Chapter 4
RNA Higher-Order Structures Within the Coronavirus 5′ and 3′ Untranslated Regions and Their Roles in Viral Replication

Pinghua Liu and Julian Leibowitz

Abstract The 5′ and 3′ untranslated regions (UTRs) of all coronaviruses contain RNA higher-order structures which play essential roles in viral transcription and replication. In this chapter we present our current knowledge of how those *cis*-acting elements were defined and their functional roles in viral transcription and replication. Cellular proteins which have been shown binding to those *cis*-acting elements and potentially support the RNA discontinuous synthesis model are also discussed. A conserved RNA structure model for the 5′ and 3′ UTRs of group 2 coronaviruses is presented with the known cellular protein binding sites.

4.1 Introduction

Coronaviruses are single-stranded, positive-sense, nonsegmented enveloped RNA viruses belonging to the family *Coronaviridae*, one of the three families in the order *Nidovirales*. They are the largest known RNA viruses with 27–31 kb genomes. Coronaviruses are classified as group 1, 2, and 3 based on serologic relatedness, genome organization and sequence similarity. Extensive phylogenetic comparisons placed the SARS coronavirus (SARS-CoV) as an early branch of the group 2 coronaviruses (Snijder et al. 2003). For all coronaviruses the 5′ two-thirds of the genome comprise the replicase gene, and the 3′ genes encode structural proteins and nonessential accessory proteins.

Coronaviruses infect cells by binding to specific receptors and enter cells by direct membrane fusion at the plasma membrane or by an endocytotic mechanism (Nash and Buchmeier 1997; Wang et al. 2008). SARS-CoV uses angiotension-

J. Leibowitz (✉)
Department of Microbial and Molecular Pathogenesis, Texas A&M University System College of Medicine, 407 Reynolds Medical Science Building, 1114 TAMUS, College Station, TX 77843-1114, USA
e-mail: jleibowitz@tamu.edu

S.K. Lal (ed.), *Molecular Biology of the SARS-Coronavirus*,
DOI 10.1007/978-3-642-03683-5_4,

converting enzyme 2 (ACE2) as its functional receptor (Li et al. 2003) and enters cells through pH- and receptor-dependent endocytosis (Wang et al. 2008). Upon entering the cytoplasm the virus particle is uncoated, releasing the RNA genome. The viral genome directs the synthesis of two large polypeptides, pp1a and pp1ab, via a frameshifting mechanism involving a pseudoknot structure (Brierley et al. 1987). The resulting polypeptide contains a conserved array of functional domains, which upon proteolytic processing results in 15–16 nonstructural proteins (nsp), many of which are likely to be involved in either RNA synthesis or proteolytic processing of the polyprotein precursors of nsp1–16 (Snijder et al. 2003). The 3′ one-third of the genome contains the genes for viral structural proteins and accessory proteins. These genes are expressed by transcription of a 3′ coterminal nested set of 7–9 mRNAs that also contain a common ~70–90 nucleotide (nt) 5′ leader identical in sequence to the 5′ end of the genome (Lai et al. 1983, 1984; Spaan et al. 1982). The 3′ end of the leader sequence contains a short (6–8 nt) sequence, the transcriptional regulatory sequence (TRS) also present in the genome just 5′ to the coding sequence for each mRNA (Budzilowicz et al. 1985).

Subgenomic negative-sense RNAs that correspond to each subgenomic RNA are found in infected cells (Sethna et al. 1989), as are replication intermediates containing subgenome-length negative strands (Sawicki and Sawicki 1990). In the currently accepted model, subgenomic mRNAs are transcribed from a complementary set of subgenome-size minus-strand RNAs, produced by discontinuous minus-strand synthesis. Molecular genetic studies with viruses containing mutations in the TRS support a model where leader-body joining takes place during synthesis of subgenomic negative-sense RNAs (Zuniga et al. 2004; Pasternak et al. 2001; van Marle et al. 1999). Sense–antisense base-pairing interactions between short conserved sequences play a key regulatory role in this process.

4.2 *cis*-Acting RNA Elements in Coronavirus Replication

The 5′ and 3′ untranslated regions (UTRs) of all coronavirus genomes contain *cis*-acting sequences required for viral transcription and replication (Chang et al. 1994; Dalton et al. 2001; Izeta et al. 1999; Kim et al. 1993). Additional *cis*-acting sequences such as packaging signals needed for assembly have been identified and mapped to internal positions in the genome. Many of these *cis*-acting sequences have been defined by studying defective interfering (DI) RNAs. These DI RNAs are extensively deleted, retain their 5′ and 3′ UTRs plus some additional genomic RNA, and are replication competent in the presence of helper virus able to provide replicase components in *trans*. Thus they retain *cis*-acting sequences needed for genome replication. DI RNAs have also been used to study the *cis*-acting signals needed for transcription (subgenomic mRNA synthesis) and for virion assembly. Recently reverse genetic systems for a number of coronaviruses have been developed, enabling the study of coronavirus *cis*-acting sequences in the context of the viral genome.

4.2.1 The Transcription Regulatory Sequence

Coronavirus RNA transcription occurs in the cytoplasm. All the coronavirus mRNAs have a common leader sequence at their 5′ ends (Spaan et al. 1982). The leader sequence contains a transcription regulatory sequence (TRS-L) at its 3′ end. This sequence motif constitutes part of the signal for subgenomic mRNA transcription. Preceding every transcription unit on the viral genomic RNA are additional transcription regulatory elements, named body transcription regulatory sequence (TRS-B) (Budzilowicz et al. 1985). All coronavirus TRSs can be divided into three sequence blocks, the core 6–8 nt sequence (CS), plus 5′ and 3′ flanking sequences (Sola et al. 2005). The most frequently used CS for group 1 coronaviruses is a hexamer (5′-CUAAAC-3′). For group 2 coronaviruses a heptameric sequence, 5′-UCUAAAC-3′ is the consensus sequence; it is almost identical to the group I CS. Interestingly, SARS-CoV has a CS (5′-ACGAAC-3′) which differs from other group 2 coronaviruses (Marra et al. 2003; Rota et al. 2003). The CS for group 3 coronaviruses is a divergent octamer, 5′-CUUAACAA-3′ (Alonso et al. 2002). The related arterivirus CS is 5′-UCAACU-3′ and partially resembles the infectious bronchitis virus (IBV) CS (van Marle et al. 1999).

Mutational analysis in a DI system found that the sequence flanking the CS-B affected the efficiency of subgenomic DI RNA transcription and that CS-B was necessary but not sufficient for the synthesis of the subgenomic DI RNA (Makino et al. 1991). Further analysis of MHV subgenomic mRNA transcription revealed that the 5′ leader sequence of MHV serves as a *cis*-acting element required for the transcription of subgenomic mRNAs (Liao and Lai 1994). Analysis of transmissible gastroenteritis virus (TGEV) mRNA synthesis in a minigenome system showed that the CS is essential for mediating a 100- to 1,000-fold increase in mRNA synthesis (Alonso et al. 2002). However, the CS flanking sequences also influenced transcription levels.

The functional importance of the TRS-L and TRS-B in the synthesis of subgenomic mRNA was shown by a mutagenesis study in equine arteritis virus (EAV), a member of the related arterivirus genus, utilizing a reverse genetic system (van Marle et al. 1999). Mutagenesis of the RNA 7 TRS-B significantly reduced its transcription. In contrast, mutagenesis of TRS-L affected all subgenomic mRNA transcription, and compensatory mutations in both TRS-L and RNA7 TRS-B restored RNA 7 transcription. This evidence strongly supports the mechanism of discontinuous minus-strand transcription. An additional comprehensive covariation mutagenesis study of several EAV TRSs demonstrated that discontinuous RNA synthesis depends not only on base-pairing between sense TRS-L and antisense TRS-B, but also upon the primary sequence of the TRS-B (Pasternak et al. 2001). While the TRS-L merely plays a targeting role for strand transfer, the TRS-B fulfills multiple functions. The sequences of mRNA leader-body junctions of TRS mutants strongly suggested the discontinuous step occurs during minus-strand synthesis. The development of reverse genetic systems for several coronaviruses has allowed a similar molecular genetic approach to investigating the role of the coronavirus

TRS rather than using DI replicons. For TGEV, analysis of the role of TRS demonstrated that the canonical CS-B was not absolutely required for the generation of subgenomic mRNAs, but its presence led to transcription levels at least 1,000-fold higher than those in its absence (Zuniga et al. 2004). A recent study in SARS-CoV rewired the TRS circuit (Yount et al. 2006). Recombinant genomes were constructed that contained mixtures of the wild-type and mutant regulatory TRS. Viable viruses were isolated from wild-type and recombinant genomes containing homogeneous transcription circuits; chimeras that contained mixed regulatory networks were invariantly lethal. In the lethal mutants the mixed TRS circuits promoted inefficient subgenomic transcription from inappropriate start sites, resulting in truncated open reading frames (ORFs) and minimized viral structural protein expression (Yount et al. 2006). This experiment provides further evidence for the discontinuous synthesis of minus-strand RNAs and the key role of TRS-L and TRS-B sequences in regulating subgenomic RNA synthesis.

In TGEV, there is a good correlation between subgenomic mRNA levels and the free energy of TRS-L and TRS-B duplex formation except for subgenomic mRNA N (Sola et al. 2005), leading the Enjuanes laboratory to seek additional regulatory sequences. Recently, a 9-nt transcriptional enhancer sequence was demonstrated 449 nts upstream of the TGEV N gene TRS core sequence (CS-N) (Moreno et al. 2008). This enhancer sequence interacts with a complementary sequence just upstream of CS-N, specifically increasing the accumulation of subgenomic mRNA N. This interaction is exclusively conserved in group 1a coronaviruses.

4.2.2 The 5′ cis*-Acting RNA Elements*

5′ *cis*-acting elements have been defined for several coronaviruses using DI deletion mapping analysis (Dalton et al. 2001; Chang et al. 1994; Izeta et al. 1999). A series of studies with MHV DIs demonstrated that as little as 466 nts at the 5′ terminus were needed for DI RNA replication (Kim et al. 1993; Kim and Makino 1995; Luytjes et al. 1996). Similar analyses with a BCoV synthetic DI RNA indicated that only the 5′ 498 nts were needed for DI RNA replication. Currently, no SARS-CoV 5′ *cis*-acting elements have been defined by a deletion analysis.

The DI RNA deletion analyses cited above defined the minimal *cis*-acting elements required for DI replication without further dissecting these sequences. Chang et al. (1994) demonstrated that the BCoV 68 nt leader is a necessary part of the 498 nt *cis*-acting signal for DI RNA replication. RNA higher-order structures contained in group 2 coronaviruses 5′ UTR were first predicted for BCoV using the Mfold algorithm (Chang et al. 1994). Subsequently, enzymatic probing and functional mutational analysis (Chang et al. 1996; Raman et al. 2003; Raman and Brian 2005) defined four stem-loops, denoted I, II, III, and IV within the first 210 nts of BCoV. The predicted stem-loop I (nts 11–42) contains three contiguous Waston–Crick base-pairs and a large 16 nt loop and is not conserved among group 2 coronaviruses (Chang et al. 1994). The predicted stem-loop II (nts 51–84) is a A–U

base-pair-rich hairpin with a low free energy that folds the TRS into the terminal loop. A poorly conserved stem-loop II homolog has been predicted in other coronaviruses, and in EAV (van der Born et al. 2004; Raman and Brian 2005). Stem-loop III (nts 97–116) is phylogenetically conserved and appears to have homologs in coronavirus groups 1 and 3; enzymatic probing and mutational analysis in DI RNA replication assays supports its existence (Raman et al. 2003). Stem-loop IV (nts 186–215), a bulged stem-loop, is also conserved amongst group 2 coronaviruses and may have a homolog in group 1 and 3. However, the predicted stem-loop IV homolog in SARS-CoV appears to be group 1-like (Raman and Brian 2005). Stem-loop IV exists as a higher-order structure based on enzymatic probing and it is required for DI RNA replication. Recently, two stem-loops, SLV (nts 239–310) and SLVI (nts 311–340), extending into the nsp1 coding region were demonstrated by RNase structure probing and sequence covariation among closely related group 2 coronaviruses. SLVI is required for DI RNA replication (Brown et al. 2007).

The recent establishment of reverse genetic systems for coronaviruses representing all of the coronavirus subgroups has facilitated the functional analysis of *cis*-acting elements in the context of the whole genome. Recently, we proposed a consensus secondary structural model of the 5′ 140 nts of the 5′ UTR based on nine representative coronaviruses (including SARS-CoV) from all three coronavirus groups (Kang et al. 2006; Liu et al. 2007). The 5′ ~140 nts of the nine coronaviruses genomes were predicted to fold into three major stem-loops, denoted SL1, SL2, and SL4 (see Fig. 4.1). Some sequences were predicted to contain a fourth stem-loop, SL3, which folds the TRS-L into a hairpin loop. SL1, SL2, and SL4 were structurally conserved amongst all coronaviruses examined. SL3 is only predicted to be stable for human coronavirus OC43 (HCoV-OC43) and SARS-CoV. It should be noted that SL1 and SL2 differ from the structures studied in BCoV pDrep 1 RNA by the Brian group (Chang et al. 1996; Raman et al. 2003; Raman and Brian 2005). However, SLIII in Brian's model is almost identical with our SL4b (Liu et al. 2007; Raman et al. 2003).

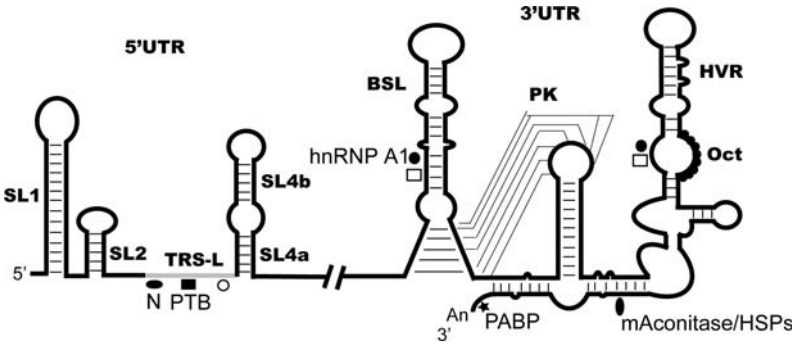

Fig. 4.1 Conserved RNA higher-order structural model within the 5′ and 3′ UTRs of group 2 coronaviruses. Proteins binding to the positive-strand RNA are shown as *solid symbols*; proteins binding to the minus-strand RNA are shown as *open symbols*

Although the full MHV and SARS-CoV 5′ UTRs are significantly different in terms of sequences and predicted secondary structures, the SARS-CoV SL1, SL2, and SL4 can functionally replace their MHV counterparts in the MHV genome and produce viable chimeric viruses (Kang et al. 2006). However, MHV chimeras containing the complete SARS-CoV 5′ UTR or the SARS-CoV SL3 were not viable. Replacing the SARS-CoV TRS with the MHV TRS in the MHV/5′ UTR SARS-CoV chimera permitted the synthesis of minus-strand genomic RNA but did not support the production of positive- or minus-strand subgenomic RNA7. This study supports the idea that SL1, SL2, and SL4 are conserved and interchangeable within the same group without affecting viral viability.

A detailed mutational and biophysical study of MHV SL1 revealed that this stem-loop is functionally and structurally bipartite. SL1 contains one or more noncanonical base-pairs in the central portion of the stem. In MHV, two pyrimidine–pyrimidine base-pairs are present in the middle of SL1, as demonstrated by NMR studies (Liu et al. 2007). These noncanonical base-pairs divide the SL1 helical stem into upper and lower segments. The upper region of SL1 is required to be base-paired; mutations that disrupted base-pairing of this region were not viable or severely impaired (Li et al. 2008). Combining both sets of mutations in the upper region of SL1 restored the base-pairing and yielded a viable virus comparable to the wild-type virus in its growth phenotype. In contrast, mutations in the lower region of SL1 that destroyed base-pairing were viable, and genomes with compensatory mutations predicted to restore base-pairing were nonviable. Deletion of a bulged or extruded A in the lower portion of the stem (mutation ΔA35), a mutation that increased the thermal stability of the lower portion of the SL1 helix, was strongly selected against. ΔA35-containing viruses were rapidly replaced by viruses containing destabilizing second-site mutations near ΔA35. Additionally, mutations that increased the stability of the lower portion of SL1 were lethal, suggesting that structural lability in the lower portion of the SL1 stem was required. Thermal denaturation and imino proton exchange experiments further demonstrated that the lower half of SL1 is unstable. SL1 second-site mutants also contained an additional second-site mutation, A29G or A78G, in their 3′ UTR, providing genetic evidence for an interaction between the 5′ and 3′ UTRs. Thus we hypothesized that the base of SL1 has an optimized lability required to mediate a physical interaction between the 5′ UTR and the 3′ UTR (Li et al. 2008). These data, plus the observed defects in subgenomic RNA synthesis in our nonviable SL1 mutants, are consistent with the genome circularization model for coronavirus transcription put forward by Zuniga et al. (2004) and suggest that replication complexes and transcription complexes have different structural requirements in the 5′ UTR.

SL2 is the most conserved secondary structure in the 5′ UTRs of all coronaviruses examined (Liu et al. 2007). Except for the core TRS leader sequence, the (C/U)UUG(U/C) sequence encompassing the predicted SL2 loop is the most conserved contiguous run of nucleotides in the entire 5′ UTR and contains features of a canonical U-turn motif, in which the middle 3 nts of the loop, UNR (U_0●N_{+1}●R_{+2}), form a triloop that stacks on a Y:Y, Y:A, or G:A noncanonical base-pair. The basic structural feature of the canonical U-turn is a sharp turn in the

phosphate backbone between U_0 and N_{+1}, with U_0 engaged in two critical hydrogen bonds: the U_0 imino proton donates a hydrogen bond to the nonbridging phosphate oxygen following R_{+2}, and the U_0 2′-OH proton donates a hydrogen bond to the N7 of R2. NMR studies of SL2 indicated that the U_0 imino proton donates a hydrogen bond in SL2, consistent with a U-turn structure. However, there was no evidence for the predicted noncanonical pyrimidine–pyrimidine base-pair between positions 47 and 51. Additional NMR studies indicate that U51 was extruded from the loop and that the Watson–Crick faces of C47 and G50 were in apposition (Li and Giedroc, unpublished). Formation of the stem was required for virus viability, although the sequence of the stem was unimportant (Liu et al. 2007). Replacing U48 with either cytosine or adenosine was lethal, consistent with a UNR loop structure for SL2 (Liu et al. 2007). However, viruses containing a U48G mutation were viable and replicated almost as well as wild-type virus. NMR studies indicated that a guanine at position 48 engaged in a hydrogen bond, similar to that observed for U48. Mutagenesis of U49 and U51 demonstrated that any nucleotide can function in these positions, whereas the G at position 50 is required (Liu et al. 2007; Liu et al. 2009). RT-PCR analyses of cells electroporated with genomes containing lethal mutations in SL2 demonstrate that SL2 is required for subgenomic RNA synthesis, as was SL1. Taken together the functional and structural data suggests that SL2 more closely resembles a YNMG-like tetraloop than a U-turn. Additional NMR studies should provide an atomic resolution structure to determine the preise geometry of the loop.

4.2.3 The 3′ cis-Acting RNA Elements

Experiments to dissect the *cis*-acting elements in the coronavirus 3′ UTR have generated a comprehensive view of *cis*-acting elements in this region. Initial DI deletion analyses found that the minimal 3′ terminus sequence required to support MHV DI RNA replication is 436 nts, a region containing part of the upstream N gene and the entire 301 nt 3′ UTR (Lin and Lai 1993; Luytjes et al. 1996). For TGEV and IBV, the minimal sequence requirements were 492 nts and 338 nts respectively, and did not include any part of the N gene (Mendez et al. 1996; Dalton et al. 2001). It was later confirmed this was also true for MHV, as a recombinant virus containing the N gene translocated into an upstream genomic position was viable (Goebel et al. 2004a). For MHV, the differing conclusions resulting from DI assays and intact virus may reflect the fact that DI assays are inherently competition assays with wild-type genomes and thus may be more sensitive at detecting minor decreases in relative fitness then assays with infectious viruses that focus on recovering viable viruses. This was clearly true in experiments in which DI RNAs carrying mutations at the 3′ end failed to replicate at detectable levels, but recombinant viruses with these same mutations were viable with only modestly impaired replication phenotypes (Johnson et al. 2005). A deletion analysis utilizing a DI RNA replicon defined the minimum sequence needed for minus-strand RNA

synthesis as the 3′-most 55 nts plus the poly(A) tail (Lin et al. 1994). Spangnolo and Hogue demonstrated that the poly(A) tail was required for DI replication, although as little as five As would suffice to initiate replication (Spagnolo and Hogue 2000).

Genetic and enzymatic probing of MHV and BCoV 3′ UTR secondary structure demonstrated the presence of three RNA secondary structures (Fig. 4.1) (Hsue and Masters 1997). The 5′-most of these, a 68 nt bulged stem-loop immediately downstream of the N gene stop codon, was predicted to be absolutely conserved in MHV, BCoV, HCoV-OC43, and bovine enteric coronavirus. This stem-loop was further characterized biochemically, and mutagenesis demonstrated that it was essential for DI RNA and for viral replication (Hsue et al. 2000). A 54 nt hairpin-type pseudoknot 3′ to the 68 nt bulged stem-loop was first found to be required for DI RNA replication in the 3′ UTR of BCoV (Williams et al. 1999). This pseudoknot is phylogenetically conserved among coronaviruses, including the SARS-CoV (Goebel et al. 2004b), both in location and in shape but only partially in nucleotide sequence. In a later study with MHV (Goebel et al. 2004a), this pseudoknot was demonstrated to partially overlap with the bulged stem-loop, such that the last part of the bulged stem-loop overlaps with stem 1 of the pseudoknot; thus these two structures cannot be formed simultaneously. This finding led to the proposal that the bulged stem-loop and the pseudoknot are components of a molecular switch that regulate viral RNA synthesis (Goebel et al. 2004a).

The third RNA secondary structure, a complex multiple stem-loop structure, is further downstream in the MHV 3′ UTR. This structure was predicted by computer-assisted analysis of the last 166 nts of the genome 3′ to the pseudoknot using the Mfold algorithm (Liu et al. 2001). Enzymatic probing of RNA secondary structure supported the existence of the predicted long bulged stem-loop encompassing nts 143–68 and with a second stem-loop from nts 67 to 52. Within the long stem-loop, a conserved bulged-stem structure (nts 142–132 and nts 79–68) also present in BCoV was identified by covariation analysis. Site-directed mutagenesis and DI RNA replication assays indicated that the long bulged-stem loop between nts 143 and 68 plays an important role in DI RNA replication. Similar assays to examine the shorter stem-loop between nts 67 and 52 failed to provide evidence for a role in DI replication (Liu et al. 2001). The long bulged stem-loop contains an octanucleotide sequence, 5′-GGAAGAGC-3′ (nts 81–74 in MHV), that is conserved in the 3′ UTR of coronaviruses from all three groups, and thus might have important biological functions.

Further analysis of the long bulged stem-loop revealed that although the octanucleotide sequence is almost universally conserved in coronaviruses, the remainder of this complex stem-loop resides in a hypervariable region (HVR) of the 3′ UTR that is poorly conserved in group 2 coronaviruses. An extensive mutational analysis of the HVR was carried out by deletion, rearrangements, and point mutations (Goebel et al. 2007). All these mutations have only modest effects on viral replication, indicating that the HVR is not essential for viral RNA synthesis. This result differs from the results obtained in DI systems by Liu et al. (2001) and by Lin et al. (1994). Since the most extensive HVR mutant deleted nts 30–170, it is clear

that not all 55 nts are required for minus-strand replication. A possible explanation for these discrepancies is the inherent competitive nature of DI replication assays greatly increasing the effects of mutations that are only moderately deleterious in the context of the intact genome. The HVR deletion mutant was highly attenuated in mice, suggesting that the HVR might play a significant role in viral pathogenesis (Goebel et al. 2007). However, it should be kept in mind that the HVR deletion virus grew to a titer 2–3-fold less than that of wild-type virus in cell culture, making the interpretation of its effect on pathogenesis difficult.

Most recently, multiple second-site revertants of the pseudoknot were recovered by characterizing an unstable mutant Alb391, with a 6 nt insertion of AACAAG in loop 1 of the pseudoknot of MHV 3′ UTR. These second-site suppressor mutations were localized to two separate regions of the genome: one group of mutations was mapped to nsp8 and nsp9 and the second group mapped to the extreme 3′ end of the genome. These observations led the authors to point out that coronavirus replicase gene products might interact with the 3′ end of the genome, and that the loop 1 of the pseudoknot has the potential to base-pair with the extreme end of the genome (Zust et al. 2008). This observation is supported by structural predictions, phylogenetic conservation of the interaction amongst all known group 2 coronaviruses, and the ability of a drastically minimized truncation mutant ΔHVR3 in which all sequences between nts 29 and 171 were replaced by a tetraloop capping the helix downstream of pseudoknot stem 2 (Zust et al. 2008). However, this base-pairing interaction has not yet been demonstrated biochemically or by mutagenesis.

Interestingly, the 3′ UTR stem-loop structures of the group 2 coronaviruses seem to be different from both group 1 and group 3 coronaviruses. All the group 1 coronaviruses contain a highly conserved pseudoknot (Williams et al. 1999), but there is no bulged stem-loop structure in the 3′ UTR. The group 3 coronaviruses have a highly conserved and functionally essential stem-loop (Dalton et al. 2001), but only a poor candidate for the pseudoknot structure can be found nearby (Williams et al. 1999). Only the group 2 coronaviruses have both the pseudoknot and the bulged stem-loop in close proximity and they all overlap in the same fashion. Despite their primary sequence divergence among the 3′ UTRs of group 2 coronaviruses, the secondary structures are all highly conserved and functionally equivalent, as shown by the replication of a BCoV DI RNA in the presence of various group 2 helper viruses (Wu et al. 2003), and by the isolation of chimeras in which the 3′ UTRs of BCoV and SARS-CoV, both group 2 coronaviruses, replaced their MHV counterpart without affecting viral viability (Hsue and Masters 1997; Goebel et al. 2004b). However, the MHV 3′ UTR cannot be replaced with either the group 1 TGEV 3′ UTR or the group 3 IBV 3′ UTR (Hsue and Masters 1997; Goebel et al. 2004b).

4.2.4 Proteins Binding to the 5′ and 3′ cis-Acting Elements

Although exactly how a coronavirus synthesizes its RNAs is still unclear, there is increasing evidence that coronavirus discontinuous synthesis of subgenomic

mRNAs is directed by *cis*-acting sequences present on the viral RNAs with the help of *trans*-acting factors encoded by the virus, as well as cellular proteins. Although it has been demonstrated that continuous protein synthesis is required for viral RNA synthesis (Sawicki and Sawicki 1986), little is known as to which viral and cellular proteins are involved in viral RNA transcription and replication.

The current discontinuous RNA synthesis model proposes that the TRS-L sequence is brought into close proximity to sequences located at the 3′ end of the genomic RNA through RNA–RNA, or RNA–protein and protein–protein interactions. Indeed, several cellular proteins have been shown to interact with the 5′ and 3′ ends of the coronavirus genome (see Fig. 4.1). At the 3′ end of the coronavirus genome, the 73 kDa cytoplasmic poly(A) binding protein (PABP) binds to the poly (A) tail (Spagnolo and Hogue 2000). A series of host proteins were found to bind to the MHV 3′-most 42 nt RNA probe using RNase protection/gel mobility shift and UV cross-linking assays (Yu and Leibowitz 1995a, 1995b). Further analysis revealed that these proteins include mitochondrial aconitase and the chaperones mitochondrial HSP70, HSP60, and HSP40 (Nanda et al. 2004; Nanda and Leibowitz 2001). DI replication assays suggested that proteins binding to both the poly(A) tail and the last 42 nts of the MHV genome had a role in replication. Mutations in the 3′(+)42 host protein binding element had a deleterious effect on the accumulation of DI RNA, and when the same mutations were introduced into the MHV genome, one mutant was found to be nonviable. This mutant had a defect in subgenomic mRNA synthesis which points to a potential role for sequences at the extreme 3′ end of the MHV genome in subgenomic RNA synthesis (Johnson et al. 2005), a finding consistent with the model proposed by the Enjuanes group (Zuniga et al. 2004). Polypyrimidine tract-binding (PTB) protein was shown to bind to two regions of the MHV 3′ UTR, a strong PTB binding site was mapped to nts 53–149, and another weak binding site was mapped to nts 270–307 on the complementary strand of the 3′ UTR (Huang and Lai 1999). Since a number of these binding sites are deleted in the replication competent HVR mutant virus discussed above, it is unlikely that most of these proteins are required for viral replication (Goebel et al. 2007).

The viral protein N binds to the TRS-L specifically and with high affinity (Nelson et al. 2000). It has been suggested that N protein binding to TRS-L favors translation of viral RNAs (Tahara et al. 1998) and may also play a role in MHV RNA synthesis (Li et al. 1999). Recently, BCoV NSP1 was shown to bind three 5′ UTR and one 3′ UTR-located *cis*-replication stem-loops and may function to regulate viral genome translation or replication (Gustin et al. 2009). Another cellular protein, the heterogeneous nuclear ribonucleoprotein (hnRNP) A1 was demonstrated to bind to the MHV minus-strand leader and TRS-B complementary sequences through immunoprecipitation after UV cross-linking and by in vitro binding assays with recombinant protein (Li et al. 1997). hnRNP A1 was shown to have two binding sites at the MHV 3′ end and these binding sites are complementary to the sites on the minus-strand RNA that bind PTB (Huang and Lai 2001). Mutations that affect PTB binding to the negative strand of the 3′ UTR also inhibited hnRNP A1 binding on the positive strand, indicating a possible

relationship between these two proteins. Furthermore, both hnRNP A1 and PTB bind to the complementary strands at the 5′ end of MHV RNA. Based on these observations, it was proposed that hnRNP A1–PTB interactions provide a molecular mechanism for potential 5′–3′ cross-talk in MHV RNA, which may be important for RNA replication and transcription. However, the role of hnRNP A1 in viral replication is controversial. Paul Masters' group (Shen and Masters 2001) tested the role of hnRNP A1 in viral transcription and replication by inserting a high-affinity hnRNP A1 binding site in place of, or adjacent to, an intergenic sequence in the MHV genome. This inserted hnRNP A1 binding site was not able to functionally replace or enhance transcription from the intergenic sequence. Additionally, MHV was able to replicate normally and synthesize normal levels of genome and subgenomic RNAs in cells lacking hnRNP A1, suggesting that hnRNP A1 is not required for MHV discontinuous transcription or genome replication. However, it was subsequently shown that other members of the hnRNP family can substitute for hnRNAP A1 (Choi et al. 2004).

A recent study also showed that a cytoplasmic host factor is indispensable for SARS-CoV in vitro RNA synthesis, although this host factor has not yet been identified (van Hemert et al. 2008). How these and other potential cellular proteins interact with the coronavirus 5′ and 3′ *cis*-acting elements to initiate and support genomic and subgenomic RNA synthesis is a long unanswered interesting question.

4.3 Future Directions

There is a paucity of data demonstrating viral proteins binding to specific *cis*-acting sequences in the coronavirus genome, with the exception of the N protein to TRS-L discussed above. Although there is genetic evidence for nsp8 and nsp9 interacting with the 3′ UTR (Zust et al. 2008), there is no data defining precisely where these proteins bind, nor is there any evidence as to how any other replicase components bind to the genome. We anticipate that research over the next few years will answer these questions and clarify how the various virus and host proteins function in viral replication.

References

Alonso S, Izeta A, Sola I, Enjuanes L (2002) Transcription regulatory sequences and mRNA expression levels in the coronavirus transmissible gastroenteritis virus. J Virol 76(3): 1293–1308

Brierley I, Boursnell ME, Binns MM, Bilimoria B, Blok VC, Brown TD, Inglis SC (1987) An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J 6:3779–3785

Brown CG, Nixon KS, Senanyake SD, Brian DA (2007) An RNA stem-loop within the bovine coronavirus nsp1 coding region is a *cis*-acting element in defective interfering RNA replication. J Virol 81:7716–7724

Budzilowicz CJ, Wilczynski SP, Weiss SR (1985) Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. J Virol 53:834–840

Chang RY, Hofmann MA, Sethana PB, Brian DA (1994) A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. J Virol 68:8223–8231

Chang RY, Krishnan R, Brian DA (1996) The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. J Virol 70(5):2720–2729

Choi KS, Mizutani A, Lai MM (2004) SYNCRIP, a member of the heterogeneous nuclear ribonucleoprotein family, is involved in mouse hepatitis virus RNA synthesis. J Virol 78(23):13153–13162

Dalton K, Casais R, Shaw K, Stirrups K, Evans S, Britton P, Brown TDK, Cavanagh, D. (2001) *cis*-Acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. J Virol 75:125–133

Goebel SJ, Hsue B, Dombrowski TF, Masters PS (2004a) Characterization of the RNA components of a putative molecular switch in the 3′ untranslated region of the murine coronavirus genome. J Virol 78(2):669–682

Goebel SJ, Taylor J, Masters PS (2004b) The 3′ *cis*-acting genomic replication element of the severe acute respiratory syndrome coronavirus can function in the murine coronavirus genome. J Virol 78(14):7846–7851

Goebel SJ, Miller TB, Bennett CJ, Bernard KA, Masters PS (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(3):1274–1287

Gustin KM, Guan B, Dziduszka A, Brian DA (2009) Bovine coronavirus nonstructural protein 1 (p28) is an RNA binding protein that binds terminal genomic *cis*-replication elements. J Virol 83(12):6087–6097

Hsue B, Masters PS (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(10): 7567–7578

Hsue B, Hartshorne T, Masters PS (2000) Characterization of an essential RNA secondary structure in the 3′ untranslated region of the murine coronavirus genome. J Virol 74(15): 6911–6921

Huang P, Lai MM (1999) Polypyrimidine tract-binding protein binds to the complementary strand of the mouse hepatitis virus 3′ untranslated region, thereby altering RNA conformation. J Virol 73(11):9110–9116

Huang P, Lai MMC (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(11):5009–5017

Izeta A, Smerdou C, Alonso S, Penzes Z, Mendez A, Plana-Duran J, Enjuanes L. (1999) Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. J Virol 73:1535–1545

Johnson RF, Feng M, Liu P, Millership JJ, Yount B, Baric RS, Leibowitz JL (2005) The effect of mutations in the mouse hepatitis virus 3′(+)42 protein binding element on RNA replication. J Virol 79(23):14570–14585

Kang H, Feng M, Schroeder ME, Giedroc DP, Leibowitz JL (2006) Putative *cis*-acting stem-loops in the 5′ untranslated region of the severe acute respiratory syndrome coronavirus can substitute for their mouse hepatitis virus counterparts. J Virol 80(21):10600–10614

Kim Y-N, Makino S (1995) Characterization of a murine coronavirus defective interfering RNA internal *cis*-acting replication signal. J Virol 69:4963–4971

Kim Y-N, Jeong YS, Makino S (1993) Analysis of *cis*-acting sequences essential for coronavirus defective interfering RNA replication. Virology 197:53–63

Lai MM, Patton CD, Baric RS, Stohlman SA (1983) Presence of leader in the mRNA of mouse hepatitis virus. J Virol 46(3):1027–1033

Lai MMC, Baric RS, Brayton PR, Stohlman SA (1984) Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. Proc Natl Acad Sci USA 81:3626–3630

Li HP, Zhang X, Duncan R, Comai L, Lai MM (1997) Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription- regulatory region of mouse hepatitis virus RNA. Proc Natl Acad Sci USA 94(18):9544–9549

Li HP, Huang P, Park S, Lai MM (1999) Polypyrimidine tract-binding protein binds to the leader RNA of mouse hepatitis virus and serves as a regulator of viral transcription. J Virol 73(1): 772–777

Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426(6965):450–454

Li L, Kang H, Liu P, Makkinje N, Williams ST, Leibowitz JL, Giedroc DP (2008) Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. J Mol Biol 377:790–803

Liao CL, Lai MMC (1994) Requirement of the 5′-end genomic sequence as an upstream *cis*-acting element for coronavirus subgenomic mRNA transcription. J Virol 68(8):4727–4737

Lin Y-J, Lai MMC (1993) Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinous sequence for replication. J Virol 67:6110–6118

Lin Y-J, Liao CL, Lai MMC (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

Liu Q, Johnson RF, Leibowitz JL (2001) Secondary structural elements within the 3′ untranslated region of mouse Hepatitis Virus strain JHM genomic RNA. J Virol 75(24):12105–12113

Liu P, Li L, Keane SC, Yang D, Leibowitz JL, Giedroc DP (2009) Mouse hepatitis virus stem-loop 2 adopts a uYNMG(U)a-like tetraloop structure that is highly functionally tolerant of base substitutions. J Virol, In press 2009, doi: 10.1128/JVI.00915-09

Liu P, Li L, Millership JJ, Kang H, Leibowitz JL, Giedroc DP (2007) A U-turn motif-containing stem-loop in the coronavirus 5′ untranslated region plays a functional role in replication. RNA 13:763–780

Luytjes W, Gerritsma H, Spaan WJM (1996) Replication of synthetic defective interfering RNAs derived from coronavirus mouse hepatitis virus-A59. Virology 216:174–183

Makino S, Joo M, Makino JK (1991) A system for study of coronavirus mRNA synthesis: a regulated, expressed subgenomic defective interfering RNA results form intergenic site insertion. J Virol 65(11):6031–6041

Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Khattra J, Asano JK, Barber SA, Chan SY, Cloutier A, Coughlin SM, Freeman D, Girn N, Griffith OL, Leach SR, Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JM, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, Vogrig R, Ward D, Watson B, Brunham RC, Krajden M, Petric M, Skowronski DM, Upton C, Roper RL (2003) The genome sequence of the SARS-associated coronavirus. Science 300(5624):1399–1404

Mendez A, Smerdou C, Izeta A, Gebauer F, Enjuanes L (1996) Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. Virology 217:495–507

Moreno JL, Zuniga S, Enjuanes L, Sola I (2008) Identification of a coronavirus transcription enhancer. J Virol 82(8):3882–3893

Nanda SK, Leibowitz JL (2001) Mitochondrial aconitase binds to the 3′-untranslated region of the mouse hepatitis virus genome. J Virol 75:3352–3362

Nanda SK, Johnson RF, Liu Q, Leibowitz JL (2004) Mitochondrial HSP70, HSP40, and HSP60 bind to the 3′ untranslated region of the murine hepatitis virus genome. Arch Virol 149(1):93–111

Nash TC, Buchmeier MJ (1997) Entry of mouse hepatitis virus into cells by endosomal and nonendosomal pathways. Virology 23(1):1–8

Nelson GW, Stohlman SA, Tahara SM (2000) High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. J Gen Virol 81 (Pt 1):181–188

Pasternak A, van der Born E, Spaan WJM, Snijder EJ (2001) Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J 20(24): 7220–7228

Raman S, Brian DA (2005) Stem-loop IV in the 5′ untranslated region is a *cis*-acting element in bovine coronavirus defective interfering RNA replication. J Virol 79(19):12434–12446

Raman S, Bouma P, Williams GD, Brian DA (2003) Stem-loop III in the 5′ untranslated region is a *cis*-acting element in bovine coronavirus defective interfering RNA replication. J Virol 77(12):6720–6730

Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300(5624):1394–1399

Sawicki SG, Sawicki DL (1986) Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. J Virol 57:328–334

Sawicki SG, Sawicki DL (1990) Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J Virol 64:1050–1056

Sethna PB, Hung S-L, Brian DA (1989) Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. Proc Natl Acad Sci USA 86:5626–5630

Shen X, Masters PS (2001) Evaluation of the role of heterogeneous nuclear ribonucleoprotein A1 as a host factor in murine coronavirus discontinuous transcription and genome replication. Proc Natl Acad Sci USA 98(5):2717–2722

Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE (2003) Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the Coronavirus group 2 lineage. J Mol Biol 331(5): 991–1004

Sola I, Moreno JL, Zuniga S, Alonso S, Enjuanes L (2005) Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. J Virol 79(4):2506–2516

Spaan WJM, Rottier PJM, Horzinek MC, van der Zeijst BAM (1982) Sequence relationships between the genome and the intracellular RNA species 1, 3, 6, and 7 of mouse hepatitis virus strain A59. J Virol 42:432–439

Spagnolo JF, Hogue BG (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(11):5053–5065

Tahara SM, Dietlin TA, Nelson GW, Stohlman SA, Manno DJ (1998) Mouse hepatitis virus nucleocapsid protein as a translational effector of viral mRNAs. Adv Exp Med Biol 440: 313–318

van der Born E, Gultyaev AP, Snijder EJ (2004) Secondary structure and function of the 5′-proximal region of the equine arteritis virus RNA genome. RNA 10:424–437

van Hemert MJ, Van den Worm SHE, Knoops K, Mommaas AM, Gorbalenya AE, Snijder EJ (2008) SARS-Coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. PLos Pathog 4(5):1–10

van Marle G, Dobbe JC, Gultyaev AP, Luytjes W, Spaan WJM, Snijder EJ (1999) Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. Proc Natl Acad Sci USA 96:3501–3506

Wang H, Yang P, Liu K, Guo F, Zhang Y (2008) SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. Cell Res 18:290–301

Williams GD, Chang RY, Brian DA (1999) A phylogenetically conserved hairpin-type 3′ untranslated region pseudoknot functions in coronavirus RNA replication. J Virol 73(10):8349–8355

Wu HY, Guy JS, Yoo D, Vlasak R, Urbach E, Brian DA (2003) Common RNA replication signals exist among group 2 coronaviruses: evidence for in vivo recombination between animal and human coronavirus molecules. Virology 315(1):174–183

Yount B, Roberts RS, Lindesmith L, Baric RS (2006) Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: engineering a recombination-resistant genome. Proc Natl Acad Sci USA 103(33):12546–12551

Yu W, Leibowitz JL (1995a) 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

Yu W, Leibowitz JL (1995b) Specific binding of host cellular proteins to multiple sites within the 3′ end of mouse hepatitis virus genomic RNA. J Virol 69:5033–5038

Zuniga S, Sola I, Alonso S, Enjuanes L (2004) Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. J Virol 78(2):980–994

Zust R, Miller TB, Goebel SJ, Thiel V, Masters PS (2008) Genetic interactions between an essential 3′ *cis*-acting RNA pseudoknot, replicase gene products, and the extreme 3′ end of the mouse coronavirus genome. J Virol 82(3):1214–1228