The 3’ cis-Acting Genomic Replication Element of the Severe Acute Respiratory Syndrome Coronavirus Can Function in the Murine Coronavirus Genome

Scott J. Goebel, Jill Taylor, and Paul S. Masters*

Wadsworth Center, New York State Department of Health, Albany, New York 12201

Received 25 February 2004/Accepted 23 March 2004

The 3’ untranslated region (3’ UTR) of the genome of the severe acute respiratory syndrome coronavirus can functionally replace its counterpart in the prototype group 2 coronavirus mouse hepatitis virus (MHV). By contrast, the 3’ UTRs of representative group 1 or group 3 coronaviruses cannot operate as substitutes for the MHV 3’ UTR.

RNA viral genomes contain cis-acting sequence and structural elements that play essential roles in RNA synthesis, gene expression, and virion assembly. Coronaviruses and other members of the nidovirus order have a unique scheme of RNA synthesis that involves both genome replication and transcription of a 3’-nested set of subgenomic (sg) mRNAs (28). Both genomic and sgRNA synthesis are thought to initiate with the production of the complementary negative-strand species, originating at the 3’ end of the genome (1, 21, 24, 29). Our laboratory and others have been studying the elements at the 3’ ends of the genomes of coronaviruses that are required for viral replication (3–6, 8, 13–15, 18, 20, 26, 27, 30–32). For the prototype coronavirus, mouse hepatitis virus (MHV), such required elements reside entirely within the 3’ untranslated region (3’ UTR) of the genome (4). In MHV and the closely related bovine coronavirus (BCoV), two highly conserved RNA structures at the upstream end of the 3’ UTR are known to be essential: a bulged stem-loop (4–6) and an adjacent pseudoknot (4, 30) (Fig. 1). These structures partially overlap, such that both cannot be formed simultaneously, and it has been proposed that they constitute components of a molecular switch, which is operative at some stage of RNA synthesis (4, 5).

Downstream of these elements, a more complex secondary structure makes up most of the rest of the 3’ UTR (15). Both the structure and sequence of this region are less well conserved between MHV and BCoV, but two features of note are found here. The first is an octanucleotide motif that is absolutely conserved in all coronavirus 3’ UTRs and is functionally critical (S. J. Goebel, T. B. Miller, and P. S. Masters, unpublished results). The second is the minimal signal required for initiation of negative-strand RNA synthesis, which for MHV falls within the final 55 nucleotides (nt) of the 3’ UTR (14).

Coronaviruses have been sorted into three phylogenetic groups, based on the relatedness of their sequences for the huge replicase polyprotein Lab gene and for the downstream-encoded canonical set of virion structural proteins, the spike glycoprotein (S), the membrane protein (M), the small envelope protein (E), and the nucleocapsid protein (N). The original sequence analyses of the newly discovered severe acute respiratory syndrome coronavirus (SARS-CoV) found this virus to be nearly equally distant from the three previously recognized groups (16, 22). It was thus proposed to establish a fourth group. However, a more recent phylogenetic analysis based on the 1b gene (which contains the RNA-dependent RNA polymerase), and which used the related toroviruses as an outgroup, concluded that SARS-CoV is most closely related to the group 2 coronaviruses (25).

In the same vein, these investigators noted that regions of the 1a gene of SARS-CoV contain domains that are unique to the group 2 coronaviruses.

Our alignment of the primary sequence of the 3’ UTR of SARS-CoV with the 3’ UTRs of representative coronaviruses (Fig. 1) revealed that this genome segment also showed no distinctive homology to any of the three groups, compared to the degree of conservation that was observed within each of the groups. Pairwise identities between SARS-CoV and the other viruses were 38% with MHV (group 2), 30% with BCoV (group 2), 26% with transmissible gastroenteritis virus (TGEV, group 1), and 34% with avian infectious bronchitis virus (IBV, group 3). By contrast, the 3’ UTRs of the two group 2 members, MHV and BCoV, are 69% identical. This alignment has been anchored at three loci of known functional importance mentioned above: the stems of the upstream pseudoknot structure, the octanucleotide motif, and the 3’ end of the genome. In some cases, higher pairwise identities could be obtained if these criteria were ignored, but this still did not provide a strong basis to assign the SARS-CoV 3’ UTR to one of the three groups. Also significant in this alignment is a region of 32 nt of near-perfect identity between the IBV and SARS-CoV 3’ UTRs (Fig. 1). This motif, which has been noted previously (9, 16), is also found near the 3’ ends of the genomes of several astroviruses and one picornavirus (19).

Although there is no compellingly preferential sequence homology between the SARS-CoV 3’ UTR and those of any particular group of coronaviruses, we think it is more relevant to compare RNA structures. As we have noted previously, the presence of both the bulged stem-loop and the overlapping pseudoknot at the upstream end of the 3’ UTR appears to be

* Corresponding author. Mailing address: David Axelrod Institute, Wadsworth Center, NYSDOH, New Scotland Ave., P.O. Box 22002, Albany, NY 12201-2002. Phone: (518) 474-1283. Fax: (518) 473-1326. E-mail: masters@wadsworth.org.
FIG. 1. Comparison of coronavirus 3′ UTR sequences and structures. Top panel, sequence alignment of the 3′ UTRs of SARS-CoV, MHV, BCoV, TGEV, and IBV. Nucleotides are numbered from the first base at the 3′ end of the genome, excluding poly(A); dashes represent gaps introduced in the alignment. Labeled and underlined features are as follows: (i) stems 1 and 2 of the pseudoknot that is common to the first four sequences and potentially present in the IBV sequence, (ii) the octanucleotide motif (oct) that is conserved in all coronavirus genomes, and (iii) a 32-nt region (IBV motif) found in the IBV and SARS-CoV 3′UTRs that is identical to segments found near the 3′ ends of the genomes of several astroviruses and an equine rhinovirus (19). The alignment of the MHV sequence (strain A59, accession no. M80644) and BCoV sequence (strain Mebus, accession no. M16620) is as presented previously (6). The SARS-CoV sequence is the Urbani strain (accession no. AY278741) (22); the TGEV sequence is the Purdue strain (accession no. M14878) (7). The IBV 3′UTR was determined from a clone, pSGIBVUTR, which we generated from an isolate of the Beaudette strain, the 505-nt sequence of which is identical to that found under accession no. AJ311362 (2), except for nt 147. The identity of this variant base in our particular isolate was confirmed from an RT-PCR product obtained from the virus. Bottom panel, the conserved, functionally essential secondary structure found at the upstream ends of the 3′UTRs of the group 2 coronaviruses MHV and BCoV (4–6, 30) and which is also predicted for the 3′UTR of SARS-CoV. The structure comprises a bulged stem-loop and a partially overlapping RNA pseudoknot. BCoV nucleotides that differ from those of MHV are circled. Numbering is the same as in the linear sequence alignment, and the N gene stop codon is boxed. Broken lines indicate alternative base pairings for the stem-loop or for pseudoknot stem 1.
a hallmark of the group 2 coronaviruses (4, 5). The group 1 coronaviruses all contain a highly conserved pseudoknot but no detectable counterpart of the bulged stem-loop in any proximity, upstream or downstream. On the other hand, the group 3 coronaviruses have a highly conserved stem-loop, which has been shown to be functionally essential (3), but only a poor candidate for the pseudoknot structure can be found nearby (30). In this respect, it was striking to notice that the upstream end of the SARS-CoV 3′/H11032 UTR could be folded into both a stem-loop and a pseudoknot in exactly the same overlapping pattern as in the MHV and BCoV 3′/H11032 UTRs (4) (Fig. 1), suggesting that this genomic segment in SARS-CoV is more closely related to those of the group 2 coronaviruses.

To test the functional significance of this observation, we asked whether the SARS-CoV 3′ UTR could replace the MHV 3′ UTR. It has previously been shown that the BCoV 3′ UTR can be substituted for that of MHV with no discernible effect on RNA synthesis or any other aspect of viral phenotype (5, 6). Moreover, it has very recently been demonstrated that replication of a BCoV defective interfering RNA could be supported by any of a number of group 2 helper viruses, including MHV (31). This suggests that not only are the 3′ cis-acting elements for RNA replication interchangeable among members of a given coronavirus group, but that the same holds true for the 5′ cis-acting elements.

To transfer the SARS-CoV 3′ UTR into the MHV genome, we used targeted RNA recombination, a reverse genetics system for coronaviruses that has been widely employed to manipulate all regions of the MHV genome downstream of the replicase 1ab gene (4–6, 11, 12, 17). This technique allows efficient recovery of mutants with severely defective phenotypes (4, 11, 12). Targeted recombination couples two intrinsic properties of coronaviruses: their propensity to carry out homologous recombination with transfected donor RNAs and the strict species specificity exhibited by many coronaviruses in tissue culture. Recombinants were generated by transfection of donor RNAs containing heterologous 3′ UTRs into feline AK-D cells that had been infected with fMHV.v2 (4), a chimeric coronavirus containing the ectodomain of the S protein of feline infectious peritonitis virus in place of the correspond-
ing region of the MHV S protein (Fig. 2). This substitution has rendered fMHV.v2 capable of growth only in feline cells. Recombinants were therefore isolated as viruses that had regained the MHV S ectodomain as a result of a crossover upstream of the S gene, concomitantly regaining the ability to grow in mouse cells (Fig. 2). The rearrangement of genes downstream of S in fMHV.v2 essentially eliminates the potential for formation of wild-type recombinants via a second crossover event downstream of the S gene (4).

A transcription vector for the synthesis of donor RNA was constructed from pSG6 (4), which encompasses the 3'-most 8.6 kb of the MHV genome and contains a coding-silent BspEI site near the end of the N gene. A cDNA of the SARS-CoV 3' UTR was generated by reverse transcription (RT)-PCR starting with RNA from infected Vero E6 cells. SARS-CoV (Urbani strain) was obtained from the Centers for Disease Control and Prevention (CDC). Its propagation for RNA isolation was performed in a Biosafety Level 3 laboratory under conditions

FIG. 3. Phenotype of the SARS-CoV 3' UTR recombinant. (A) Comparison of plaques of the SARS-CoV 3' UTR recombinant (Alb424) and the wild type (Alb240). Plaque titrations were done at 37°C on mouse L2 cells; monolayers were stained with neutral red at 48 h postinfection and were photographed 18 h later. (B) Single-step growth kinetics. Confluent monolayers of 17Cl1 cells were infected with Alb424 and Alb240 at a multiplicity of 0.6 PFU per cell. At the indicated times postinfection, aliquots of medium were removed and infectious titers were determined on mouse L2 cells. Open and solid squares (Alb424) and circles (Alb240) represent results from two independent experiments. (C) RNA synthesis. Infected or mock-infected 17Cl1 cells were metabolically labeled with [33P]orthophosphate in the presence of actinomycin D, and RNA was isolated and electrophoretically analyzed in 1% agarose containing formaldehyde as described previously (17). gRNA, genomic RNA.
meeting or exceeding CDC requirements (www.cdc.gov/ncidod/sars/lab/biosafety.htm). Amplification of cDNA was carried out with a negative-sense megaprimer running from the polyliner region of pSG6 through a 118-nt stretch of poly(dT) and a positive-sense primer containing the 3’ end of the MHV N gene (including the BspEi site) joined to the 5’ end of the SARS-CoV 3’ UTR. The resulting PCR product was digested with BspEi and Pacl and was used to replace the corresponding fragment of pSG6 to yield the vector pSGSARUTR (Fig. 2).

We also constructed transcription vectors pSGTGEVUTR and pSGIBVUTR, containing the 3’ UTRs of TGEV (group 1) and IBV (group 3), by using the same strategy. The template for construction of pSGTGEVUTR was pFG5, a clone of the 3’ end of the genome of the Purdue strain of TGEV (7) generously provided by David Brian. The template for construction of pSGIBVUTR was a 1967 isolate of the Beaudette strain of IBV. It should be noted that the IBV 3’ UTR, which is much larger than those of the other coronaviruses (Fig. 1), is bipartite. The upstream end is hypervariable, both in size and sequence, among different IBV strains (3, 23). However, the region shown to be essential for viral RNA replication is contained within the very highly conserved 338-nt downstream end (3).

In two separate trials, targeted RNA recombination with pSGSARUTR-derived donor RNA yielded MHV mutants containing the SARS-CoV 3’ UTR. Recovered mutant viruses were plaque purified in mouse L2 cells and amplified in mouse 17Cl1 cells, and the presence of the entire incorporated substitution was confirmed by RT-PCR and sequencing (Fig. 2). From two independent sets of such recombinants, one isolate, designated Alb424, was chosen for further analysis. Plaques formed by Alb424 (Fig. 3A), as well as by the other isolates (data not shown), were markedly smaller than those of wild-type MHV. Alb424 also had somewhat delayed growth kinetics by comparison with the wild type and grew to maximal titers roughly sixfold lower than those of the wild type (Fig. 3B). Metabolic labeling of RNA in infected cells with Alb424 revealed that it produced the same set of sgRNA species as wild-type MHV and in the same relative proportions (Fig. 3C). For the smaller of these species, sgRNA4 through sgRNA7, a slightly slower mobility relative to the wild type was detected, owing to the larger size of the SARS-CoV 3’ UTR with respect to the MHV 3’ UTR (Fig. 1). Overall, these results indicated that the SARS-CoV 3’ UTR was able to functionally replace its counterpart in the MHV genome. This substitution did not lead to any noticeable aberrations in viral RNA synthesis, although its function was clearly suboptimal compared to that of the native MHV 3’ UTR.

By contrast, we could not select MHV recombinants harboring either the TGEV 3’ UTR or the IBV 3’ UTR following seven independent trials with pSGTGEVUTR-derived donor RNA or four independent trials with pSGIBVUTR-derived donor RNA, respectively. In parallel with these attempts, positive control targeted recombination experiments with pSG6-derived donor RNA yielded robust numbers of reconstructed wild-type recombinants. We have previously shown that the host-range-based selection enabled by fMHV and fMHV.v2 can recover extremely impaired mutants (11, 12), including some with lesions constructed in the 3’ UTR (4). This argues strongly that replacement of the MHV 3’ UTR with either the TGEV 3’ UTR or the IBV 3’ UTR is lethal.

Taken together with previous demonstrations of the mutual recognition of cis-acting genomic elements among the group 2 coronaviruses (5, 6, 31), our results lend experimental support to the proposed classification of SARS-CoV as a group 2 coronavirus (25). We have shown that the RNA synthetic machinery of a group 2 coronavirus, MHV, is able to recognize and use the 3’ cis-acting structures and sequences of the SARS-CoV genome. Moreover, the MHV replicate cannot function with the corresponding elements from either a group 1 coronavirus, TGEV, or a group 3 coronavirus, IBV. In light of this observation, the presence in the SARS-CoV 3’ UTR of a 32-nt motif found in IBV (9, 16) (Fig. 1) is particularly intriguing. However, since this motif is also found in members of two other viral families (19), we favor the interpretation that it was horizontally acquired by both SARS-CoV and IBV (16) and that it does not indicate that these two coronaviruses have a unique common ancestor.

As has been pointed out by others (31), intragroup compatibility between coronavirus RNA polymerases and cis-acting genomic sequences makes recombination between viral species more likely in the event of mixed infection. Given the current lack of a clear candidate for the natural host for SARS-CoV and the near ubiquity of MHV in mouse populations, our study has potential implications for the further evolution of coronaviruses in the wild.

We are grateful to William Bellini and Thomas Klarekz of the CDC for providing SARS-CoV, and we thank David Brian, University of Tennessee, for TGEV clone pFG5. We also thank the Molecular Genetics Core Facility of the Wadsworth Center for oligonucleotide synthesis and DNA sequencing. We acknowledge Stuart Lehrman of the Wadsworth Center Computerized Photography and Illustration Unit for essential assistance in the preparation of figures. This work was supported in part by Public Health Service grants AI 45695 and AI 39544 from the National Institutes of Health and in part by Cooperative Agreement number U50/CCU212415 from the CDC. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the CDC.

REFERENCES
