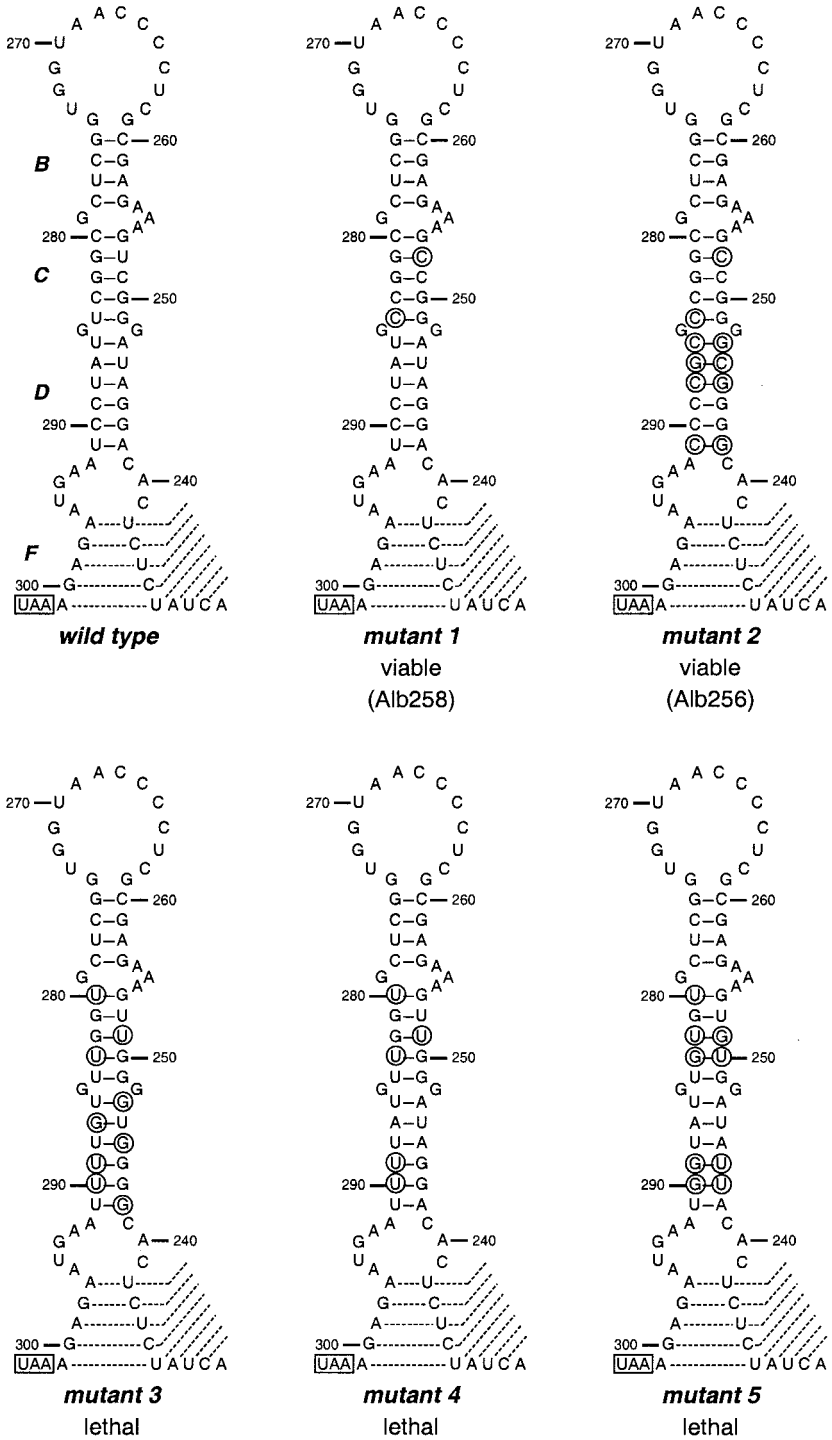

FIG. 3. Stabilization or destabilization of the bulged stem-loop. For each mutant, nucleotides that were changed from those of the wild-type sequence are circled. Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A). Broken lines indicate alternative base pairings for stem segment F or pseudoknot stem 1. Unless indicated otherwise, viable mutants have a wild-type phenotype.

segment B and the loop (Fig. 4, mutant 6), which resulted in a phenotypically wild-type virus, as did the additional disruption of a single base pair in stem segment C (mutant 7). This demonstrated that the terminal components of the bulged stem-loop structure are nonessential, in accord with previous results (6), and it also ruled out the possibility that the loop participates in a pseudoknot structure elsewhere in the genome. Destabilization of subsets of the G-C base pairs of stem segments C and D suggested that changing the G243-C290 base pair to a G-U base pair was lethal (compare mutants 8

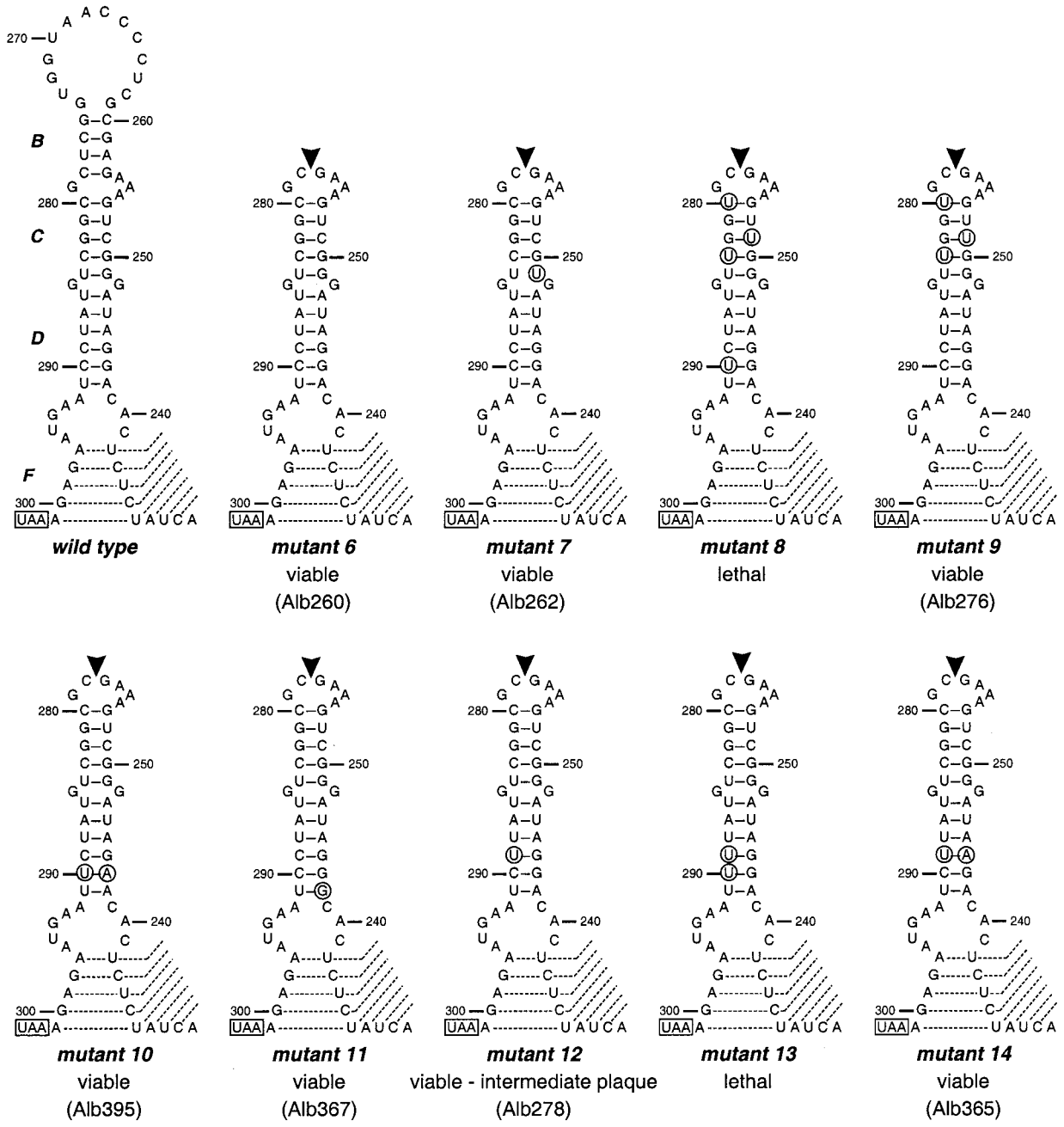

FIG. 4. Truncation and more localized destabilization of the bulged stem-loop. For each mutant, nucleotides that were changed from those of the wild-type sequence are circled. Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A), and the N-gene stop codon is boxed. Broken lines indicate alternative base pairings for stem segment F or pseudoknot stem 1. The black arrowhead indicates a deletion. Unless indicated otherwise, viable mutants have a wild-type phenotype.

and 9). However, alteration of this same base pair to A-U gave a wild-type phenotype (mutant 10). More localized mutagenesis of the region of stem D adjacent to the internal loop (mutants 11 through 14) revealed that the threshold of minimal stability was reached upon changing the G244-C289 base pair to a G-U base pair (mutant 12), which resulted in a virus with a smaller plaque size than that of the wild type.

Finally, we investigated the effect of changing the identity of one or both of the bulge bases (G248 and G285) between stem segments C and D. This analysis was performed because features that disrupt the continuity of perfect A-form RNA helices often serve as specific sites of recognition for RNA-binding proteins (22). However, we found that mutation of these two G residues to A or C, in any combination, had no effect (Fig. 5, mutants 15 through 18), and the closure of the bulge by converting the two G residues to a G-C base pair had no effect (mutant 19).

The pseudoknot in the MHV 3′ UTR is essential. Coincident with our group's finding of the bulged stem-loop in the MHV

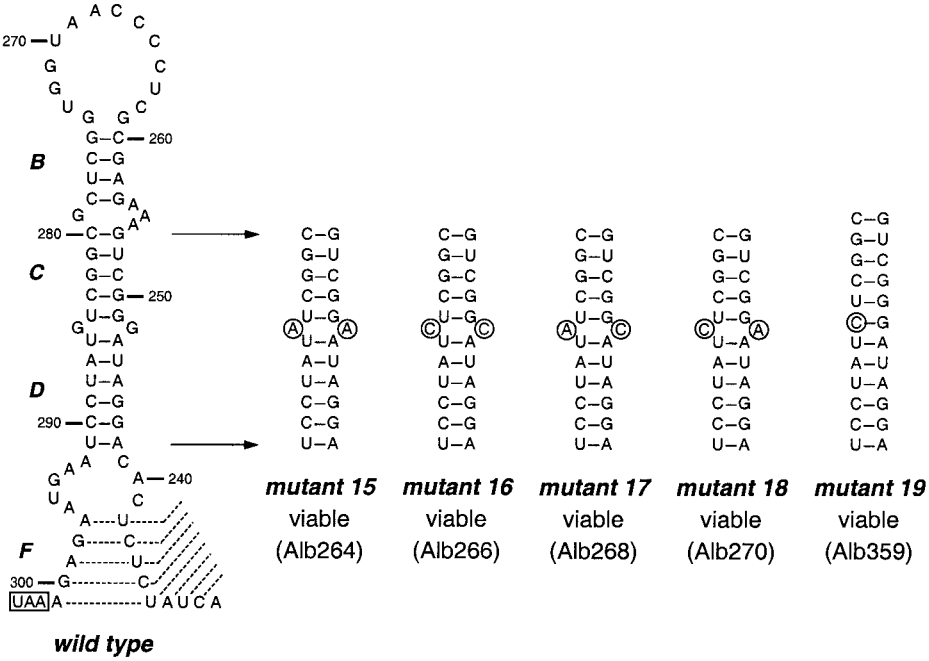

FIG. 5. Alteration of the bulge bases between stem segments C and D. For each mutant, nucleotides that were changed from those of the wild-type sequence are circled. Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A), and the N-gene stop codon is boxed. Broken lines indicate alternative base pairings for stem segment F or pseudoknot stem 1.

and BCoV 3′ UTRs (6, 7), the Brian laboratory discovered a nearby RNA pseudoknot, which they showed to be essential for BCoV DI RNA replication (37). Although their analysis almost certainly pertained equally to the nearly identical structure in the MHV 3′ UTR, it remained to be formally determined whether the pseudoknot is also operative in MHV and whether it is required in situ in genomic RNA as it is in DI RNA. To verify this, we generated classical replacement mutants in which each base of one arm of stem 2 of the pseudoknot was substituted by its complementary base. In this manner, it was found that replacement of either the downstream arm (Fig. 6, mutant 20) or the upstream arm (mutant 21) of stem 2 was lethal. However, the simultaneous replacement of both the upstream and downstream arms by their complements restored base pairing and resulted in a wild-type phenotype (mutant 22). This was also true for a mutant containing a single mismatch in the restored stem 2 (mutant 23), which had arisen serendipitously during the construction of the donor RNA vector for mutant 22. These results showed that stem 2 of the pseudoknot is indispensable for MHV replication, but as with stem segments C and D of the bulged stem-loop, stem 2 is important to maintain the pseudoknot's secondary structure, not its primary sequence.

We also sought to selectively destabilize stem 2 of the pseudoknot by changing multiple G-C base pairs to G-U base pairs. Replacement of all four of the G-C base pairs was lethal (mutant 24), but the same donor RNA yielded, via a secondary crossover between bases 206 and 211, a mutant containing the two upstream C-to-U substitutions (mutant 25). It was also possible to create three G-U base pair substitutions spaced along the length of the stem (mutant 26), but the concentration of three such changes at one end of the stem was not allowed (mutant 27). Thus, just as with the bulged stem-loop, pseudoknot stem 2 has a minimal threshold of required stability that cannot be crossed.

Structural requirements of the pseudoknot. In seeking ways to affect the function of the 3′ UTR pseudoknot, we drew on results from studies of some of the most completely characterized RNA pseudoknots, those involved in ribosomal frameshifting (4, 10). In this class of structures, overall conformation is extremely sensitive to the composition of the junction between the two stems, which form a quasicontinuous helix. We thus focused on A205, the interhelical bulge base of the 3′ UTR pseudoknot. Surprisingly, deletion of this base did not have any apparent effect on viral replication (Fig. 7, mutant 28). A wild-type phenotype was also retained when A230, in loop 1, was changed to U (mutant 29), an alteration that would allow it to base pair with A205, thereby eliminating the interhelical bulge base and simultaneously increasing stem 1 from 8 to 9 bp. A measurable effect was obtained only by changing A205 to C (mutant 30), which allowed it to base pair with G195, thereby extending stem 2 from 10 to 11 bp. The smaller plaque phenotype of this mutant likely resulted from the concomitant reduction of loop 2 to a single base, which would be expected to distort the angle between the two stems (4).

Another strategy that we followed was to make insertions into loop 1 of the pseudoknot, between bases 226 and 227, thereby creating a greater spacing between the two arms of stem 1. Since the upstream arm of stem 1 is alternately a component of the terminal stem segment of the bulged stem-loop (segment F), we expected that such insertions might differentially affect the two structural elements. Initially, insertions of 5, 10, and 15 bases were made (mutants 31, 39, and 40, respectively); the smallest of these was tolerated, but the larger

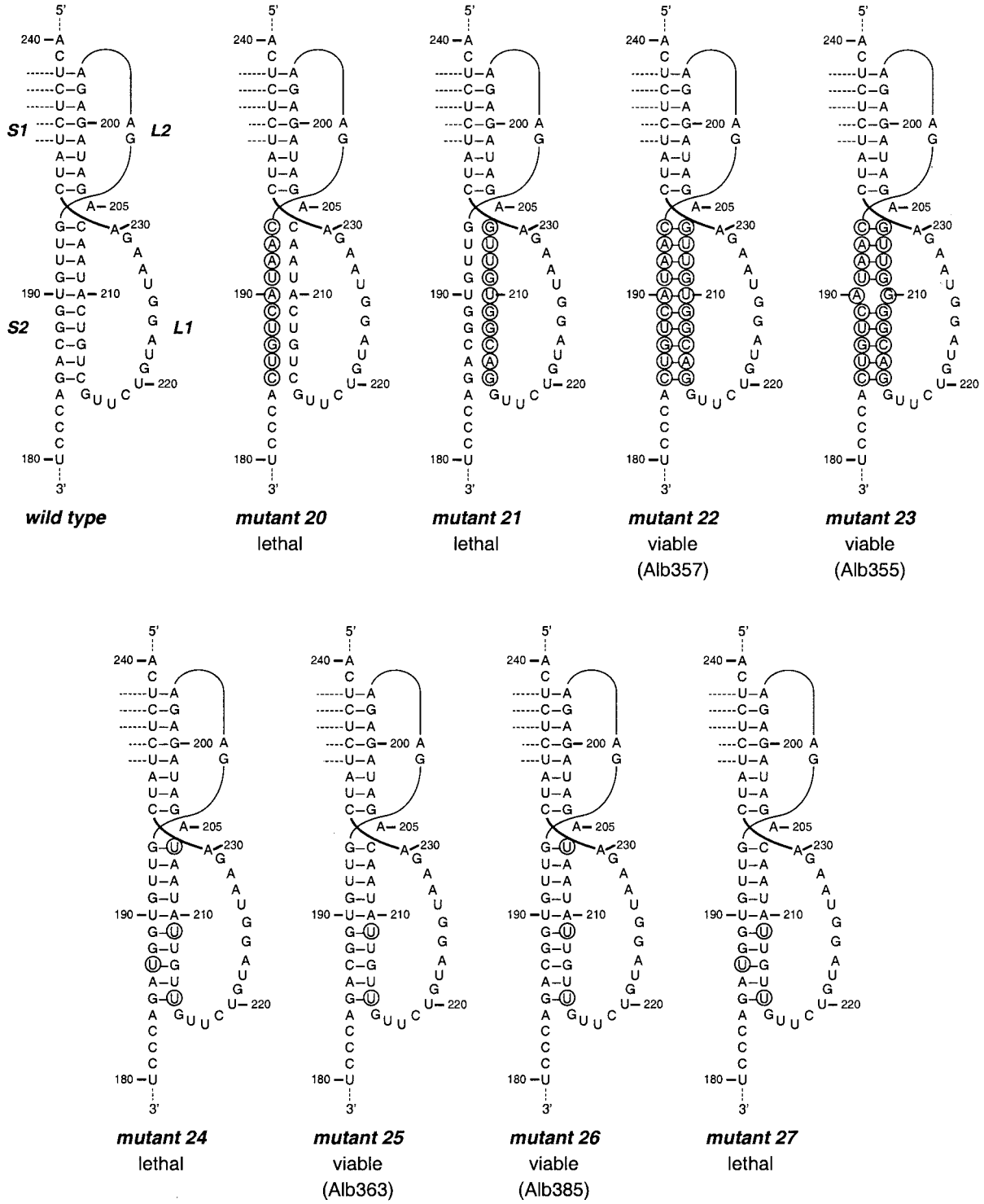

FIG. 6. Disruption or destabilization of pseudoknot stem 2. To clarify the relationships among the components of the pseudoknot, we show a more standard representation of pseudoknot stem 1 here (different from the schematic in Fig. 1). For each mutant, nucleotides that were changed from those of the wild-type sequence are circled. Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A). Broken lines indicate alternative base pairings for stem segment F or pseudoknot stem 1. Unless indicated otherwise, viable mutants have a wild-type phenotype.

two were lethal. A stepwise retreat downward from 10 inserted bases revealed that as few as six bases of the original insertion, AACAAC, were lethal (mutants 35 through 38). Notably, however, when the 3′ base of the six-base insertion was changed to each of the other three possible alternatives, we found that insertion of AACAAA (mutant 32) or AACAAG (mutant 33) was viable, whereas AACAAU (mutant 34) was lethal.

The AACAAA insertion mutant (mutant 32), like the five-

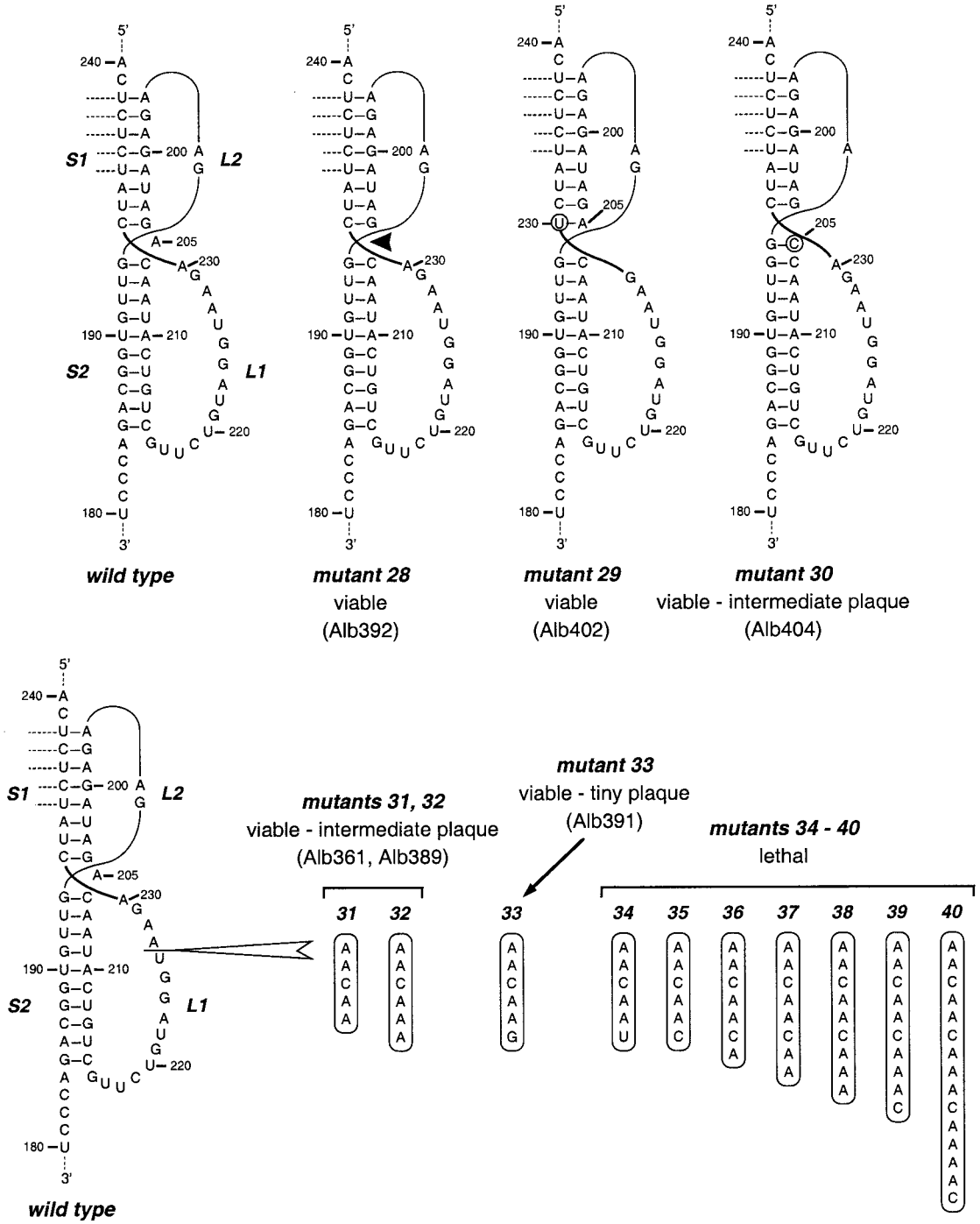

FIG. 7. Destabilization of pseudoknot components. For each mutant, nucleotides that were changed from those of the wild-type sequence are circled. Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A). Broken lines indicate alternative base pairings for stem segment F or pseudoknot stem 1. The black arrowhead (mutant 28) indicates a deletion. The white arrowhead (mutants 31 through 40) indicates an insertion of variable size. Unless otherwise indicated, viable mutants have a wild-type phenotype.

base insertion mutant (mutant 31), formed smaller plaques than did the wild type. The plaques of AACAAG insertion mutant 33, however, were tiny compared to those of the wild type (Fig. 8), displaying the most dramatic phenotype of any 3′ UTR mutant that we have isolated to date. The plaque size difference was manifested similarly at 33 and 39°C as at 37°C. Moreover, the AACAAG insertion mutant was found to be very unstable, exhibiting a high frequency of revertants arising in a single passage or even within the duration of a plaque assay. Many of the larger plaques of the mutant in Fig. 8 are those that have begun to be outgrown by a revertant. Possibly because of the severity of this mutant's defect, we have thus far

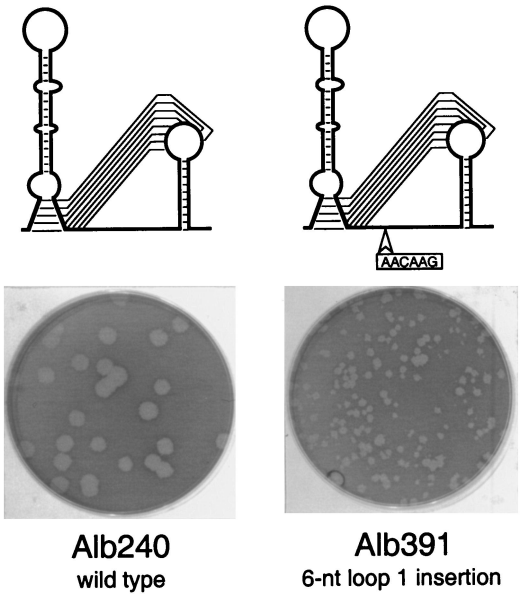

FIG. 8. Plaque phenotype of loop 1 insertion mutant Alb391 compared to wild type. Plaque titrations were performed on mouse L2 cells at 37°C. Monolayers were stained with neutral red at 48 h postinfection and photographed 24 h later.

been able to obtain it in only one independent isolate (Alb391).

We have begun to analyze a small number of independent revertants of Alb391. One of these has lost the first five bases of the insertion, leaving a footprint of a single G residue insertion. This finding argues strongly that the lesion in Alb391 is indeed due to the constructed six-base insertion and that it is not the result of some other spontaneous mutation. Interestingly, the other large-plaque revertants examined thus far retain the six-base insertion and therefore must contain one or more second-site reverting mutations. Remarkably, none of these map anywhere within the 3′ UTR. For one particular revertant, the suppressor mutation is also not found within bases 30 through 510 of the *cis*-acting replication signal at the 5′ end of the genome (11, 16, 34), nor is it located within the region of the packaging signal (26), at bases 20196 through 20455 of the genome. Further analysis of these revertants has the potential to reveal a genetic interaction between the pseudoknot and a more distant *cis*-acting element or gene product encoded by the MHV genome.

The bulged stem-loop and pseudoknot, although mutually exclusive, are both essential. We and others have presented considerable evidence that the 3′ UTR bulged stem-loop and pseudoknot, independently, are essential for MHV or BCoV replication (6, 7, 37). However, bases 234 through 238 of the MHV 3′ UTR (equivalent to bases 222 through 226 of the BCoV 3′ UTR) participate in both of these structures, forming either stem segment F of the bulged stem-loop or part of stem 1 of the pseudoknot. We previously showed that replacement of each base in the upstream arm of stem F by its complement (Fig. 9, mutant 41) was lethal to viral replication, as a result of disruption of stem F. Similarly, replacement of the downstream arm of stem F by complementary bases (mutant 42) was lethal, due to disruption of stem F, and as we later postulated, disruption of stem 1 of the pseudoknot. Moreover, in contrast to other classical stem replacement mutants, the simultaneous replacement of both the upstream and downstream arms of stem F by complementary bases (mutant 43) was also lethal, despite restoring the base pairing of stem F. This result was best explained by the fact that 5 of the 8 bp of pseudoknot stem 1 remained disrupted. We tested this notion by similarly disrupting the same 5 bp through replacement of the corresponding bases in the downstream arm of stem 1 (mutant 44); as expected, this alteration was also found to be lethal. However, the simultaneous replacement of all three segments—the upstream arm of stem F, the downstream arm of stem F, and the overlapping region of stem 1—gave rise to a viable virus (mutant 45), albeit one with plaques somewhat smaller than those of the wild type. This indicated that there is a strict requirement for both sets of alternative base pairings to be formed during the course of viral replication. It must be noted that this triple stem substitution mutant (Alb387) has been independently isolated only once. Thus, at present, we cannot rule out the possibility that Alb387 also harbors some compensating second-site mutation, although we have established that no such mutation occurs in the 3′ UTR or in the *cis*-acting replication signal at the 5′ end of the genome.

To qualitatively assess the net effect on RNA synthesis in the two most striking mutants in this study, we metabolically labeled RNA in cells infected with the triple stem substitution mutant (Alb387) and the AACAAG pseudoknot loop 1 insertion mutant (Alb391). For a control, we used Alb240, a wild-type MHV recombinant that had been reconstructed from fMHV and pMH54 donor RNA (14). No significant difference in RNA synthesis by the two viral mutants, compared to the wild type, was noted when equal amounts of labeled RNA were electrophoretically analyzed (Fig. 10). Neither mutant exhibited a gross change in the amount of any individual sgRNA or in the relative amounts of sgRNA and genomic RNA. Since the mutants have a lower (Alb387) or much lower (Alb391) growth rate than that of the wild type, we thus infer that the net effect of the constructed mutations must be to attenuate the kinetics of RNA synthesis. One caveat to this conclusion is that Alb391 has such a high rate of reversion that we cannot be sure that the bulk of labeled RNA for this sample represented RNA of the mutant rather than of revertants. More subtle assays will need to be developed to specifically pinpoint the step(s) of viral RNA synthesis that is impaired in the mutants. At this point, our data do not allow us to distinguish whether negative-strand or positive-strand RNA synthesis or both are impaired in the mutants.

DISCUSSION

A principal goal of this study was to delineate the relationship between the two essential RNA secondary structures that have been discovered at the upstream ends of the 3′ UTRs of group 2 coronaviruses. Another goal was to design mutants with defects in these RNA structures that would potentially present an opportunity to unravel the functional significance of these *cis*-acting elements. Toward these ends, we have dissected out nonessential regions of the bulged stem-loop and pseudoknot, and we have defined the extent to which essential

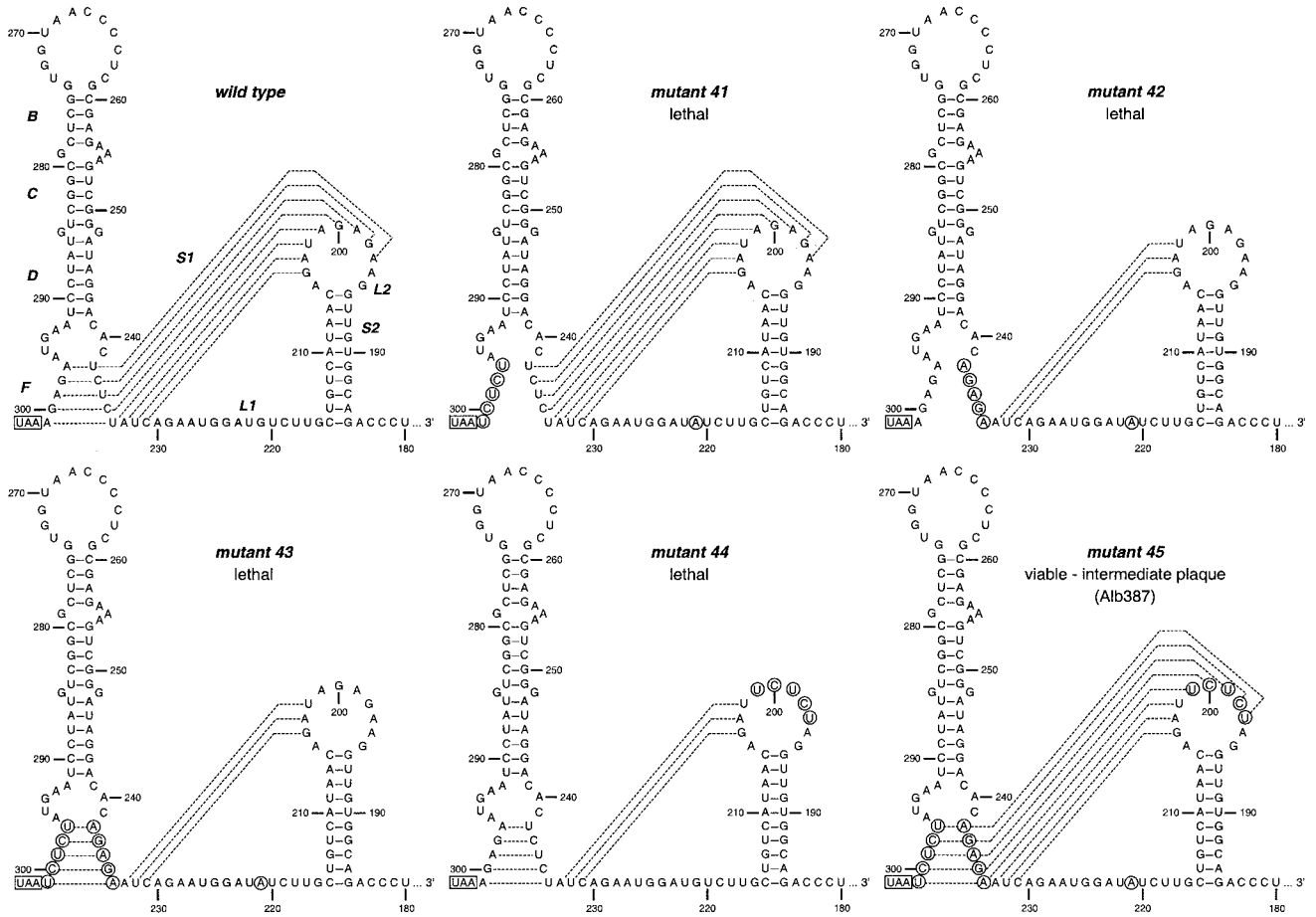

FIG. 9. Essential nature of the overlap of stem segment F and pseudoknot stem 1. Mutants 41 through 43 have been reported previously (designated mutants MFL, MFR, and MFLR, respectively) (6). For each mutant, nucleotides that were changed from those of the wild-type sequence are circled. Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A), and the N-gene stop codon is boxed. Broken lines indicate alternative base pairings for stem segment F or pseudoknot stem 1. Unless indicated otherwise, viable mutants have a wild-type phenotype.

regions of each element can be destabilized or distorted prior to loss of function.

For the bulged stem-loop, results presented here provide an extension of our previous work (6, 7). The loop of this structure, as well as the adjacent stem segment B, can be truncated without affecting viral replication in tissue culture (Fig. 4). This finding is consistent with the previous demonstration that disrupting stem B or introducing sequence variation into the loop was without consequence (6). Conversely, the next two stem segments, C and D, had previously been found to be essential, although the sequences of their arms could be exchanged as long as base pairing was maintained (6). We have now found that numerous sequence changes in these stem segments are tolerated, providing that some key positions near the bottom of stem D are not destabilized by mutation to G-U base pairs (Fig. 3 and 4). Surprisingly, the identities of the symmetric bulge bases between stems C and D are not critical (Fig. 5).

For the pseudoknot, the base pairing, but not the primary sequence, of stem 2 was shown to be essential. Similar to stems C and D, stem 2 tolerated a limited degree of destabilization by introduction of G-U base pairs (Fig. 6). However, in contrast to ribosomal frameshifting pseudoknots (4, 10), the 3′ UTR pseudoknot was almost entirely insensitive to the deletion or introduced base pairing of the interhelical bulge base (Fig. 7). As noted by Williams and colleagues (37), it is highly unlikely that the 3′ UTR pseudoknot is involved in ribosomal frameshifting. Additionally, unlike ribosomal frameshifting pseudoknots, the 3′ UTR pseudoknot has a large loop 1 and a small loop 2. This inverted configuration may suggest that, whereas ribosomal frameshifting pseudoknots are approached by ribosomes from the upstream side, the apparatus that interacts with the 3′ UTR pseudoknot approaches it from the downstream side.

One region of the 3′ UTR pseudoknot that is remarkably sensitive to alteration is the 15-base loop 1. We found that insertion of as few as six additional bases into this stretch could be lethal (Fig. 7). For the larger insertions of 10 or 15 bases, the insertion may have introduced an unallowed spacing, or the extra bases may have occluded some tertiary interaction. However, for the smaller insertions, there was an apparent sequence dependence: a six-base insertion of AACAAA or AA CAAG was viable, but insertion of AACAAC or AACAAU

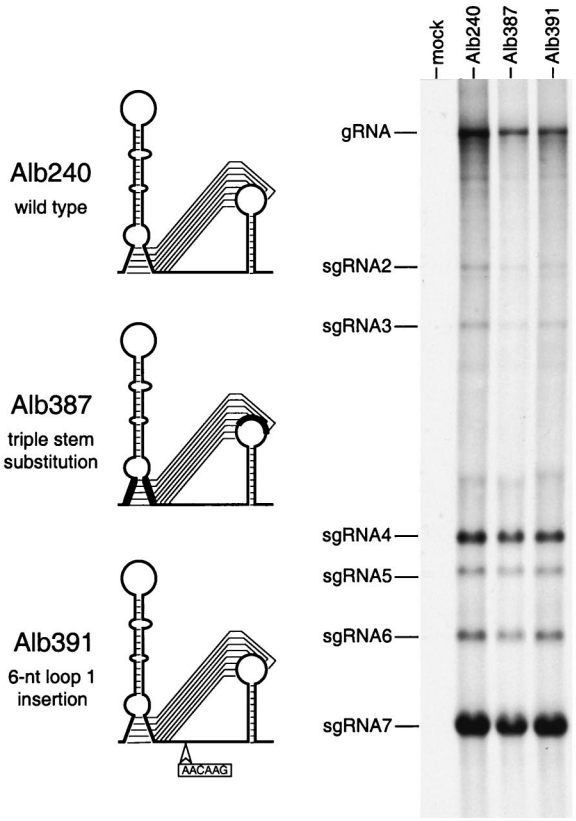

FIG. 10. RNA synthesis by 3′ UTR mutants Alb387 (mutant 45 in Fig. 9) and Alb391 (mutant 33 in Fig. 7) compared to wild-type Alb240. Infected or mock-infected cells were metabolically labeled with $^{33}P_i$, and RNA was isolated and analyzed as described in Materials and Methods. gRNA, genomic RNA.

was lethal. This observation raises the possibility that loop 1 forms essential base pairs with some other site in the genome, but as yet we have found no tenable candidate partner for such a relationship. The striking phenotype of the AAACAG insertion mutant (Alb391 [Fig. 7 and 8]), however, raises the possibility that we can uncover a longer-range genomic interaction through revertant analysis.

The most salient result from the work presented here is our confirmation of the previously hypothesized interrelationship of the bulged stem-loop and pseudoknot. The isolated disruption of stem segment F of the bulged stem-loop was lethal (Fig. 9, mutant 41), and the isolated disruption of stem 1 of the pseudoknot was also lethal (mutants 43 and 44). Similarly, the disruption of the shared segment of both elements was lethal (mutant 42). Viral viability was regained only upon the simultaneous restoration of both elements, through the interchange of all three affected stem segments (mutant 45). The essential nature of two sets of mutually exclusive structural interactions supports our previous suggestion that the stem-loop and pseudoknot are components of a molecular switch. Our working model is that one of these elements provides a critical function at one stage of viral replication, while the other acts in a regulatory role, to turn off that function at another stage of viral replication. An open question, then, is the nature of the putative function governed by this switch. The minimal signal required for initiation of MHV negative-strand RNA synthesis has been mapped to the last 55 bases of the 3′ UTR plus some extent of poly(A) (17). It is possible, however, that the upstream molecular switch could be involved in a decision of whether to elongate the nascent negative-strand RNA. Alternatively, the switch might function in circularization of the genome, regulation of genome translation versus replication, or localization of genomic RNA to the site of RNA synthesis. Also yet to be determined are the viral and cellular components with which the switch interacts. The Leibowitz laboratory has described a complex RNA secondary structure that occupies most of the 3′ UTR downstream of the pseudoknot (18) (Fig. 1). It is currently not clear whether this structure represents a separate module, or whether it and the switch elements are integrated in some way that remains to be recognized. Additionally, future work must be done to define the roles of RNA-binding proteins that have been found to recognize various regions of the 3′ UTR and the poly(A) tail (8, 9, 19, 25, 33, 38, 39).

An ancillary conclusion that emerges from our study is that the 3′ *cis*-acting region required for MHV replication extends no further upstream than the 3′ UTR. Previous results from a number of MHV DI RNA experiments have suggested that, in addition to the 3′ UTR, viral RNA replication requires some 77 to 162 nt of the 3′ end of the adjacent N gene (11, 16, 34). However, the replicative competence of fMHV.v2 and its MHV counterpart, MHV-SmNEM (3), clearly shows there is no essential *cis*-acting region in the N gene within the intact viral genome. Or, if such a region does exist, it can act at a distance of more than 1,457 nt. The above observation is consistent with results obtained with the group 1 and group 3 coronaviruses transmissible gastroenteritis virus and infectious bronchitis virus (2, 23), for which there are replicating DI RNAs that are totally devoid of their respective N genes. By contrast, it has been shown that 3′ *cis*-acting replication signals in both the equine (24) and porcine arteriviruses (36) spread from their relatively small 3′ UTRs into adjacent coding regions.

An additional phylogenetic contrast is found in a comparison of the three coronavirus groups. The genomes of all group 1 coronaviruses (exemplified by transmissible gastroenteritis virus) have a highly conserved pseudoknot at the upstream end of their 3′ UTRs (37), but no counterpart for the bulged stem-loop structure can be found anywhere nearby. Conversely, the genomes of the group 3 coronaviruses (exemplified by infectious bronchitis virus) have a highly conserved stem-loop structure at the upstream end of their 3′ UTRs (2), but only a weak candidate for the pseudoknot structure can be found in a neighboring position (37). Paradoxically, only the group 2 coronaviruses have both elements, and, as we have shown, both are essential. It is not clear how the requirement for one of the two elements has been overcome in the group 1 and group 3 coronaviruses, but this contrasting pattern appears to suggest that divergent pathways have evolved to perform the same required functions. Most noteworthy in this regard, we have found that the 3′ UTR of the genome of the recently emerged severe acute respiratory syndrome (SARS)-associated coronavirus (SARS-CoV) contains both the bulged stem-loop and pseudoknot, and they overlap in precisely the same

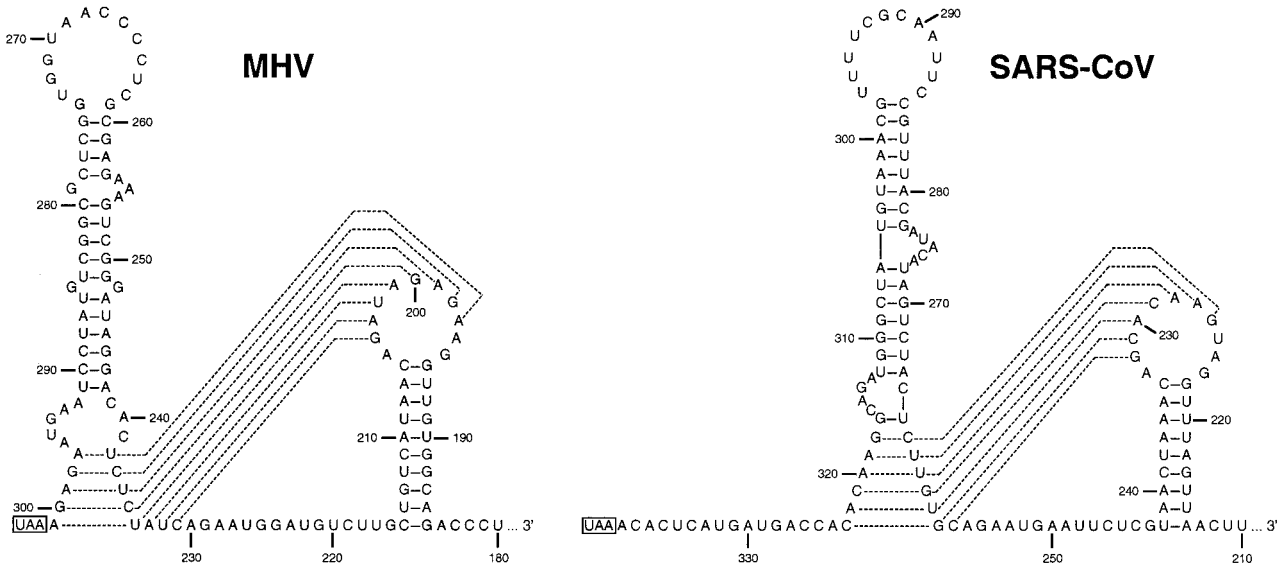

FIG. 11. Conservation of the 3′ UTR bulged stem-loop and overlapping pseudoknot structure between group 2 coronaviruses (exemplified by MHV) and the SARS-associated coronavirus (SARS-CoV).

manner as in MHV and BCoV (Fig. 11). The earliest analyses of the genome of the SARS-CoV suggested that it represents the first identified member of a new group (group 4) of the coronavirus family (20, 29), but a recent work places this virus closer to the group 2 lineage (32). Our finding would lend support to the latter assignment, and it further compels us to attempt to understand the functional role of the 3′ UTR RNA secondary structures.

ACKNOWLEDGMENTS

We are grateful to Lili Kuo for critical advice during the course of this work. We thank the Molecular Genetics Core Facility of the Wadsworth Center for oligonucleotide synthesis and DNA sequencing.

This work was supported in part by Public Health Service grants AI 45695 and AI 39544 from the National Institutes of Health.

REFERENCES

1. **Baric, R. S., and B. Yount.** 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. J. Virol. **74:**4039–4046.
2. **Dalton, K., R. Casais, K. Shaw, K. Stirrups, S. Evans, P. Britton, T. D. K. Brown, and D. Cavanagh.** 2001. *cis*-acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. J. Virol. **75:**125–133.
3. **de Haan, C. A. M., H. Volders, C. A. Koetzner, P. S. Masters, and P. J. M. Rottier.** 2002. Coronaviruses maintain viability despite dramatic rearrangements of the strictly conserved genome organization. J. Virol. **76:**12491–12493.
4. **Giedroc, D. P., C. A. Theimer, and P. L. Nixon.** 2000. Structure, stability, and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. J. Mol. Biol. **298:**167–185.
5. **Horton, R. M., and L. R. Pease.** 1991. Recombination and mutagenesis of DNA sequences using PCR, p. 217–247. *In* M. J. McPherson (ed.), Directed mutagenesis, a practical approach. IRL Press, New York, N.Y.
6. **Hsue, B., T. Hartshorne, and P. S. Masters.** 2000. Characterization of an essential RNA secondary structure in the 3′ untranslated region of the murine coronavirus genome. J. Virol. **74:**6911–6921.
7. **Hsue, B., and P. S. Masters.** 1997. A bulged stem-loop structure in the 3′ untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. J. Virol. **71:**7567–7578.
8. **Huang, P., and M. M. C. Lai.** 1999. Polypyrimidine tract-binding protein binds to the complementary strand of the mouse hepatitis virus 3′ untranslated region, thereby altering conformation. J. Virol. **73:**9110–9116.
9. **Huang, P., and M. M. C. Lai.** 2001. Heterogeneous nuclear ribonucleoprotein A1 binds to the 3′ untranslated region and mediates potential 5′-3′-end cross talks of mouse hepatitis virus RNA. J. Virol. **75:**5009–5017.
10. **Kim, Y.-G., L. Su, S. Maas, A. O'Neill, and A. Rich.** 1999. Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. Proc. Natl. Acad. Sci. USA **96:**14234–14239.
11. **Kim, Y.-N., Y. S. Jeong, and S. Makino.** 1993. Analysis of *cis*-acting sequences essential for coronavirus defective interfering RNA replication. Virology **197:**53–63.
12. **Kuo, L., G.-J. Godeke, M. J. B. Raamsman, P. S. Masters, and P. J. M. Rottier.** 2000. Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. J. Virol. **74:**1393–1406.
13. **Kuo, L., and P. S. Masters.** 2002. Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus. J. Virol. **76:**4987–4999.
14. **Kuo, L., and P. S. Masters.** 2003. The small envelope protein E is not essential for murine coronavirus replication. J. Virol. **77:**4597–4608.
15. **Lai, M. M. C., and D. Cavanagh.** 1997. The molecular biology of coronaviruses. Adv. Virus Res. **48:**1–100.
16. **Lin, Y.-J., and M. M. C. Lai.** 1993. Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontiguous sequence for replication. J. Virol. **67:**6110–6118.
17. **Lin, Y.-J., C.-L. Liao, and M. M. C. Lai.** 1994. Identification of the *cis*-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription. J. Virol. **68:**8131–8140.
18. **Liu, Q., R. F. Johnson, and J. L. Leibowitz.** 2001. Secondary structural elements within the 3′ untranslated region of mouse hepatitis virus strain JHM genomic RNA. J. Virol. **75:**12105–12113.
19. **Liu, Q., W. Yu, and J. L. Leibowitz.** 1997. A specific host cellular protein binding element near the 3′ end of mouse hepatitis virus genomic RNA. Virology **232:**74–85.
20. **Marra, M. A., S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattra, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girn, O. L. Griffith, S. R. Leach, M. Mayo, H. McDonald, S. B. Montgomery, P. K. Pandoh, A. S. Petrescu, A. G. Robertson, J. E. Schein, A. Siddiqui, D. E. Smailus, J. M. Stott, G. S. Yang, F. Plummer, A. Andonov, H. Artsob, N. Bastien, K. Bernard, T. F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Stroher, G. A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R. C. Brunham, M. Krajden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper.** 2003. The genome sequence of the SARS-associated coronavirus. Science **300:**1399–1404.
21. **Masters, P. S., C. A. Koetzner, C. A. Kerr, and Y. Heo.** 1994. Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis virus. J. Virol. **68:**328–337.
22. **Mattaj, I. W.** 1993. RNA recognition: a family matter? Cell **73:**837–840.
23. **Méndez, A., C. Smerdou, A. Izeta, F. Gebauer, and L. Enjuanes.** 1996. Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. Virology **217:**495–507.

24. **Molenkamp, R., H. van Tol, B. C. D. Rozier, Y. van der Meer, W. J. M. Spaan, and E. J. Snijder.** 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. J. Gen. Virol. **81:**2491–2496.
25. **Nanda, S. K., and J. L. Leibowitz.** 2001. Mitochondrial aconitase binds to the 3′ untranslated region of the mouse hepatitis virus genome. J. Virol. **75:** 3352–3362.
26. **Narayanan, K., and S. Makino.** 2001. Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging. J. Virol. **75:**9059–9067.
27. **Parker, M. M., and P. S. Masters.** 1990. Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. Virology **179:**463–468.
28. **Pasternak, A. O., E. van den Born, W. J. M. Spaan, and E. J. Snijder.** 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J. **20:**7220–7228.
29. **Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science **300:**1394–1399.
30. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Sawicki, S. G., and D. L. Sawicki.** 1998. A new model for coronavirus transcription. Adv. Exp. Med. Biol. **440:**215–219.
32. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. **331:**991–1004.
33. **Spagnolo, J. F., and B. G. Hogue.** 2000. Host protein interactions with the 3′ end of bovine coronavirus RNA and the requirement of the poly(A) tail for coronavirus defective genome replication. J. Virol. **74:**5053–5065.
34. **van der Most, R. G., W. Luytjes, S. Rutjes, and W. J. M. Spaan.** 1995. Translation but not the encoded sequence is essential for the efficient propagation of defective interfering RNAs of the coronavirus mouse hepatitis virus. J. Virol. **69:**3744–3751.
35. **van Marle, G., J. C. Dobbe, A. P. Gultyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. Proc. Natl. Acad. Sci. USA **96:**12056–12061.
36. **Verheije, M. H., R. C. Olsthoorn, M. V. Kroese, P. J. M. Rottier, and J. J. Meulenberg.** 2002. Kissing interaction between 3′ noncoding and coding sequences is essential for porcine arterivirus RNA replication. J. Virol. **76:**1521–1526.
37. **Williams, G. D., R. Y. Chang, and D. A. Brian.** 1999. A phylogenetically conserved hairpin-type 3′ untranslated region pseudoknot functions in coronavirus RNA replication. J. Virol. **73:**8349–8355.
38. **Yu, W., and J. L. Leibowitz.** 1995. Specific binding of host cellular proteins to multiple sites within the 3′ end of mouse hepatitis virus genomic RNA. J. Virol. **69:**2016–2023.
39. **Yu, W., and J. L. Leibowitz.** 1995. A conserved motif at the 3′ end of mouse hepatitis virus genomic RNA required for host protein binding and viral RNA replication. Virology **214:**128–138.