JOURNAL OF VIROLOGY, Feb. 2008, p. 1214–1228 Vol. 82, No. 3
0022-538X/08/$08.00+0 doi:10.1128/JVI.01690-07

Genetic Interactions between an Essential 3′ *cis*-Acting RNA Pseudoknot, Replicase Gene Products, and the Extreme 3′ End of the Mouse Coronavirus Genome▽

Roland Züst,[1] Timothy B. Miller,[2] Scott J. Goebel,[2] Volker Thiel,[1] and Paul S. Masters[2]*

Research Department, Kantonal Hospital St. Gallen, St. Gallen, Switzerland,[1] and Wadsworth Center, New York State Department of Health, Albany, New York 12201[2]

Received 3 August 2007/Accepted 11 November 2007

The upstream end of the 3′ untranslated region (UTR) of the mouse hepatitis virus genome contains two essential and overlapping RNA secondary structures, a bulged stem-loop and a pseudoknot, which have been proposed to be elements of a molecular switch that is critical for viral RNA synthesis. It has previously been shown that a particular six-base insertion in loop 1 of the pseudoknot is extremely deleterious to the virus. We have now isolated multiple independent second-site revertants of the loop 1 insertion mutant, and we used reverse-genetics methods to confirm the identities of suppressor mutations that could compensate for the original insertion. The suppressors were localized to two separate regions of the genome. Members of one class of suppressor were mapped to the portions of gene 1 that encode nsp8 and nsp9, thereby providing the first evidence for specific interactions between coronavirus replicase gene products and a *cis*-acting genomic RNA element. The second class of suppressor was mapped to the extreme 3′ end of the genome, a result which pointed to the existence of a direct base-pairing interaction between loop 1 of the pseudoknot and the genomic terminus. The latter finding was strongly supported by phylogenetic evidence and by the construction of a deletion mutant that reduced the 3′ UTR to its minimal essential elements. Taken together, the interactions revealed by the two classes of suppressors suggest a model for the initiation of coronavirus negative-strand RNA synthesis.

RNA virus genomes contain *cis*-acting sequences and structures that interact with viral and cellular components to initiate a variety of essential functions. Such interactions can mediate RNA replication and transcription, the selective translation of viral transcripts, and the specific packaging of progeny genomic RNA (gRNA) into assembling virions. Initiation of RNA virus genome replication occurs at the extreme 3′ end of gRNA in order to produce a strand of complementary polarity. For positive-strand RNA viruses, the initiation of negative-strand RNA synthesis requires the specific recognition of the viral gRNA template and the juxtaposition of replicative proteins and the 3′ genomic terminus. *cis*-acting sequences are believed to play pivotal roles in these processes. An intriguing example of how *cis*-acting sequences can coordinate the assembly and targeting of a replication complex to the initiation site of negative-strand RNA synthesis has been provided by studies of poliovirus replication (4, 23, 57). A crucial initial event in this system is the formation of a ribonucleoprotein complex around a cloverleaf RNA structure at the 5′ end of the poliovirus genome. This complex consists of the poliovirus protease-polymerase precursor 3CD and the cellular poly(rC) binding protein, both of which bind to the 5′ cloverleaf (1, 2, 7, 17). In order to target poliovirus 3CD to the initiation site of negative-strand RNA synthesis, another cellular protein, poly(A) binding protein, becomes involved. Poly(A) binding protein is associated with the poly(A) tail of the poliovirus genomic RNA (which also serves as a *cis*-acting replication element) and bridges the 5′ and 3′ genomic termini by binding to the poly(rC) binding protein and 3CD (4, 23). Thus, in this example, *cis*-acting elements provide not only the basis for the discrimination between viral and cellular RNA but also a platform for the assembly of a negative-strand replication initiation complex.

Coronaviruses are a family of positive-strand RNA viruses possessing a complex, multicomponent RNA synthetic machinery and a sophisticated program of gene expression that is characterized by the transcription of a 3′-nested set of subgenomic RNAs (sgRNAs), in addition to gRNA replication (41). These features are shared with other members of the nidovirus order, which also includes arteriviruses, toroviruses, and roniviruses. For these viruses, the synthesis of both gRNA and sgRNA begins with the initiation of negative-strand RNA at the 3′ end of the genome (49). A prerequisite for this event is the translation of the replicase gene from the positive-strand RNA genome, resulting in two large polyproteins. Extensive self-processing by viral proteinases residing within the polyproteins eventually generates a number of processing intermediates and 15 to 16 mature end products, termed nonstructural proteins (nsp's) 1 to 16 (nsp1 to nsp16) (67). Many nsp's colocalize, together with the structural nucleocapsid (N) protein, at double-membrane vesicles to form the membrane-associated replication complex (8, 10, 14, 22, 53). Although this complex is believed to interact with *cis*-acting elements in order to replicate and transcribe viral RNAs (9), our knowledge about specific interactions that may be involved in the various

* Corresponding author. Mailing address: David Axelrod Institute, Wadsworth Center, NYSDOH, New Scotland Avenue, P.O. Box 22002, Albany, New York 12201-2002. Phone: (518) 474-1283. Fax: (518) 473-1326. E-mail: masters@wadsworth.org.
▽ Published ahead of print on 21 November 2007.

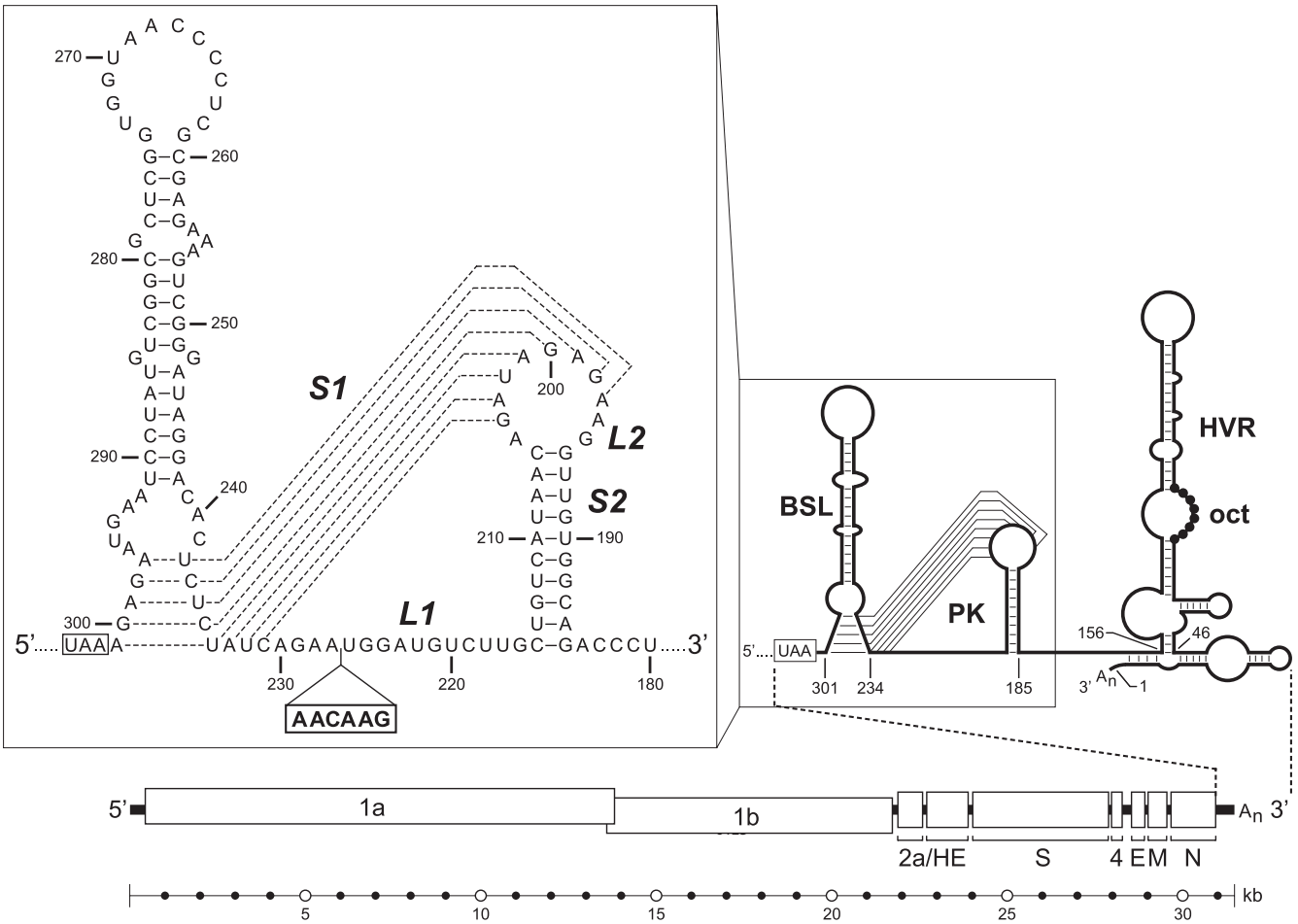


FIG. 1. Structure of the MHV 3′ UTR. The organization of the 31.3-kb MHV genome is shown at the bottom. Above this, at the right, is a schematic of the current model of the RNA secondary structure of the entire 301-nt 3′ UTR. The 3′ UTR comprises an essential bulged stem-loop (BSL; nt 234 through 301) (18, 26, 27) and overlapping pseudoknot (PK; nt 185 through 238) (18, 60, 63). Downstream of these structures is a nonessential HVR (nt 46 through 156), which contains the highly conserved coronavirus octanucleotide motif (oct), 5′-GGAAGAGC-3′ (19, 39). The arrangements shown for the segment downstream of the HVR (nt 1 through 67) and for part of the segment between the HVR and the pseudoknot (nt 143 through 160) are those proposed by Liu and coworkers (38) but are modified in the work reported herein. At the upper left is a detailed view of the upstream end of the 3′ UTR, denoting the pseudoknot stems (S1 and S2) and loops (L1 and L2). Broken lines represent alternative base pairings either for the bottom stem segment of the bulged stem-loop or for pseudoknot stem 1. Also indicated is the 6-nt insertion (AACAAG) in pseudoknot loop 1 that was constructed in the mutant Alb391 (18). All nucleotide numbering is from the first base at the 3′ end of the genome, excluding poly(A), and the N-gene stop codon is boxed.

steps of coronavirus RNA synthesis is extremely limited. RNA-binding activity has been ascribed to a number of replicase gene products (16, 32, 55, 56, 66) and to the N protein (41). However, currently available structural and biochemical data do not yet allow the assignment of any of these RNA-binding proteins to defined interactions with *cis*-acting elements. Additionally, it is not known whether cellular proteins that have been shown to bind to viral RNA, such as PTB (28), hnRNP A1 (29), and a complex of mitochondrial aconitase and chaperones (45, 46, 64, 65), are critically involved in coronavirus RNA synthesis.

A fairly comprehensive view of *cis*-acting elements located at the 3′ end of the coronavirus genome has come from studies of mouse hepatitis virus (MHV) and its close relative, bovine coronavirus (BCoV) (9, 42). Both MHV and BCoV are members of the second of the three phylogenetic groups into which coronaviruses are sorted. For MHV, work with viral mutants has demonstrated that the 3′ genomic *cis*-acting elements essential for RNA synthesis lie entirely within the 301-nucleotide (nt) 3′ untranslated region (3′ UTR) (13, 18), even though earlier studies with defective interfering (DI) RNAs had suggested a potential role for a portion of the adjacent N gene (33, 36, 40). The current picture of the MHV 3′ UTR, shown in Fig. 1, began with the definition of structures at its upstream end. Immediately following the N-gene stop codon is a bulged stem-loop (27). Adjacent to this is a classical hairpin-type pseudoknot (60), first discovered in BCoV. Exhaustive genetic analyses, in addition to chemical and enzymatic probing, have confirmed the existence and the essential nature of these two structures both in DI RNAs and in the intact viral genome (18, 26, 27, 60). Most significantly, the bottommost stem of the bulged stem-loop and stem 1 of the pseudoknot overlap, and thus, the two cannot simultaneously fold to completion. Consequently, we have previously proposed that these alternative

structures are components of a molecular switch regulating a transition between individual steps of viral RNA synthesis (18, 26).

Overlapping bulged stem-loops and pseudoknots are maintained in the upstream 3′ UTRs of all group 2 coronaviruses, including severe acute respiratory syndrome coronavirus (SARS-CoV) (18, 20), human coronavirus HKU1 (61), and the recently discovered bat coronaviruses HKU4, HKU5, and HKU9 (62). (Although it is stated otherwise in reference 62, HKU9 clearly harbors a pseudoknot, running from nucleotide 28898 through 28947, and two candidate overlapping bulged stem-loops can be found upstream of it.) This structural conservation, despite considerable primary sequence divergence, strongly implies functional conservation. More to the point, it has been shown that the entire 3′ UTR of either BCoV (27) or SARS-CoV (20) can functionally replace the MHV 3′ UTR. Similarly, multiple group 2 coronaviruses closely related to MHV and BCoV are competent to serve as helper viruses for the replication of a BCoV DI RNA (63).

Most of the downstream remainder of the MHV 3′ UTR is composed of a hypervariable region (HVR) (Fig. 1); the HVR is highly divergent in both sequence and structure, even among closely related coronaviruses like MHV and BCoV. The structure of the MHV HVR has been confirmed by enzymatic probing (38). Detailed genetic analysis has shown that unlike the RNA elements in the upstream portion of the 3′ UTR, the HVR can be mutated or entirely removed from the viral genome with little effect on virus replication in tissue culture (19). The nonessential nature of the HVR is surprising, since the region harbors the octanucleotide 5′-GGAAGAGC-3′, which is almost universally conserved among coronaviruses and is always situated some 70 to 80 nt from the 3′ end of the genome. Although deletion of the HVR greatly attenuates MHV virulence in the mouse host, the HVR's dispensability with respect to in vitro replication greatly simplifies consideration of the basic mechanism of RNA synthesis.

In our previous mutational analysis of the bulged stem-loop and pseudoknot, we observed that loop 1 of the pseudoknot is extremely sensitive to insertions of various sizes (18). A threshold of viability was seen to occur at an insertion size of six nucleotides: an insertion one base smaller, AACAA, was fully tolerated, whereas an insertion one base larger, AACAACA, was lethal. One particular 6-nt insertion, AACAAA, had only a minimal effect on the viral phenotype. By contrast, a mutant designated Alb391, with a 6-nt insertion of AACAAG, was extremely unstable. Alb391 formed very small plaques compared to those of the wild type, and it gave rise to large-plaque revertants within a single passage. In the work described here, we carried out an analysis of multiple revertants of Alb391. The results of this study provide the first indication of an interaction between a 3′ *cis*-acting element required for viral RNA synthesis and components of the coronavirus replicase. In addition, our results point to an interaction between stem 1 of the pseudoknot and the extreme 3′ terminus of the MHV genome. These two sets of interactions suggest a model for the initiation of viral negative-strand RNA synthesis.

MATERIALS AND METHODS

Cells and viruses. Wild-type MHV-A59, MHV mutants, and revertants were propagated in mouse 17 clone 1 (17Cl1) cells. MHV plaque assays and plaque purifications were carried out in mouse L2 or L929 cells, which were overlaid with 1% Noble agar and stained with neutral red at 48 h postinfection. The interspecies chimeric virus fMHV.v2 (18) was grown in feline FCWF cells. Vaccinia viruses were propagated in CV-1 and BHK-21 cells, and plaque purifications were performed on CV-1 and D980R cells (11).

Vaccinia virus-based MHV reverse genetics. The recombinant MHV mutants rAlb391, rN8-Alb391, rN9-Alb391, rN8, and rN9 were derived from vaccinia virus vMHV-inf-1, which contains a cloned full-length MHV-A59 cDNA (GenBank accession number AY700211). Mutagenesis was done using the reverse-genetics system described previously (11, 24). Briefly, two rounds of vaccinia virus-mediated homologous recombination were performed using the *Escherichia coli* guanine-phosphoribosyltransferase (GPT) gene as a selection marker. First, the target region within the MHV cDNA was replaced by the GPT gene, and second, the GPT gene was replaced by a mutated version of the target region. To construct recombinant vaccinia viruses encoding the Alb391-specific AACAAG insertion in the MHV 3′ UTR, we used the plasmid pRec5 for recombination with vaccinia virus vMHV-inf-1. The GPT gene in pRec5 is flanked on its left by MHV-A59 nt 28510 to 29024 and on its right by 500 bp of vaccinia DNA that is encoded downstream of the MHV-A59 3′ terminus in vMHV-inf-1. The resulting recombinant vaccinia virus, vRec5, was used for recombination with the plasmid pRec7 to insert the Alb391-specific 3′ terminus. The plasmid pRec7 contains MHV Alb391 nt 28510 to 31335, followed by a synthetic poly(A) sequence and 500 bp of vaccinia virus DNA. The resulting recombinant vaccinia virus clone was designated vMHV-rAlb391. To mutagenize nsp8 and nsp9, the region encoding MHV nsp7 to nsp11 was replaced by GPT, using vaccinia virus-mediated homologous recombination of vMHV-inf-1 or vMHV-rAlb391 with plasmid pRec3. The plasmid pRec3 encodes the GPT gene flanked on its left by MHV-A59 nt 11464 to 11956 and on its right by MHV-A59 nt 13594 to 14151. The resulting recombinant vaccinia viruses were then recombined with reverse transcription (RT)-PCR-derived DNA fragments obtained from RNA isolated from the Alb391 revertant Rev2 or Rev3. The resulting recombinant vaccinia viruses thus produced were designated vMHV-rN8 (carrying the Rev2 mutations in nsp8 and the wild-type MHV 3′ UTR), vMHV-rN9 (carrying the Rev3 mutation in nsp9 and the wild-type MHV 3′ UTR), vMHV-rN8-Alb391 (carrying the Rev2 mutations in nsp8 and the Alb391 3′ UTR), and vMHV-rN9-Alb391 (carrying the Rev3 mutation in nsp9 and the Alb391 3′ UTR). The identities of plasmids and recombinant vaccinia viruses were confirmed by sequencing analysis of the mutated regions. Further cloning details, plasmid maps, and sequences are available from the authors upon request.

Recombinant coronaviruses were rescued from cloned cDNA using purified, EagI-cleaved vaccinia virus DNA as a template for the in vitro transcription of recombinant full-length MHV genomic RNA, which was electroporated into BHK-MHV-N cells as described previously (11). Following electroporation, the transfected BHK-MHV-N cells were mixed with a fourfold excess of murine 17Cl1 cells and cultured at 37°C. At days 1 and 2 postelectroporation, tissue culture supernatants were taken for further analysis.

Targeted RNA recombination. For the construction of the ΔHVR3 mutant and the reconstruction of the 3′ UTR A2G mutation, targeted RNA recombination with host range-based selection was used, as described in detail previously (13, 18, 19, 20, 34, 35, 43). In brief, monolayers of feline FCWF cells were infected with fMHV.v2 and were then transfected by electroporation with capped, T7 polymerase-transcribed donor RNAs (mMessage mMachine transcription kit; Ambion). Templates for donor RNAs were derived from pSG6 (18) via PCR-based mutagenesis methods, which replaced plasmid segments between unique restriction sites in the 3′ UTR or the downstream polylinker. pSG6 contains a cDNA segment of the 5′-most 0.5 kb of the MHV genome linked to the 3′-most 8.6 kb of the MHV genome. Progeny virus from infected and transfected feline cells were selected on the basis of having regained the ability to form plaques on murine L2 cell monolayers, and the resulting recombinant candidates were purified through two rounds of plaque titration. For mutants constructed by targeted RNA recombination, the wild-type virus used for comparisons was Alb240, a well-characterized isogenic recombinant that was previously reconstructed from fMHV and pMH54 donor RNA (35).

Characterization of mutants and revertants. For analysis of the sequences of constructed mutants and isolated revertants, total cellular RNA was extracted from infected monolayers and reverse transcribed with avian myeloblastosis virus (Life Sciences) or Superscript II (Invitrogen) reverse transcriptase using random hexanucleotide (Boehringer Mannheim) or MHV sequence-specific primers. Regions of the genome were then amplified by PCR under standard conditions using AmpliTaq polymerase (Roche), and PCR products were purified with Quantum-prep columns (Bio-Rad) prior to automated sequencing. For viruses for which the entire genome was sequenced (Rev2, Rev5, and Rev6), the extreme 5′-end sequence was obtained through rapid amplification of cDNA ends (5′/3′

RACE kit; Roche). In order to resolve the mixed sequence obtained for the mutant rAlb391 isolate S3, we cloned RT-PCR products using the pGEM-T Easy Vector System (Promega), and individual plasmids were subsequently sequenced.

For the measurement of growth kinetics, confluent monolayers of 17Cl1 cells were inoculated at a multiplicity of 0.01 or 5.0 PFU per cell (for the ΔHVR3 mutant) or at a multiplicity of 1.0 PFU per cell (for all other mutants) for 2 h at 37°C. Inocula were removed, monolayers were washed three times, and incubation was continued in fresh medium at 37°C. Sample aliquots of medium were then withdrawn at various times from 2 to 32 h postinfection, and infectious titers were subsequently determined on mouse L2 cells.

RNA structural predictions were carried out with the *mfold* software (44, 68), available at http://www.bioinfo.rpi.edu/applications/mfold.

Radiolabeling of viral RNA. Intracellular viral RNA was metabolically labeled in 17Cl1 cells as described previously (18–20, 27). In brief, 20-cm^2 confluent monolayers were infected with the ΔHVR3 mutant or wild-type MHV at a multiplicity of 5.0 PFU per cell at 37°C. Infected cells were starved from 2 h through 8 h postinfection in Eagle's minimal essential medium containing 5% dialyzed fetal bovine serum with 1/10 of the normal concentration of phosphate. Cells were then labeled from 8 h through 10 h postinfection in 1 ml phosphate-free medium containing 100 μCi per ml [^{33}P]orthophosphate (MP Biomedicals), 5% dialyzed fetal bovine serum, and 20 μg of actinomycin D (Sigma) per ml. Total cytoplasmic RNA was purified, and samples were analyzed by electrophoresis through 1% agarose containing formaldehyde and were visualized by fluorography.

RESULTS

Initial revertant analysis of a 3′ UTR pseudoknot mutant. In a previous study, we used targeted RNA recombination to construct a number of mutants of the MHV 3′ UTR bulged stem-loop and pseudoknot in order to obtain an understanding of the structural requirements of these two elements and their interrelationship (18). An additional goal was to generate severely impaired viral mutants that might direct us to genetic interactions between the 3′ UTR and other regions of the genome. A mutant that met the latter criterion was Alb391, in which an insertion of 6 nt, AACAAG, was made between bases 226 and 227 of pseudoknot loop 1 (Fig. 1). Alb391 had a markedly defective phenotype, exhibiting slow growth and forming much smaller plaques than did wild-type virus at 33°C, 37°C, and 39°C. Revertants of this mutant were rapidly selected in as few as one or two passages. Six independent revertants (Rev1 to Rev6), which formed plaques identical to or slightly smaller than those of the wild type, were isolated and analyzed by partial or total genomic sequencing (Table 1). One of these, Rev1, had lost the first five bases of the insertion, retaining a footprint of only the final G residue. This finding constituted strong evidence that the 6-nt insertion, rather than some extraneous mutation, was indeed responsible for the Alb391 phenotype.

All of the other revertants retained the complete 6-nt insertion, indicating that second-site suppressor mutations had to account for their observed reversion. For one of these, Rev2, we determined the sequence of the entire viral genome and compared it to the particular database reference sequence (GenBank accession number NC001846) that is closest to the laboratory strain from which Alb391 had been derived. A total of nine mutations, two synonymous and seven nonsynonymous, were found relative to the reference sequence. However, all except two of these base changes also occurred in the corresponding wild-type virus or in the immediate fMHV precursor to Alb391, and they were therefore not associated with reversion. The two remaining mutations produced a pair of closely spaced coding changes in nsp8 (Table 1), which is one of a group of four small processed polypeptide products encoded by the downstream-most end of replicase gene 1a. Similarly, sequencing of the nsp7-through-nsp10 region of both Rev3 and Rev4 revealed single coding mutations in nsp9 and in nsp8, respectively (Table 1). Notably, both nsp8 and nsp9 have been described as RNA-binding proteins, and nsp8 has been shown to possess an RNA-dependent RNA polymerase (RdRp) activity (16, 30, 56, 66). By contrast, Rev5 and Rev6 were found to have no mutations in the nsp7-through-nsp10 region, but each had the same A-to-G transition at the penultimate nucleotide preceding the poly(A) tail (denoted A2G in the numbering system used for the 3′ UTR) (Table 1). Initially we did not consider this mutation to be significant, since A2G is distant from the pseudoknot and since the same base change occurs naturally in MHV strain 3 (47). Accordingly, we sequenced the entire genomes of Rev5 and Rev6 and found no other base changes. Thus, our initial revertant analysis detected three classes of candidate reverting mutations for the 3′ UTR pseudoknot insertion in Alb391: one mapping to the original site of the insertion, a second falling within the replicase gene, and the last near the 3′ genomic terminus.

TABLE 1. Sequence analysis of initial set of candidate reverting mutations of Alb391[a]

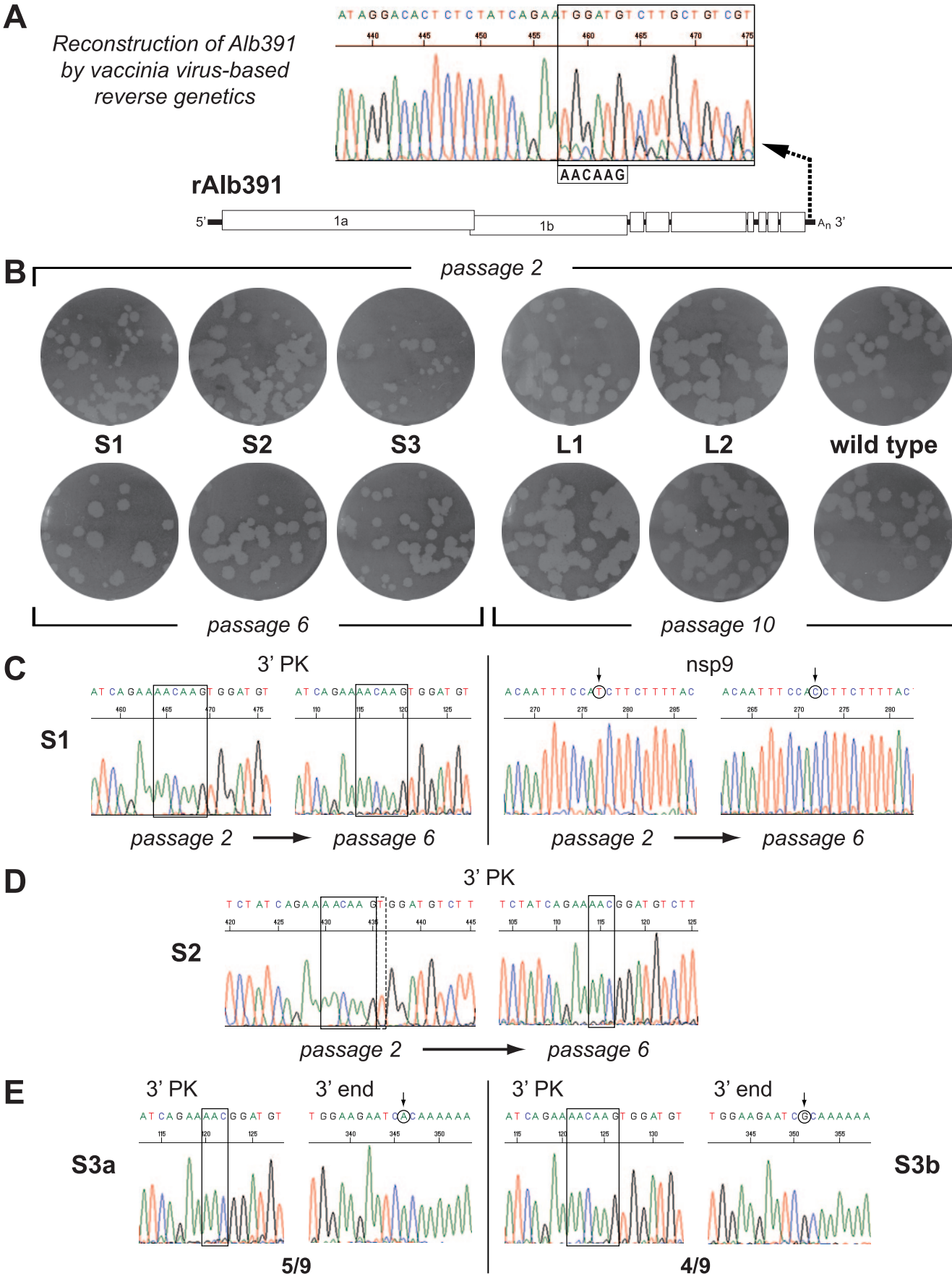
Mutant	Gene or region	Nucleotide change(s)[b]	Amino acid change(s)[c]
Rev1	3′ UTR	Deletion of nt 1 through 5 of the 6-nt insertion between bases 226 and 227	None
Rev2	nsp8	G12439A, C12452U	D62N, T66I
Rev3	nsp9	U13017G	N60K
Rev4	nsp8	A12344G	Q30R
Rev5	3′ UTR	A2G	None
Rev6	3′ UTR	A2G	None

[a] For Rev2, Rev5, and Rev6, the entire genomes were sequenced. For Rev1, Rev3, and Rev4, the genomic regions sequenced were as follows: the 5′-most 510 bases, the segment of gene 1a encoding nsp7 through nsp10, and the entire 3′ UTR.

[b] Nucleotide numbering is from the 5′ end of the genome for replicase (gene 1) mutations and from the 3′ end of the genome, excluding poly(A), for 3′ UTR mutations. All mutants, except Rev1, retained the 6-nt insertion (AACAAG) between bases 226 and 227 of the 3′ UTR.

[c] Amino acid numbering indicates the residue position in each processed replicase protein product.

Reconstruction of Alb391 by full-length cDNA cloning. To test whether replicase gene mutations, mapping some 20 kb distant from the original lesion, were responsible for reversion of the Alb391 mutation, we aimed to reconstruct a subset of the revertants. To accomplish this, we decided to use the vaccinia virus-based complete reverse genetics system for MHV (11, 58), which, unlike targeted RNA recombination (43), allows access to the huge coronavirus replicase gene. However, we first needed to clarify whether the Alb391-specific 6-nt insertion would confer the same phenotype in the genetic background of the recombinant MHV-inf-1 molecular clone. Therefore, we generated the 6-nt pseudoknot insertion mutation in the MHV-inf-1 genome and analyzed the rescued mutant, designated rAlb391. In accord with previous results, the reengineered mutant was found to be extremely unstable, even at the earliest stages of recovery. As shown in Fig. 2A, rAlb391 virus obtained from N protein-expressing cells that had been



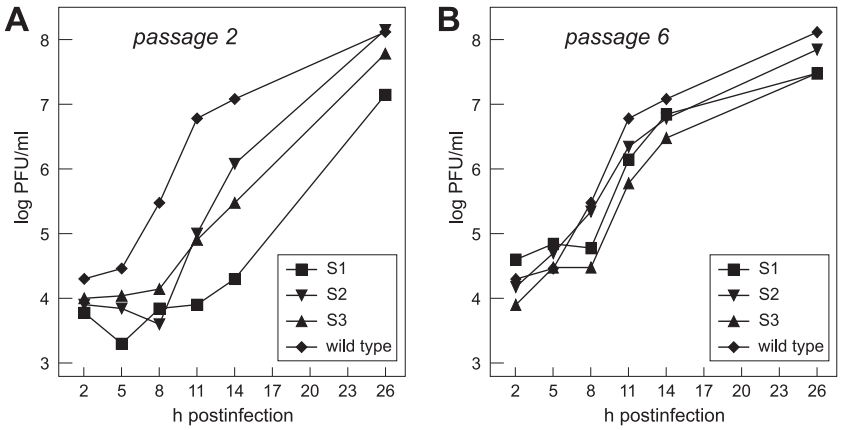


FIG. 3. Growth kinetics for small-plaque isolates S1 to S3 of rAlb391, compared with a wild-type control. Cultures were inoculated with passage-2 virus (A) or passage-6 virus (B) at a multiplicity of 1.0 PFU per cell. Data shown are a set from one representative of duplicate experiments, which gave highly similar results.

transfected with synthetic mutant genomic RNA had a mixture of mutant and wild-type sequence at the position of the insertion (and consequently at all points downstream in a positive-strand sequence). Plaques of rAlb391 exhibited considerable heterogeneity, ranging from very small to wild-type size (data not shown). To analyze this in more detail, we picked 3 of the smallest plaques (S1 through S3) and 2 of the largest plaques (L1 and L2) and propagated them for 6 and 10 passages, respectively. Isolates L1 and L2 formed uniformly large plaques at passages 2 and 10 (Fig. 2B), and sequence analysis revealed that they had exactly lost the 6-nt insertion (data not shown); this identified them as precise primary-site revertants.

By contrast, isolates S1 through S3 had a variety of plaque sizes by passage 2. By passage 6, they formed fairly uniform plaques that were the same size, or nearly the same size, as those of the wild type (Fig. 2B). Sequencing showed that this stability had been achieved through the acquisition of one of a number of new mutations. In the stock of isolate S1, the 6-nt 3′ UTR pseudoknot insertion was retained. However, by passage 6, a previously unseen nsp9 mutation had been selected (A13010G), resulting in the amino acid change D58G (Fig. 2C). For isolate S2, the passage-6 stock was composed of a new type of primary site footprint mutant, in which the last three bases of the 6-nt insertion (AAG), as well as the next downstream base (U), were deleted (Fig. 2D). There were no additional changes in either the nsp7-through-nsp10 region or the remainder of the 3′ UTR of S2 passage 6. Finally, bulk sequencing of the passage-6 stock of isolate S3 gave a mixture of sequences, which were resolved through TA cloning of RT-PCR products from the 3′ UTR. Five out of nine TA clones were found to have the same 3-nt footprint of the original insertion that had been found in the S2 isolate by passage 6 (Fig. 2E, left), with no other mutations in the 3′ UTR. On the other hand, the remaining four TA clones showed an intact 6-nt insertion, accompanied by the same A2G mutation that had been found in Rev5 and Rev6 (Fig. 2E, right). The S3 passage-6 stock had no additional changes in the nsp7-through-nsp10 region of the genome. In accord with the apparently strong selective pressure against the original Alb391 mutation, the growth of passage-2 stocks of the S1 through S3 isolates was greatly impaired relative to that of the wild type for at least the first 11 h of infection (Fig. 3A). Later, by 26 h postinfection, S1 through S3 mutant titers appeared to approach that of the wild type, but this may have reflected the emergence and overrunning of the cultures by revertants. By passage 6, the small-plaque isolates replicated with kinetics comparable to that of the wild type (Fig. 3B). As expected, the growth kinetics for either the passage-2 or passage-10 stocks of the L1 and L2 isolates were indistinguishable from the kinetics for the wild type (data not shown).

Taken together, these results showed that the AACAAG insertion mutation in pseudoknot loop 1 conferred the same phenotype in the MHV-inf-1 genetic background as had been seen in the initial Alb391 mutant (18). Furthermore, the analyses of the Alb391 and rAlb391 revertants revealed the same three classes of compensatory mutations, as summarized in Fig. 4A. First, there were primary-site revertants, in which there occurred loss of the entirety or a portion of the AACAAG insertion, as in rAlb391 isolates L1 and L2 or in the two types of footprint mutations in Rev1 and rAlb391 isolates

FIG. 2. Reconstruction of the Alb391 3′ UTR insertion mutant by using the vaccinia virus-based reverse-genetics system for MHV. (A) Sequence of RT-PCR product from the 3′ UTR of the initially recovered mutant, rAlb391. The position of the AACAAG insertion is denoted, and mixed sequence downstream of the site of the insertion is boxed. (B) Plaques from passages 2 and 6 of three small-plaque isolates of rAlb391 (S1 to S3), compared with plaques from passages 2 and 10 of two large-plaque isolates of rAlb391 (L1 and L2) and a wild-type control. (C) Sequences of RT-PCR products from the pseudoknot region (3′ PK) of the 3′ UTR and nsp9 from passages 2 and 6 of the S1 isolate; note that the nsp9 sequence is the negative sense. (D) Sequences of RT-PCR products from the pseudoknot region of the 3′ UTR from passages 2 and 6 of the S2 isolate. (E) Sequences of the pseudoknot region and the downstream terminus (3′ end) of the 3′ UTR of representative TA clones from passage 6 of the S3 isolate. In C to E, the insertion or its remnant is boxed, and mutant nucleotides or their wild-type counterparts are circled.

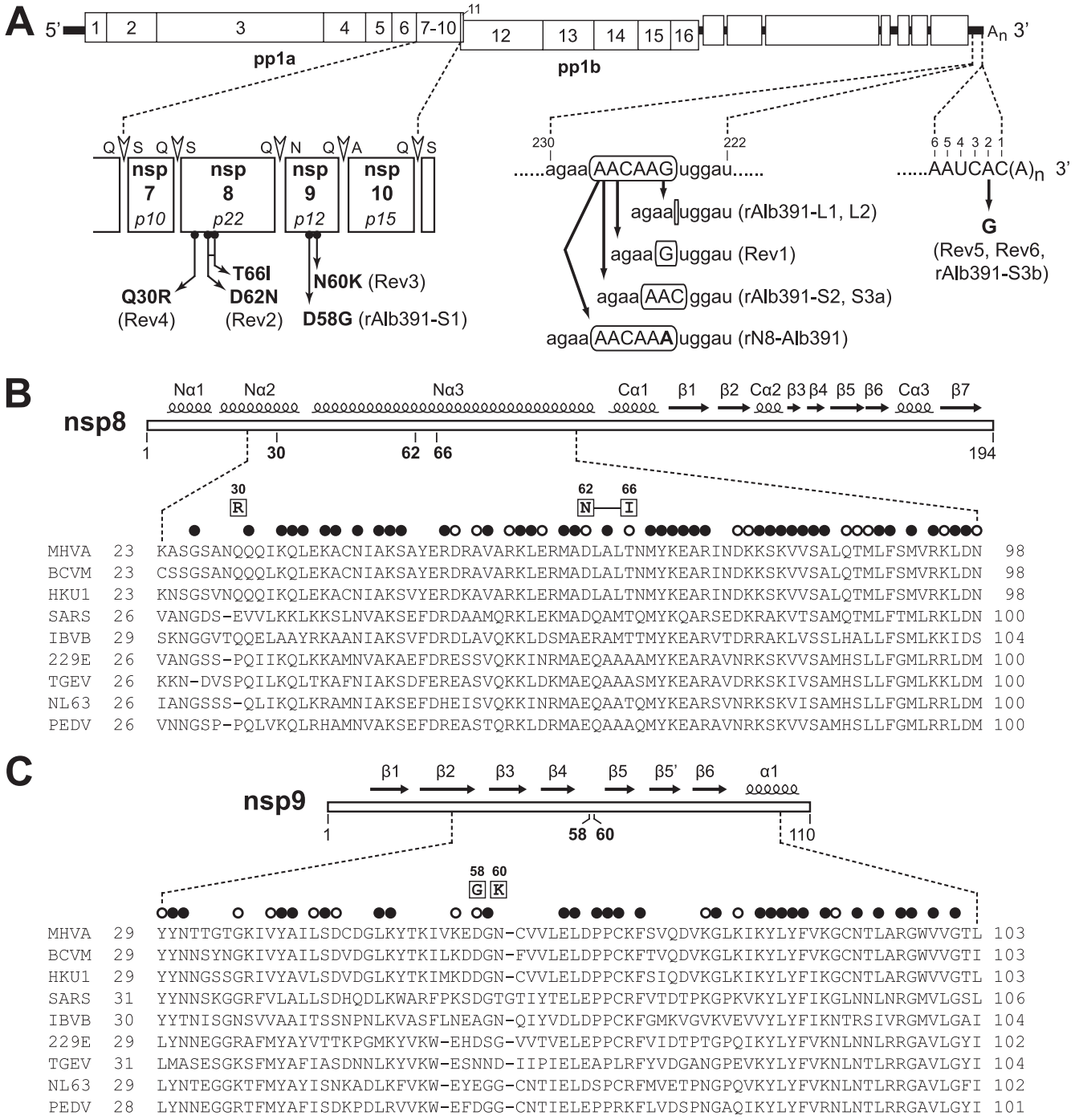


FIG. 4. Genetic cross talk between loop 1 of the 3′ UTR pseudoknot and both gene 1a and the 3′ end of the MHV genome. (A) Summary of all types of reverting mutations from the original Alb391 mutant and the reconstructed mutant rAlb391. In the expansion of the region of the genome encoding nsp7 to nsp10, main protease cleavage sites (8, 52) are denoted by open arrowheads and the flanking amino acid residues are indicated; alternate names for the MHV nsp7-to-nsp10 polypeptides are given in italics. In the expansions of regions of the 3′ UTR, nucleotide numbering is from the first base at the 3′ end of the genome, excluding poly(A). The original 6-nt insertion and remnants thereof are boxed; mutated nucleotides are shown in bold. For each type of reverting mutation, the designations of independently obtained isolates are given in parentheses. (B) Loci of the reverting mutations in nsp8. The schematic of the secondary structure of nsp8 of SARS-CoV corresponds to the crystal structure reported by Zhai and coworkers (66). (C) Loci of the reverting mutations in nsp9. The schematic of the secondary structure of nsp9 of SARS-CoV corresponds to the crystal structure reported by Egloff and coworkers (16), but it is also substantially the same as that reported by Sutton and coworkers (56). Beneath the schematics in panels B and C are alignments for portions of nsp8 and nsp9 of coronaviruses spanning all three phylogenetic groups. GenBank accession numbers for the sequences shown are as follows: MHVA, MHV strain A59, AY700211; BCVM, bovine coronavirus strain Mebus, U00735; HKU1, human coronavirus strain HKU1, AY597011; SARS, SARS coronavirus strain Urbani, AY278741; IBVB, infectious bronchitis virus strain Beaudette, M94356; 229E, human coronavirus strain 229E, AF304460; TGEV, transmissible gastroenteritis virus strain Purdue, AJ271965; NL63, human coronavirus strain NL63, AY567487; and PEDV, porcine epidemic diarrhea virus strain CV777, AF353511. MHV mutant Alb391 reverting mutations are boxed above each alignment. Filled circles mark residues that are identical among at least seven of the nine sequences; open circles mark additional residues that are identical among all group 2 coronaviruses. Amino acid numbering indicates the residue position in each processed replicase protein product.

S2 and S3. Second, four independent reverting mutations, Rev2 through Rev4 and rAlb391 isolate S1, mapped to three loci in nsp8 and nsp9, altering moderately conserved residues in these RNA-binding proteins. Finally, three independent revertants, Rev5, Rev6, and rAlb391 isolate S3, each had a single point mutation, changing the second residue of the 3′ UTR from A to G.

Interactions between the 3′ UTR pseudoknot and replicase gene products. Having shown that Alb391 and the reconstructed rAlb391 mutants displayed the same phenotype, we next examined the effects of reconstruction of two of the gene 1a candidate reverting mutations in both the presence and the absence of the 3′ UTR pseudoknot AACAAG insertion. For this purpose, we chose the pair of nsp8 mutations from Rev2 (G12439A and C12452U) and the single nsp9 mutation from Rev3 (U13017G) (Table 1). We analyzed three independent isolates bearing the nsp8 mutations in the presence of the 3′ UTR insertion (mutant rN8-Alb391) and two isolates bearing the nsp8 mutations alone (mutant rN8). Likewise, two isolates each were analyzed for the nsp9 mutation in the presence (mutant rN9-Alb391) or the absence (mutant rN9) of the 3′ UTR insertion (Fig. 5A). The rN8-Alb391 and rN9-Alb391 reconstructed revertants produced plaques that were only slightly smaller than those of the wild type (Fig. 5B), demonstrating that the nsp8 and nsp9 mutations were indeed responsible for the suppression of the Alb391 phenotype in Rev2 and Rev3. In general, the rN8-Alb391 and rN9-Alb391 mutants were stable for up to 10 passages: the nsp8 or nsp9 mutation remained unchanged, and the 6-nt insertion in the 3′ UTR was retained. However, an exception to this trend was that for 2 of the rN8-Alb391 isolates, the 6th residue of the 3′ UTR insertion (G) was gradually replaced by A over the course of 10 passages. An example of this replacement is shown in Fig. 5C. This suggested that the nsp8 mutations could not optimally compensate for the distortion of the 3′ UTR pseudoknot and that there was some residual selective pressure to mitigate the effect of the AACAAG insertion. The suppression of the Alb391 phenotype was also evident in the growth kinetics for the rN8-Alb391 and rN9-Alb391 mutants, which closely resembled the kinetics for the wild type at both passage 2 and passage 10 (Fig. 5D and E). At early times postinfection, both mutants lagged slightly behind the wild type, but they were far more robust than the passage-2 small-plaque isolates of rAlb391 (Fig. 3).

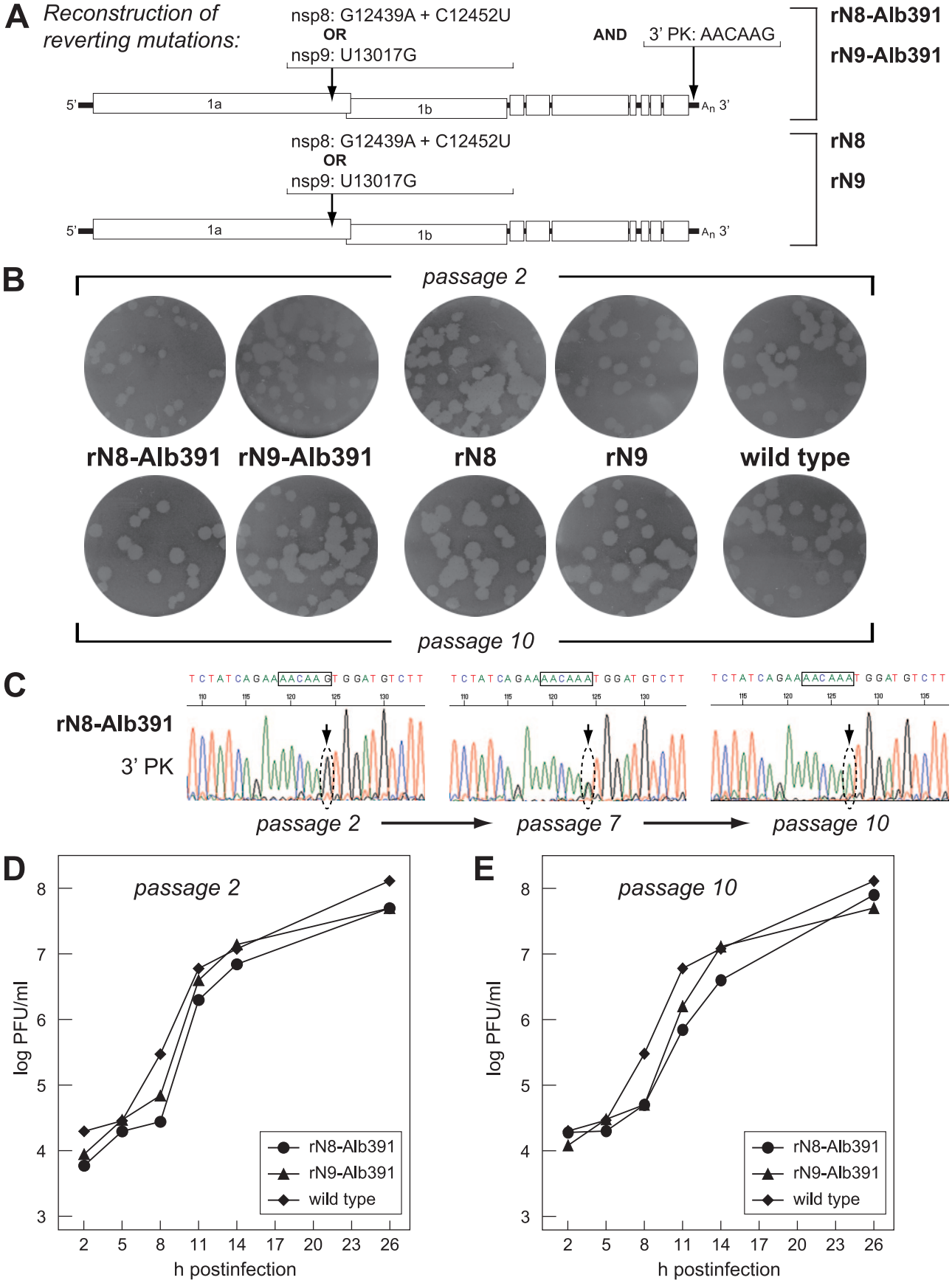
Somewhat differently, plaques of the rN8 and rN9 mutants were of fully wild-type size, indicating that the suppressor mutations, in isolation, were not noticeably deleterious to the virus (Fig. 5B). Sequence analysis confirmed this inference. The nsp8 and nsp9 mutations were stable over 10 passages, and their presence did not select for any compensating changes in the 3′ UTRs of these viruses. In addition, the growth kinetics for the rN8 and rN9 mutants were highly similar to that for the wild type (data not shown).

Collectively, these data demonstrated that the mutated replicase gene products nsp8 and nsp9 are still functional in the context of both the wild-type 3′ UTR and the Alb391 3′ UTR. More importantly, it was directly proven, by reconstruction of the Alb391 revertants Rev2 and Rev3, that the replicase gene mutations mapping to nsp8 and nsp9 could compensate for the markedly defective phenotype of Alb391, thus making it very likely that the other two replicase gene mutations, found in revertants Rev4 and rAlb391 isolate S1, were also genuine suppressors of the severe growth defect mediated by the 6-nt pseudoknot insertion. These data were indicative of an interaction of two replicase gene products, the RNA-binding proteins nsp8 and nsp9, with an essential *cis*-acting RNA pseudoknot located at the MHV 3′ UTR.

Interactions between the pseudoknot and the 3′ end of the genome. The last class of revertant, the A2G mutation, was very surprising, because the poly(A)-adjacent end of the genome was not obviously connected with the 3′ UTR pseudoknot in previous RNA structural analyses. To confirm whether A2G was truly a reverting mutation, we reconstructed this base change in both the presence and absence of the 6-nt pseudoknot insertion, using targeted RNA recombination. As we had observed for the replicase gene mutations, the A2G mutation, on its own, produced no detectable phenotype (as is also the case for MHV strain 3 [47]). Similarly, recombinants containing A2G together with the pseudoknot insertion formed plaques that were only slightly smaller than plaques of the wild type and were rescued at a much higher frequency than Alb391 (data not shown). This proved conclusively that the apparently simple A2G mutation could also suppress the severe impairment caused by the AACAAG pseudoknot insertion.

This result also provided the first indication of the existence of a direct interaction between pseudoknot loop 1 and the 3′ genomic terminus. The folding of these regions that is shown in Fig. 1 is derived from a study (38) in which a structure running from nt 166 to the 3′ end of the genome was predicted with the *mfold* software (44, 68). At the outset, the chosen 5′ upper limit of nt 166 would have excluded potential interactions with further-upstream portions of the 3′ UTR, including the pseudoknot. Indeed, we found that the lowest-free-energy *mfold* predictions obtained for the distal end of the 3′ UTR, running instead from the start of pseudoknot loop 1 (nt 230) through the 3′ terminus, support a different RNA secondary structure, which is shown in Fig. 6A. In this case, the 3′-most 29 nt of the genome are seen to form two helices with further upstream regions, one with pseudoknot loop 1 and one with the segment immediately distal to pseudoknot stem 2. The likelihood of this alternative structure is further supported by its absolute phylogenetic conservation among all other known group 2 coronaviruses, including the closely related BCoV and human coronavirus OC43 as well as the more divergent SARS-CoV, human coronavirus HKU1, and bat coronaviruses HKU4, HKU5, and HKU9 (Fig. 6B). Thus, not only do members of this coronavirus subfamily strictly maintain the 3′ UTR bulged stem-loop and pseudoknot but they also appear to be required to preserve base pairing between loop 1 of the pseudoknot and the 3′ end of the genome. Similar loop 1–3′-end structures potentially occur in the genomes of the group 1 and group 3 coronaviruses, but as yet no experimental evidence supports their existence.

To further test the loop 1–3′-end interaction in MHV, we designed a mutant harboring a large deletion within the 3′ UTR that would retain the two putative helices of the new model while eliminating the entire HVR and all of the adjacent region that was formerly thought to base pair with the 3′ end of the genome. We had previously constructed a deletion mu-



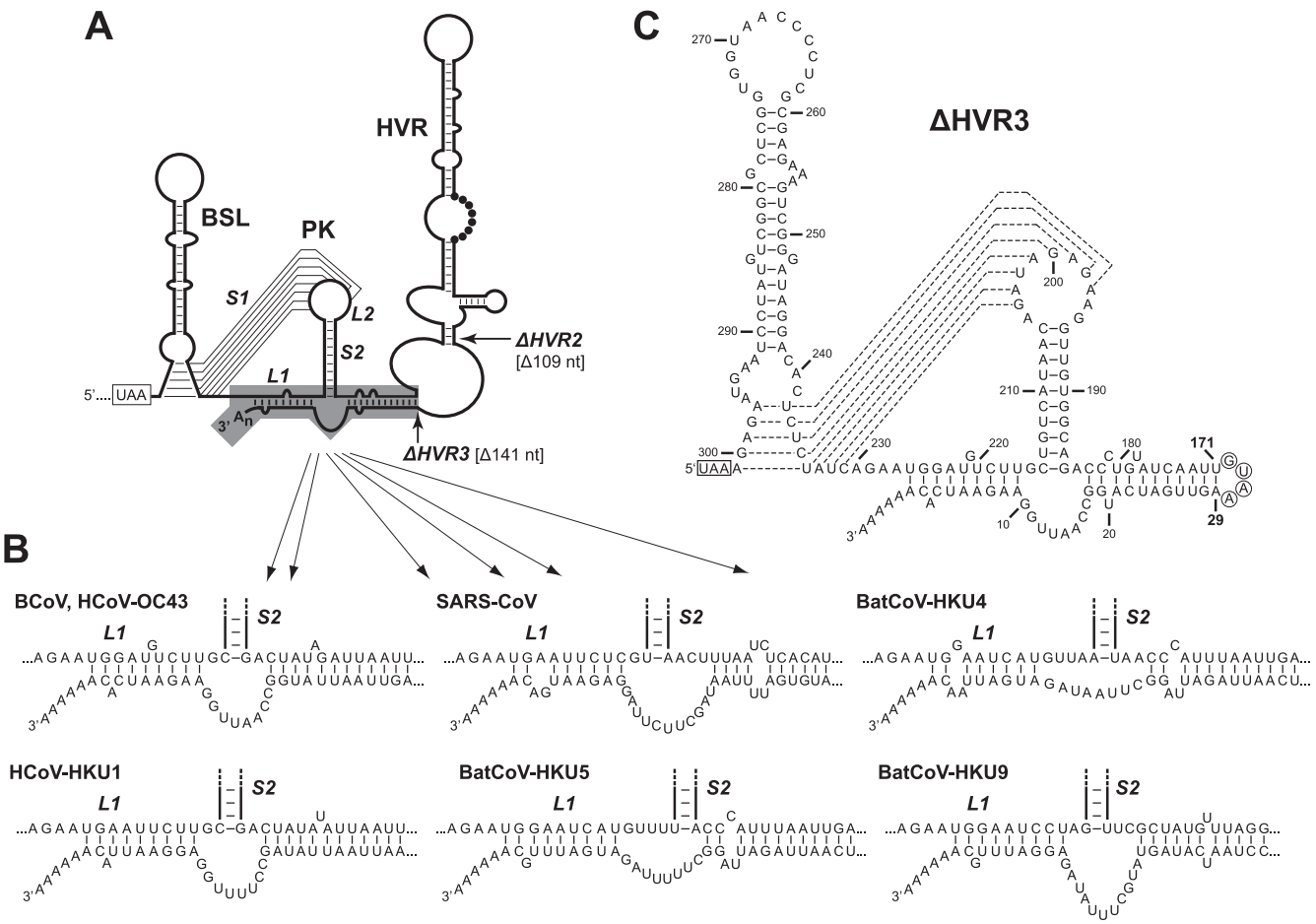


FIG. 6. Interaction between loop 1 (L1) of the 3′ UTR pseudoknot (PK) and the 3′ end of the genome. (A) Refinement of the complete structure of the MHV 3′ UTR, incorporating base pairing of the 3′ end of the genome with both pseudoknot loop 1 and the region downstream of pseudoknot stem 2 (S2); other abbreviations are the same as those used in Fig. 1. Arrows indicate points at which truncations were introduced in the previously characterized ΔHVR2 deletion mutant (19) and in a new mutant, the ΔHVR3 mutant. (B) Phylogenetic conservation of the loop 1–3′-end interaction among all group 2 coronaviruses. GenBank accession numbers for the sequences shown, other than those given in the legend to Fig. 4, are as follows: HCoV-OC43, human coronavirus strain OC43, AF523847; and bat coronavirus (BatCoV) strains HKU4, HKU5, and HKU9, respectively, EF065505, EF065509, and EF065513. (C) Detailed view of the complete 164-nt 3′ UTR of the ΔHVR3 mutant. Broken lines represent alternative base pairings either for the bottom stem segment of the bulged stem-loop or for pseudoknot stem 1. Nucleotides of the tetraloop introduced between positions 29 and 171 are circled. Nucleotide numbering is from the first base at the 3′ end of the genome, excluding poly(A), and the N-gene stop codon is boxed.

tant, the ΔHVR2 mutant (Fig. 6A), in which 109 nt of the 3′ UTR were excised with little consequence to the in vitro growth of the virus (19). In the new mutant, designated the ΔHVR3 mutant, we created an even more extensive deletion, of 141 nucleotides, in which all sequence between bases 29 and 171 was replaced by a tetraloop capping the helix downstream of pseudoknot stem 2 (Fig. 6C). The ΔHVR3 mutant was not only viable but formed plaques that were 80% of the size of wild-type plaques and were larger than those of the previously characterized ΔHVR2 virus. Moreover, the ΔHVR3 mutant exhibited growth kinetics comparable to those of the wild type in single-step, high-multiplicity infections (Fig. 7A) and in multicycle, low-multiplicity infections (Fig. 7B). Under both conditions, growth of the ΔHVR3 mutant lagged behind that of

FIG. 5. Reconstruction of viruses harboring candidate gene 1a reverting mutations of Alb391. (A) Recombinants rN8-Alb391 and rN9-Alb391, respectively, contain the nsp8 and nsp9 mutations from revertants Rev2 and Rev3 in the presence of the original Alb391 3′ UTR insertion mutation. Recombinants rN8 and rN9, respectively, contain the Rev2 and Rev3 mutations in the presence of the wild-type 3′ UTR. (B) Plaques from passages 2 and 10 of rN8-Alb391, rN9-Alb391, rN8, and rN9, compared with a wild-type control; the wild-type control plaques are the same as those shown in Fig. 2. (C) Sequences of RT-PCR products from the pseudoknot region (3′ PK) of the 3′ UTR from passages 2, 7, and 10 of one rN8-Alb391 isolate. The variable base of the AACAAG insertion is denoted. (D and E) Growth kinetics for reconstructed revertants rN8-Alb391 and rN9-Alb391, compared with a wild-type control. Cultures were inoculated with passage-2 virus (D) or passage 10 virus (E) at a multiplicity of 1.0 PFU per cell. Data shown for each mutant are a set from one representative of duplicate experiments, which gave highly similar results. Also, duplicate sets from two additional independent isolates of rN8-Alb391 were essentially identical to the data shown.

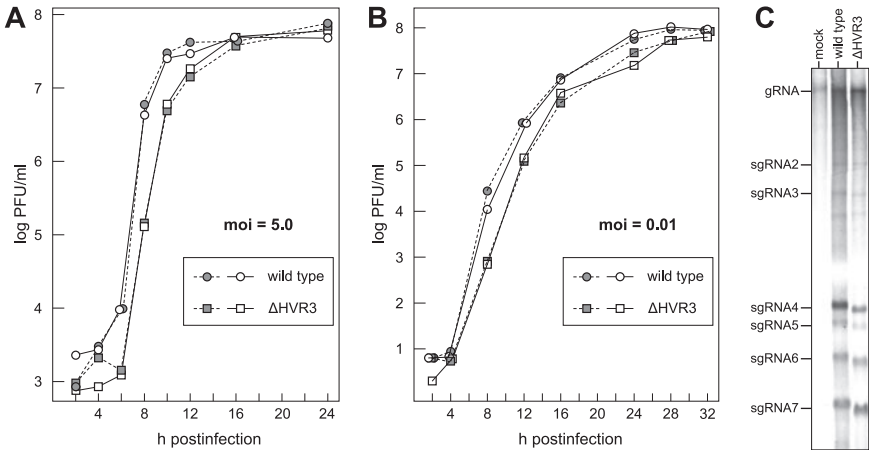


FIG. 7. Replicational competence of the ΔHVR3 mutant. Growth kinetics of the ΔHVR3 mutant relative to that of the wild type at a high multiplicity of infection (5.0 PFU per cell) (A) or a low multiplicity of infection (0.01 PFU per cell) (B). Open and solid symbols represent results from two independent experiments. (C) RNA synthesis by the ΔHVR3 mutant and the wild type. Infected or mock-infected 17Cl1 cells were metabolically labeled with [^{33}P]orthophosphate in the presence of actinomycin D, and RNA was isolated and electrophoretically separated in 1% agarose containing formaldehyde as described in Materials and Methods.

the wild type by 2 h but ultimately reached the same maximal titers. Additionally, metabolic labeling of infected cells revealed no significant difference in gRNA or sgRNA synthesis between the mutant and the wild type (Fig. 7C). Taken together, these results showed that the 3′ genomic *cis*-acting elements essential for MHV RNA synthesis are restricted to two separate segments of the 3′ UTR, nucleotides 1 through 29 and nucleotides 171 through 301. Surprisingly, therefore, almost half of the 3′ UTR is dispensable with respect to functions required for growth in tissue culture.

DISCUSSION

All RNA viruses harbor *cis*-acting sequences and structures in their genomes that are critical for replication. Within the order *Nidovirales*, 3′ *cis*-acting elements are currently best characterized for the group 2 coronaviruses, in which an overlapping bulged stem-loop and pseudoknot in the 3′ UTR have been found to be essential for RNA synthesis (Fig. 1) (9, 42). In the group 3 coronaviruses, a functionally essential stem-loop is located at the upstream end of the 3′ UTR (12), but the status of different nearby candidate pseudoknots is unresolved (20, 60). Conversely, the group 1 coronaviruses all contain a very highly conserved pseudoknot, but a neighboring potential counterpart of the bulged stem-loop (15) is not clearly identifiable in every member of this group. Over a greater phylogenetic distance, in the arteriviruses, which have much smaller 3′ UTRs, certain required *cis*-acting RNA structures have been described that extend into the upstream N gene (5, 6, 59). It is not obvious at present how closely analogous to coronavirus 3′ elements these RNA structures will turn out to be, but the genomes of members of the two virus families might be expected to interact in similar ways with their respective replicases, which are highly homologous (21).

Interactions between the pseudoknot and the 3′ end of the genome. In the work presented here, we analyzed revertants of a highly defective MHV mutant containing a particular insertion, AACAAG, in loop 1 of the 3′ UTR pseudoknot (summarized in Fig. 4A). Our finding of six independent primary-site reverting mutations that deleted or altered part or all of the insertion confirmed that the insertion was, in actuality, the basis of the deleterious and unstable genotype of the original mutant. More informative, however, were two groups of second-site reverting mutations that were obtained in both the original mutant, Alb391, which had been constructed by targeted RNA recombination, and rAlb391, which was reconstructed by vaccinia virus-based reverse genetics. These two classes of suppressors mapped to distinct regions of the genome, thereby revealing two types of previously unsuspected interactions. One class, independently isolated three times, was represented by a single type of revertant, in which the second 3′ UTR residue upstream of the poly(A) tail was changed from an A to a G. That this single base change could rescue the Alb391 insertion mutant suggested that the 3′ end of the genome must somehow communicate with loop 1 of the pseudoknot. This conclusion led us to identify a longer-range secondary structure in which the 3′ end of the genome folds into two helices through base pairing with pseudoknot loop 1 and with the region downstream of pseudoknot stem 2 (Fig. 6). Additional evidence in support of this interaction came from the following: (i) *mfold*-based structural predictions; (ii) the total conservation of the interaction among all known group 2 coronaviruses (and possibly also among group 1 and group 3 coronaviruses); (iii) our construction of a mutant, the ΔHVR3 mutant, which preserves this base pairing in a drastically minimized version of the 3′ UTR; and (iv) our previous finding of the lethality of a more extensive deletion in the 5′-ward direction, ΔHVR1, which suggested that at least some subset of nucleotides 157 through 184 must be essential (19).

It is noteworthy that in our redefined folding of the 3′ end of the genome, base A2 does not participate in a Watson-Crick base pair (Fig. 6). Thus, it is not immediately obvious how the A2G mutation relieves the problem caused by the AACAAG insertion. We favor the explanation that the insertion in loop 1 forces an alternative folding of the pseudoknot-3′ end interac-

tion and that the A2G mutation destabilizes this alternative structure, thereby allowing the wild-type fold to predominate. This notion is supported by the fact that the A2G mutation, in the context of an otherwise wild-type 3′ UTR, has no detectable effect on the viral phenotype. A potentially related observation is that the mutation A2U, which is also neutral by itself, is lethal in combination with the AACAAG insertion (data not shown). Further mutational analysis will be needed to more completely define the sequence and structural requirements for the interaction between the extreme 3′ end of the genome and pseudoknot loop 1, as well as the region downstream of pseudoknot stem 2.

The viability and nearly normal phenotype of the ΔHVR3 mutant indicate that the minimal 3′ *cis*-acting element required for MHV RNA synthesis is limited to no more than 160 nt of the 3′ UTR. This means that nearly half of the 3′ UTR, bases 30 through 170, can be discarded without affecting the basic mechanism of viral RNA synthesis. The deleted region in the ΔHVR3 mutant encompasses in vitro binding sites for a number of host cell proteins, including PTB (27), hnRNP A1 (28), and a complex of mitochondrial aconitase and the chaperones mtHSP70, HSP60, and HSP40 (45, 46, 64, 65). It follows that these binding sites cannot be strictly required for RNA synthesis, although they may play ancillary roles yet to be defined. Additionally, the region of the 3′ UTR that was eliminated in the ΔHVR3 mutant encompasses some RNA segments that were previously proposed to base pair with the 3′ end of the genome (38) (Fig. 1). Our redefined folding of the 3′ UTR is also at variance with an alternative *mfold*-predicted structure proposed for solely the 3′-terminal 42 nt of the genome (31). Notably, our results prompt reconsideration of the conclusions reached by a previous study, which mapped the promoter for negative-strand RNA synthesis based on RNase protection assays specific for the negative strand of a heterologous RNA that was connected to progressively decreasing lengths of the MHV 3′ UTR (37). That work concluded that the promoter for negative-strand initiation constitutes the last 45 to 55 nt of the genome. However, it is now clear that the essential sequence at the 3′ end of the genome is no more than 29 nt, together with a poly(A) tail of at least 5 to 10 residues (54). Moreover, we suggest that nt 1 to 29, on its own, is not sufficient for the initiation of negative-strand RNA synthesis, but it likely acts in association with the pseudoknot. Such an association would normally occur in *cis*, but in the promoter mapping experiments (37), it must have occurred in *trans*, between the heterologous RNA and the helper virus genome.

Interactions between the pseudoknot and gene 1a replicase products. The second class of suppressor mutations found in our genetic analysis was represented by four independently isolated revertants that mapped to nsp8 or nsp9. These replicase components fall within a group of small proteins, nsp7 through nsp10 (each 10 to 22 kDa), that are encoded at the downstream end of gene 1a (Fig. 4A) (8, 52). All four are processed from the replicase polyproteins 1a and 1ab by the main viral proteinase (nsp5), and they localize in discrete perinuclear and cytoplasmic foci in infected cells as parts of the membrane-associated replication complex (8, 10, 14, 22, 53).

Structural studies of each of the highly conserved SARS-CoV homologs of nsp7 through nsp10 have recently been reported. SARS-CoV nsp9, the crystal structure of which was separately solved by two groups, has been shown to have nonspecific RNA-binding activity and a fold not previously seen in any RNA-binding protein (16, 56). Nsp9 crystallizes in two dimeric forms, and biophysical evidence has also been presented for the existence of an interaction between nsp9 and nsp8 (56). A hexadecameric complex of SARS-CoV nsp8 and nsp7 has been found to bind to double-stranded RNA. The crystal structure of this complex (66), partially confirmed by the nuclear magnetic resonance solution structure of nsp7 (48), is a hollow cylinder potentially able to encircle double-stranded nucleic acid. Within this ring, which resembles a processivity clamp, nsp8 monomers are folded in two different conformations. Finally, there are two independently solved structures of SARS-CoV nsp10, which reveal that it represents a novel class of zinc finger protein which assembles into a dodecameric complex having a weak affinity for single- and double-stranded RNA (32, 55). A classical MHV mutant, *ts*LA6, which has the phenotype of shutoff of negative-strand RNA synthesis at the nonpermissive temperature (3, 51), has been mapped to nsp10 (50), and reverse-genetics analysis of MHV nsp10 has also indicated that this protein plays an essential role in viral RNA synthesis (14).

In the absence of further information, such as would be provided by RNA cocrystals, we cannot yet use the abundance of available structural data to draw strong conclusions about the nature of the reverting mutations that suppress the Alb391 pseudoknot insertion mutation. Each of the four nsp8 or nsp9 reverting mutations introduces a more basic charge into its respective molecule (Fig. 4B and C). The exact loci of the nsp8 mutations Q30R and D62N-T66I suggest that the mutations are most likely to affect the packing between monomers at the interface of the two nsp8 molecules in different conformations in the asymmetric unit of the nsp7-nsp8 hexadecamer. Likewise, the nsp9 mutations D58G and N60K may affect intersubunit beta-sheet interactions that stabilize one type of dimer formation (56). In both cases, these alterations could indirectly influence the RNA-binding specificity or affinity of nsp8 or nsp9. Alternatively, the conformations of the nsp's in a polyprotein precursor or in an RNA-bound form may differ from that revealed in the crystal structure, and thus, the increased positive charge caused by the reverting mutations might then directly enhance the affinity of nsp8 or nsp9. We are currently seeking in vitro biochemical data to corroborate these ideas.

Very surprisingly, purified SARS-CoV nsp8 has recently been found to possess RdRp activity, and it has been proposed to function as a primase for the canonical coronavirus RdRp, nsp12 (30). The in vitro enzymatic activity of SARS-CoV nsp8 has a dependence on manganese ion, and it exhibits a strong preference for cytidine-rich RNA templates, synthesizing small oligoribonucleotides of six or fewer residues. We have confirmed these same properties for the MHV homolog of nsp8 (S. J. Goebel and P. S. Masters, unpublished data). Imbert and coworkers (30), who discovered the RdRp activity of nsp8, have suggested that nsp8 could act at multiple points throughout the genomic template, although it must be noted that the characteristics of this enzyme may be substantially altered in the company of nsp7, nsp9, nsp10, and the authentic viral RNA substrate. In view of our suppressor results, we favor the idea that nsp8 acts as a primase during the initiation of negative-

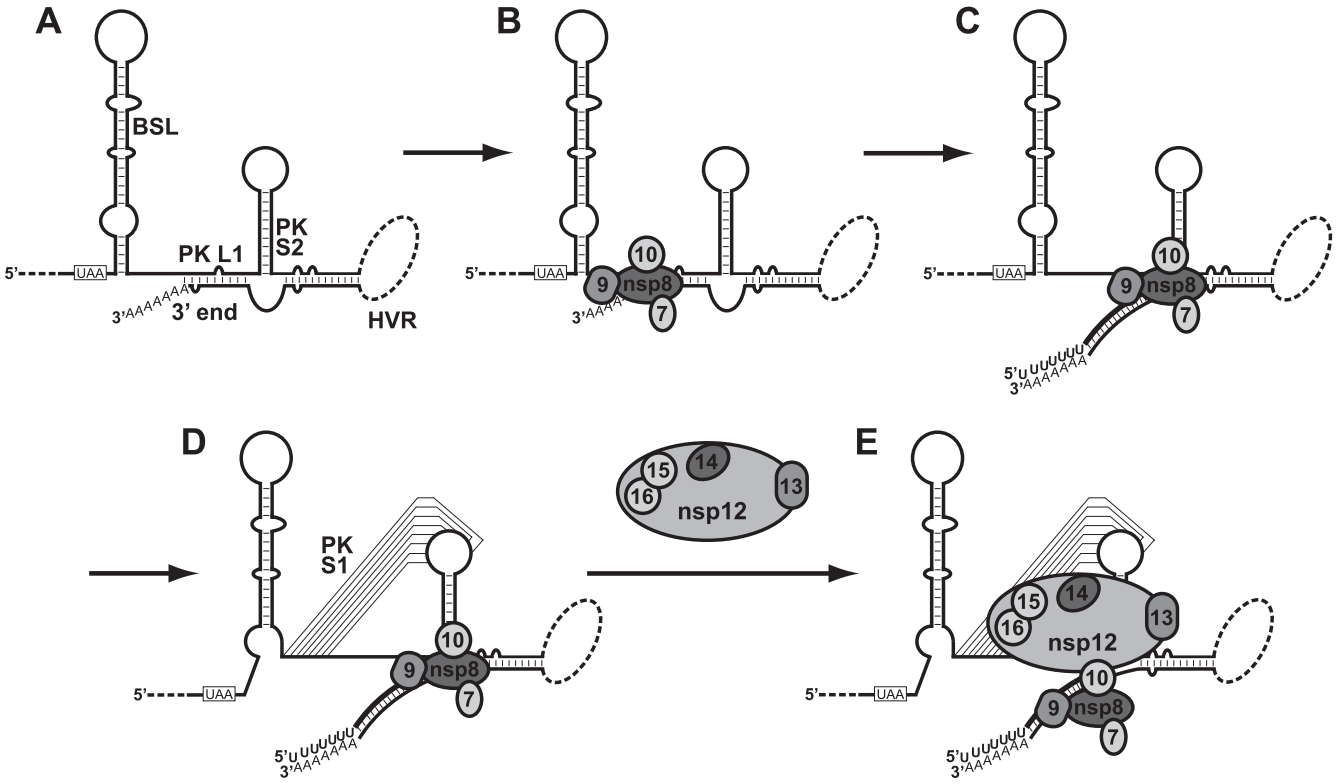


FIG. 8. Hypothetical model of initiation of coronavirus negative-strand RNA synthesis. (A) At the outset, formation of the pseudoknot is precluded by (i) closure of the bottom segment of the bulged stem-loop (BSL) and (ii) the interaction between pseudoknot (PK) loop 1 (L1) and the 3′ end of the genome. (B) The nsp8 primase, in conjunction with nsp7, nsp9, and nsp10, binds to the stem formed by loop 1 and the 3′ end. (C) Synthesis of the 5′ end of negative-strand RNA by the nsp8 primase separates the 3′ end of the genome from loop 1, thereby allowing the pseudoknot to form (D). The nsp12 RdRp then binds to the pseudoknot, together with nsp13 (helicase), nsp14 (ExoN nuclease), nsp15 (NendoU nuclease), and nsp16 (methyltransferase). (E) The nsp12 complex carries out elongation of the negative-strand primer. Not shown are possible roles played by other nsp's, the poly(A) binding protein, and the viral N protein.

strand synthesis. The interactions of the MHV 3′ pseudoknot loop 1 with the extreme 3′ genomic terminus and with nsp8 and nsp9 would position a template RNA secondary structure and particular replicase components very close to the start site of negative-strand RNA synthesis, which initiates with production of a tract of oligo(U) some 9 to 26 residues into the poly(A) tail (25). It is thus conceivable that the primase activity of nsp8 is involved in the initiation of coronavirus negative-strand RNA synthesis before the canonical RdRp activity, encoded by the nsp12 gene, can resume negative-strand RNA synthesis by elongating the nsp8-derived primer molecules. This idea is supported by the fact that nsp12 can be targeted to replication complexes when expressed in *trans* as a fusion protein with green fluorescent protein in MHV-infected cells; moreover, nsp12 can be coimmunoprecipitated with nsp8 and nsp9 from infected cell lysates (10).

Model for initiation of MHV negative-strand RNA synthesis. Our suppressor results, integrated with the finding of primase activity of nsp8, thus lead us to propose a model for initiation of MHV negative-strand RNA synthesis, which is shown in Fig. 8. We speculate that the most stable conformation of the 3′ UTR RNA is one in which the bulged stem-loop is fully folded and the 3′ end of the genome is annealed to pseudoknot loop 1. These two folding alternatives would preclude formation of stem 1 of the pseudoknot. The stem formed by loop 1 and the 3′ end of the genome is postulated to serve as the site of assembly of a complex that includes at least nsp8 and nsp9 and probably also nsp7 and nsp10. These replicase constituents may act as processed monomers, as shown. Alternatively, they may be part of a much larger nsp4-nsp11 precursor, which has been suggested to be the relevant species for negative-strand RNA synthesis (50). Thus, the MHV 3′ pseudoknot in conjunction with the 3′ end of the genome may constitute a *cis*-acting element that allows the specific recognition of the gRNA as the template for initiation of minus-strand RNA synthesis. Following substrate recognition, initiation of negative-strand RNA synthesis by the nsp8 primase complex leads to separation of the 3′ end of the genome from loop 1, thereby allowing formation of stem 1 of the pseudoknot. Finally, the completely folded pseudoknot, in conjunction with nsp8 and nsp9, then recruits the nsp12 RdRp and the associated activities contained in nsp13 through nsp16, and the resulting complex resumes the elongation of negative-strand RNA.

We are aware that many mechanistic details of this process remain to be elucidated. The proposed role of the MHV pseudoknot in association with the 3′ genomic end bears some resemblance to the function of the 5′ *cis*-acting element required for the assembly and positioning of the poliovirus preinitiation RNA synthesis complex discussed above (4, 23).

However, it remains to be established whether *cis*-acting elements located at the 5′ end of the coronavirus genome also participate in this process. Although the notion of genome circularization has been previously invoked (29, 69), there is currently no clear evidence that coronavirus 5′ and 3′ genomic ends are bridged by viral or cellular proteins or by heretofore-unidentified RNA-RNA interactions. Congruent with the analogy to the poliovirus system, however, it has been demonstrated that the length of the poly(A) tail of coronavirus DI RNA affects the efficiency of replication (54). Finally, the exact composition of the coronavirus preinitiation complex remains to be determined; the complex conceivably assembles cotranslationally and involves more nsp's than are pictured in our model. We hope that our discovery of an interaction between an essential 3′ *cis*-acting pseudoknot, specific replicase gene products, and the extreme 3′ end of the MHV genome will provide a starting point for further genetic, structural, and biochemical analyses aimed at establishing a more detailed understanding of the earliest events of coronavirus RNA synthesis.

ACKNOWLEDGMENTS

We are grateful to Joachim Jaeger for providing valuable help in localizing the nsp8 and nsp9 mutations in the reported X-ray crystal structures. We thank the Molecular Genetics Core Facility of the Wadsworth Center for some of the DNA sequencing reported here.

This work was supported by the Swiss National Science Foundation (V.T.) and by Public Health Service grants AI 45695 and AI 060755 from the National Institutes of Health (P.S.M.).

REFERENCES

1. **Andino, R., G. E. Rieckhof, and D. Baltimore.** 1990. A functional ribonucleoprotein complex forms around the 5′ end of poliovirus RNA. Cell **63:**369–380.
2. **Andino, R., G. E. Rieckhof, P. L. Achacoso, and D. Baltimore.** 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5′-end of viral RNA. EMBO J. **12:**3587–3598.
3. **Baric, R. S., and B. Yount.** 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. J. Virol. **74:**4039–4046.
4. **Barton, D. J., B. J. O'Donnell, and J. B. Flanegan.** 2001. 5′ cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. EMBO J. **20:**1439–1448.
5. **Beerens, N., and E. J. Snijder.** 2006. RNA signals in the 3′ terminus of the genome of Equine arteritis virus are required for viral RNA synthesis. J. Gen. Virol. **87:**1977–1983.
6. **Beerens, N., and E. J. Snijder.** 20 June 2007. An RNA pseudoknot in the 3′ end of the arterivirus genome has a critical role in regulating viral RNA synthesis. J. Virol. doi:10.1128/JVI.00747-07.
7. **Blyn, L. B., K. M. Swiderek, O. Richards, D. C. Stahl, B. L. Semler, and E. Ehrenfeld.** 1996. Poly(rC) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5′ noncoding region: identification by automated liquid chromatography-tandem mass spectrometry. Proc. Natl. Acad. Sci. USA **93:**11115–11120.
8. **Bost, A. G., R. H. Carnahan, X. T. Lu, and M. R. Denison.** 2000. Four proteins processed from the replicase gene polyprotein of mouse hepatitis virus colocalize in the cell periphery and adjacent to sites of virion assembly. J. Virol. **74:**3379–3387.
9. **Brian, D. A., and R. S. Baric.** 2005. Coronavirus genome structure and replication. Curr. Top. Microbiol. Immunol. **287:**1–30.
10. **Brockway, S. M., C. T. Clay, X. T. Lu, and M. R. Denison.** 2003. Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. J. Virol. **77:**10515–10527.
11. **Coley, S. E., E. Lavi, S. G. Sawicki, L. Fu, B. Schelle, N. Karl, S. G. Siddell, and V. Thiel.** 2005. Recombinant mouse hepatitis virus strain A59 from cloned, full-length cDNA replicates to high titers in vitro and is fully pathogenic in vivo. J. Virol. **79:**3097–3106.
12. **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.
13. **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.
14. **Donaldson, E. F., A. C. Sims, R. L. Graham, M. R. Denison, and R. S. Baric.** 2007. Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. J. Virol. **81:**6356–6368.
15. **Dye, C., and S. G. Siddell.** 2005. Genomic RNA sequence of feline coronavirus strain FIPV WSU-79/1146. J. Gen. Virol. **86:**2249–2253.
16. **Egloff, M. P., F. Ferron, V. Campanacci, S. Longhi, C. Rancurel, H. Dutartre, E. J. Snijder, A. E. Gorbalenya, C. Cambillau, and B. Canard.** 2004. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc. Natl. Acad. Sci. USA **101:**3792–3796.
17. **Gamarnik, A. V., and R. Andino.** 1997. Two functional complexes formed by KH domain containing proteins with the 5′ noncoding region of poliovirus RNA. RNA **3:**882–892.
18. **Goebel, S. J., B. Hsue, T. F. Dombrowski, and P. S. Masters.** 2004. Characterization of the RNA components of a putative molecular switch in the 3′ untranslated region of the murine coronavirus genome. J. Virol. **78:**669–682.
19. **Goebel, S. J., T. B. Miller, C. J. Bennett, K. A. Bernard, and P. S. Masters.** 2007. A hypervariable region within the 3′ *cis*-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis. J. Virol. **81:**1274–1287.
20. **Goebel, S. J., J. Taylor, and P. S. Masters.** 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.
21. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. *Nidovirales*: evolving the largest RNA virus genome. Virus Res. **117:**17–37.
22. **Gosert, R., A. Kanjanahaluethai, D. Egger, K. Bienz, and S. C. Baker.** 2002. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. J. Virol. **76:**3697–3708.
23. **Herold, J., and R. Andino.** 2001. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. Mol. Cell **7:**581–591.
24. **Hertzig, T., E. Scandella, B. Schelle, J. Ziebuhr, S. G. Siddell, B. Ludewig, and V. Thiel.** 2004. Rapid identification of coronavirus replicase inhibitors using a selectable replicon RNA. J. Gen. Virol. **85:**1717–1725.
25. **Hofmann, M. A., and D. A. Brian.** 1991. The 5′ end of coronavirus minus-strand RNAs contain a short poly(U) tract. J. Virol. **65:**6331–6333.
26. **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.
27. **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.
28. **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.
29. **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.
30. **Imbert, I., J. C. Guillemot, J. M. Bourhis, C. Bussetta, B. Coutard, M. P. Egloff, F. Ferron, A. E. Gorbalenya, and B. Canard.** 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. EMBO J. **25:**4933–4942.
31. **Johnson, R. F., M. Feng, P. Liu, J. J. Millership, B. Yount, R. S. Baric, and J. L. Leibowitz.** 2005. Effect of mutations in the mouse hepatitis virus 3′(+)42 protein binding element on RNA replication. J. Virol. **79:**14570–14585.
32. **Joseph, J. S., K. S. Saikatendu, V. Subramanian, B. W. Neuman, A. Brooun, M. Griffith, K. Moy, M. K. Yadav, J. Velasquez, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2006. Crystal structure of nonstructural protein 10 from the severe acute respiratory syndrome coronavirus reveals a novel fold with two zinc-binding motifs. J. Virol. **80:**7894–7901.
33. **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.
34. **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.
35. **Kuo, L., and P. S. Masters.** 2003. The small envelope protein E is not essential for murine coronavirus replication. J. Virol. **77:**4597–4608.
36. **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.
37. **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.
38. **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.

39. **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.
40. **Luytjes, W., H. Gerritsma, and W. J. M. Spaan.** 1996. Replication of synthetic interfering RNAs derived from coronavirus mouse hepatitis virus-A59. Virology **216:**174–183.
41. **Masters, P. S.** 2006. The molecular biology of coronaviruses. Adv. Virus Res. **66:**193–292.
42. **Masters, P. S.** 2007. Genomic *cis*-acting elements in coronavirus RNA replication, p. 63–78. *In* V. Thiel (ed.), Coronaviruses: molecular and cellular biology. Caister Academic Press, Norwich, United Kingdom.
43. **Masters, P. S., and P. J. M. Rottier.** 2005. Coronavirus reverse genetics by targeted RNA recombination. Curr. Top. Microbiol. Immunol. **287:**133–159.
44. **Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner.** 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. **288:**911–940.
45. **Nanda, S. K., R. F. Johnson, Q. Liu, and J. L. Leibowitz.** 2004. Mitochondrial HSP70, HSP40, and HSP60 bind to the 3′ untranslated region of the Murine hepatitis virus genome. Arch. Virol. **149:**93–111.
46. **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.
47. **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.
48. **Peti, W., M. A. Johnson, T. Herrmann, B. W. Neuman, M. J. Buchmeier, M. Nelson, J. Joseph, R. Page, R. C. Stevens, P. Kuhn, and K. Wuthrich.** 2005. Structural genomics of the severe acute respiratory syndrome coronavirus: nuclear magnetic resonance structure of the protein nsP7. J. Virol. **79:**12905–12913.
49. **Sawicki, S. G., D. L. Sawicki, and S. G. Siddell.** 2007. A contemporary view of coronavirus transcription. J. Virol. **81:**20–29.
50. **Sawicki, S. G., D. L. Sawicki, D. Younker, Y. Meyer, V. Thiel, H. Stokes, and S. G. Siddell.** 2005. Functional and genetic analysis of coronavirus replicase-transcriptase proteins. PLoS Pathog. **1:**e39.
51. **Schaad, M. C., and R. S. Baric.** 1994. Genetics of mouse hepatitis virus transcription: evidence that subgenomic negative strands are functional templates. J. Virol. **68:**8169–8179.
52. **Schiller, J. J., A. Kanjanahaluethai, and S. C. Baker.** 1998. Processing of the coronavirus MHV-JHM polymerase polyprotein: identification of precursors and proteolytic products spanning 400 kilodaltons of ORF1a. Virology **242:** 288–302.
53. **Sims, A. C., J. Ostermann, and M. R. Denison.** 2000. Mouse hepatitis virus replicase proteins associate with two distinct populations of intracellular membranes. J. Virol. **74:**5647–5654.
54. **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.
55. **Su, D., Z. Lou, F. Sun, Y. Zhai, H. Yang, R. Zhang, A. Joachimiak, X. C. Zhang, M. Bartlam, and Z. Rao.** 2006. Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. J. Virol. **80:** 7902–7908.
56. **Sutton, G., E. Fry, L. Carter, S. Sainsbury, T. Walter, J. Nettleship, N. Berrow, R. Owens, R. Gilbert, A. Davidson, S. Siddell, L. L. Poon, J. Diprose, D. Alderton, M. Walsh, J. M. Grimes, and D. I. Stuart.** 2004. The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. Structure **12:**341–353.
57. **Teterina, N. L., D. Egger, K. Bienz, D. M. Brown, B. L. Semler, and E. Ehrenfeld.** 2001. Requirements for assembly of poliovirus replication complexes and negative-strand RNA synthesis. J. Virol. **75:**3841–3850.
58. **Thiel, V., and S. G. Siddell.** 2005. Reverse genetics of coronaviruses using vaccinia virus vectors. Curr. Top. Microbiol. Immunol. **287:**199–227.
59. **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.
60. **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.
61. **Woo, P. C. Y., S. K. P. Lau, C.-M. Chu, K.-H. Chan, H.-W. Tsoi, Y. Huang, B. H. L. Wong, R. W. S. Poon, J. J. Cai, W.-K. Luk, L. L. M. Poon, S. S. Y. Wong, Y. Guan, J. S. M. Peiris, and K.-Y. Yuen.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. **79:**884–895.
62. **Woo, P. C. Y., M. Wang, S. K. P. Lau, H. Xu, R. W. S. Poon, R. Guo, B. H. L. Wong, K. Gao, H.-W. Tsoi, Y. Huang, K. S. M. Li, C. S. F. Lam, K.-H. Chan, B.-J. Zheng, and K.-Y. Yuen.** 2007. Comparative analysis of twelve genomes of three novel group 2c and group 2d coronaviruses reveals unique group and subgroup features. J. Virol. **81:**1574–1585.
63. **Wu, H.-Y., J. S. Guy, D. Yoo, R. Vlasak, E. Urbach, and D. A. Brian.** 2003. Common RNA replication signals exist among group 2 coronaviruses: evidence for in vivo recombination between animal and human coronavirus molecules. Virology **315:**174–183.
64. **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.
65. **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.
66. **Zhai, Y., F. Sun, X. Li, H. Pang, X. Xu, M. Bartlam, and Z. Rao.** 2005. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. Nat. Struct. Mol. Biol. **12:**980–986.
67. **Ziebuhr, J.** 2005. The coronavirus replicase. Curr. Top. Microbiol. Immunol. **287:**57–94.
68. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415.
69. **Zuniga, S., I. Sola, S. Alonso, and L. Enjuanes.** 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J. Virol. **78:**980–994.