To the Graduate Council:

I am submitting herewith a dissertation written by Kortney Michele Gustin entitled "*cis*-Acting Determinants of Coronavirus Genome Translation and Replication." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Microbiology.

David A. Brian, Major Professor

We have read this dissertation
and recommend its acceptance:

Chunlei Su

Igor B. Jouline

Albrecht von Arnim

Accepted for the Council:

Carolyn R. Hodges,
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

cis-Acting Determinants of Coronavirus Genome Translation and Replication

A Dissertation
Presented for the
Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Kortney Michele Gustin
December 2008

Acknowledgements

I would like to thank my family and friends for their encouragement and support. A special thanks goes to Jeremiah, we started and finished our graduate careers together, and although it was a long and arduous journey, we did it together. I would also like to thank my advisor, David Brian for his guidance, support and enthusiasm for science. For all of their advice, help, and friendship, I thank all of the current and past Brian lab members, Hung-Yi, Agnieszka, Bo-Jhih, and Kim. Finally, thanks to my committee members Dr. Su, Dr. von Arnim, and Dr. Jouline, for their support and interest in my research.

Abstract

Coronaviruses are a family of positive-sense, single-stranded, 5'-capped and 3'-polyadenylated RNA viruses that replicate entirely in the cell cytoplasm. Replication of the viral genome requires translation to produce proteins used for RNA synthesis and virion assembly. The 5'- and 3'- untranslated regions of the coronavirus genome have been found to contain *cis*-acting elements that are required for replication of the genome and a defective interfering RNA. Presumably, both viral and cellular proteins interact with these elements and serve as *trans*-acting factors in genome translation and replication. Of interest is the functional significance of a 5'-proximal *cis*-acting 397-nucleotide region which includes the 210-nucleotide 5' untranslated regions and the 5' proximal 187 nucleotides of the coding region for the first nonstructural protein, p28.

Here, research examining the structural requirements in this region and in the 3' untranslated region for bovine coronavirus replication is presented. The following features were discovered: (i) A second higher order structure within the coding region of p28 is required for replication but not translation of the defective interfering RNA. Although no viral proteins were found to bind this region, two cellular proteins of ~60 and ~100kDa were found which might prove to be essential *trans*-acting factors. (ii) p28, the first synthesized protein in the genome is an RNA binding protein that interacts with *cis*-replication stem-loops in the 5' and 3' untranslated regions. p28 may therefore function through these interactions to regulate genome translation or replication. (iii) In constructs with a renilla luciferase reporter containing only the genomic 5' and 3' untranslated regions, *in vitro* translation in the absence of viral proteins revealed a 5' and 3' end interaction that mediates a repression of cap-dependent translation. Therefore, translation of the coronavirus genome may require the assistance of viral proteins.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

aa	amino acid
BCoV	bovine coronavirus
CoV	coronavirus
DI RNA	defective interfering RNA
EtBr	ethidium bromide
hnRNP	heterogeneous nuclear ribonucleoprotein
HCV	hepatitis C virus
HCoV	human coronavirus
IBV	infectious bronchitis virus
kb	kilobase
kcal/mol	kilocalories/mole
kDa	kilodalton
MHV	mouse hepatitis coronavirus
Min	Minute
N	nucleocapsid protein
Nsp	nonstructural protein
Nt	nucleotide
ORF	open reading frame
PABP	poly(A) binding protein
PCBP	poly(C) binding protein
PTB	polypyrimidine tract binding protein
PV	Poliovirus
RdRp	RNA-dependent RNA-polymerase
RRL	rabbit reticulocyte lysate
RLuc	Renilla Luciferase
SARS	severe acute respiratory syndrome
SARS-CoV	SARS coronavirus
sgmRNA	subgenomic messenger RNA
SL	stem-loop
TGEV	transmissible gastroenteritis virus
Timept	timepoint
UTR	untranslated region
VP1/(2)	virus passage to fresh cells (indicates packaging of progeny viral RNA)

Chapter I. Literature Review

Background

Coronavirus History

The Coronaviridae belong to the order Nidovirales and were formally characterized as a family around 1975 (122). However, the viruses that make up this family were first isolated as early as the 1930s and new viruses continue to be isolated from birds, mammals, and humans. As shown in Table 1-1, coronaviruses have been categorized into three groups based on antigenic differences and genome sequence (106). Group 3 includes infectious bronchitis virus (IBV), a significant bird pathogen, as well as other avian coronaviruses that cause respiratory and enteric diseases in birds (112). Recently, the first non-avian group 3-like coronavirus was isolated from a beluga whale (110). Group 1 includes both animal and human pathogens such as transmissible gastroenteritis virus (TGEV) and two human coronaviruses, HCoV-229E and HCoV NL63. Infection with a group 1 virus typically results in acute respiratory and gastroenteric disease (156). Group 2 coronaviruses are perhaps the most well studied of the coronaviruses partially due to the appearance of SARS in 2003 and also because both mouse hepatitis virus (MHV) and bovine coronavirus (BCoV) have been used as models for the study of coronavirus replication and pathogenesis. Group 2 contains the majority of known human coronaviruses including HCoV-OC43, HCoV 4408, and HCoV HKU1, all of which have been associated with respiratory infections (77). Other group 2 viruses infect animals and are of significant economic importance in terms of farming and agriculture and also cause both respiratory and gastroenteric diseases (112).

Recently, coronaviruses have been the subject of intense research due to the appearance of SARS-CoV. The emergence of SARS in early 2003 resulted in 8,098 reported cases and 774 deaths (~10%) worldwide (24). Infection most commonly resulted in severe pneumonia and spread via respiratory droplets (24). Since that time the number of new

Table 1-1. Coronavirus Group Members.

Coronavirus	Year Isolated	Host[a]	Disease	Gen.[b]	Genbank[c]
Group 1					
TGEV	1946 (46)	Pig	Gastroenteritis	28586	AJ271965
HCoV 229E	1966 (61)	Human	Respiratory	27317	AF304460
Feline CoV	1970 (163)	Cat	Enteritis, Peritonitis	29147	DQ010921
Canine CoV	1974 (13)	Dog	Enteritis		AY796289
PEDV	1978 (28)	Pig	Diarrhea	28033	AF353511
Rabbit CoV	1980 (91)	Rabbit	Cardiomyopathy		
HCoV NL63	2004 (52, 157)	Human	Respiratory	27553	AY567487
Bat CoV	2005 (126)	Bat	Unknown		AY864196
Raccoon dog CoV	2007 (161)	Rac. Dog	Unknown		EF192157
Group 2					
Mouse hepatitis virus	1949 (29)	Mouse	Enteritis, Hepatitis	31526	NC_006852
PHEV	1962 (59)	Pig	Vomiting	30480	DQ011855
Human CoV OC43	1965 (155)	Human	Respiratory	30738	AY391777
Rat CoV	1970 (121)	Rat	Sialodacryoadenitis		AF207551
Bovine CoV	1973 (145)	Cow	Diarrhea, Dysentery	31028	U00735
Puffinosis CoV	1982 (119)	Bird	Unknown		AJ544718
Equine CoV	1983 (69)	Horse	Diarrhea	30992	EF446615
Human CoV 4408	1994 (174)	Human	Enteritis		AF523848
Canine Respiratory CoV	2003 (50)	Dog	Respiratory		DQ682406
SARS CoV	2003 (88, 123)	Human	Severe pneumonia	29751	AY274119
Bat SARS CoV	2005 (99)	Bat	Unknown	29736	DQ071615
Bat SARS HKU3	2005 (93)	Bat	Unknown	29728	DQ022305
Human CoV HKU1	2005 (167)	Human	Respiratory	29926	AY597011
HCoV type 5	2006 ***	Human	Unknown		AY860197
Bat CoV HKU4	2006 (168)	Bat	Unknown	30286	EF065505
Bat CoV HKU5	2006 (168)	Bat	Unknown	30482	EF065509
Bat CoV HKU9	2007 (169)	Bat	Unknown	29114	EF065513
Group 3					
IBV	1937 (9)	Chicken	Infectious Bronchitis	27608	M95169
Turkey CoV	1957 (154)	Turkey	Gastroenteritis	27657	EU095850
Pheasant CoV	2002 (23)	Bird	Respiratory		AJ619579
Coronavirus SW1	2008 (110)	Whale	Unknown	31686	EU111742
Unclassified					
Duck CoV	2005 (76)	Duck	Unknown		AJ854130
Goose CoV	2005 (76)	Goose	Unknown		AJ854114
Pigeon CoV	2005 (76)	Pigeon	Unknown		AJ854131
Parrot CoV	2006 (57)	Parrot	Unknown		DQ233651
Human CoV NO	2006 ***	Human	Unknown		AM261820
Chicken enteric CoV	2006 ***	Chicken	Unknown		DQ401279
Quail CoV	2007 (35)	Quail	Unknown		EF446155
Asian leopard cat CoV	2007 (43)	Leop. Cat	Unknown		EF584908
Chinese bamboo rat CoV	2007 (43)	Rat	Unknown		EF584902
Chinese ferret badger CoV	2007 (43)	Fer.Bad.	Unknown		EF584914
Flying squirrel CoV	2007 (43)	Squirrel	Unknown		EF584907
Lesser Indian civet CoV	2007 (43)	Civet	Unknown		EF584903
Masked palm civet CoV	2007 (43)	Civet	Unknown		EF584905
Siberian weasel CoV	2007 (43)	Weasel	Unknown		EF584906
Yellow-bellied weasel CoV	2007 (43)	Weasel	Unknown		EF584904

TGEV, transmissible gastroenteritis virus; PEDV, porcine epidemic diarrhea virus; PHEV, porcine hemagglutinating encephalomyelitis virus; IBV, infectious bronchitis virus

[a] host denotes animals from which virus was first isolated.

[b] Genome length in nucleotides

[c] Genbank numbers refer to representative sequence. If not available, a partial sequence is given.

***Unpublished sequences from Genbank.

coronaviruses detected has more than doubled and new coronaviruses representing all three groups have been isolated from birds, bats, humans, and other mammals. For many of the newly discovered and unclassified isolates it is not known if infection actually results in disease or if the virus was obtained from a reservoir host. In all, over 20 complete genomes have been sequenced (Table 1-1).

Coronavirus Replication

BCoV like all coronaviruses is a positive sense, single-stranded RNA virus with a 5' cap and 3' poly(A) tail. The genome is 31kb, making it one of the largest viral RNA genomes known (106). Coronaviruses replicate entirely in the cytoplasm and after uncoating the genome is initially translated to produce viral proteins for RNA synthesis. These replicase proteins are used to replicate the genome and transcribe a set of 5' and 3' coterminal subgenomic messenger RNAs (sgmRNA) which are translated to yield the structural proteins necessary for virion assembly and other accessory proteins (106). In BCoV there are eight sgmRNAs (66, 135) (Fig. 1-1). sgmRNAs all have a common 5' end, the viral genomic leader sequence (nts 1-65), encoded only at the genome and added by discontinuous transcription. The sequences of the 5' UTRs are unique to each sgmRNA and for BCoV range from 70 to 193 nucleotides (nt) in length (135). The genomic 3' UTR and poly(A) tail are also present on each sgmRNA (106). BCoV sgmRNA 7 which encodes the nucleocapsid protein and is the most abundantly translated RNA species, has one of the shortest 5' UTRs, 77nt, and contains the 65nt leader sequence (135). Following replication, genomic RNA is coated to form the nucleocapsid which is then packaged into virus particles by budding through the endoplasmic reticulum or Golgi membranes. Virions collect in vesicles that fuse with the plasma membrane via the exocytic pathway and are released outside the cell (51).

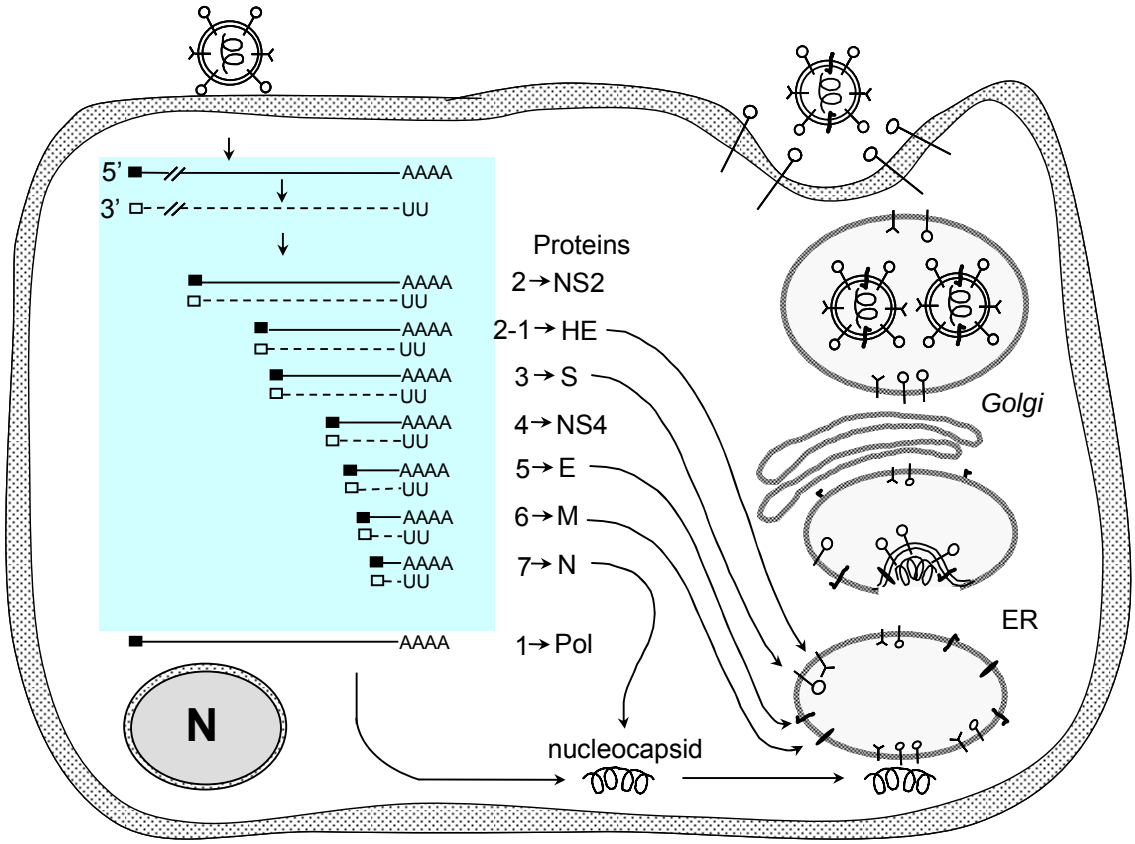


Figure 1-1. Coronavirus Replication Scheme. Adapted from K. Holmes and M. Lai, Fields *Virology*, 1996; Sethna, Hung, Brian. 1989. PNAS 86:5626.

Defective Interfering RNA (DI RNA)

To facilitate identification and characterization of *cis*-acting elements in BCoV, a defective-interfering RNA (DI RNA) has been used. The naturally occurring BCoV DI RNA is composed of the termini of the genome. The 5'-terminal 495 nucleotides (210-nt 5' UTR, 285-nt nsp 1) is fused in frame to the fourth nucleotide upstream of the nucleocapsid (N) start codon and contains the entire N sequence and the 288 nucleotide 3' UTR plus a poly(A) tail (26). To date all characterized coronavirus DI RNAs contain the entirety of the 5' and 3' UTRs and a portion of the first nonstructural protein coding region (14, 15). Replication of the BCoV DI RNA presumably reflects replication of the genome and contains the *cis*-acting elements necessary for genomic replication. BCoV DI RNA replication is accomplished with the help of BCoV trans-acting factors (Fig. 1-2). The BCoV DI RNA and sgmRNA 7, which codes for the nucleocapsid protein are almost identical in composition. The difference is a 421 nucleotide region that includes 136 nucleotides of the 5' UTR and part of the nsp 1 coding region. This is interesting as sgmRNAs do not replicate, leaving the 421 nucleotide region as the determining factor for replication. The region that confers the ability to replicate includes *cis*-acting stem loops III and IV in the 5' UTR, and 285 nucleotides of nsp 1, 187 of which are required for DI RNA replication, including *cis*-acting stem loop VI (15, 20) (Fig. 1-2). These elements are described in more detail below.

cis-Acting Replication Elements

To date six *cis*-acting elements in the 5' and 3' untranslated regions (UTR) have been identified and shown to be required for DI RNA replication. In the 5' terminal 90nt region of the 210nt 5' UTR, two stem-loops, SLI (nt11 to 42) and SLII (nt51 to 84) were predicted to exist by the Tinoco algorithm and confirmed by structure probing (25, 26). Recently in MHV, two different stem-loops in this same region (nt 1 to 51) were described and one of these was shown to be a *cis*-acting replication element (102). The second *cis*-acting element identified in the 5' UTR, SLIII (nt 97 to 116), is a conserved element in group 2 coronaviruses and has been

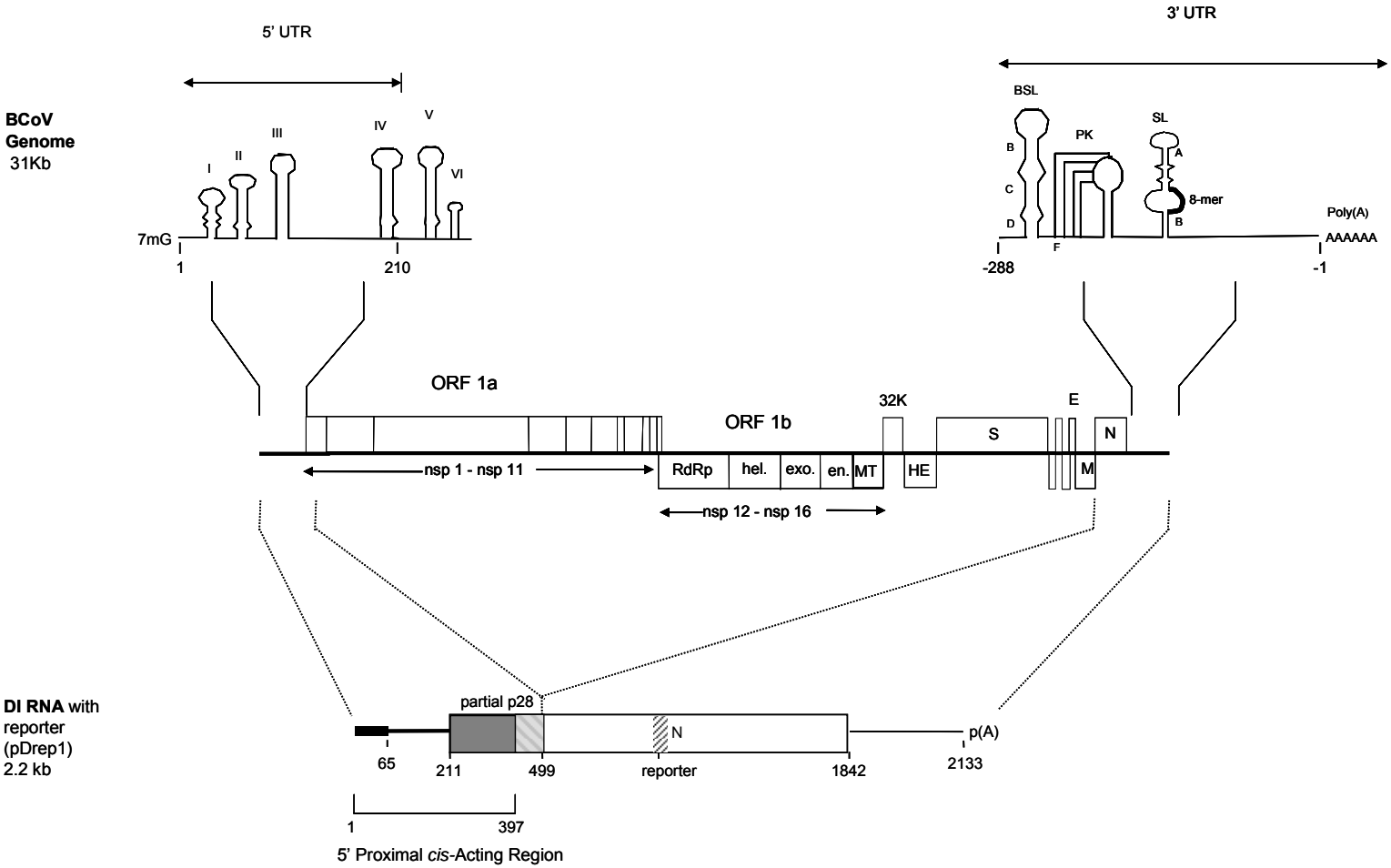


Figure 1-2. BCoV Genome and DI RNA. Sixteen nonstructural proteins are encoded within ORF 1a/b and structural genes including hemagluttin-esterase (HE) and the spike glycoprotein (S) are encoded at the 3' terminus of the genome. *cis*-Acting secondary structures are found in both the 5' and 3' UTRs. Two cis-acting structures are also found in the 5' terminal coding region of nsp 1 (SLV-VI). The BCoV DI RNA is composed of only the termini of the genome, including the 5' proximal *cis*-acting region, the entire nucleocapsid coding region, and the 3' UTR. A 30nt reporter sequence has been inserted for use in replication assays.

confirmed by RNase structure probing (129). The third and final *cis*-acting element identified in the 5' UTR, SLIV (nt 186 to 215), is located immediately upstream of the coding region and contains the start codon for the ORF 1a/1b polyprotein. This stem-loop is conserved among group 2 coronaviruses except for SARS-CoV (130). Three elements within the 3' UTR have been identified as *cis*-acting replication elements. A bulged stem-loop and associated pseudoknot located just downstream of the nucleocapsid coding region, a second bulged stem-loop located near the 3' terminus that contains a conserved octamer sequence GGAAGAGC, and a poly(A) tail (55, 67, 68, 143, 166). In addition to these six elements, a seventh *cis*-acting stem-loop has been identified immediately downstream of the 5' UTR within the coding region of the first nonstructural protein (nsp 1), p28 (Fig. 1-2) (20). This is the second *cis*-acting element discovered in the coding region of coronaviruses, the first being the frame-shifting pseudoknot located at the ORF 1a/1b junction (16, 17). Although these described higher-order elements have been shown to function in DI RNA replication, the exact roles they play in replication of the genome remain unknown.

Translation and Proteins

Translation of the capped and polyadenylated viral genome and sgmRNAs is believed to follow the typical cellular mRNA model and occur in a closed loop or circular complex through interactions between the 5' cap structure and 3' poly(A) tail (15). After adsorption and uncoating, the genomic RNA must be translated in order to produce the viral proteins that are necessary for RNA synthesis and replication. In coronaviruses, the 5' and 3' UTRs have been shown to be important for replication but it is not known how or if these elements function specifically in translation. In the closely related MHV, the leader sequence at the 5' terminus has been described as a translational enhancer in virus infected cells and mutation of stem-loop II in this region decreases translation efficiency *in vitro* (102, 148). In addition, the viral N

protein has been identified as enhancing translation of a reporter RNA in virus infected cells (149). The N protein also binds to the 5' leader sequence; however it is unclear if it functions as a translation enhancer through RNA-protein interactions with the leader sequence (117, 147).

The genomic RNA of coronaviruses is translated into a polyprotein (ORF 1a/1b) that is cleaved co- and post-translationally into sixteen nonstructural proteins (nsps) by two viral proteases (56). The functions of a majority of the nonstructural proteins have been predicted, but only a few have been confirmed with experimental data. ORF 1a contains 11 predicted protein end products (nsp 1-11) and ORF 1b contains five putative protein end products (nsps 12-16), identified as an RNA-dependent RNA-polymerase (RdRp), a helicase, an exonuclease, an endonuclease, and a 2'-O-methyltransferase (Fig. 1-2) (34, 141). Of the putative proteins in ORF 1b, function has been demonstrated for nsps 14 and 15, an exonuclease and endonuclease, respectively, in SARS-CoV (11, 75, 111). In SARS-CoV ORF 1a, nsp 8, has been shown to be a primase, and nsp 9 has been shown to be an RNA binding protein of unknown function (49, 74). Other nonstructural proteins with known functions include nsp 3, which contains two papain-like protease domains and nsp 5, a 3C-like protease (106).

Structural and accessory proteins are encoded in the 3' proximal one-third of the genome and are translated from polycistronic sgmRNAs. Although each sgmRNA contains the coding sequence for downstream ORFs, and are therefore structurally polycistronic, only the 5'-most ORF is translated, making the sgmRNAs functionally monocistronic. All coronaviruses have sgmRNAs that code for the spike glycoprotein (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. Other ORFs code for accessory proteins that are not necessarily conserved among coronaviruses. Some of those identified are group specific, such as the hemagglutinin-esterase (HE) which is found in some group 2 viruses (Fig. 1-2) (15). Accessory proteins, many of which have unknown functions, can be structural or nonstructural proteins (106).

Identified Protein Factors

To facilitate virus replication, both viral and cellular proteins interact with viral RNA and *cis*-acting RNA elements. In coronaviruses, the identification of viral proteins that bind *cis*-acting elements has not been fully explored. However, the identification of RNA-protein interactions with cellular proteins has been better documented. Cellular proteins have been found to bind both the 5' and 3' UTRs. Two identified cellular proteins have been found to bind to the 5' UTR of MHV, hnRNP I or polypyrimidine tract-binding protein (PTB) and hnRNP Q3 (33, 96). In addition to poly(A)binding protein (PABP), mitochondrial aconitase along with three associated heat shock proteins, and hnRNP A1 have been shown to bind the viral 3' UTR (97, 113, 114, 144). Several unidentified cellular proteins ranging from 30 to 142 kDa have also been found and likely include one or more of the identified proteins (173). Binding of these proteins to the UTRs is believed to be important for virus replication. In support of this, PTB and hnRNP Q3 were also found to bind to the complementary strand of the 3' UTR, thus binding to both termini of the genome, and therefore likely participants in RNA synthesis (33, 96). hnRNPA1 has also been found to bind both the leader sequence and anti-leader sequence (complementary leader sequence), and therefore also binds both ends of the genomic RNA (70).

These and other hnRNP family members have been found to be involved in both RNA synthesis and translation of RNA virus genomes such as those for poliovirus and hepatitis C virus (1, 22, 62, 79). For coronaviruses, there is evidence to suggest the identified hnRNP proteins are involved in RNA synthesis. In MHV, PTB-hnRNPA1 binding interactions have been proposed to circularize the genome and provide a circular closed loop for replication (70). Overexpression or depletion of these proteins resulted in altered MHV growth kinetics (32, 33, 138). Alternatively, these proteins do not appear to be involved in translation as the depletion or

downregulation of PTB or hnRNP Q3 has been shown to be inconsequential in cap-dependent translation of an MHV reporter construct (32, 33).

Several of the unidentified cellular proteins have also been found to bind to the UTRs of both BCoV and MHV. No less than eleven cellular proteins have been found to bind to the 3' UTR ranging from 30kDa to 142kDa, some of which may include the already identified proteins (143, 173). In the 5' UTR, six cellular proteins have been found to bind to *cis*-acting SL III and nine cellular proteins have been found to bind *cis*-acting SL IV (128, 130). These fifteen unidentified proteins range from approximately 33kDa to 61kDa and could also include one or all of the currently identified proteins. How these proteins may function in virus replication is unknown.

Nonstructural Protein 1/p28

The first nonstructural protein (nsp 1) encoded by ORF1a is called p28 due to its approximate size of 28kDa. p28 is cleaved from the ORF 1a polyprotein cotranslationally by the viral protease nsp 3 soon after its translation (38, 44, 45, 72, 73). In MHV, mutation of the nsp 1-nsp 2 cleavage site results in reduced replication indicating p28 is important for replication and cleavage from the polyprotein is important for its function (39). In BCoV DI RNA replication, it has been shown that deletion of the partial nsp 1 coding region results in a loss of DI RNA replication (20). In studies with an MHV infectious clone, deletion of this region in the viral genome was also detrimental to virus replication (18). Two stem-loops have been predicted to exist within the 5' proximal coding region of nsp 1 and confirmed to exist by RNase structure probing. One of these, SLVI has been shown to be a *cis*-acting element in DI RNA replication (20). Interestingly, this N-terminal portion of p28 is found in all characterized DI RNAs (14, 15). In addition to the requirement for the RNA secondary structure, the translation product of this coding region is also required for replication. In MHV, deletion of p28 using a full length infectious clone resulted in deficient or no virus replication (18). Alternatively, C-terminal

truncations did not affect replication (18, 177). Also, replication of BCoV DI RNA was inhibited when the start codon was mutated to a stop codon or frameshift mutations were introduced to change the sequence of the partial p28 coding region (20). Thus there exists a separate requirement for both the RNA secondary structure and the protein product of the N-terminal region of p28.

It is postulated that p28 is part of the viral replication complex, and it was found to associate with viral RNA and viral proteins such as the helicase, RdRp, N, and matrix protein. p28 has furthermore been shown to directly bind two viral nonstructural proteins, nsp 7 and nsp 10 (19). The function of p28 in virus replication remains unknown; however effects on cellular processes have been reported. Cell cycle arrest has been reported to occur when MHV p28 is overexpressed in the absence of other viral proteins and this has more recently been observed with the SARS-CoV homolog (30, 164). Other cellular effects reported for both the MHV and SARS-CoV nsp 1 protein when expressed in cells, include cellular mRNA degradation and translation shutoff (78, 115, 164, 177). In addition SARS-CoV nsp 1 has been shown to induce cytokine dysregulation (94).

Observations That Led To Research

Previous research has focused on *cis*-acting replication elements that exist within a 5' proximal 421 nucleotide region of the BCoV DI RNA. This region was initially defined as the difference between the replicating DI RNA and the nonreplicating sgmRNA 7. The 421 nucleotide region includes nucleotides 74 through 210 of the 5' UTR and nucleotides 211 through 495 of the nsp 1 coding region (15). Thus, it was hypothesized that this region contained elements required for replication and therefore sgmRNAs that code for the viral structural genes do not replicate.

Experiments using the BCoV DI RNA have found that there is a separate requirement for both the RNA secondary structure and translation of the partial nsp 1 coding region in the DI

RNA. Translationally silent mutations of a stem-loop structure (SL VI) mapped by RNase enzyme probing resulted in a loss of DI RNA replication. Compensatory mutations that restored the stem structure but altered the primary nucleotide sequence restored replication to wildtype levels (20). This implicated that a higher order stem-loop structure within the coding region of the first nonstructural protein was an important feature in replication of the DI RNA. In addition, experiments in which the translation product of the partial nsp 1 (p28) coding region was altered by a frameshift mutation resulted in a loss of DI RNA replication (20). From these results it was proposed that either the translation product or the act of translation was a replication requirement (20, 25).

In the research described in Chapter II, a second stem-loop structure (SL V) which was mapped immediately upstream of the cis-acting stem-loop VI is shown to be a *cis*-acting higher order structure. Also, the separate requirements for RNA secondary structure and translation led us to hypothesize that the translation product (p28) may be autoregulatory. Chapter II furthermore shows that although p28 does not appear to bind to the *cis*-acting structures within its own coding region, it does interact with other *cis*-acting elements within the 5' and 3' UTRs. Experimental data presented in Chapter II and elsewhere (described in the Introduction) has demonstrated that p28 binds viral RNA and proteins and appears to have an effect on cellular and viral mRNAs. Using this experimental data in combination with a bioinformatic analysis of BCoV p28, putative RNA and protein binding motifs are presented in Chapter IV.

To extend our hypothesis that a viral protein may be regulatory, we focused on translation regulation and the involvement of the untranslated regions. The only well characterized translation regulatory mechanism for coronaviruses is the -1 frameshifting pseudoknot at the ORF 1a/1b junction. In coronaviruses there is the potential for translation regulation to occur in at least three other stages: translation of the viral polyprotein, subgenomic mRNAs, and the translation to replication switch. Currently the only viral protein shown to have

an effect on translation is the N protein although a mechanism has not been elucidated (149). Furthermore, the leader sequence of the MHV 5' UTR has also been shown to influence the efficiency of translation of a reporter construct in virus infected cells (148). However, overall there has been little investigation into the factors (cellular or viral) that regulate the translation of viral proteins and how translation is inhibited and RNA synthesis begins.

Thus, we hypothesize that viral proteins are necessary for efficient translation, regulation, and may modulate the switch from translation to RNA synthesis. The first step in investigating this is to determine how or if the UTRs affect canonical translation. Here, an initial look at the roles of the untranslated regions in translation in the absence of viral proteins is investigated using an *in vitro* system and renilla luciferase reporter constructs. Interestingly, a BCoV specific 5'-3' UTR mediated repression of cap-dependent translation was observed. Characterization of this phenomenon is described in Chapter III.

Chapter II. Bovine Coronavirus p28 (nsp1) is an RNA Binding Protein that Binds Terminal Genomic Cis-Replication Elements But Not Cis-Acting Stem-Loops V and VI within the p28 Cistron

This chapter has been submitted for publication in the *Journal of Virology*. Kortney M. Gustin,[1] Bo-Jhih Guan,[1] Agnieszka Dziduszko,[2] and David A. Brian[1,2,*], Departments of Microbiology[1] and Pathobiology[2] , University of Tennessee College of Veterinary Medicine, Knoxville, Tennessee 37996-0845. *Corresponding Author. Work performed by coauthor Bo-Jhih Guan includes all experiments concerning SLIII and p28. Gel-shift experiments with the 3' UTR and p28 were performed by coauthor Agnieszka Dziduszko. The coauthor's contributions are noted in figure legends.

Introduction

Coronaviruses (142) cause primarily respiratory and gastroenteric diseases in birds and mammals (89, 165). In humans, they most commonly cause mild upper respiratory disease but the recently-discovered human coronaviruses (HCoVs) HCoV-NL63 (157), HCoV-HKU1 (167), and SARS-CoV (105) cause serious diseases in the upper and lower respiratory tract. The SARS-CoV causes pneumonia with an accompanying high (~10%) mortality rate (162). The ~30 kilobase positive-strand coronavirus genome, the largest known among RNA viruses, is 5'-capped, 3'-polyadenylated, and replicates in the cytoplasm (106). As with other characterized cytoplasmically-replicating positive-strand RNA viruses (6), translation of the coronavirus genome is an early step in replication, and terminally-located cis-acting RNA signals regulate translation and direct genome replication (106). How these happen mechanistically in coronaviruses is only beginning to be understood.

In the highly-studied group 2 mouse hepatitis coronavirus model (MHV-A59 strain) and its close relative the bovine coronavirus (BCoV-Mebus strain), five higher-order cis-replication signals have been identified in the 5' and 3' untranslated regions. These include two in the 5' UTR required for BCoV DI RNA replication (Fig. 1A) described as stem-loop (SL) III (129) and SLIV (130). Recently, the SLI region in BCoV (26) has been re-analyzed along with the homologous region in MHV and is now described as comprising SL1 and SL2 (Fig. 1A), of

which SL2 has been shown to be a cis-replication structure in the context of the MHV genome (102). In the 3' UTR, two higher-order cis-replication structures have been identified that function in both DI RNA and MHV genome. These are a 5'-proximal bulged stem-loop and adjacent pseudoknot that potentially act together as a unit (55, 67, 68, 166) and a 3'-proximal octamer-associated bulged stem-loop (103, 172) (Fig. 1A). In addition, the 5'-terminal 65-nt leader and the 3'-terminal poly(A) tail have been shown to be cis-replication signals for BCoV DI RNA (26, 143).

In coronaviruses, the 5'-proximal open reading frame of ~20 kilobases (called ORF 1) comprising the 5' two-thirds of the genome is translated to overlapping polyproteins of ~500 and ~700 kDa, respectively named pp1a and pp1ab (106). pp1ab is formed by a -1 ribosomal frameshift event at the ORF1a-ORF1b junction during translation (106). pp1a and pp1ab are proteolytically processed into potentially 16 nonstructural protein (nsp) end products or partial end products that are proposed to function together as the replicase (56). ORF 1a encodes nonstructural proteins 1-11 which include papain-like proteases (nsp 3), a 3C-like main protease (nsp5), membrane anchoring proteins (nsps 4 and 6), a potential primase (nsp8), and RNA-binding proteins (nsp 7/8 complex, nsps 9 and 10) of imprecisely understood function (41, 42, 56, 58, 74, 108, 125, 175). ORF 1b encodes nonstructural proteins 12-16 which function as an RNA-dependent RNA polymerase, a helicase, an exonuclease, an endonuclease, and a 2'-O-methyltransferase, respectively (12, 36, 56, 111). 3' proximal genomic ORFs encoding structural and accessory proteins are translated from a 3'-nested set of subgenomic mRNAs (106).

The N-terminal ORF 1a protein, nsp1, in the case of BCoV and MHV is named p28 to identify the cleaved 28 kilodalton product (37). The precise role of nsp1 in virus replication has not been determined, but it is known that a sequence encoding an N-proximal p28 region in MHV (nt 255-369 in the 738-nt coding sequence) cannot be deleted from the genome without

loss of productive infection (18). Nsp1 also directly binds nsp 7 and 10 (19), and by confocal microscopy is found associated with the membranous replication complex (18, 158) and virus assembly sites (19). The amino acid sequence of nsp1 is poorly conserved among coronaviruses indicating it may be a protein that interacts with cellular components (2, 141). In the absence of other viral proteins, MHV p28 induces general host mRNA degradation (177) and cell cycle arrest (30). The SARS-CoV nsp1 homolog, a 20 kDa protein, has been reported to cause mRNA degradation (78, 115), inhibition of host protein synthesis (78, 115, 164), inhibition of interferon signaling (164, 177), and cytokine dysregulation in lung cells (94).

In this study, we examine the RNA-binding properties of BCoV p28 with the hypothesis that it is a potential regulator of translation or replication. The rationale for this hypothesis stems from five observations: (i) In the BCoV DI RNA, the 5'-terminal ~one third of the nsp1 coding region and the entire N protein coding region together comprise the single contiguous ORF in the DI RNA, and most of both coding regions appear required for DI RNA replication (26). (ii) The partial p28 cistron in the DI RNA must be translated in cis for DI RNA replication in helper virus-infected cells (20, 25). (iii) A similar part of the nsp1 ORF is found in the genome of all characterized naturally-occurring group 1 and 2 coronavirus DI RNAs described to date (14, 15). (iv) A cis-acting stem-loop named SLVI is found within the partial p28 cistron in the BCoV DI RNA (20). (v) Translation, which involves a 5'$\rightarrow$3' transit of ribosomes, and (-)-strand synthesis, which involves a 3'$\rightarrow$5' transit of the RNA-dependent RNA polymerase, cannot simultaneously occur on the same molecule with a single ORF (7, 80). Thus, to enable genome replication an inhibition of translation at least early in infection for cytoplasmically-replicating (+)-strand RNA viruses is required (7, 8, 54, 81). Mechanisms of translation inhibition have been described for Qβ viral genome wherein the viral replicase autoregulates translation by binding an internal cis-replication element (81), and for the poliovirus genome wherein genome circularization inhibits the early translation step (8, 54). Therefore, since p28 is synthesized

early and also contains a cis-replication element in its coding region, we postulated that it is autoregulatory with RNA binding properties.

Here we (i) demonstrate by mutagenesis analysis that the 72-nt SLV, mapping immediately upstream of SLVI and within the partial nsp1 ORF, is also a cis-acting DI RNA replication element, (ii) show by gel-shift and UV-crosslinking analyses that there is likely no binding of an intracellular viral protein to SLV-VI, but there is binding of unidentified cellular proteins of ~60 and 100 kDa molecular weight, and (iii) show by gel-shift analysis that recombinant p28 purified from *E. coli* does not bind SLV-VI, but does bind SLs I-IV in the 5' UTR and also the 3'-terminal bulged stem-loop in the 3' UTR suggesting a possible regulatory role at these sites. Notably, specific binding with ~2.5 μM affinity of p28 to SLIII and its flanking regions in the 5' UTR was observed. Additionally, we show that under conditions of p28 overexpression from a sgmRNA there is greatly reduced DI RNA levels but only mild transiently reduced levels of viral RNA species. These results together indicate p28 is an RNA-binding protein that may function as a regulator of viral translation or replication, but not through its binding of cis-acting stem-loops V and VI within the p28 cistron.

Materials and Methods

Cells and viruses. A DI RNA-free stock of the Mebus strain of BCoV (GenBank # U00735) at 4.5×10^8 PFU/ml (25, 26) was grown on human rectal tumor cell line HRT-18 (153) as described (26). Other genome sequences analyzed were those of MHV-A59 (GenBank # NC_001846) and SARS-CoV-Toronto (GenBank # NL_004718).

RNA structure predictions. *Mfold* (http://mfold.bioinfo.rpi.edu/) (107, 176) was used to predict RNA secondary structures.

Plasmid constructs. To make BCoV DI RNA with mutations in SLV, PCR mutagenesis (132) was done using pDrep1 DNA. pDrep1 is a pGEM3Zf(-) (Promega)-based plasmid containing a cDNA clone of a naturally occurring 2.2 kb DI RNA of BCoV modified to

carry a 30-nt in-frame reporter (26). For mutagenesis, overlapping primers bordering SLV and containing the desired base changes (Table 2-1) and Hpa I and Xba I sites (occurring in pDrep 1 at nt 150 and 595, respectively) were used for directional ligation into wt pDrep1.

To make BCoV DI RNA for expressing wt p28 from subgenomic mRNAs, a modification of pDrep12.7 (171), which is itself a modification of pDrep1(171), was used and named pDrep-ISp28. For this, a 756-nt fragment containing the upstream 7-nt template-switching core and flanking sequences for N mRNA (AATAATCTAAACTTTAAGG, a total of 18 nt called the intergenic sequence [IS] (87)) and the nsp1 cistron (BCoV genomic nt 211-948, a total of 738 nt) flanked by BamH I and Kpn I restriction enzyme sites was used to replace the 198-nt BamH I (nt 1667)-Kpn I (nt 1865) fragment in pDrep12.7. For pDrep-ISp28$^{1\text{-}95}$ (abbreviated N-term), genomic nt 211-495 (encoding nsp1 aa 1-95) were used in place of the full-length nsp1 cistron to make the construct, and for pDrep-ISp28$^{95\text{-}246}$ (abbreviated C-term), genomic nt 493-948 (encoding nsp1 aa 95-246) were used. To make pDrep-IS-2'O-MT for expressing the 2'-O-MT (nsp 16), BCoV genomic nt 20596-21498 replaced genomic nt 211-948 in pDrep-ISp28 between the BamH I and Kpn I sites.

To make RNA probes for gel-shift analyses, primers bordering the cis-replication elements in pDrep 1 were used that had incorporated EcoR I and Hind III sites for directional ligation into Eco RI/Hind III-linearized pGEM3Zf(-) (Promega).

To make p28 in *E. coli*, the nsp1 coding region from BCoV (nt 211-948) was subcloned with BamH I and EcoR I site-containing primers into the BamH I/EcoR I site of pET28a(+) vector (Novagen). This construct yielded p28 with a 33-nt upstream 6His/thrombin site/T7-tag-containing sequence.

RNA synthesis. For synthesis of uncapped (+)-strand transcripts of SLV DI RNA mutants for testing mutational effects on DI RNA replication, 2 µg of Mlu I-linearized plasmid DNA were transcribed in a 100 µl reaction mix containing 40 U of T7 RNA polymerase

Table 2-1. Oligonucleotides Used for SLV-VI Studies.

Oligonucleotide[a]	Pol[b]	Sequence (5'→3')	Binding Region in pDrep1[c]
pRight-5' 306/309	+	GGATAACCCTAGcAGcTCAGAGGT	294-317
pRight-3' 306/309	-	ACCTCTGAgCTgCTAGGGTTATCC	294-317
pLeft-5' 240/243	+	TCGAgCTgCACTGGGCTCCAGAATTTCC	236-263
pLeft-3' 240/243	-	GGAAATTCTGGAGCCCAGTGcAGcTCGA	236-263
pRtLt-5'	+	CTCGAgCTgCACTGGGCTCCAGAATTTCCcTGGATG	235-270
pRtLt-3'	-	CATCCAgGGAAATTCTGGAGCCCAGTGcAGcTCGAG	235-270
pRtLt (-loops)-5'	+	AGAATTTCCATGGATGTTTGAGGACGCAGA	255-284
pRtLt (-loops)-3'	-	TCTGCGTCCTCAAACATCCATGGAAATTCT	255-284
pLoops-5'	+	CCAGAATTTCCcTGGATGTTcGAaGACGCA	253-282
pLoops-3'	-	TGCGTCtTCgAACATCCAgGGAAATTCTGG	253-282
pRtLoops-L-5'	+	CACTGGGCTCCAGAATTTCCcTGGATG	244-270
pRtLoops-L-3'	-	CATCCAgGGAAATTCTGGAGCCCAGTG	244-270
pRtLoops-R-5'	+	GTTcGAaGACGCAGAGGAGAAGTTGGATAACCCTAGcAGcTCA	270-312
pRtLoops-R-3'	-	TGAgCTgCTAGGGTTATCCAACTTCTCCTCTGCGTCtTCgAAC	270-312
150-HpaI-5'	+	GTTAACAGCTTTCAGCCAGGGACGTGTTG	150-178
600-XbaI-3'	-	TCTAGATTGGTCGGACTGATCGGCC	576-600
BCoV leader	+	GATTGTGAGCG	1-11
TGEV8(+)	-	CATGGCACCATCCTTGGCAACCCAGA	1098-1123
HSVgD(+)	-	GAGAGAGGCATCCGCCAAGGCATATTTG	1854–1885
3'UTR (+)	-	CAGCAAGACATCCATTCTGATAGAGAGTG	2226-2254

[A]The positive and negative symbols in the oligonucleotide name indicate the polarity of the nucleic acid to which the oligonucleotide anneals.
[B]Polarity of the oligonucleotide relative to the positive strand DI RNA
[C]For probe binding to the negative-strand sequence the numbers correspond to complementary positive-strand sequence.

(Promega) following the Promega protocol for synthesis of uncapped nonlabeled RNA. The mix was treated with 3 U of RQ1 RNase-free DNase (Promega) and RNA was purified by silica matrix affinity using the RNaid kit (MP Biochemicals) and quantified spectrophotometrically.

For synthesis of (+)-strand radiolabeled RNA probes, 2 μg of Hind III-linearized plasmid DNA were transcribed in a 50 μl reaction mix containing 40 U of T7 RNA polymerase (Promega) and 120 μCi α^{32}P-UTP (3000 Ci/mmole) (MP Biochemicals). The mix was treated with 3 U of RQ1 RNase-free DNase (Promega) and probes were purified by electrophoresis on a 5 or 9% polyacrylamide/7% urea gel. RNA visualized by autoradiography was eluted from the excised gel band overnight in elution buffer (1mM EDTA, 0.5M ammonium acetate), and ethanol precipitated. Cerenkov counts were determined by scintillation counting in water.

Northern assay for DI RNA replication. Northern blot assays for detecting reporter-containing DI RNAs were performed essentially as described (137, 171). Briefly, cells (~6 × 10^6) at ~80% confluency in a 35-mm dish were infected with BCoV at a multiplicity of 10 PFU per cell and transfected 1 h later with 500 ng of RNA using Lipofectin (Invitrogen). At the indicated times postinfection, RNA was extracted with Trizol (Invitrogen) and stored in water at -20^0C. For passage of progeny virus, supernatant fluids were harvested at 48 hpi and 300 μl was used to infect freshly confluent cells (~8 × 10^6) in a 35-mm dish (called virus passage 1, VP1) from which supernatant fluids at 48 hpi were used to infect new cells (called VP2). For VP1 and VP2, RNA was extracted at 48 hpi. For electrophoretic resolution in a formaldehyde-agarose gel, one quarter of the RNA from each dish was used per lane (where one quarter was ~14, 30, and 44 μg for 1, 24 and 48 hpi, and ~100 μg each for VP1 and VP2). Approximately 1 ng of RNA transcript, identified as RNA in the Northern blot figures, was used as a marker. RNA was transferred to Hybond N+ nylon membrane (Amersham) by vacuum blotting and membranes were UV irradiated. Blots were probed with 20 pmoles of γ-^{32}P-end labeled oligonucleotide probe at ~5×10^5 cpm/pmol. Probes used were TGEV8(+) or HSVgD(+) for

detecting DI RNA progeny and 3'UTR(+) for detecting DI RNA, viral genome and viral sgmRNAs. Probed blots were exposed to Kodak XAR-5 film for 1 to 7 days at -80^{0}C for imaging and counts were quantified with a PhosphorImager (Molecular Dynamics).

Sequence analysis of progeny from mutant DI RNA replicons. Reverse transcription with Superscript II Reverse Transcriptase (Invitrogen) was done on RNA extracted at VP1 and VP2, and PCR was done with primers BCoV leader(-) and TGEV(+). PCR products were sequenced directly.

In vitro translation. 500 ng of wt or mutant DI RNA as transcripts from plasmids were translated in a 50 µl reaction mix of rabbit reticulocyte lysate (Promega) containing 20 µCi ^{35}S-Met. At 90 min, proteins in a 5 µl sample were resolved by SDS-PAGE in a gel of 12% polyacrylamide and radioactivity was quantified on the dried gel by phosphor imaging.

Cell lysate preparation. BCoV-infected and mock-infected cell lysates were prepared essentially as described by Thomson et al (152). Briefly, cells from a 75 cm^2 flask were collected by scraping and pelleted cells were resuspended in 100 µl lysis buffer (10 mM HEPES [pH 7.5], 3 mM $MgCl_2$, 14 mM KCl, 5% glycerol, 0.5% NP40, 1 mM DTT, 1×Complete Protease Inhibitor [Roche]) and incubated on ice for 20 min. Lysates were clarified by microfuge centrifugation at 3000 rpm for 4 min at 4^{0}C and aliquots were stored at -80^{0}C. For infected cell lysates, cells were infected with an MOI of 10 viruses per cell and lysates were prepared at 6 hpi. Protein concentrations were determined by the Bradford assay.

Protein expression and purification from *E. coli*. BL21 (DE3) *E. coli* cells transformed with p28(+)-nsp1 were grown in LB medium containing 34 µg/ml chloramphenicol and 25 µg/ml kanamycin until ~0.6 OD_{600}. Protein expression was induced with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) and cells were grown an additional 15 hrs at 15°C. Cells (~10 g) were pelleted at 5000×G, flash frozen in liquid N_2, and resuspended with sonication in

30 ml lysis buffer (50 mM Tris [pH 8.0], 0.3 M NaCl, 10 mM imidazole, 5% glycerol, 1 mM βME, 1 mM PMSF [phenylmethanesulphonylfluoride], 1× EDTA-Free Complete Protease Inhibitor [Roche], and 25 U/ml Benzonase endonuclease [Novagen]). Fusion protein was purified by chromatography on Ni-NTA superflow resin (Qiagen) and p28 was released in column by thrombin cleavage and concentrated by filtration on YM-10 Centriprep/Centricons cellulose membranes (Millipore).

Gel-shift assays. The method of Andino et al (5) as modified by Silvera et al (140) and as described here was used. Briefly, ~4×10^4 cpm of α-^{32}P-labeled RNA probe (1-5 ng) was incubated in a 15 μl binding reaction mix containing 5 mM HEPES (pH 7.9), 40 mM KCl, 2 mM $MgCl_2$, 4% glycerol, 2 mM DTT, 0.25 μg/μl heparin, 1 U/μl RNasin [Promega], 0.5 μg/μl yeast tRNA, 1× EDTA-Free Complete Protease Inhibitor [Roche] and either 20 μg HRT cell lysate protein, or purified *E.coli*-expressed fusion protein to make [5-30 μM] final. The mix was preincubated without probe for 10 min at 30^{0}C and then for an additional 10 min after adding probe. 2 μl of 50% glycerol was added and reaction products were electrophoretically separated on a 5% nondenaturing polyacrylamide gel for approximately 4 hours at 4^{0}C. Gels were dried and exposed to Kodak XAR-5 film and results were visualized by autoradiography and quantified by phosphor imaging.

UV cross-linking. The method essentially as described by Pestova et al (124) was used. Briefly, ~3×10^5 cpm of α-^{32}P-labeled RNA probe in 15 μl binding reaction mix with cell lysate was incubated as described for the gel-shift assay and irradiated on ice at 3 mwatts/cm^2 for 30 min using a UV Stratalinker 1800. Irradiated reaction mixes were treated with 2.5 U RNase A, 100 U RNase T1, and 0.1 U RNase CV1 (all from Ambion) for 30 min at 37^{0}C and proteins were resolved by SDS-PAGE. Dried gels were exposed to Kodak XAR-5 film for autoradiography.

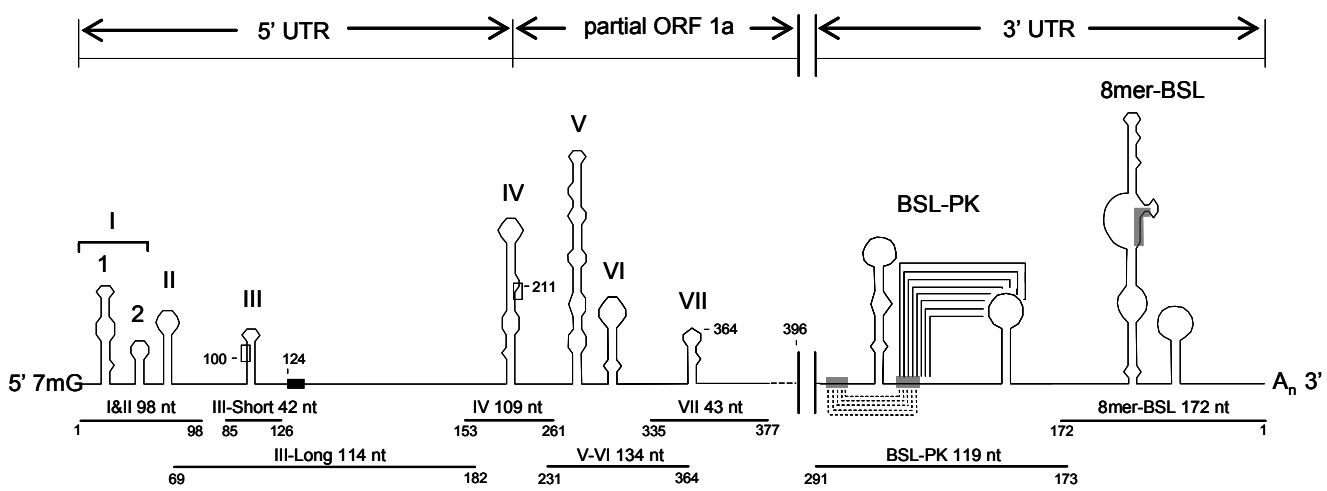
Bioinformatics. Potential RNA and protein binding domains in BCoV and MHV nsp1 were predicted with the SMART program (http://smart.embl-heidelberg.de/) (95, 133). Sequence alignments were made with Vector NTI Suite 8 (Invitrogen).

Results

Stem-loop V mapping within the p28 coding region immediately upstream of SLVI is a required cis-acting signal for DI RNA replication.

A 30-nt stem-loop, SLVI, mapping at nt 101-130 within the 5'-terminal partial p28 (nsp1) coding sequence (genomic nt 311-340) (Figs. 2-1A and 2A) was recently identified as a cis-acting replication element for BCoV DI RNA (20). Although its cis-replication feature remains to be tested in the context of the full-length viral genome, this stem-loop along with the frame-shifting pseudoknot at the ORF 1a/1ab junction (16) are to our knowledge the only higher-order cis-replication elements known within coronavirus ORF 1 (20). Just upstream of SLVI is a structure-mapped 72-nt stem-loop, SLV at nt 29-100 within ORF 1a (genomic nt 239-310) (Fig. 2- 2A) (20), whose function as a cis-acting higher-order structure has not yet been tested. To examine its potential cis-acting role in BCoV DI RNA replication, synonymous base substitutions designed to disrupt base-pairings in a helical region and compensatory mutations designed to restore base-pairings were tested. Silent helix-disrupting substitutions were feasible only in the downstream side of the bottom-most stem for which changes U306C and U309C were made to form mutant pRt and transcripts were tested for DI RNA replication. Following transfection into BCoV-infected cells, pRt DI RNA accumulation as determined by a visible band in Northern analysis was undetectable at 48 hpi and was present only at low levels in virus passages (VP) 1 and VP2 when compared to wt DI RNA levels (Fig. 2-2B). Sequence analysis of progeny RNA showed only wt revertants in VP2 indicating that the mutant had replicated poorly, if at all, and revertants with wt sequence, likely arising through recombination with the 5' end of the helper virus genome, were selected (data not shown). However, when base pairings were restored by

A.



B.

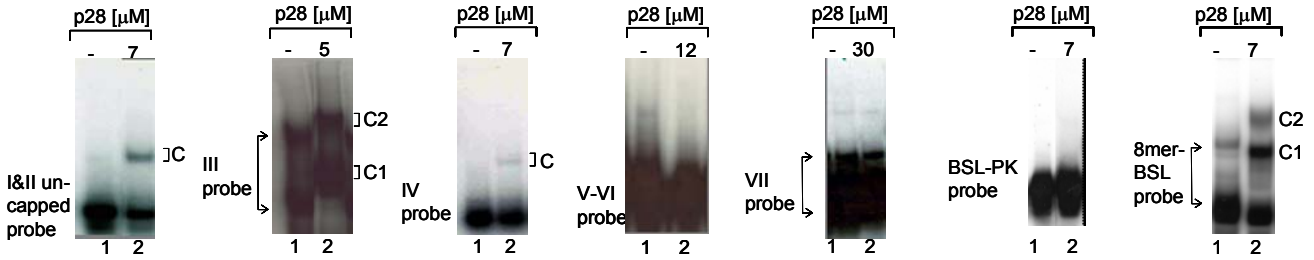
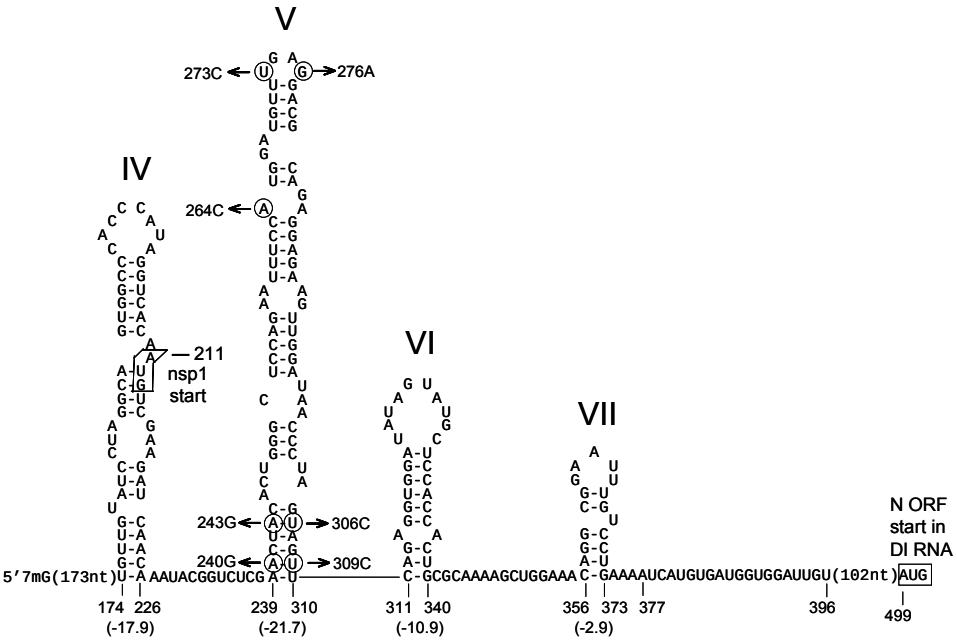


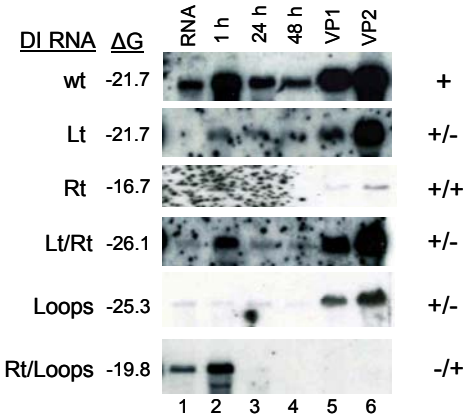
Figure 2-1. RNA Structures in the BCoV Genome Tested for p28 Binding. A. BCoV 5'-terminal and 3'-terminal cis-acting RNA stem-loop structures and flanking sequences identified for BCoV DI RNA replication. Regions of the genome are identified and stem-loop cis-replication elements are identified schematically. Open boxes at nt 100 and 211 identify AUG start codons for the short upstream ORF and ORF 1, respectively. A closed box at nt 124 identifies the UAG stop codon for the short upstream ORF. Shown below the stem-loop structures are the RNA segments used as [32P]-labeled probes in the gel-shift assays. **B.** Gel-shift assays for probes when used with purified p28. Protein-RNA complexes (C) identifying a shifted probe are labeled (coauthor contributions: SLIII by B.Guan, BSL-pk and 8mer-BSL by A.Dziduszko).

Figure 2-2. Mutations in Stem-Loop V that Affect Replication of wt BCoV DI RNA. A. Mapped structures of SLIV-VI and predicted structure of SLVII in wt BCoV DI RNA (as produced from pDrep1). Note that in mutant pΔ397-498 (Brown et al, 2007), nt 397-498 are deleted which resulted in no loss of replicating ability. In mutant pΔ397-498 DI RNA, the N ORF begins at nt 397. The mutations in SLV are as follows: Left, A240G, A243G; Right, U306C, U309C; Loops, A264C, U273C, G276A; Rt/Loops, U306C, U309C, A264C, U273C, G276A. **B.** Northern analysis of wt and SLV mutant DI RNAs. The calculated free energies (ΔG) of wt and mutated SLV's are shown. The Northern analyses used a [32P]-radiolabeled probe specific for the reporter sequence (30-nt TGEV reporter) on extracted RNA at the indicated times post-infection. VP1, virus passage 1; VP2, virus passage 2. Replication is based on the presence of DI RNA by Northern analysis in VP2, and reversion is judged from the sequence of the RT-PCR product in VP2 RNA. If only wt (i.e., reverted sequence) is found in VP2, the mutant is considered dead. **C.** In vitro translation of wt and mutant DI RNA transcripts. Upper panel: [^{35}S]-radiolabeled translation product. Lower panel: EtBr-stained RNA transcript used for in vitro translation.

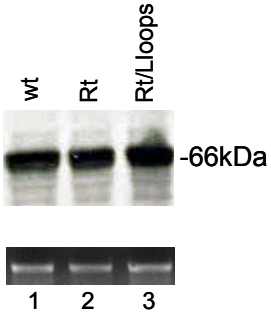
A.


V
IV
VI
VII
273C
276A
264C
211
nsp1
start
243G
306C
240G
309C
5′7mG(173nt)U-A AAUACGGUCUCG A-U C-GCGCAAAAGCUGGAAA C-GAAAAUCAUGUGAUGGUGGAUUGU(102nt)AUG
N ORF
start in
DI RNA
174 226
(-17.9)
239 310
(-21.7)
311 340
(-10.9)
356 373 377
(-2.9)
396
499

B.


DI RNA
ΔG
RNA
1 h
24 h
48 h
VP1
VP2
wt -21.7 +
Lt -21.7 +/-
Rt -16.7 +/+
Lt/Rt -26.1 +/-
Loops -25.3 +/-
Rt/Loops -19.8 -/+
1 2 3 4 5 6

C.


wt
Rt
Rt/Loops
-66kDa
1 2 3

compensatory mutations A240G and A243G in mutant pLtRt, replication as evidenced by accumulation of pLtRt DI RNA in Northern blots was restored to near wt levels in VP1 and VP2 (Fig. 2-2B) and mutations on both sides of the stem were retained in VP2 as revealed by sequence analysis (data not shown). A240G and A243G alone in mutant pLt were not predicted to disrupt basepairing in the stem and, consistent with this structure, their presence alone did not block DI RNA replication as evidenced by DI RNA accumulation in Northern blots (Fig. 2-2B) and sequence analysis. These results together indicate that optimal replication of DI RNA requires base pairing in the lowest helical region of SLV. Furthermore, when three silent point mutations were made in the upper loops of SLV, namely A264C, U273C and G276A (Fig. 2-2A), to form the mutant named pLoops, the overall *mfold*-predicted structure of SLV was only slightly changed (data not shown) and progeny DI RNA from pLoops, although slow in accumulating, retained the parent mutations in VP2. However, when the mutations in pLoops were combined with those in pRt to form pRt/Loops, there was additional structure disruption as predicted by *mfold* and accumulation was further impaired over pRt (Fig. 2-2B). That is, accumulation of pRt/Loops was not visible by Northern analysis at 48 hpi or in VP1 or VP2 (Fig. 2-2B) even though small amounts of wt progeny and progeny with reversions only at 306 and 309 were found by sequencing RT-PCR products (data not shown). These results indicate that the overall higher-order structure of SLV, in addition to integrity in the lower stem, contributes to optimal BCoV DI RNA replication.

Translation of the partial 1a ORF has been shown to be a cis-requirement for DI RNA replication (20, 25). To rule out that the observed loss of replication for SLV mutants was a result of unforeseen blocked translation arising from altered secondary structure (85), capped transcripts of wt and two selected mutants that did not replicate were translated in vitro in rabbit reticulocyte lysate and the products compared. Proteins translated in the presence of ^{35}S-methionine from equal amounts of RNA (Fig. 2-2C) were evaluated by SDS-PAGE and the

results showed that wt, pRt, and pRt/Loops translated at comparable levels (Fig. 2-2C). Thus, the observed decreases in RNA accumulation can be attributed to defects in RNA synthesis and not defects in translation.

p28 binds cis-replication structures in the 5' and 3' untranslated regions, but not cis-acting stem-loops V and VI in the 5' partial p28 coding region.

To test our hypothesis that p28 binds the 5'-proximal cis-replication elements SLV and SLVI within its ORF, two approaches were taken. In the first, evidence of viral protein binding to the SLV-VI probe was sought by gel-shift experiments with lysates of infected cells. In this case, when ^{32}P-labeled SLV-VI probe was tested three retarded protein complexes were found that were also present with lysates from uninfected cells (Fig. 2-3A). Thus, no complexes were found that were unique to BCoV-infected cells suggesting that SLV and VI are bound by cellular but not viral proteins (described further below).

In the second approach, binding was sought with recombinant p28 purified from *E. coli*. When purified p28 was incubated with ^{32}P-labeled SLV-VI, SLV, SLVI, or (putative) SLVII probes, no gel shifting was observed, although shifting was found with other terminal cis-acting structures noted below indicating by this second method that p28 is probably not binding cis-acting SLV and VI (Fig. 2-1B).

Inasmuch as coronaviral proteins other than p28 have been shown to bind cis-replication structures, namely, nucleocapsid protein (N) binds the 5'-proximal SLII region (116, 147, 148), 2'-O-methyltransferase (nsp 16) binds the 3'-terminal octamer-associated bulged stem-loop (Dziduszko, et al, unpublished), and an unknown viral protein(s) from infected cell lysates binds SLIII (Raman and Brian, unpublished), we systematically sought the binding of p28 to the three 5' UTR-located cis-replication elements SLI-II, SLIII and SLIV, and two 3'-UTR-located cis-replication elements, the bulged stem-loop-pseudoknot and the octamer-associated bulged stem-loop. Binding of p28 to these would suggest other possible sites for p28 regulation of

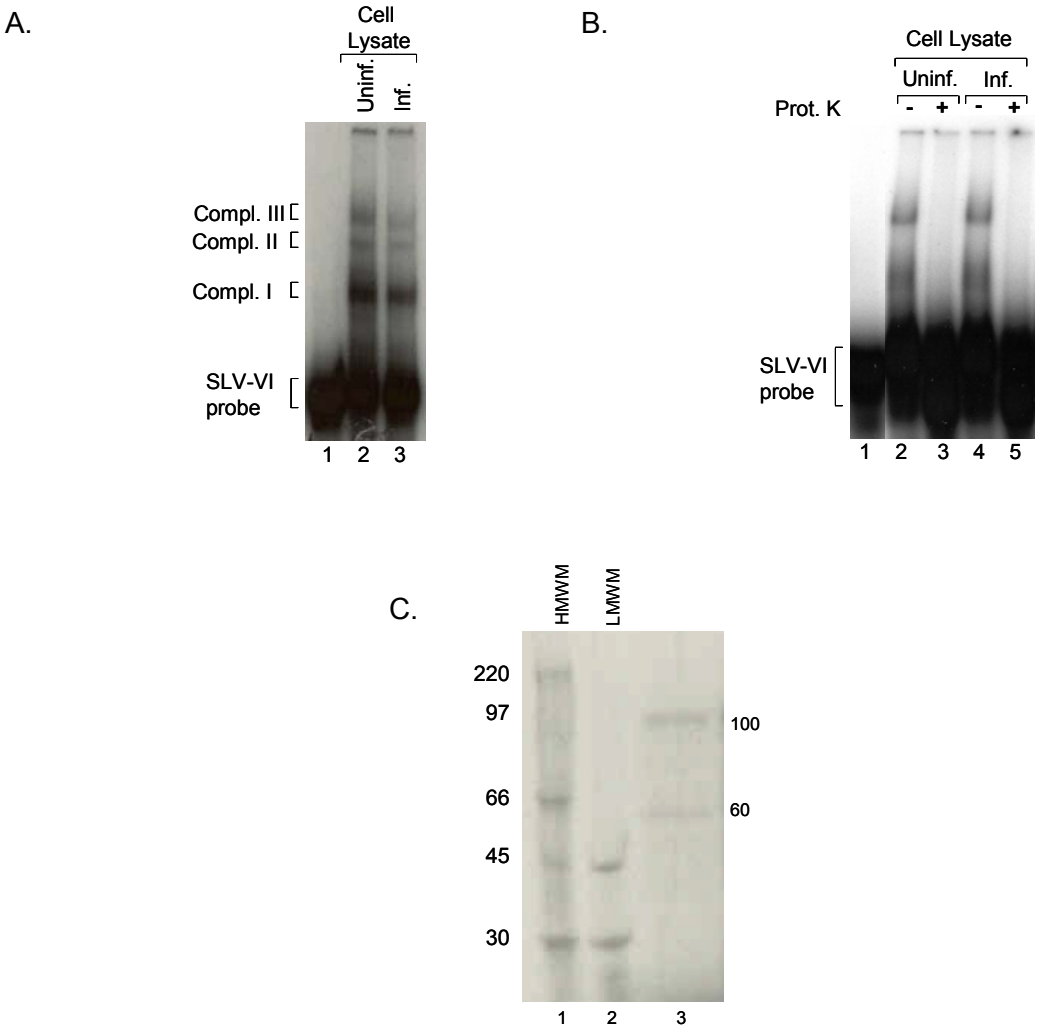


Figure 2-3. Cellular Proteins but Apparently No Viral Proteins From Cell Lysates Bind the SLV-VI Probe. A. Three gel-shift complexes are formed with lysates of uninfected and infected cells with SLV-VI probe. **B.** Protein-RNA complexes with SLV-VI probe are destroyed by proteinase K treatment. **C.** Proteins of 100 and 60 kDa from uninfected cell lysate UV cross-link to the SLV-VI probe.

translation or RNA synthesis. In gel-shift experiments, p28 was not found to bind the 3'-proximal bulged stem-loop-pseudoknot, but was found to weakly bind 5'-proximal SLI-II and SLIV, to moderately bind the 3'-terminal octamer-associated bulged stem-loop, and to quantitatively shift the probe representing 5' SLIII with its extended flanking sequences (SLIII-Long) (Fig. 2-1B). Additionally, it was not found to bind the 5'-proximal uncharacterized putative SLVII in the partial p28 coding region (Fig. 2-1B). These results indicate p28 is an RNA-binding protein that may function to regulate RNA translation or replication through its interaction with 5' or 3' cis-replication elements, but not through binding to SLV-VI.

p28 binds SLIII and its flanking sequences with specificity and micromolar affinity.

Since p28 caused a quantitative shift in the SLIII-Long probe, we sought to establish both its specificity and affinity of binding to SLIII-Long (SLIII with 28-nt upstream and 67-nt downstream flanking sequences), as described in Figure 2-1, and SLIII-short (SLIII with only 10-nt upstream and 10-nt downstream flanking sequences) which nearly represents SLIII proper. Figure 2-4A, lane 1, shows that the SLIII-Long probe migrates as upper and lower bands, probably representing two folded forms of the probe. Lane 2 shows approximately 50% of the lower form and all of the upper form of the probe fully shifting positions with 2.5 μM p28 indicating a K_d of ~2.5 μM for p28 binding to SLIII-Long. Fig. 2-4A, lanes 5-7, show a complete inhibition of SLIII-Long shifting with 5, 10 and 15 μg of unlabeled probe, representing 2,500, 5,000, and 7,500 molar excess over labeled probe, respectively, whereas no inhibition was seen with 5, 10 and 15 μg of tRNA. These results together show specific binding of p28 to SLIII-Long. Fig. 2-4B, lane 1, shows the SLIII-Short probe migrating as a single band. Lanes 2-4 show far less than quantitative shifting of the probe to an upper band and thus a binding affinity was not determined by these experiments. Lanes 5-7, however, show that a band shifting with 2.5 μM protein was completely inhibited by the addition of 4, 8 and 12 μg unlabeled SLIII-Short

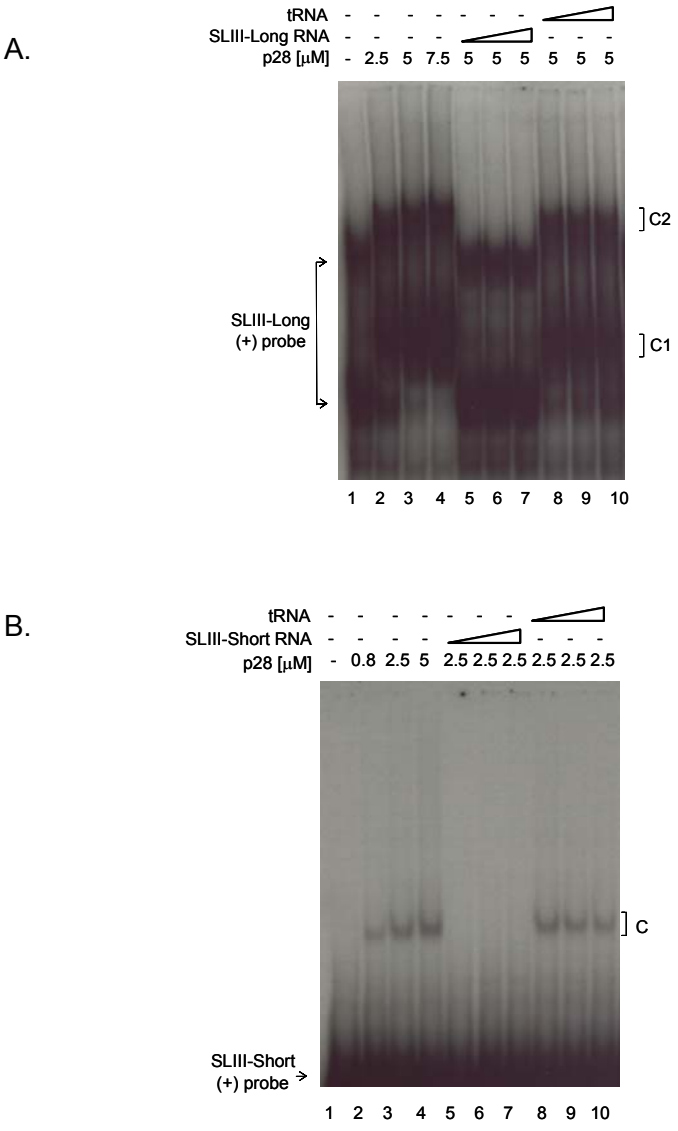


Figure 2-4. Gel-Shift Analysis of Purified BCoV p28 Binding to SLIII. A. SLIII-Long. Lane 1, probe alone showing upper and lower forms. Lanes 2-4, shifted bands into lower (C1) and upper (C2) complexes that form with the indicated amounts of protein. Lanes 5-7, competition of protein binding with 5, 10, and 15 μg of unlabeled SLIII-Long RNA. Lanes, 8-10, competition of protein binding with 5, 10, and 15 μg of tRNA over and above the tRNA in the binding buffer present in all lanes. **B.** SLIII-Short. Lane 2, probe alone. Lanes 2-4, shifted band with the indicated amounts of protein. Lanes 5-7, competition of protein binding with 4, 8, and 12 μg of unlabeled SLIII-Short RNA. Lanes 8-10, competition of protein binding with 4, 8, and 12 μg tRNA over and above the tRNA in the binding buffer present in all lanes (All work presented in this figure was performed by coauthor B.Guan).

probe, respectively, whereas lanes 8-10 show no inhibition of binding with 4, 8, and 12 μg of tRNA, respectively. These results show weaker but still specific binding of p28 to SLIII-Short. Together, the results show that whereas SLIII per se is probably a specific target for p28, the affinity of binding is greatly increased with the presence of its extended flanking sequences.

Unidentified cellular proteins of ~60 and 100 kDa molecular weight bind *cis*-acting stem-loops V and VI.

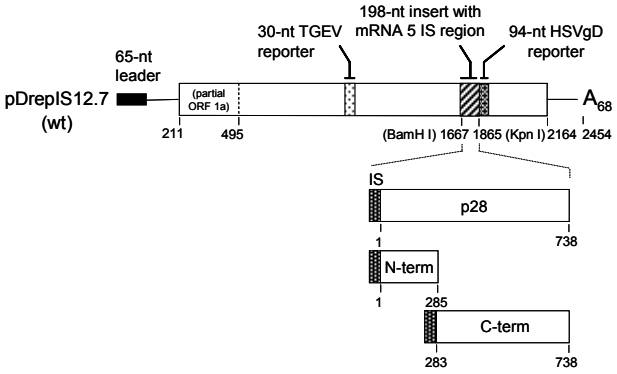
With the use of ^{32}P-labeled SLV-VI probe, three retarded protein-RNA complexes were found by gel-shift experiments that appeared the same in both infected and uninfected cells (Fig. 2-3A). In each case, complex formation was protein-dependent since all were destroyed by protease treatment (Fig. 2-3B). To determine the number and molecular weights of bound cellular proteins, UV-crosslinking followed by RNase digestion and SDS-PAGE was done (Fig. 2-3C). Proteins UV-crosslinked to the ^{32}P-labeled RNA would be labeled by virtue of covalently bound short ^{32}P-labeled RNA tags. Proteins of ~60 kDa and ~100 kDa molecular weight were found to cross-link (Fig. 2-3C). The identity of these proteins and the significance of their interactions with SLV and VI remain to be determined.

The p28 ORF positioned for sgmRNA overexpression from DI RNA leads to long-term inhibition of DI RNA accumulation but to only transient inhibition of viral RNA.

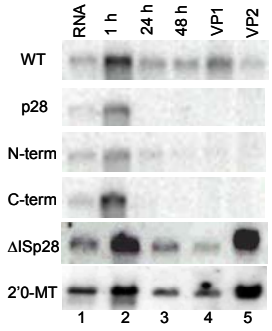
To determine whether higher than normal levels of p28 expression have an effect on BCoV replication as might be expected of a regulatory protein, BCoV nsp1 cistron was cloned into a characterized BCoV DI RNA vector at a site known to express subgenomic RNA in 3-fold molar excess over viral genome (Fig. 2-5A) (171). This method of expression does not involve a non-coronaviral expression system and the expressed protein would presumably be available to the viral replication machinery (136). When BCoV-infected cells were transfected with transcripts of DI RNA pDrep12.7 (wt) (171), pDrep-ΔISp28 (which is a p28 cistron-containing DI RNA construct missing the 18-nt IS transcription start signal), or with transcripts of pDrep-IS-2'-

Figure 2-5. Detrimental Effects of p28 Expression From Subgenomic mRNA in DI RNA. **A.** Map of DI RNA constructs for expressing the full-length p28, the N-terminal 95 aa, and the C-terminal 152 aa of p28. The intergenic sequence (IS) used to induce sgmRNA transcripts for p28 expression is the 18-nt sequence upstream of the BCoV N mRNA (sgmRNA 7) (see text). **B.** Northern analysis of in vivo-generated transcripts of wt and truncated p28 constructs shown in panel A as identified by a probe specific for DI RNA (HSVgD reporter). Results from control constructs are also shown. ΔISp28 represents a construct identical to pDrep-ISp28 except that the 18-nt IS is deleted. 2'O-MT represents a construct identical to pDrep-ISp28 except that the nsp 16 ORF replaces the nsp1 ORF. **C.** Quantitative summary of PhosphorImager data collected from data in panel B. **D.** Northern analysis of viral RNA probed with a 3'end-specific probe (3'UTR(+)). Subgenomic mRNAs 6 and 7 are identified. **E.** and **F.** Quantitative summary of phosphor imaging data shown in panel D for sgmRNA 6 and sgmRNA 7, respectively. Asterisks identify the relative quantities of sgmRNA 6 and sgmRNA 7 for the p28 and wt constructs.

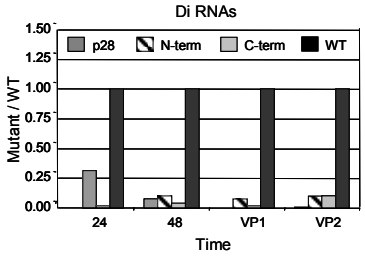
A.


198-nt insert with
mRNA 5 IS region
30-nt TGEV
reporter
94-nt HSVgD
reporter
65-nt
leader
pDrepIS12.7
(wt)
(partial
ORF 1a)
A68
211
495
(BamH I) 1667 1865 (Kpn I) 2164 2454
IS
p28
1
738
N-term
1
285
C-term
283
738

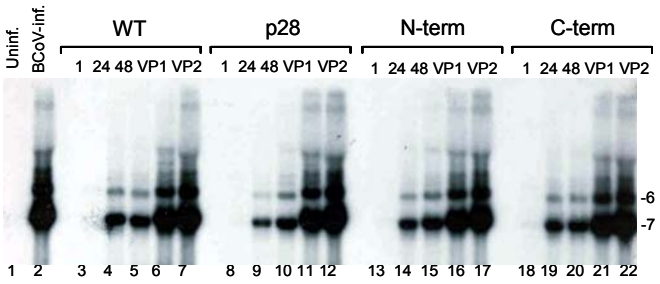
B.


RNA
1 h
24 h
48 h
VP1
VP2
WT
p28
N-term
C-term
ΔISp28
2'0-MT
1 2 3 4 5

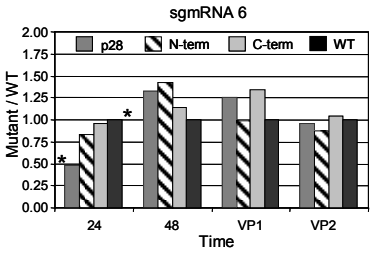
C.


Di RNAs
p28
N-term
C-term
WT
Mutant / WT
1.50
1.25
1.00
0.75
0.50
0.25
0.00
24
48
VP1
VP2
Time

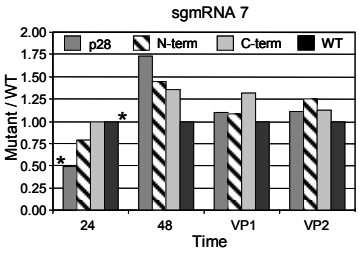
D.


Uninf.
BCoV-inf.
WT
p28
N-term
C-term
1 24 48 VP1 VP2
-6
-7
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

E.


sgmRNA 6
p28
N-term
C-term
WT
Mutant / WT
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
*
*
24
48
VP1
VP2
Time

F.


sgmRNA 7
p28
N-term
C-term
WT
Mutant / WT
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
*
*
24
48
VP1
VP2
Time

O-MT (designed to express nsp 16 from an identical IS signal) and evaluated by Northern analysis, abundant levels of DI RNA for each was observed with a probe that detects the downstream reporter sequence (HSVgD) (Fig. 2-5B). However, with RNA from infected cells transfected with DI RNA designed to express the full-length p28 cistron from a sgmRNA (i.e., pDrep-ISp28), no DI RNA progeny were detected at any timepoint (Fig. 2-5B and C). This result was not an effect of the p28-coding RNA in the DI RNA since progeny were made from pDrep-ΔISp28 which contains the same p28 coding sequence. Interestingly, reduction in viral genome and sgmRNA levels with p28-expressing DI RNA was small and short-lived when compared to expression levels in pDrep12.7-transfected cells (Fig. 2-5D and E). That is, a 2-fold reduction in viral sgmRNAs 6 and 7 was observed at 24 h post-transfection (note asterisks in Fig. 2-5E) but after this time wt levels were reached. We interpret these results to mean that p28 was expressed from sgmRNA and that translation or replication of the DI RNA was more sensitive to the effects of p28 than that of viral genome or sgmRNAs. Interestingly, inhibition of DI RNA accumulation was observed not only with expression of full-length p28 from a sgmRNA, but also with expression of the C-terminal region of p28 (genomic nt 493-948, aa 95-246) representing a part of the genome that can be deleted without loss of viral viability (Fig. 2-5B and C) (18). Less but still significant inhibition was seen with the N-terminal region of p28 (genomic nt 211-495, aa 1-95) at 24 h postinfection, which contains a 30-nt part of the MHV nsp1 coding region (nt 255-369) required for MHV genome replication (18, 47). These results together suggest that the RNA encoding p28, or components of p28, or both, have a detrimental effect in *trans* on DI RNA translation or DI RNA synthesis when over-expressed.

Discussion

Here, we test the hypothesis that p28 (nsp1) of BCoV, the first replicase protein to be synthesized and proteolytically released, is an RNA-binding protein that binds the cis-acting SLVI replication element mapping within its coding sequence. Binding of p28 to this structure

would identify a potential means for autoregulation of gene expression. The rationale for this hypothesis is stated in the Introduction. Although in this study we found no evidence of p28 binding to the intra-cistron SLVI (20), or to a newly-identified SLV cis-acting replication signal mapping immediately upstream of SLVI within the p28 cistron (this study), binding was found at other sites within the 5' and 3' UTRs that might serve to regulate translation or replication. Notably, binding to a long form of cis-acting stem-loop III in the 5' UTR with ~2.5 μM affinity was found.

To our knowledge, this is the first report of an RNA-binding property for coronavirus nsp1. While we tested binding here to only specific terminal regions of genomic RNA, the presence of multiple binding sites suggests possible broad spectrum binding properties that might be important for binding other sites in the viral genome, and might be mechanistically relevant to the variety of cellular responses to coronavirus infection noted in the Introduction. Binding of nsp1 to cellular RNAs, for example, might induce pathogenic cellular responses. To date, however, only protein-protein interactions have been evaluated in these processes.

With regard to direct involvement of p28 in virus replication, several possibilities can be envisioned that will need testing. (i) p28 bound to SLIII (or SLsIII and IV) might repress translation of ORF 1. The moderately high affinity of p28 for the long form of SLIII in the 5' UTR, and perhaps the lesser but still specific affinity to SLIV, is consistent with results of an earlier study showing a 12-fold repression of translation in BCoV-infected cells for a reporter mRNA carrying the 210-nt genomic 5' UTR but not one carrying the 77-nt sgmRNA 7 5' UTR (135). That is, SLIII and SLIV are present in the 210-nt genomic 5' UTR that experienced repression but not in the 77-nt sgmRNA 5' UTR that experienced no repression (135). The molecules tested in these experiments had identical 3' UTRs (135). If the observed repression is a function of a viral protein as postulated (135) then p28 becomes a candidate for this function. It is intriguing that SLIII contains some portion of a short upstream ORF found in virtually all

coronavirus genomes (129). Although the function of the upstream ORF has yet to be found, conceivably it, or its translated product, plays a role in regulating translation or RNA replication, and p28 by binding SLIII may influence these processes. Based on precedents in other well-studied (+)-strand RNA viruses, an inhibition of translation might facilitate (-)-strand synthesis in a couple of ways. (a) p28 binding might block ribosomal transit and enable initiation of (-)-strand synthesis. (b) p28 might interact with SLs I-IV and also the 3'-end, or other 3' end-binding components, to form a circular molecule that functions to inhibit translation and enable initiation of (-)-strand synthesis as has been shown for picornaviruses (8, 54). (ii) p28 might interact with a membrane anchoring protein to bring the genome into the membrane-protected replication compartments where the viral replication complex assembles and presumably functions in the (mostly) absence of the cellular translation machinery (136, 159). This mechanism too has precedent in nodaviruses and several plant viruses (160). (iii) p28 by binding directly to the very 3' end of the genome might take part in replication complex formation and initiation of (-)-strand synthesis. (iv) p28 through its affinity for SLIII in the 5' UTR may play a role in RdRp template switching events that direct leader formation at the 3' end of sgmRNA-length templates for sgmRNA synthesis (170, 171). (v) Finally, p28 might function in the initiation of nascent (+)-strands from the 3' end of the (-)-strand antigenome. For this, a systematic study has not been made but we have learned that SLIII does specifically, although weakly, bind the (-)-strand counterparts to SLIII-Long and SLIII-Short (data not shown), suggesting a possible role for p28 in (+)-strand synthesis.

At this time we are unable to explain how an overexpression of p28 affects the accumulation of DI RNA more than genome and sgmRNAs, but it is possible that the DI RNA is simply more sensitive to p28 function, or that it is in a different membranous compartment than the genome and sgmRNAs and therefore experiences a higher concentration of p28. The fact that both the N-terminal and C-terminal regions affect accumulation suggests there are many

functional domains on p28 for regulation. Nevertheless, a nearly complete repression of DI RNA accumulation and a 2-fold repression of viral sgmRNA abundance, even if short-lived, is evidence that p28 has an effect on viral mRNA expression.

Both the identity and function of the ~60 and 100 kDa cellular proteins that bind SLV-VI remain to be determined. Curiously, six cellular proteins with molecular masses of 25 to 58 kDa but no viral proteins were found to bind SLIV in the 5' UTR (130). The arrangement of three cis-acting elements in close proximity suggests they might function together as a larger complex. Although a relationship between cellular proteins binding SLV-VI or SLIV-VI and translation may exist, it is also possible that the cellular proteins play a role in targeting the viral genome to the membranes making up the replication compartment (159).

Where is the RNA binding domain in p28? Bioinformatics analysis of the BCoV p28 amino acid sequence predicts no obvious RNA binding domain but does reveal a potential Zn finger with similarity to RNA binding motifs (21, 60). This domain (aa 34-75), common to both BCoV and MHV, is highlighted in Figure 2-6. It has also been noted that an 8-amino acid C-proximal domain of unknown function, also highlighted in Fig. 6, is conserved among all group 2 coronaviruses including the SARS-CoV and the positively-charged surface this domain could be an RNA binding site (2). The functional RNA binding domain in p28, however, remains to be empirically determined.

We conclude that together the results indicate that p28 is an RNA-binding protein that may function as a regulator of viral translation or replication, but not through the binding of 5'-proximal cis-acting stem-loops V and VI within its coding region. How and where p28 functions as an RNA-binding protein, and how SLV and VI function as cis regulatory structures are questions that remain to be answered.

Figure 2-6. Aligned Amino Acid Sequences for nsp 1 in BCoV (p28), MHV (p28), and SARS-CoV (p20). The arrowhead identifies the separation site of BCoV p28 into the N-terminal and C-terminal parts studied in Fig. 5. The potential zinc finger-like RNA-binding domain in BCoV and MHV (nt 34-75) is identified by highlighting. A highly similar 8-aa sequence in nsp1 of all three viruses identifying a potential positively-charged region of the protein is identified by highlighting. Numbers refer to amino acid positions in the BCoV sequence. Regions in the proteins coinciding with the coding regions harboring SLV and SLVI are indicated.

Acknowledgements

We thank Kim Nixon and Hung-Yi Wu for many helpful discussions. This work was supported by Public Health Service grant AI014367 from the National Institute of Allergies and Infectious Diseases and by funds from the University of Tennessee, College of Veterinary Medicine, Center of Excellence Program for Livestock Diseases and Human Health.

Chapter III. Cap-Dependent Translation of mRNA with Coronavirus 5'- and 3'-Untranslated Regions is Repressed in the Absence of Viral Proteins

Introduction

Coronaviruses primarily cause respiratory and gastroenteric diseases in birds and mammals and have been the subject of increased research since the appearance of SARS in 2003 (89, 165). Since that time new coronaviruses have been isolated from many different hosts including bats and humans. Unlike SARS-coronavirus, which causes severe lower respiratory tract disease in humans, most human coronavirus infections cause mild upper respiratory disease (77, 162).

Bovine coronavirus (BCoV) like all coronaviruses is a positive-sense, single-stranded RNA virus with a 5' cap and 3' poly(A) tail (106). Translation of the coronavirus genome is thought occur via a closed loop or circular complex through interactions with the 5' cap binding complex eIF4F and 3' poly(A) tail-poly(A) binding protein (PABP) interaction, the same as many cellular mRNAs, in order to produce viral proteins necessary for RNA synthesis and replication (15). With over twenty potential protein products, regulation of translation to produce both structural and nonstructural proteins is likely an important process in the replication cycle. Regulation of protein synthesis, or translational control, can be expected to occur in at least three stages of virus replication: genome translation to produce the replicase polyprotein, subgenomic mRNA translation to produce structural and accessory proteins, and a translation to replication switch. A fourth stage of translation regulation during coronavirus replication is known to occur during translation of the replicase polyprotein. A -1 frameshifting pseudoknot located at the ORF1a/1b junction results in translation of the ORF 1b replicase proteins (16, 17).

In coronaviruses, *cis*-acting replication elements have been found within the untranslated regions (UTR) and may function in translation. To date six *cis*-acting elements in the 5' and 3' UTRs have been identified and shown to be important for genome or DI RNA

replication (15). There are three elements in the 5' UTR including the leader region containing SLI and II, SLIII, and SLIV (27, 129, 130). In the 3' UTR a bulged stem-loop and adjacent pseudoknot, a terminal stem-loop with a conserved octamer sequence, and a poly(A) tail have been shown to be *cis*-acting (68, 103, 143, 144, 166).

Currently, only the 5' UTR leader sequence has been shown to be important for translation in MHV infected cells. The entire 5' leader sequence has been shown to enhance translation of a reporter construct in virus infected cells and a higher order structure, SLII, within the leader sequence has been shown to be necessary for efficient translation of a reporter construct (102, 148). Because the leader sequence is present on the genomic RNA and all subgenomic mRNAs (sgmRNAs), it appears to be one element important for translation of all viral RNAs. (148). The viral nucleocapsid protein (N) is currently the only viral protein shown to enhance translation in infected cells (149). This protein has also been shown to bind the leader sequence, potentially exerting its effect on translation directly through RNA-protein interactions, although this has not been confirmed experimentally (117, 147).

As the first step in a systemic study of translational control of coronaviruses and to understand the effects of the viral UTR sequences on translation of genomic and subgenomic mRNAs, luciferase reporter constructs with the BCoV 5' and 3' UTRs were translated *in vitro* and in the absence of all viral proteins. In this system, both the 5' and 3' UTRs were found to independently enhance translation. Unexpectedly, however, the translational advantage conferred by cap-dependent translation initiation was repressed when both UTRs were used, suggesting an inhibitory 5'-3' mechanism. Further analysis demonstrated the repressive effects to be localized within the viral leader sequence and a higher-order structure in the 3' UTR. These results show a repression of translation that is (i) independent of viral proteins, (ii) implicates a 5'-3'-end interaction that is likely mediated by cellular proteins, and (iii) suggests cap-dependent translation may require the assistance of viral protein(s).

Materials and Methods

Viral Sequence. Viral sequences used were from bovine coronavirus Mebus strain (66, 92), Genbank # U00735.

Constructs. Renilla luciferase (RLuc) reporter sequence was obtained using AmpR [696] pBS-35S:hRLUC (Packard version) (a gift from A.von Arnim Lab). RLuc constructs were cloned into plasmid pGEM3zf- (Promega) using NgoM IV and Hind III restriction sites (Fig. 3-1A). The parent construct, containing the complete viral 5' and 3' UTRs was generated using overlap PCR in which the BCoV DI RNA was used as the template DNA. This construct, **UTR210-288**, contains the entire 210nt 5' UTR, 288nt 3' UTR, and a poly(A) tail (~65As) (Fig. 3-1B). The T7 promoter sequence was placed downstream of the NgoM IV restriction site using the PCR primer to ensure the first nucleotide transcribed was the authentic viral 5' base. The coding region start codon was placed immediately upstream of a BamH I restriction site in order to maintain the exact viral 5' UTR. Designing the constructs in this manner resulted in every construct having the same base, G at the +4 position, altering the BCoV native Kozak context (T at +4, nt214). Similarly, the stop codon at the end of the coding region was placed downstream of a Hind III restriction site. 5' and 3' UTR modifications were made by replacing the viral wildtype sequence with PCR generated fragments using Ngom IV and BamH I (5' UTR) or double Hind III restriction sites (3' UTR) (Fig. 3-1A). Polyadenylated 3' UTRs were made by using 2 rounds of overlap PCR, one to generate the 3' UTR sequence with a poly(A) tail and another to fuse the UTR with the rest of the construct. All viral sequences were cloned using BCoV pDrep 1 as the template DNA (26). The sequence for 5'OPT.UTR was designed based on the optimal sequence context from Kozak (82). 5'-GAATACTCCCCGCCGCCACC-3'. The sequence for the 40nt 3' UTR is vector sequence from the pGEM vector cloning site with some modifications to eliminate unwanted restriction sites or start codons, 5'-GAATACTCAGCTTGCCCGCTGCAGGTCGACTCTAGAGTGC-3'. Construct names are given

A.

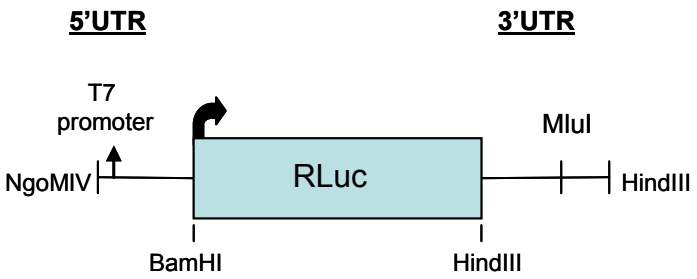
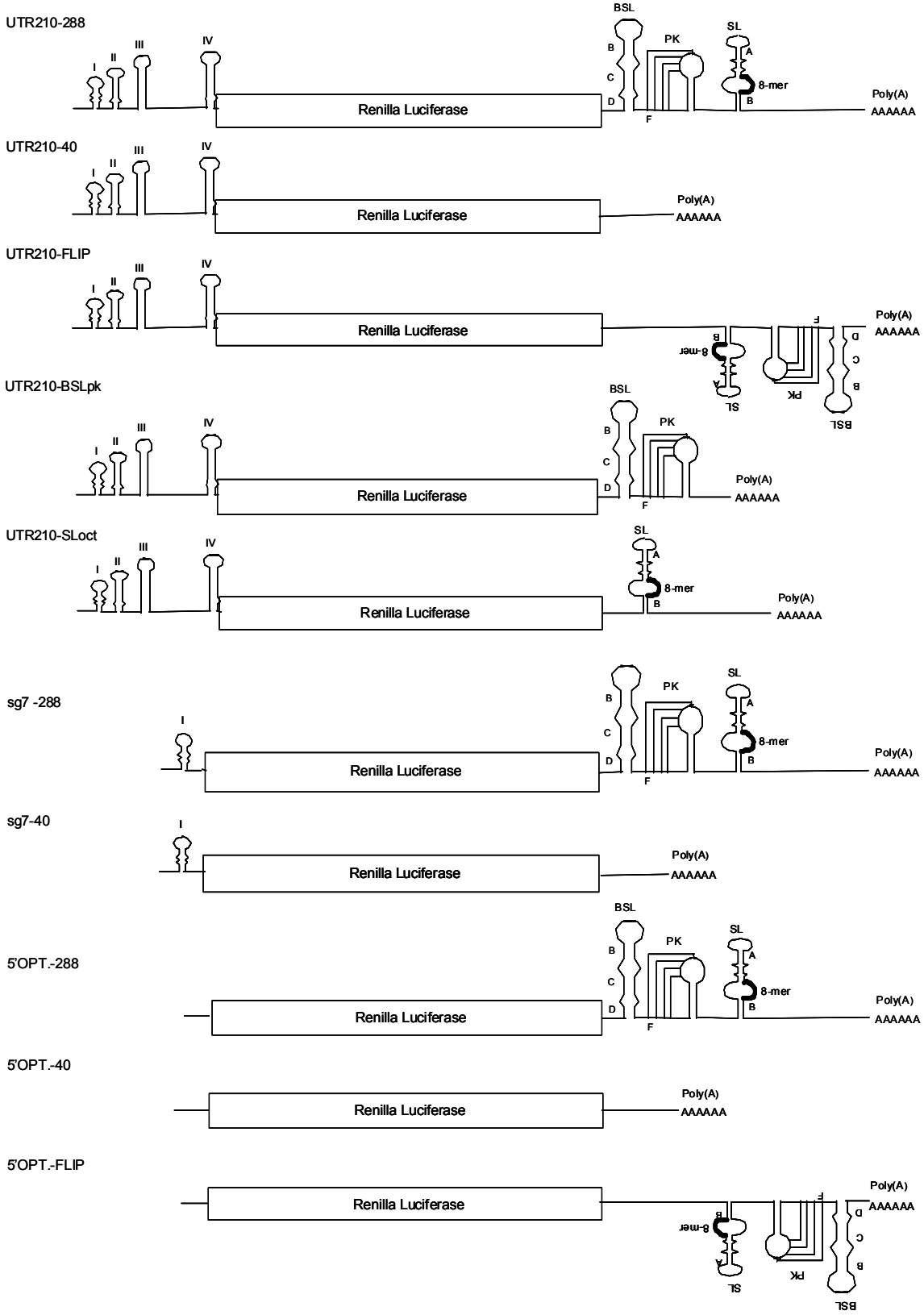


Figure 3-1. Renilla Luciferase Constructs. A. Renilla Luciferase Construct Design. BCoV 5' and 3' UTRs were exchanged using Ngom IV-BamH I or Hind III-Hind III restriction sites. Constructs were linearized with Mlu I and RNA transcribed from a T7 promoter. **B.** Renilla Luciferase Construct Diagrams. Construct names are given as: 5' UTR sequence **–** 3' UTR sequence.

Figure 3.1, cont. Part B.



as 5' UTR sequence – 3' UTR sequence and diagrams are shown in Figure 3-1B and Appendix II, Figure A2-5). Oligonucleotides used for constructs are listed in Table 3-1.

RNA synthesis. 2µg of Mlu I linearized DNA was used for RNA transcription. Capped RNAs were synthesized using T7 Message Machine (Ambion) and uncapped RNAs were synthesized using T7 RNA polymerase (Promega) following the manufacturer's protocols. 2 U of RNase-free DNase was added to reaction mixtures and RNA was purified using the RNaid RNA purification kit (MP Biomedicals). Cap incorporation (m7G(5')ppp(5')G) under these conditions is estimated to be 80% (Ambion). RNAs were quantitated using a spectrophotometer and formaldehyde gel electrophoresis.

***in vitro* translation**. Translations of RLuc RNAs were performed using rabbit reticulocyte lysate (Promega). 100ng of RNA was added to each 50µl reaction and translation reactions were performed following the manufacturer's protocol (Promega). Reactions were incubated at 30°C for 90 minutes and 2.5µl aliquots were taken at the indicated times and used in the luciferase assay. Independent translation reactions for each construct were performed a minimum of three times.

Luciferase assay. Renilla luciferase assay reagent (Promega) was added to each 2.5µl aliquot and luciferase activities determined using a TD20/20 luminometer. Results are presented as the average of three readings per sample and at least three independent experiments. Translation results for each construct are shown in Appendix II, Table A2-1. p-values were calculated using the Student's t-test.

Radioactive *in vitro* translation. Translation reactions using 20µCi ^{35}S-Met and 500ng of RNA in a 50µl reaction were performed following the manufacturer's protocol (Promega) as described above. A 5µl aliquot was taken at 90 minutes and used in SDS-PAGE. Gels were dried and results were visualized and quantitated using a phosphorimager.

Table 3-1. Oligonucleotides Used for Renilla Luciferase Containing Constructs.

Oligonucleotide	Pol.[a]	Sequence (5'→3')	Purpose
NgoMIV-T7-1	+	GGGC **GCC GGC TAA TAC GAC TCA CTA TAG** GAT TGT GAG CGA TTT GC	T7 prom +BCoV nt1
5'UTR/RLUC-5'	+	GGC CCA CCC ATA GGT CAC AAT G**GG ATC C**GC TTC CAA GGT GTA CGA CC	BCoV 5'UTR overlap with RLuc
5'UTR/RLUC-3'	-	GGT CGT ACA CCT TGG AAG C**GG ATC C**CA TTG TGA CCT ATG GGT GGG CC	BCoV 5'UTR overlap with RLuc
3'UTR/RLUC-5'	+	GAA CGA GCA GAG ATC TAT C**AA GCT T**TA AGA GAA TGA ACC TTA TGT CG	BCoV 3'UTR overlap with RLuc
3'UTR/RLUC-3'	-	CGA CAT AAG GTT CAT TCT CTT A**AA GCT T**GA TAG ATC TCT GCT CGT TC	BCoV 3'UTR overlap with RLuc
sg7-3'	-	GGGC **GGA TCC** CAT CCT TAA AGT TTA G	5'UTR 77nt leader sequence
BCov-396-3'	-	GCCC **GGATCC** ACA ATC CAC CAT CAC ATG A	5' to include nsp 1 nts 211-396
40nt-Δp(A)-5'	+	GGGC **AAGCTT** TAA GAA TAC TCA GCT TGC	40nt 3'UTR
40nt-Δp(A)-3'	-	GGGC **AAGCTT** GCATGC **ACGCGT** GCA CTC TAG AGT CG	40nt 3'UTR
40nt-p(A)OVLP-5'	+	GCAGGTCGACTCTAGAGTGCAAAAAAAAAAAAAAA	40nt + p(A) (3'UTR only)
40nt -p(A)OVLP-3'	-	TTTTTTTTTTTTTTTGCACTCTAGAGTCGACCTGC	40nt + p(A) (3'UTR only)
40-p(A)OVLP 2-5'	+	GAACGAGCAGAGATCTATCAAGCTTTAAGAATACTC AGCTTG	40nt + p(A)
40-p(A)OVLP 2-3'	-	CAAGCTGAGTATTCTTAAAGCTTGATAGATCTCTGC TCGTTC	40nt + p(A)
FLIP 3'utr -5'	+	GGGC **AAGCTT** TAA GTG ATT CTT CCA ATT GG	Reversed BCoV 3'UTR
FLIP 3'utr -3'	-	GCCC **AAGCTT** GCA TGC **ACG CGT** GAG AAT GAA CCT TAT GTC GGC	Reversed BCoV 3'UTR
FLIP 3'utr –ovlp5'	+	CCGACATAAGGTTCATTCTCAAAAAAAAAAAAAAA	FLIP + p(A)
FLIP 3'utr –ovlp3''	-	TTTTTTTTTTTTTTTGAGAATGAACCTTATGTCGG	FLIP + p(A)
SLoct ovlp5'	+	GAACGAGCAGAGATCTATCAAGCTTTAATTGAAAGT TTTGTGTGGTAATG	BSLoct 3'UTR
SLoct ovlp3'	-	CATTACCACACAAAACTTTCAATTAAAGCTTGATAG ATCTCTGCTCGTTC	BSLoct 3'UTR
BSLpk ovlp5'	+	ATAGCAGACTATAGATTAATTAGAAAAAAAAAAAAA AAAA	BSLpk 3'UTR
BSLpk ovlp3'	-	TTTTTTTTTTTTTTTTTCTAATTAATCTATAGTCTGCT AT	BSLpk 3'UTR
m.uAUG-5'	+	CAC TCC CTG TAT TCT A**G**G CTT GTG GGC **C**TA GA	uAUG to AGG
m.uAUG-3'	-	TCT A**G**G CCC ACA AGC **C**TA GAA TAC AGG GAG TG	uAUG to AGG
m.124STOP-5'	+	GTG GGC GTA GAT TTT TCA T**C**G TGG TGT C	A125C overlap
m.124 STOP -3'	-	GAC ACC AC**G** ATG AAA AAT CTA CGC CCA C	A125C overlap
m.124/184 STOP -5'	+	GGA CGT GTT GTA TCC T**C**G GCA GTG GC	A185C overlap
m.124/184 STOP -3'	-	GCC ACT GCC **G**AG GAT ACA ACA CGT CC	A185C overlap
uORF-fusion-3'	-	GCCC **GGATCC CC**T TGT GAC CT**G** TGG GTG GGC CAC	U202C,U212G overlap

[a]Polarity of the oligonucleotide relative to the positive strand of the construct. Mutated bases are in bold and underlined. Restriction sites are in bold.

Results

Capped transcripts with genomic 5' and 3' UTR sequences show no translational advantage when translated *in vitro*.

For the purpose of observing the intrinsic effects of the viral UTR sequences on translation, we first analyzed the translation efficiency of a RLUC reporter construct with the BCoV 210nt 5' UTR, 288nt 3' UTR and a poly(A) tail, **UTR210-288**, and compared it to a control construct with putative optimal, nonviral UTRs. The control construct, **5'OPT.-40**, contains a 20 nt 5' UTR with an optimal Kozak sequence, a 40nt 3' UTR (vector sequence) that is long enough so as to not sterically hinder translation, and a poly(A) tail (82, 150). Transcripts were translated for 90 minutes and samples were taken at the indicated times. Luciferase activity was measured and the results are presented as the average of at least three independent experiments. Comparing the translation levels of capped transcripts of the viral wildtype construct **UTR210-288** with the capped nonviral control **5'OPT.-40**, reveals there is virtually no difference in luciferase activity (Fig. 3-2). Thus, it appears as though the viral 210nt 5' UTR and 288nt polyadenylated 3' UTR are able to support translation levels equal to that of a much shorter 20nt 5' UTR with an optimal translation context and a 40nt polyadenylated 3' UTR. Therefore, capped transcripts with the viral UTRs do not appear to have a *cis*-acting translational advantage over an optimal RNA control, at least in the absence of viral proteins.

Because small differences in translation may be masked by the translational advantage given by the 5' cap, uncapped RNAs were compared. Translation levels of uncapped transcripts of the optimal control construct, **5'OPT-40**, were 50% decreased compared to the translation levels of capped transcripts. Interestingly, the translation levels of uncapped **UTR210-288** were the same as the levels observed for capped **UTR210-288** (Fig. 3-2). These results indicate that translation levels of **UTR210-288**, which contains both the viral 5' and 3' UTRs are the same, regardless of whether or not the input RNA is capped or uncapped. This

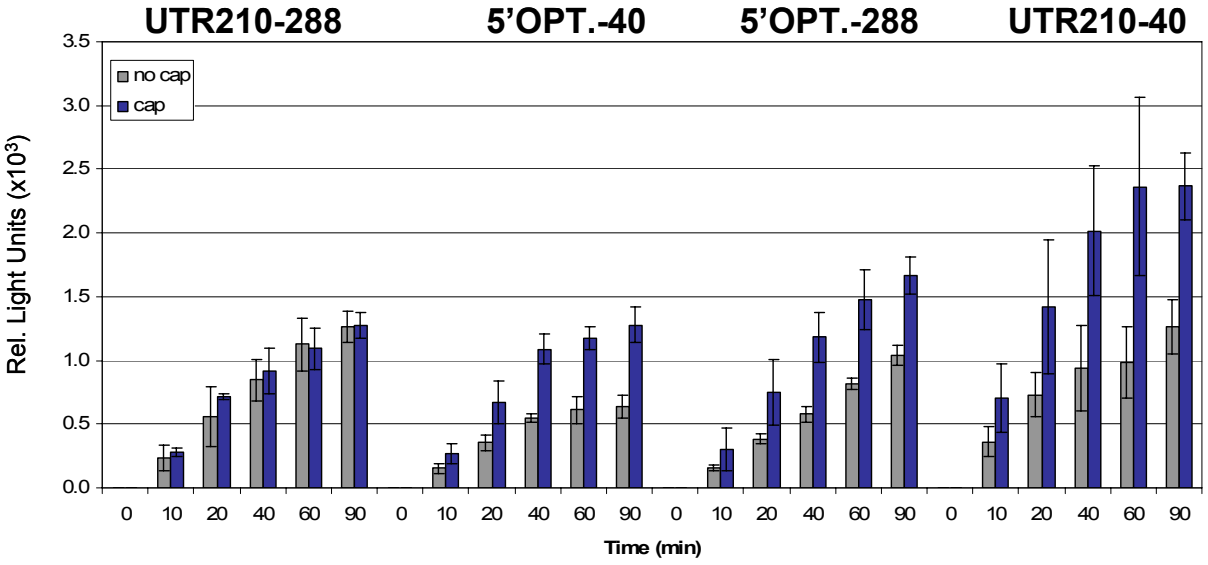


Figure 3-2. Capped Transcripts With Genomic 5' and 3' UTR Sequences Show No Translational Advantage When Translated *in vitro*. Translation levels of RNAs with both the viral 210nt 5' and 288nt 3' UTR (**UTR210-288**), only the viral 210nt 5' UTR (**UTR210-40**), only the viral 288nt 3' UTR (**5'OPT.-288**), or neither (**5'OPT.-40**).

result was unexpected, as it appears that cap-independent translation is as efficient as cap-dependent translation of transcripts with both viral UTRs. Following current translation models for naturally capped mRNAs, translation levels of the capped RNA should be increased compared to translation levels of the uncapped RNA (53), as demonstrated with **5'OPT.-40**. Coronavirus RNAs are capped and thought to be translated via cap-dependent initiation and ribosome scanning (134, 135). Thus, assuming coronavirus RNAs are translated via this mechanism, two questions can be asked; (i) are translation levels of capped RNA repressed, and (ii) are translation levels of uncapped RNA increased?

Viral genomic 5' and 3' UTRs both contribute to the repression of cap-dependent translation.

To address these proposed questions, we examined translation levels of transcripts with differing combinations of 5' and 3' UTRs. Reference basal levels of translation were established by translating uncapped transcripts. Comparing translation levels of **UTR210-288** with **5'OPT-288**, and **UTR210-40**, an increase in translation of capped transcripts was observed with both **5'OPT-288** and **UTR210-40** compared with translation levels of uncapped transcripts (Fig. 3-2). These results suggests that translation of capped **5'OPT.-288** and **UTR210-40** are both more efficient than **UTR210-288** and that by replacing one of the viral UTRs with a nonviral sequence, cap-dependent translation levels can be increased. Therefore, cap-dependent translation appears to be repressed when both viral 5' and 3' UTRs are present.

The results of these experiments indicate there are elements within the viral 5' and 3' UTRs collectively contributing to an overall repression of cap-dependent translation. Furthermore, the results suggest that there may be a 5'-3'-end interaction that contributes to a mechanism of cap-dependent translation repression. To more clearly understand the roles of the viral 5' and 3' UTRs in translation and how they may function together to repress cap-dependent translation, the contributions of each UTR were analyzed separately.

Elements within the viral 3' UTR contribute to a repression of cap-dependent translation.

From the above experiments, two effects were evident from the control 40nt 3' UTR in the context of the genomic 210nt 5' UTR: (i) translation of capped **UTR210-40** was increased compared to **UTR210-288**, (ii) translation of uncapped **UTR210-40** was decreased compared to capped **UTR210-40**, indicating there was no repression of cap-dependent translation. To determine if the length or an element within the 3' UTR contributed to these effects, a construct with an equally long UTR was designed. For this construct, **UTR210-FLIP**, the negative strand of the 3' UTR replaced the wildtype UTR. Translation levels of capped **UTR210-FLIP** were decreased by 20% compared to the levels of **UTR210-288** at 90 minutes ($p=0.023$) (Fig. 3-3A). In addition, translation levels of capped **UTR210-288** and **UTR210-FLIP** were both decreased approximately 50% at 90 minutes compared to **UTR210-40** ($p\leq0.002$) (Fig. 3-3A). This indicates that the short 40nt 3' UTR confers more efficient translation than either the viral 288 3' UTR or FLIP 3' UTR, and that the length of the 3' UTR may negatively affect translation. However, translation levels of uncapped **UTR210-FLIP** were 30% decreased compared to the capped transcript, suggesting there was no repression of cap-dependent translation as observed with **UTR210-288**. These results therefore indicate that the repression of cap-dependent translation does not reside in the FLIP sequence and while the length of the 3' UTR may contribute to lower translation levels, it does not contribute to the repression observed with the wildtype genomic 3' UTR. Hence, it is likely repression of cap-dependent translation is a function of the sequence or higher-order structures of the 3' UTR.

To determine if the 288nt 3' UTR sequence represses cap-dependent translation in the context of a non-viral 5' UTR, the effects with different combinations of 5' UTRs were tested. In Figure 3-3B capped transcripts with the 5'OPT. 5' UTR and three different 3' UTRs are shown. Translation levels of capped **5'OPT.-288** were greater than **5'OPT.-40** and **5'OPT.-FLIP** ($p\leq0.028$, 90min.). Therefore, the viral 288nt 3' UTR appears to support increased translation

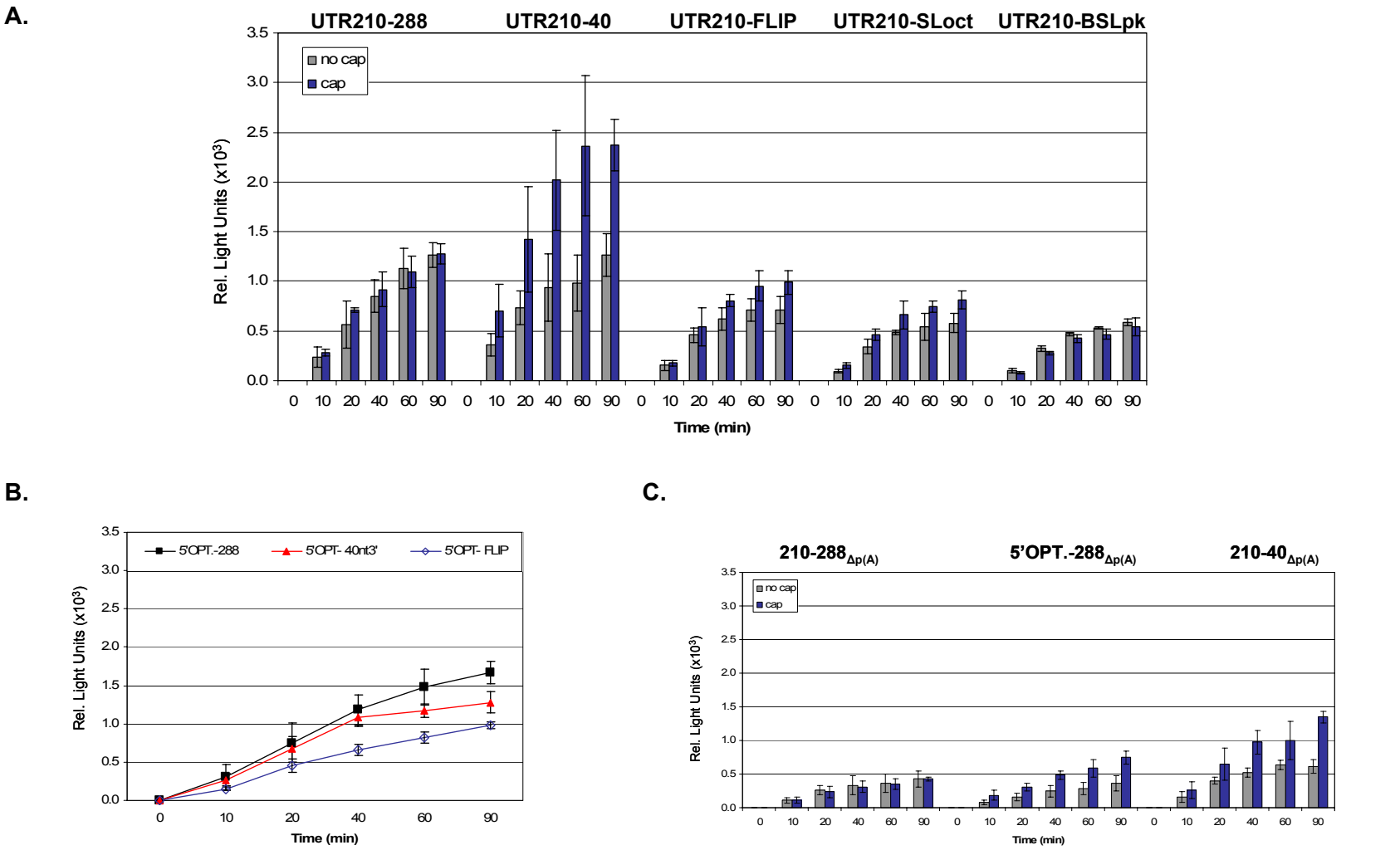


Figure 3-3. The Viral 3' UTR Does Not Independently Repress Cap-Dependent Translation **A.** Translation levels of transcripts with the viral 210nt 5' UTR and five different 3' UTRs (288nt viral UTR, 40nt control sequence, FLIP: negative strand of the 288nt viral UTR), SLoct (nts 1-172) viral 3' UTR sequence, BSLpk (nts 173-288) viral 3' UTR sequence. **B.** Translation levels of RNAs with the 5'OPT. UTR and three different 3' UTRs. **C.** Translation levels of nonpolyadenylated RNAs.

levels when paired with the nonviral optimal sequence 5' UTR. These data indicate that the viral 3' UTR does not independently contribute to the observed repression of cap-dependent translation and supports a 5'-3' UTR repressive interaction that may involve a specific element within the 3' UTR.

The 3' UTR bulged stem-loop and adjacent pseudoknot contribute to a repression of cap-dependent translation.

To identify the features in the 288nt 3' UTR involved in repressing translation, the viral 3' UTR was divided into two sections. In addition to the poly(A) tail, two *cis*-acting replication elements have been identified within the viral 3' UTR. A bulged-stem-loop and adjacent pseudoknot (BSLpk) that likely function together is located immediately downstream of the coding region stop codon. The second *cis*-acting stem-loop structure (SLoct) is located near the 3' UTR terminus and contains a conserved 8-mer sequence that is found in all known coronavirus 3' UTRs (15). To determine if one of these two higher-order structures was specifically responsible for the observed translation repression, each element was cloned separately with the viral 210nt 5' UTR to yield constructs **UTR210-SLoct** (nts 1-172 from 3' end) and **UTR210-BSLpk** (nts 173-288 from 3' end). The results show that translation levels of both transcripts were lower than **UTR210-288** (Fig. 3-3A). However, translation levels of capped and uncapped **UTR210-BSLpk** were equal, suggesting cap-dependent translation was repressed, potentially by the same mechanism observed with **UTR210-288** (Fig. 3-3A). Alternatively, translation levels of capped **UTR210-SLoct** were increased compared to uncapped transcripts, indicating cap-dependent translation was not repressed. Therefore, the *cis*-acting bulged stem-loop and pseudoknot structures may aid in cap-dependent translation repression.

The observed repression of cap-dependent translation is independent of the 3' poly(A) tail.

To eliminate the potential influence on translation from the 5' cap-3' poly(A) tail interaction and to look only at the influence of the UTRs, RNAs without poly(A) tails were tested.

Although the translation levels of the non-polyadenylated RNAs were lower, the overall trend remains the same (Fig. 3-3C, compare to Figure 3-2). When both the viral 5' and 3' UTRs were present, (**UTR210-288$_{\Delta p(A)}$**), translation levels were lower compared to when only one viral UTR was present (**5'OPT.-288$_{\Delta p(A)}$**, **UTR210-40$_{\Delta p(A)}$**). Furthermore, translation levels of capped and uncapped **UTR210-288$_{\Delta p(A)}$** were equal, indicating cap-dependent translation was repressed independently of the 5' cap-3' poly(A) tail interaction (Fig. 3-3C). This observation suggests there is a specific, 5'-3' interaction that is responsible for the repression of cap-dependent translation and that it is independent of the 5' cap-3' poly(A) interaction.

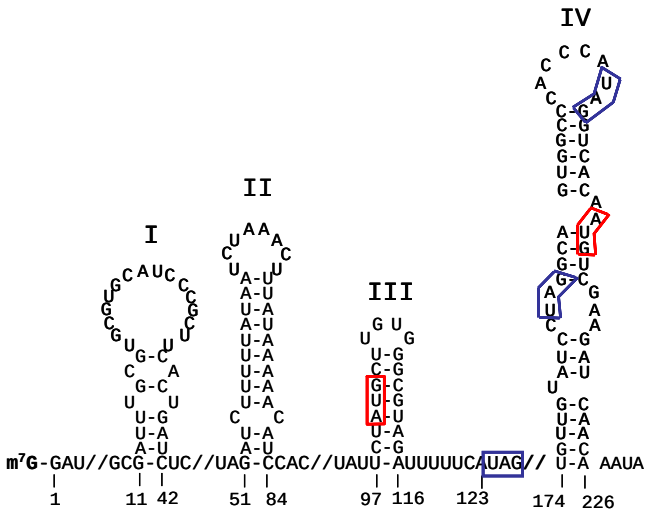
An uAUG in the 5' UTR does not contribute to the repression of cap-dependent translation.

Regulatory elements have been found in the 5' UTRs of mRNAs that may increase or decrease translation, typically at the rate-limiting step, translation initiation (63). One *cis* element within the 5' UTR that may attenuate translation is an upstream AUG (uAUG). Regulation of these elements can furthermore be dependent on interactions with the 3' UTR (109). Ribosomes typically initiate translation at the first AUG and an uAUG can decrease translation initiation events of the main open reading frame because ribosomes instead initiate translation upstream of the main coding region (84). A single uAUG in virtually all coronaviruses and in the BCoV 5' UTR is found at nucleotide 100 (Fig. 3-4A). If translated, the open reading frame (uORF) would yield an eight amino acid peptide.

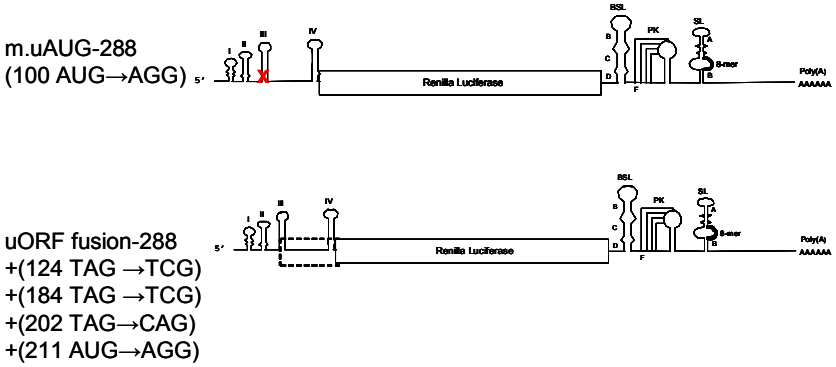
To determine if the uAUG was being used to initiate translation and therefore contributing to a repression of cap-dependent translation, it was necessary to extend the uORF coding region in order to detect the translation product. Construct **uORF fusion-288** was made by mutating the three in-frame stop codons at nucleotides 124, 184, and 202 to amino acid codons. Also, the main 211 AUG was mutated to AGG (Fig. 3-4A, B). With this RNA, initiation should only occur at the uAUG, making a slightly larger RLuc fusion protein. The results show that translation levels of capped **uORF fusion-288** were only 30% of capped **UTR210-288**

Figure 3-4. An uAUG in the 5' UTR Does Not Contribute to the Repression of Cap-Dependent Translation. A. BCoV 5' UTR. The uAUG is within *cis*-acting SLIII and if translated would yield an 8aa peptide. The uORF is in-frame with the main 211AUG. There are two additional in-frame stop codons at nts 124 and 184. **B.** Construct diagrams of mutant uAUG (**m.uAUG-288**) in which the uAUG has been changed to a non-start codon, AGG (U101G). One additional mutation (G112C) was made to maintain the base pairing in the stem structure. **uORF fusion-288** has the indicated codon changes in which the three native stop codons and main 211AUG have been mutated to amino acid codons, and translation initiation will only occur at the uAUG and yield an RLuc fusion translation product. **C.** Translation levels of capped **UTR210-288** and **m.uAUG-288** compared to **uORF fusion-288**. **D.** ^{35}S-Met. translation products from capped RNA show translation from the uAUG is minimal and the major translation product is from a downstream coding region AUG.

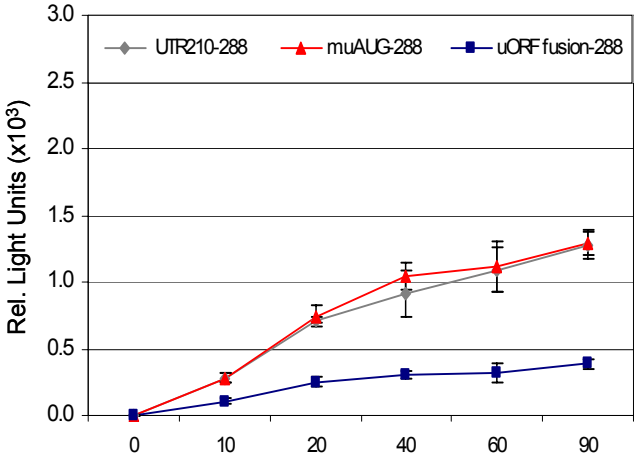
A.


I
II
III
IV
m7G-GAU//GCG-CUC//UAG-CCAC//UAUU-AUUUUUCAUAG// U-A AAUA
1
11 42
51 84
97 116
123
174 226

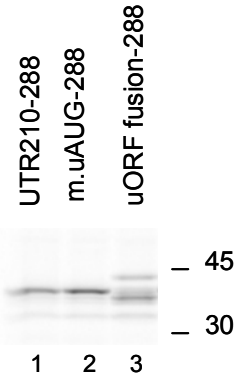
B.


m.uAUG-288
(100 AUG→AGG)
Renilla Luciferase
uORF fusion-288
+(124 TAG →TCG)
+(184 TAG →TCG)
+(202 TAG→CAG)
+(211 AUG→AGG)
Renilla Luciferase

C.


UTR210-288
muAUG-288
uORF fusion-288
Rel. Light Units (x10³)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
10
20
40
60
90

D.


UTR210-288
m.uAUG-288
uORF fusion-288
45
30
1
2
3

levels at 90 minutes, indicating the uAUG is not used frequently (Fig. 3-4C). In a second method of analysis, translation of these RNAs using ^{35}S-Met and SDS-PAGE was done (Fig. 3-4D). Confirming the luciferase data, translation of **uORF fusion-288** resulted in one light upper band which is representative of initiation from the uAUG. A second band, the major translation product, represents translation from an AUG within the RLuc coding region. Based on phosphorimager quantitation, the ratio of the bands representing translation from the uAUG and the coding region AUG, indicate the uAUG is used ~20% of the time. These results show that the uAUG is used as a start codon, however, leaky scanning past the uAUG appears to occur most of the time. This may in part be due to the poor Kozak context of the uAUG.

In order to determine if the uAUG could have a regulatory effect on translation, a mutant construct was designed, (**m.uAUG-288**), in which the uAUG was changed to the non-start codon AGG (Fig. 3-4A, B). In addition, since this codon is within a *cis*-acting stem-loop, one additional compensatory mutation was made to maintain the stem-loop structure as predicted by *mfold*. As shown in Figure 3-4C and 3-4D, there was no change in translation levels of capped **m.uAUG-288**, compared to **UTR210-288**. This result was surprising as it was expected that the uAUG would be a regulatory element since it is highly conserved and is found in all coronaviruses sequenced to date with only one exception (Brian, D.A., Gustin, K.M., unpublished data). From these results it appears as though the uAUG is not a translation regulatory element due to its inefficient usage and therefore not likely to be involved in an overall repression of translation.

Repression of cap-dependent translation by the 5' UTR is localized to the 65 nucleotide viral leader sequence.

To further characterize the features of the 5' UTR that contribute to the repression of cap-dependent translation, the 77nt 5' UTR of subgenomic mRNA 7 (**sg7-288**) was tested and compared to **UTR210-288**. The first 65nt of **sg7-288** and **UTR210-288** are identical, called the

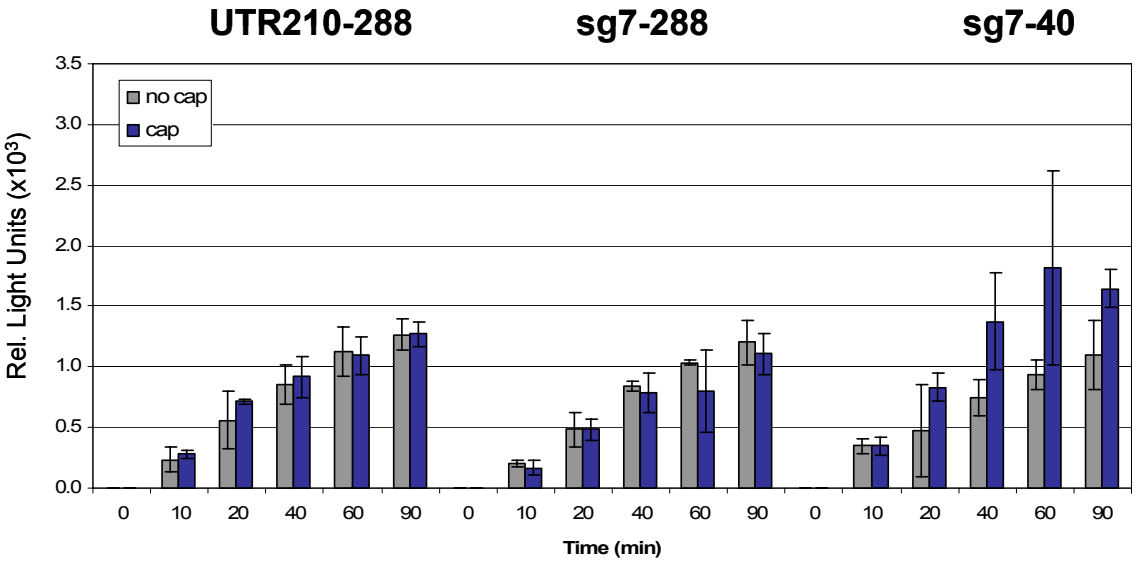
leader sequence, and present on all viral genomic and subgenomic messages (135). Translation levels of **sg7-288** are consistent with those of **UTR210-288**, indicating there is no difference in translation between the 210nt and 77nt 5' UTRs. Interestingly, translation levels of capped and uncapped **sg7-288** were also the same, indicating a similar cap-dependent translation phenomenon as observed with **UTR210-288** (Fig. 3-5A). Therefore, it appears that cap-dependent translation is also repressed when only a short 77nt viral 5' UTR sequence is present with the viral 3' UTR. The observed effect on cap-dependent translation is thus localized to within the common 65nt viral leader sequence of the 5' UTR. As with the 210nt genomic 5' UTR, there was no repression observed with the 40nt 3' UTR and the 77nt 5' UTR, (**sg7-40**), and translation levels of capped transcripts were ~30% increased compared to **sg7-288** at 90 minutes (Fig. 3-5A). Furthermore, translation levels of capped **sg7-40,** were increased over levels of the uncapped transcript, supporting the observation that repression of cap-dependent translation is also dependent on the 3' UTR as shown with the full length 210nt 5' UTR (Fig. 3-5A).

In addition, in the absence of the viral 3' UTR, the 210nt 5' UTR does not appear to negatively affect translation levels and may instead contain elements that enhance translation since translation levels of capped **UTR210-40** are increased compared to those of **sg7-40** and **5'OPT.-40** (Fig. 3-5B). Because both **UTR210-40** and **sg7-40** share the viral leader sequence, the proposed enhancing elements may be within the downstream 5' UTR sequence which contains *cis*-acting stem-loops III and IV. These results indicate that the viral 5' UTR does not independently repress cap-dependent translation levels and supports a 5'-3' UTR interaction that is localized to within the viral leader sequence.

Discussion

In this study we report the observation that *in vitro* and in the absence of viral proteins, a luciferase reporter RNA with the bovine coronavirus 5' and 3' UTRs does not display enhanced

A.



B.

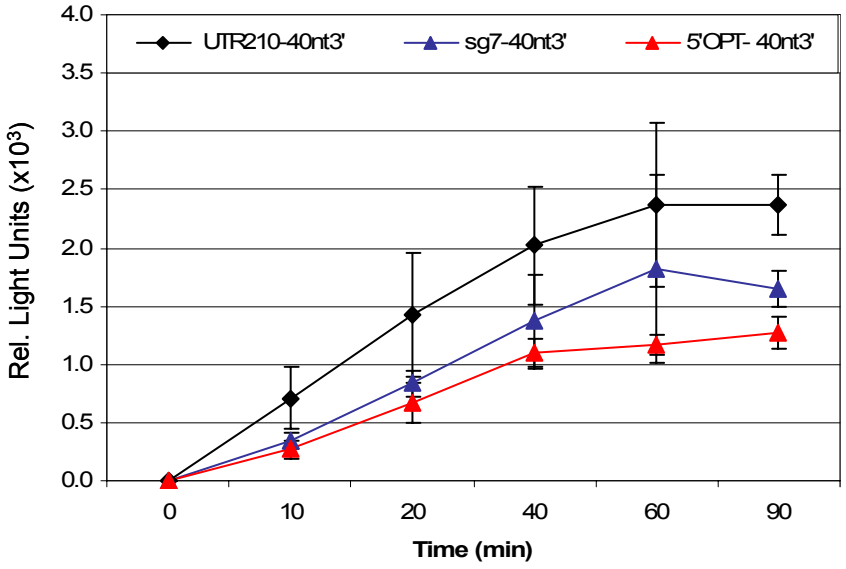


Figure 3-5. Cap-Dependent Translation Repression is Localized to Within the Viral 65 nt Leader Sequence. A. Translation levels of **UTR210-288**, **sg7-288**, and **sg7-40**. **B.** Translation levels of transcripts in the absence of the viral 288nt 3' UTR.

translation as a function of a 5' cap structure. Since the coronavirus genome and sgmRNAs are 5' capped and 3' polyadenylated, it has been presumed that they utilize a cap-dependent ribosomal entry mechanism for translation initiation. Comparisons with control constructs with nonviral UTRs demonstrated that the translational enhancement conferred via cap-dependent translation was repressed only in the presence of both viral UTRs. Analyses mapped the repressive elements to the 65nt leader in the 5' UTR and to the bulged stem-loop-pseudoknot structure in the 3' UTR. The significance of these findings are two-fold. First, there must be a mechanistic explanation for repression. Second, experiments addressing the influences of viral proteins on translation are necessary.

While an enhancement of translation via a 5' cap-3' poly(A) tail interaction is believed to occur during translation of naturally capped cellular mRNAs, additional interactions in the BCoV UTRs appear to have the opposite effect. The mechanism through which the observed translation repression occurs could in theory involve RNA-RNA interactions between the UTRs. However no strong association or complementarity has been found based on sequence analysis. Recently in MHV a direct 5'-3' RNA interaction has been proposed to occur between bases in a stem-loop in the 5' UTR and 3' UTR terminal region based on the recovery of 5' UTR mutant viruses with 3' UTR second site suppressor mutations (98). However, it is unclear if this region is representative of a true 5'-3' interaction or if it could occur in BCoV. Therefore, protein-protein interactions or RNA-protein interactions likely cause the observed translation suppression. Since this study was performed in the absence of viral proteins, cellular proteins are the only potential mediators.

Translation regulation of this type could be virus or cell-mediated and control the rate at which viral proteins are produced. Alternatively, repression of translation initiation could be a precursor to the switch from translation to genome replication which likely involves cellular proteins. Cellular proteins have been found to bind to both the 5' and 3' UTRs. Two cellular

proteins have been found to bind to the 5' UTR of MHV, hnRNP Q3 and polypyrimidine tract-binding protein (PTB) or hnRNP I (33, 96). In addition, mitochondrial aconitase, three heat shock proteins, and hnRNP A1 have been shown to bind to the viral 3' UTR (97, 113, 114).

The hnRNP family of proteins have been found to bind to the UTRs of RNA viruses and are involved in both translation and RNA synthesis (90). Both hnRNP E (poly C-binding protein (PCBP)), and hnRNP C have been shown to be important for poliovirus (PV) RNA synthesis (4, 22). In addition hnRNP A1 is involved in hepatitis C virus (HCV) replication (79). PTB has been shown to interact with RNA in poliovirus, hepatitis C virus, hepatitis A virus, HIV, and human T-cell leukemia virus (90). Although the significance of these interactions is not known in all cases, PTB has been shown to be necessary for IRES dependent translation in viruses such as PV and HCV (1, 62, 118).

In MHV, PTB-hnRNPA1 binding interactions have been proposed to circularize the genome and provide a circular closed loop for replication (70). These two components of a pre-mRNA splicing complex were initially co-purified and therefore may together form part of a viral RNA transcription complex (90). Depletion of PTB or hnRNPQ3 has been shown to be inconsequential in cap-dependent translation of an MHV reporter construct, indicating that these proteins do not function in translation (32, 33). However, overexpression or depletion of these factors resulted in altered growth kinetics in MHV implicating a role in RNA synthesis (32, 139).

Because both PTB and hnRNPA1 bind to the 5' and 3' UTRs respectively, and have been shown to interact with each other (70), they could be suppressing cap-dependent translation by blocking access to the 5' UTR. In this scenario, cap-dependent translation may be inhibited through the association of protein factors necessary for RNA transcription. However, in the absence of viral proteins, as in our experiments, there is no transcription (Fig. 3-6). This model is applicable to both the genomic and subgenomic RNAs since translation was suppressed with only the 77nt 5' UTR. Based on our data we can speculate that *in vitro* and in

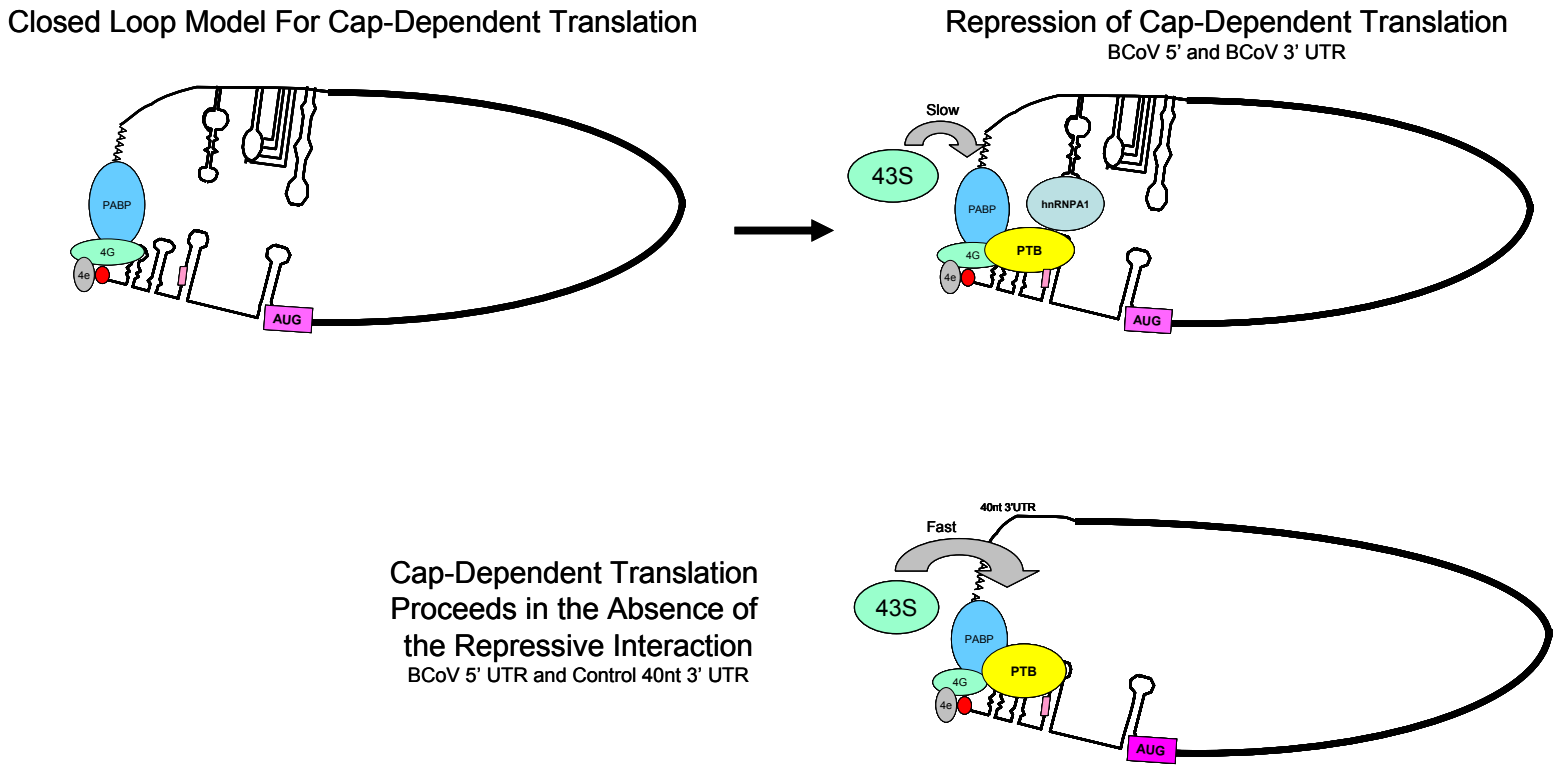


Figure 3-6. Translation Repression Model. Translation of coronaviruses is believed to occur by the canonical cap-dependent translation initiation via a closed loop formed through interactions with the 5' UTR cap binding complex eIF4F and PABP. However, an additional 5'-3' UTR interaction formed through interactions with cellular proteins (possibly proteins involved in RNA synthesis) could result in inefficient recruitment or binding of initiation factors, ribosome loading, or ribosome scanning. In the absence of this interaction, there is no repression.

the absence of viral proteins, all viral RNAs are in an intermediate complex conformation and that additional viral protein factors involved in either promoting translation or the initiation of negative strand RNA synthesis are missing.

In infected cells, viral proteins also likely differentially mediate translation of both the genomic RNA and the sgmRNAs. It has previously been reported that in infected cells the N protein enhances translation of a reporter construct with the MHV UTRs (149). Also, in BCoV infected cells, translation of transfected RNAs with the 210nt 5' UTR were decreased compared to RNAs with only the 77nt sgmRNA 7 5' UTR (135). These data indicate that not only do viral proteins regulate translation, but that in the presence of viral proteins, mRNA species with different 5' UTRs may be regulated differently. Viral proteins could upregulate or downregulate translation through interactions directly with the RNA or through interactions with cellular proteins depending on the stage of replication.

This study provides a detailed look at the effects of the viral UTRs in coronavirus translation. In this limited *in vitro* system we observed no evidence of a *cis*-acting translational advantage that might be necessary for a productive virus infection, but instead found seemingly repressed levels of cap-dependent translation. The proposed model suggests the repressed translation could be indicative of an intermediate stage in virus replication. However, further studies will be necessary to investigate the mechanism. Our initial observations led us to pose two questions, (i) is translation of capped transcripts repressed, or (ii) is translation of uncapped transcripts enhanced. Although we addressed the former question, the latter remains unanswered and intriguing. During coronavirus infection, synthesis of viral proteins continues even after host protein synthesis has been inhibited by the viral infection (10, 64, 127). Therefore we can ask: if in an environment in which cap-dependent translation is inhibited, is there a secondary mechanism through which coronavirus RNAs are translated? There are reports of secondary mechanisms of translation when cap-dependent translation is inhibited

including IRES mediated translation of cellular RNAs and an apparent novel mechanism that involves a 5'-3' UTR interaction in dengue virus RNAs which circumvents the need for eIF4E (48, 104). In addition there is precedent for alternate translation mechanisms in coronaviruses, including reports of IRES translation of coronavirus multicistronic ORFs, ORF 3c in IBV, ORF 5b in MHV, and ORF 3b in TGEV (100, 120, 151). Also, a weak IRES has been identified in the 5' UTR of the IBV genome (101). Future studies will be necessary to investigate this possibility.

Overall, the underlying mechanism of translation and its regulation in coronaviruses remains elusive. Future studies are needed to address the influence of viral proteins on translation and regulation. Furthermore, it is necessary to determine if the observed repression of cap-dependent translation occurs *in vivo* and if there is a secondary mechanism for the preferential translation of viral RNAs when cap-dependent translation is inhibited. Also, since each viral UTR was shown to enhance translation, identification of translation enhancer elements within the viral UTRs would provide useful information regarding translation regulation in coronaviruses.

Chapter IV. Coronavirus Nsp 1 Bioinformatic Analysis

Introduction

Coronaviruses have been isolated from birds, mammals, and humans. Most coronavirus infections are associated with gastroenteric and respiratory infections (89, 165). Severe acute respiratory syndrome coronavirus (SARS-CoV) however, causes severe respiratory disease and the 2003 outbreak resulted in an ~10% mortality rate worldwide (24). Since the emergence of SARS-CoV, the number of new coronaviruses and coronavirus isolates has dramatically increased. Coronaviruses have been classified into three groups based on antigenic differences and genome sequence (106). More recently further classification has divided groups into subgroups. Group 2 coronaviruses are perhaps the most studied of the three groups of coronaviruses, partially due to the fact that SARS-CoV is a member of this group. Group 2 members bovine coronavirus (BCoV) along with mouse hepatitis coronavirus (MHV) have been subdivided into group 2a, and SARS-CoV has been subdivided into group 2b (161, 169). More recent group 2 members have been placed in subgroups 2c and 2d and currently include only bat coronaviruses (169).

All coronavirus nonstructural proteins are translated and cleaved from a large polyprotein encoded in the 5'-proximal two-thirds of the ~30kb genome. In group 2 viruses there are sixteen predicted protein products from the ORF 1a/1b polyprotein. The first nonstructural protein (nsp 1) is cleaved cotranslationally almost immediately after it is translated (38). This protein is approximately 28kDa in group 2a coronaviruses and is also referred to as p28. In SARS-CoV and group 2c and 2d the protein is smaller, approximately 20kDa.

A definitive function for nsp 1 in virus replication has not been determined but it has been shown to be required for replication. It is known that the N-terminal part of the coding region has a separate requirement for both the RNA secondary structure and the translation product (Chapter II) (18, 20). Two *cis*-acting stem-loops are known to exist within this region of

nsp 1 in BCoV and are required for DI RNA replication (Chapter II) (20). In MHV, deletion of nsp 1 N-terminal segments or N-terminal alanine mutagenesis using a full length infectious clone results in impaired or no virus replication (18). Considering the requirement for the translation product, replication of BCoV DI RNA is inhibited when the nsp 1 start codon is mutated to a stop codon or the coding sequence is altered by the introduction of a frameshift mutation (20). It is postulated that nsp 1 is part of the viral replication complex, as it was found to associate with viral RNA and viral proteins that include the helicase, RdRp, nucleocapsid, and matrix proteins (19). Nsp 1 has furthermore been shown to directly bind two viral nonstructural proteins nsp 7 and nsp 10 (19). The functions of these two nonstructural proteins in virus replication are not known.

Although there is a replication requirement for nsp 1 and it interacts with both RNA and other proteins, the significance of these interactions and how they function in virus replication is unknown. In MHV and SARS-CoV, nsp 1 has recently been shown to affect cellular processes including cell cycle arrest, mRNA decay and translation shutoff (78, 115, 164, 177). Furthermore in BCoV, p28 overexpression negatively affects viral RNA levels and could potentially affect cellular levels as well, implying a conserved function may exist between these viral proteins (Chapter II). It is possible these observed effects are due to interactions between nsp 1 and other unidentified factors.

Due to the close relationship between group 2a coronaviruses BCoV and MHV, the *cis*-acting function of the nsp 1 5'-terminal region and protein functions of nsp 1 are likely conserved elements. However, it is not known if the more divergent SARS-CoV nsp 1 or other coronavirus nsp 1 sequences share some or all of these elements. In this analysis, the predicted secondary structures and the nsp 1 amino acid sequences have been compared in order to characterize and identify elements similar to those in BCoV nsp 1. Analysis of the secondary structure of the nsp 1 5' terminal regions from selected group 2 coronaviruses showed most to be BCoV like

and revealed SARS-CoV to be more group 1 like. Amino acid alignments also showed most sequences having high similarity to BCoV. However, SARS-CoV showed some similarities with group 1 coronaviruses. In addition, although RNA and protein binding have been experimentally demonstrated for BCoV and MHV nsp 1, respectively, no specific binding domains have been determined. Therefore, potential binding motifs have been predicted by bioinformatic analysis of BCoV nsp 1. A total of three putative RNA and protein interaction motifs have been predicted in BCoV and may be partially conserved in MHV. Only one of these regions may be conserved among BCoV, MHV, and SARS-CoV, and could potentially contribute to a conserved interaction domain and possibly a conserved functional domain.

Materials and Methods

Genbank Accession Numbers. BCoV #U00735, HECV-4408 #AF523844, HCoV-OC43 #NC_005147, PHEV #DQ011855, ECoV #NC_010327, MHV A59 #AY700211, SARS-CoV#AY291315, BatSARS-HKU3-1 #DQ022305, BatCoV HKU4-1 #EF065505, BatCoV HKU9-1 #EF065513, FCoV #NC_007025, TGEV #NC_002306, HCoV-229E #AF304460, PEDV #NC_003436

RNA structure predictions. *Mfold* (http://mfold.bioinfo.rpi.edu/) was used to predict RNA secondary structures (107, 176).

Bioinformatics. Nsp 1 amino acid sequence alignments and trees were constructed using Vector NTI Suite 8 (Invitrogen) Align X. Phylogenetic trees were calculated from the multiple alignment by the Neighbor-Joining method (Saitou and Nei) (131). Distance values representing the degree of divergence between sequences are shown in parentheses. Protein binding domains were predicted with the SMART database (http://smart.embl-heidelberg.de/) (95, 133) using the full length p28 sequence as well as ~100 amino acid segments. The putative binding domains were then used for PSI-BLAST (Position-Specific Iterated BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). (3, 146). PSI-BLAST iterations were done until a non-

coronavirus match was found. This sequence was then used in SMART to determine if the same domains were predicted as in BCoV. 1. For BCoV p28 35-75: *Physcomitrella patens* (EDQ51842) was identified and SMART identified putative binding domains including zinc fingers. 2. For BCoV p28 100-160: *Burkholderia pseudomallei* (Genbank #ZP_02403227) was identified and SMART identified putative protein binding domain. 3. For BCoV p28 90-190: *Polaribacter dokdonensis* (EAQ41748) was identified and SMART identified a transmembrane domain. No matches were found for BCoV p28 190-246 (Appendix III). Transmembrane domains were predicted using Tmpred. Scores above 500 are significant (Appendix III) (http://www.ch.embnet.org/software/TMPRED_form.html) (65).

Results

RNA Secondary Structure

In BCoV, stem-loops V and VI in the immediate nsp 1 coding region are necessary for DI RNA replication (Chapter II) (20). These structures have been confirmed by RNase enzyme probing and disruption of either stem-loop structure results in a loss of DI RNA replication (Chapter II) (20). The secondary structures of the nsp 1 5'-terminal *cis*-acting region of group 2a viruses such as MHV, human enteric coronavirus HECV 4408, human coronavirus HCoV-OC43, and porcine hemagglutinating encephalomyelitis virus (PHEV) are all predicted by *mfold* to be similar to that for BCoV (nucleotide sequence identity ≥60%), with an elongated stem-loop (SLV) downstream of the ORF 1a start codon that is immediately followed by a shorter stem-loop (SLVI) (Fig. 4-1), BCoV and MHV shown). Although the *cis*-replication functions of these elements in all group 2a coronaviruses have not been confirmed, based on sequence identity and secondary structure predictions, it is expected that these regions will function as *cis*-replication elements as shown in BCoV (sequence alignment shown in Appendix III, Figure A3-1A).

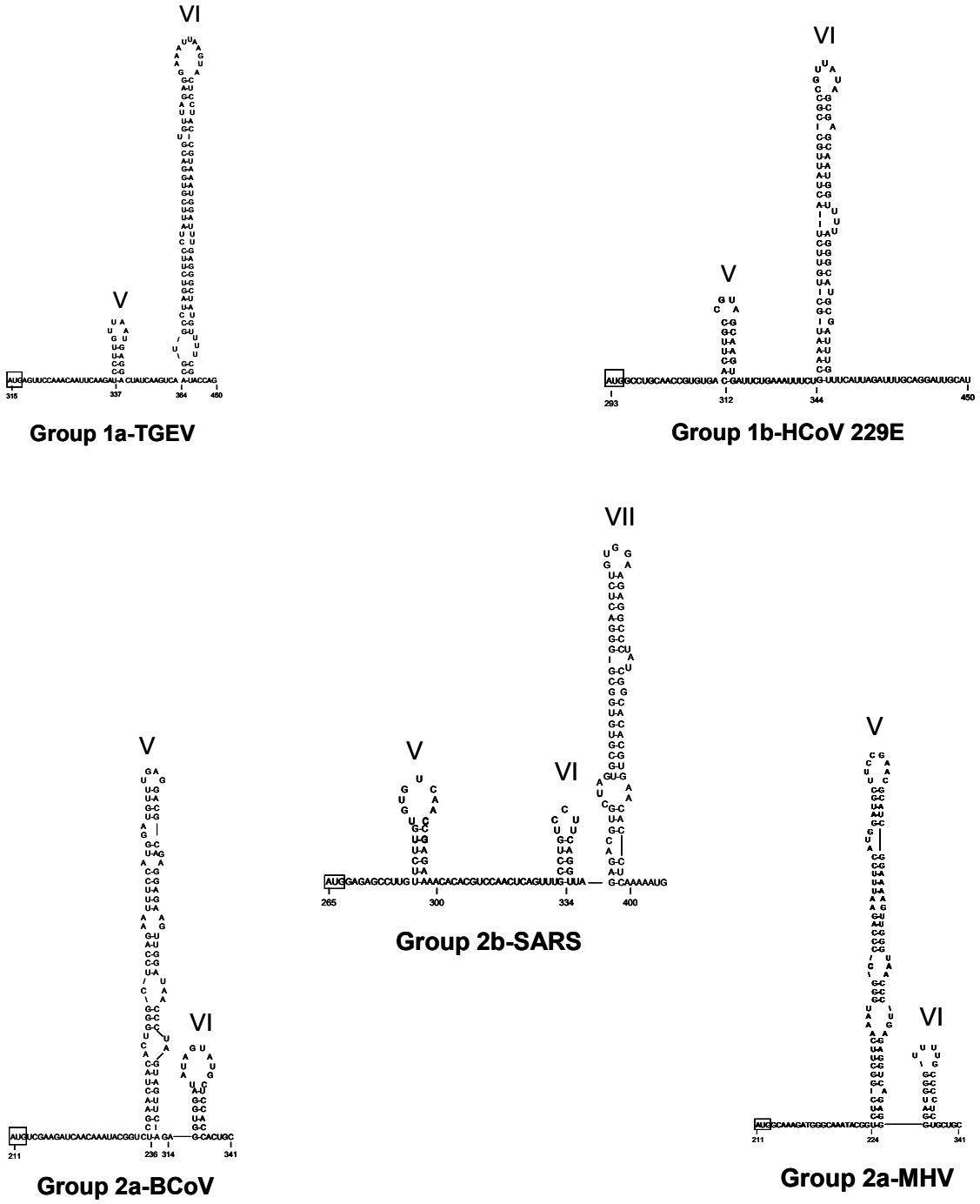


Figure 4-1. 5'-Terminal nsp 1 Secondary Structure *mfold* Predictions for Group 1 and Group 2 Coronaviruses. Predicted secondary structures of group 1 TGEV and HCoV229E, group 2b SARS-CoV and group 2a BCoV and MHV. The secondary structures of SLV and SLVI in BCoV have been confirmed by RNase probing.

Group 2b member SARS-CoV is predicted to have three stem-loops within the 5'-terminal nsp 1 coding region (Fig. 4-1). Although SARS-CoV has been assigned to group 2, these predicted secondary structures appear more similar those in group 1 coronaviruses. The nsp 1 sequence of group 1a FCoV and TGEV and group 1b members HCoV-229E and PEDV, are all predicted to have similar structures in the 5'-terminal nsp 1 coding region, a short stem-loop followed by an elongated stem-loop (Fig. 4-1, TGEV and HCoV-229E shown). The predicted stem-loops in SARS-CoV follow this arrangement with one additional upstream stem-loop structure (sequence alignment shown in Appendix III, Figure A3-1B). The predicted secondary structures of group 2c and 2d sequences are different from both group 1 and group 2 and are not predicted to have similar stem-loop structures in the 5' terminal nsp 1 region (data not shown).

Interestingly, the SARS-CoV predicted stem-loop arrangement is more similar to the predicted structures in the 5' terminal nsp 1 sequences of group 1 coronaviruses. This is intriguing because stem-loop IV within the 5' UTR of SARS-CoV has also been predicted to be more similar to group 1 coronaviruses (130). This mix of sequence identities may reflect the recombinant origin of SARS-CoV and represent a link between nsp 1 of group 1 and group 2 coronaviruses.

Nsp 1 Amino Acid Sequence

To analyze the nsp 1 amino acid sequences of the coronaviruses of interest, the entire nsp 1 coding regions were aligned and the results are presented as the percentage of similar residues. A sequence alignment and phylogenetic tree of group 2 coronavirus nsp 1 amino acids is shown in Figure 4-2 and shows that HCoV-OC43, PHEV and HCoV-4408 are BCoV like and are more than 95% similar while equine coronavirus (ECoV) is approximately 75% similar. MHV is also closely related and approximately 60% similar with BCoV, while the more distant SARS-CoV is only 24% similar (Fig. 4-2). Bat-SARs HKU3-1 nsp 1 sequence is virtually

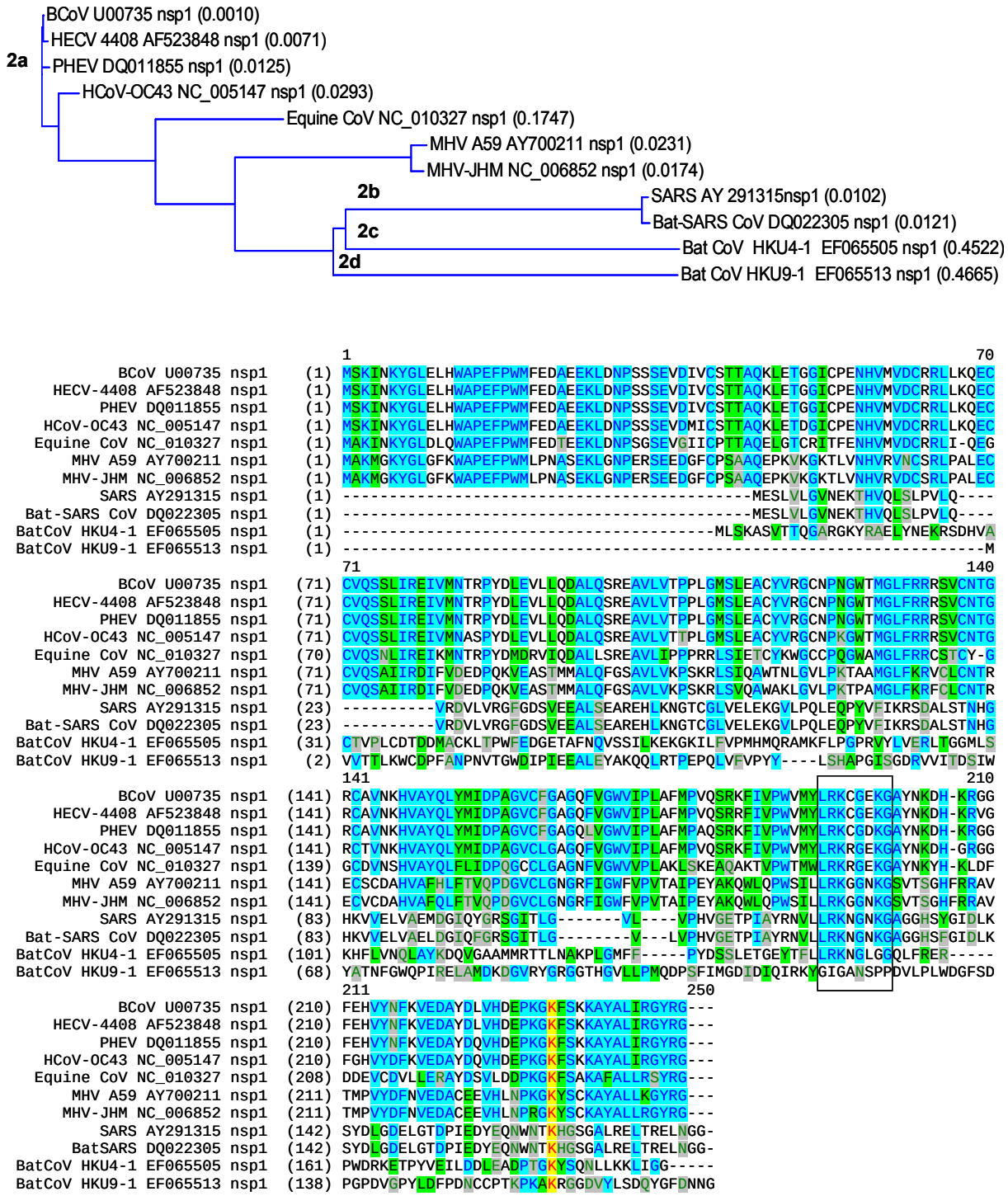


Figure 4-2. Group 2 Coronavirus nsp 1 Amino Acid Alignment and Phylogenetic Tree. Neighbor-Joining tree distance scores are shown in parentheses. Conserved LRKxGxKG region is boxed.

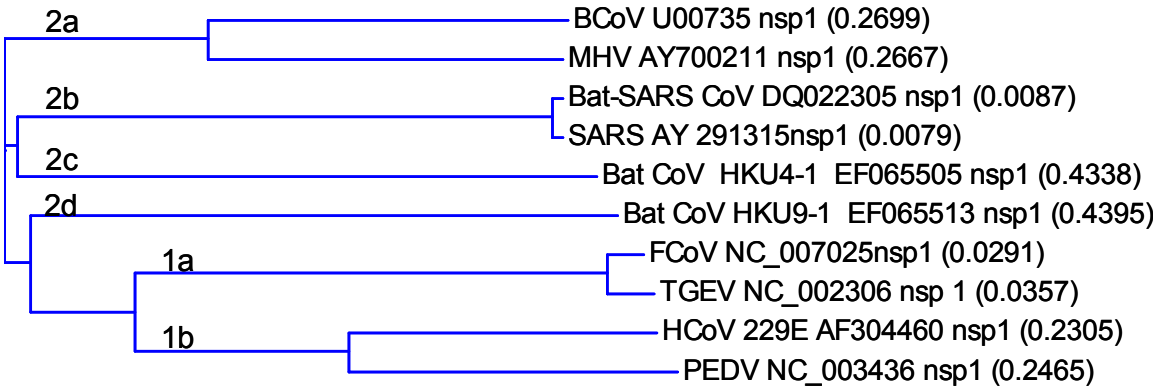
identical to the SARS-CoV sequence (98% identity) (Fig. 4-2). Group 2c and 2d bat coronaviruses are approximately 30% similar to each other and are more closely related to SARS-CoV than group 2a with 30% similarity (Fig. 4-2). One region of amino acids, LRKxGxKG, (BCoV aa 192-199, SARS-CoV, aa 123-130), is conserved among all group 2 coronaviruses analyzed except BatCoV 4-1 which has the sequence LRKxGxGG and BatCoV 9-1 which does not have this sequence at all (Fig. 4-2). This region consists primarily of charged residues and was mentioned by Almeida, et.al, as a potential interaction domain (2).

A phylogenetic tree of nsp 1 sequences from group 1 and group 2 coronaviruses is shown in Figure 4-3 (alignment shown in Appendix III, Figure A3-2). Group 1a members including FCoV and TGEV have nsp 1 sequences that are 94% similar. HCoV229E and PEDV, both group 1b members have nsp 1 sequences that are 64% similar. Interestingly, as observed with the nsp 1 RNA secondary structure predictions, SARS-CoV nsp 1 amino acid sequence is both group 2-like and group 1-like with 24% similarity to group 2a BCoV and 21% similarity to group 1a FCoV (Fig. 4-3). This apparent intergroup relationship may provide insight into both the origin and conserved functional properties of the nsp 1 proteins.

Predicted Binding Motifs in BCoV Nsp 1

In BCoV and MHV, nsp 1 has been shown to interact with both viral RNA and proteins. In addition, experimental data suggests a conserved function may exist among nsp 1 proteins (30, 78, 115, 164, 177). However, potential RNA or protein interaction regions through which nsp 1 may exert its function(s) have not been identified. Bioinformatic analysis of BCoV nsp 1 using the SMART domain prediction database (95, 133), has revealed potential binding motifs (Table 4-1). A comparison with nsp 1 in MHV and SARS-CoV sequences reveals potential conservation with MHV but less so with SARS-CoV.

A domain search using SMART with the BCoV nsp 1 sequence identified three potential binding regions. No interaction domains were confidently predicted, but all



SARS-CoV nsp 1 is group 2 like

```
                    1                                                    70
BCoV nsp1     (1)   MSKINKYGLELHWAPEFPWMFEDAEEKLDNPSSSEVDIVCSTTAQKLETGGICPENHVMVDCRRLLKQEC
 MHV nsp1     (1)   MAKMGKYGLGFKWAPEFPWMLPNASEKLGNPERSEEDGFCPSAAQEPKVKGKTLVNHVRVNCSRLPALEC
SARS nsp1     (1)   ------------------------------------------MESLVLGVNEKTHVQLSLPVLQVRDV
                    71                                                   140
BCoV nsp1    (71)   CVQSSLIREIVMNTRPYDLEVLLQDALQSREAVLVTPPLGMSLEACYVRGCNPNGWTMGLFRRRSVCNTG
 MHV nsp1    (71)   CVQSAIIRDIFVDEDPQKVEASTMMALQFGSAVLVKPSKRLSIQAWTNLGVLPKTAAMGLFKRVCLCNTR
SARS nsp1    (27)   LVR-G------FGDSVEEALSEAREHLKNGTCGLVELEK----------GVLPQLEQPYVFIKR----SD
                    141                                                  210
BCoV nsp1   (141)   RCAVNKHVAYQLYMIDPAGVCFGAGQFVGWVIPLAFMPVQSRKFIVPWVMYLRKCGEKGAYNKDHKR-GG
 MHV nsp1   (141)   ECSCDAHVAFHLFTVQPDGVCLGNGRFIGWFVPVTAIPEYAKQWLQPWSILLRKGGNKGSVTSGHFRRAV
SARS nsp1    (76)   ALSTNHGHKVVELVAEMDGIQYGRS----GITLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLK
                    211                                 249
BCoV nsp1   (210)   FEHVYNFKVEDAYDLVHDEPKGKFSKKAYALIRGYRG--
 MHV nsp1   (211)   TMPVYDFNVEDACEEVHLNPKGKYSCKAYALLKGYRG--
SARS nsp1   (142)   SYDLGDELGTDPIEDYEQNWNTKHGSGALRELTRELNGG
```

SARS-CoV nsp 1 is group 1 like

```
                    1                                                    70
FCoV nsp1     (1)   MSSKQFKILVNEDYQVNVPSLPFRDALQ-----------EIKYCYRNGFDGYVFVPEYRRDLVDCNRKD
TGEV nsp1     (1)   MSSKQFKILVNEDYQVNVPSLPIRDVLQ-----------EIKYCYRNGFEGYVFVPEYCRDLVDCDRKD
SARS nsp1     (1)   MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEAREHLKNGTCGLVELEKGVLPQLEQPYVF
                    71                                                   140
FCoV nsp1    (59)   HYVIGVLGNGISDLKPVLLTEPSVMLQG---------------------FIVRANCN----GVLEDFDL
TGEV nsp1    (59)   HYVIGVLGNGVSDLKPVLLTEPSVMLQG---------------------FIVRANCN----GVLEDFDL
SARS nsp1    (71)   IKRSDALSTNHGHKVVELVAEMDGIQYGRSGITLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDL
                    141                                 180
FCoV nsp1   (103)   KFARTGNG---------------------------------
TGEV nsp1   (103)   KIARTG-----------------------------------
SARS nsp1   (141)   KSYDLGDELGTDPIEDYEQNWNTKHGSGALRELTRELNGG
```

Figure 4-3. Phylogenetic Tree of Group 1 and Group 2 Coronaviruses. Neighbor-Joining tree distance scores are shown in parentheses. SARS-CoV nsp 1 sequence is approximately 20% similar to both group 2 and group 1 coronaviruses.

Table 4-1. Putative Motifs in BCoV nsp 1 as Determined by SMART.

Name	Aa		*E-Value	Description
Putative Motif I				
AN1 ZnF	40	67	1.92e+02	Zinc finger; involved in protein binding
DED	15	95	2.03e+02	Death effector domain; involved in protein binding
C2HC ZnF	52	64	1.12e+03	Zinc finger; found in the Nucleocapsid protein of retroviruses and eukaryotic RNA/ssDNA binding proteins.
C4 ZnF	34	79	2.44e+03	Zinc finger; involved in DNA-binding.
Zpr1-type ZnF	51	165	4.24e+03	ZPR1-type zinc finger domain; involved in protein binding
Putative Motif II				
Thaumatin	106	241	2.71e+01	Primarily in plants; involved in pathogenesis; possible inhibitor
EGF Lam	115	161	1.27e+03	Laminin-type epidermal growth factor-like domain; binds components of basement membranes
Kringle	114	144	1.50e+03	Protein binding domain found in serine proteases
Bowman-Birk	120	161	1.99e+03	Primarily in plants; found in serine protease inhibitors
Putative Motif III				
4.1m	190	208	5.95e+02	Putative band 4.1 homologues' binding motif; Found in neurexins, syndecans and glycophorin C intracellular C-termini.

* E-value=score representing the probability that two sequences are matched by chance. A lower number indicates higher similarity.

matches with above threshold E-values fit within three regions of the nsp 1 sequence (Fig. 4-4). Consistent with this data, a PSI-BLAST search using the putative motifs identified non-coronavirus sequences that produced similar SMART hits. However, because all hits are above levels considered significant, nsp 1 most likely contains novel interaction domains that need to be experimentally defined.

An N-terminal motif located between amino acids S34 and S75 has identity to five different protein and RNA binding domains. Four are zinc finger domains, two of which are known to be protein binding zinc fingers and two are nucleotide binding domains (Table 4-1). Analysis of MHV nsp 1 sequence also reveals identity to a nucleic acid binding zinc-finger domain in the same region (Fig. 4-4 and data not shown). In BCoV this region contains five cysteine residues, four of which are conserved in MHV, and a conserved histidine residue, which could collectively make up a binding motif (Fig. 4-4) (21, 60). Since it is known from experimental data that BCoV nsp 1 is an RNA binding protein (Chapter II), this proposed N-terminal region is therefore a good candidate for the identified RNA binding interactions with the *cis*-acting elements in the viral UTRs and may also be conserved in MHV. Also, this domain is within the N-terminal region of the protein necessary for efficient virus replication, suggesting one function of nsp 1 in replication may require binding to viral RNA. Alternatively, analysis of SARS-CoV nsp 1 did not find any potential binding motif in this region and sequence conservation with BCoV and MHV within this region is low (data not shown and Fig. 4-3).

The second region found to be a potential binding motif is located between amino acids E114 and D156 (Fig. 4-4). Four different types of protein binding domains showed some similarities with this region in nsp 1 (Table 4-1). In MHV, using deletion and truncation mutants it has been reported that nsp 1 interacts with viral proteins nsp 7 and nsp 10 between amino acids E84 and R166 (T84 and Q166 in BCoV), which fits within the BCoV putative motif (19) (Fig. 4-4). This region of nsp 1 is therefore a potential conserved binding motif among group 2a

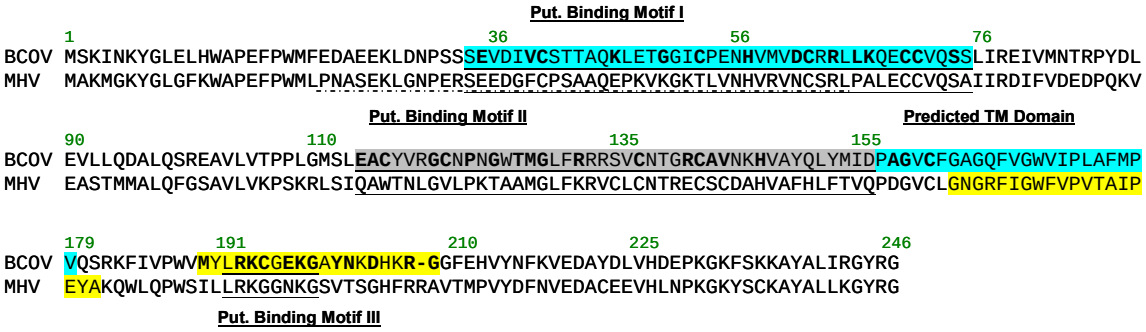


Figure 4-4. BCoV nsp 1 Putative Interaction Motifs. Results from SMART. **Putative binding motif I** shows similarities with nucleic acid and protein binding zinc fingers. The MHV nsp 1 sequence contains some conserved residues in this region. The sequence in MHV found to have similarity to a zinc finger is noted with a dotted underline. **Putative binding motif II** has conserved residues with MHV and resides within the experimentally defined protein binding region for MHV nsp 1(aa E84-R166). **Putative binding motif III** is similar to a binding domain known to interact with a cell cycle regulatory protein. This motif contains the most conserved region between group 2 nsp 1 sequences, LRKxGxKG. The predicted transmembrane (TM) domain in BCoV was only weakly predicted for MHV and therefore may only occur in BCoV or BCoV-like nsp 1 sequences. Residues that were matches to the majority of the predicted motifs for each region (Table 4-1) are in bold.

coronaviruses and is likely a protein interaction region. A similar binding motif was not predicted to exist in SARS-CoV (data not shown).

In BCoV a transmembrane domain was confidently predicted at positions P157-V179 by SMART and the transmembrane prediction program TMpred (Fig. 4-4). In comparison, in MHV, a transmembrane domain is only predicted in the same region, amino acids 163-181, using TMpred. Although significant, the overall score is much lower than in BCoV so it is unclear if this is a true transmembrane domain (Appendix III). In SARS-CoV, no transmembrane domain was predicted (data not shown). Hypothesizing that nsp 1 is an integral part of the replication complex, a transmembrane domain could aid in localizing replication complexes to membranes.

A third binding motif is predicted to exist at the C-terminus of BCoV nsp 1. This C-terminal predicted protein binding region exists downstream of the transmembrane domain (Fig. 4-4). Only a single known motif was found to have some similarities with the nsp 1 sequence in this region, band 4.1 homologues' binding motif (Table 4-1). This motif has been found on the intracellular region of transmembrane glycoproteins such as glycophorin C. Protein 4.1 binds the motif in glycophorin C and is involved in red blood cell shape and stability (40). More recently, protein 4.1 has been found to be more ubiquitously expressed and a component of centromeres and mitotic spindles in nucleated cells. Depletion of the protein resulted in cell cycle arrest by causing G_1 accumulation and preventing progression into the S phase (86). Interestingly, in MHV, nsp 1 has been shown to arrest the cell cycle in the Go/G_1 phase (30). In addition, a similar affect has been reported with SARS-CoV nsp 1 (164). The predicted domain is within the highly conserved region containing LRKxGxKG (Fig. 4-4, boxed amino acids), making it plausible that all three viral nsp 1 proteins have this binding motif and are able to bind protein 4.1, or a related 4.1/FERM superfamily protein, resulting in cell cycle dysregulation.

Discussion

Analysis of nsp 1 sequences revealed both RNA secondary structure and amino acid sequence similarities. In terms of protein sequence similarity, group 2a coronaviruses are highly conserved with greater than 60% similarity at the amino acid level. While experimental data has shown that nsp 1 binds both viral RNA and proteins, the specific domains responsible for these interactions remain elusive. The predicted putative interaction motifs are not significant enough to conclusively define, however the data are consistent with the experimental observations. Putative RNA and protein binding motifs and possibly a transmembrane domain are therefore also likely conserved features of group 2a nsp 1 proteins. Based on amino acid sequence alignments of the ORF 1b polyprotein, SARS-CoV was placed into group 2 as a distant relative of MHV and BCoV (141). Considering the nsp 1 predicted RNA secondary structure and amino acid sequence, it appears that the distantly related SARS-CoV is both group 1-like and group 2-like. Similarity between SARS-CoV and group 1a is equal to that observed for group 2a, and therefore may have functions that are both group 1 and group 2 like and could represent a link between the two groups.

The most conserved region in BCoV, MHV, and SARS-CoV, the sequence LRKxGxKG, is within a potential binding region that could bind a member of the protein 4.1/FERM superfamily, such as protein 4.1, a protein recently found to be necessary for cell division (86). The C-terminal portion of the protein has been implicated as the effector of mRNA decay and translation shutoff. In MHV and SARS-CoV, deletion or mutation of the C-terminal portion of nsp 1 resulted in virus attenuation and a decrease in the observed cellular effects (115, 164, 177). In our truncation experiments, overexpression of only the C-terminal portion of p28 resulted in more severely diminished accumulation of DI RNA compared to expression of only the N-terminal portion of p28 (Chapter II). Thus, this may be a conserved function localized to the C-terminal portion of p28, and possibly due to interactions involving binding domains in that

region. Interestingly, group 1 coronavirus HCoV229-E nsp 1 has been reported to affect cellular mRNA levels when expressed from a plasmid (177). This observation was not extensively studied and it is unclear if the C-terminal region of the protein is responsible as shown for SARS-CoV and MHV, but brings up the possibility of a conserved function among group 1 and group 2 nsp 1 proteins.

Although nsp 1 protein sequences among the coronaviruses are dissimilar in many respects, they may have similar functions. It will be interesting to explore the functional similarities among different nsp 1 sequences and how these features differentially affect virus replication and pathogenesis. As SARS-CoV nsp 1 appears to contain features of both group 1 and group 2 coronaviruses, it may provide a common link between these divergent proteins and perhaps help identify functional relationships.

Chapter V. Future Directions

Introduction

The worldwide outbreak of SARS in 2003 resulted in an expanded interest in coronaviruses. The sudden appearance, rapid spread, and high mortality caused by SARS-CoV highlights the need for more efficient and more effective diagnosis, treatment, and prevention of viral diseases. In order to achieve this, a better understanding of virus replication is required.

In this dissertation, fundamental research addressing the requirements for bovine coronavirus replication has been presented. A second *cis*-acting higher order structure within the coding region of the nonstructural protein p28 was identified and shown to be required for replication of a DI RNA and proposed to also be required for genome replication. In addition, although the replication requirement for p28 is not understood, it was shown to be an RNA binding protein that may function through interactions with *cis*-acting elements in the 5' and 3' untranslated regions. Also, a novel 5' and 3' UTR dependent repression of cap-dependent enhancement of translation has been identified. This apparent regulatory mechanism may have important ramifications for the regulation of viral translation and replication. An analysis of the contributions of both the 5' and 3' UTRs in translation without the influence of viral proteins may provide a basis for re-evaluating the basic tenets of translation and regulation.

The identification of replication requirements may provide a basis for the development of specific virus-targeted therapeutics. This research opens up future directions in the study of virus replication that will lead to a more complete understanding of how to better diagnose, treat, prevent, and perhaps cure illnesses caused by coronavirus infections.

Identification of Cellular Proteins That Bind the SLV And SLVI *cis*-Acting Region

Cellular proteins of approximately 60 and 100 kDa were found to bind *cis*-acting replication stem-loops V and VI (Chapter II). Since no viral proteins have been found to bind

specifically to this region, the unidentified cellular proteins may have a role in the *cis*-acting function of SLV and VI. Identification of these proteins would further expand our understanding of coronavirus replication and how SLV and VI function in replication.

Experimentally Define the RNA Binding Domain of p28

The first nonstructural protein, p28, was shown for the first time to be an RNA binding protein. Although bioinformatics has established a putative RNA binding domain in p28, this needs to be experimentally tested. If our hypothesis is correct, the N-terminal portion of p28, which is found in all known coronavirus DI RNAs and required for virus replication, contains the RNA binding domain. Therefore, gel shift experiments with only the N-terminal portion of p28 would still be expected to bind the *cis*-acting stem-loop structures identified in Chapter II. Experimental elucidation of this domain would help define the replication requirement for the N-terminal portion of p28 separate from the secondary structure *cis*-requirement of SLV and VI. The dual replication requirement of the N-terminal portion of p28 secondary structure and protein presents a unique targeting opportunity that should be further explored.

Expand Translation Studies

Translation studies *in vitro* using RRL and a Renilla Luciferase reporter construct fused to the BCoV 5' and 3' UTRs showed a repression of cap-dependent translation. It is hypothesized this repression is mediated by a 5'-3' UTR interaction via cellular proteins (Chapter III). This may represent a transcriptionally poised but inactive complex. To begin to investigate the primary hypothesis that viral proteins are involved in translation through interactions with the *cis*-acting elements within the UTRs, three approaches might be taken. First, *in vitro* translation using BCoV infected and prepared lysates. This approach will not elucidate specific factors that influence translation but will show how translation proceeds with viral proteins in a limited system and in the absence of competition with cellular mRNAs. Second, translation of reporter constructs in both uninfected and BCoV infected cells. This

approach will take the previous step further and look at translation with the influence of virus infection in cells. It is expected that in a competitive and more complex environment the factors that influence translation will be varied. These studies will provide information on how the genomic RNA is initially translated in the presence of cellular mRNAs and cellular factors. Thirdly, using *in vitro* translation, individually expressed and purified viral proteins can be added to translation reactions to test how these regulate translation. Although this approach will not determine a mechanism of action, the effects of individual viral proteins on translation can be studied. More extensive experiments including both RNA and protein binding domain mutants and how they affect translation will be necessary to show a direct interaction results in translation modulation.

References

References

1. **Ali, N., and A. Siddiqui.** 1995. Interaction of polypyrimidine tract-binding protein with the 5' noncoding region of the hepatitis C virus RNA genome and its functional requirement in internal initiation of translation. J Virol **69:**6367-75.

2. **Almeida, M. S., M. A. Johnson, T. Herrmann, M. Geralt, and K. Wuthrich.** 2007. Novel beta-barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. J Virol **81:**3151-61.

3. **Altschul, S., T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. **25:**3389-3402.

4. **Andino, R., N. Boddeker, D. Silvera, and A. V. Gamarnik.** 1999. Intracellular determinants of picornavirus replication. Trends Microbiol **7:**76-82.

5. **Andino, R., G. E. Rieckhof, and D. Baltimore.** 1990. A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. Cell **63:**369-80.

6. **Ball, L. A.** 2001. Replication strategies of RNA viruses, p. 105-118. *In* D. M. Knipe (ed.), Fields Virology, Fourth ed, vol. 1. Lippincott Williams and Wilkins, Philadelphia.

7. **Barton, D. J., B. J. Morasco, and J. B. Flanegan.** 1999. Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. J Virol **73:**10104-12.

8. **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:**1439-48.

9. **Beaudette, F. R., Hudson, C.B.** 1937. Cultivation of the virus of infectious bronchitis. J Am Vet Med Assoc **90:**51-60.

10. **Bechill, J., Z. Chen, J. W. Brewer, and S. C. Baker.** 2008. Coronavirus infection modulates the unfolded protein response and mediates sustained translational repression. J Virol **82:**4492-501.

11. **Bhardwaj, K., L. Guarino, and C. C. Kao.** 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. J Virol **78:**12218-24.

12. **Bhardwaj, K., S. Palaninathan, J. M. Alcantara, L. L. Yi, L. Guarino, J. C. Sacchettini, and C. C. Kao.** 2008. Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. J Biol Chem **283:**3655-64.

13. **Binn, L. N., E. C. Lazar, K. P. Keenan, D. L. Huxsoll, R. H. Marchwicki, and A. J. Strano.** 1974. Recovery and characterization of a coronavirus from military dogs with diarrhea. Proc Annu Meet U S Anim Health Assoc**:**359-66.

14. **Brian, D. A., and Spaan, W.J.M.** 1997. Recombination and coronavirus defective interfering RNAs. Seminars in Virology **8:**101-111.

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

16. **Brierley, I., M. E. Boursnell, M. M. Binns, B. Bilimoria, V. C. Blok, T. D. Brown, and S. C. Inglis.** 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. Embo J **6:**3779-85.

17. **Brierley, I., P. Digard, and S. C. Inglis.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell **57:**537-47.

18. **Brockway, S. M., and M. R. Denison.** 2005. Mutagenesis of the murine hepatitis virus nsp1-coding region identifies residues important for protein processing, viral RNA synthesis, and viral replication. Virology **340:**209-23.

19. **Brockway, S. M., X. T. Lu, T. R. Peters, T. S. Dermody, and M. R. Denison.** 2004. Intracellular localization and protein interactions of the gene 1 protein p28 during mouse hepatitis virus replication. J Virol **78:**11551-62.

20. **Brown, C. G., K. S. Nixon, S. D. Senanayake, and D. A. Brian.** 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:**7716-24.

21. **Brown, R. S.** 2005. Zinc finger proteins: getting a grip on RNA. Curr Opin Struct Biol **15:**94-8.

22. **Brunner, J. E., J. H. Nguyen, H. H. Roehl, T. V. Ho, K. M. Swiderek, and B. L. Semler.** 2005. Functional interaction of heterogeneous nuclear ribonucleoprotein C with poliovirus RNA synthesis initiation complexes. J Virol **79:**3254-66.

23. **Cavanagh, D., K. Mawditt, B. Welchman Dde, P. Britton, and R. E. Gough.** 2002. Coronaviruses from pheasants (Phasianus colchicus) are genetically closely related to coronaviruses of domestic fowl (infectious bronchitis virus) and turkeys. Avian Pathol **31:**81-93.

24. **CDC** 2005, posting date. Centers for Disease Control and Prevention. http://www.cdc.gov/ncidod/sars/factsheet.htm. [Online.]

25. **Chang, R. Y., and D. A. Brian.** 1996. cis Requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. J Virol **70:**2201-7.

26. **Chang, R. Y., M. A. Hofmann, P. B. Sethna, and D. A. Brian.** 1994. A cis-acting function for the coronavirus leader in defective interfering RNA replication. J Virol **68:**8223-31.

27. **Chang, R. Y., R. Krishnan, and D. A. Brian.** 1996. The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. J Virol **70:**2720-9.

28. **Chasey, D., and S. F. Cartwright.** 1978. Virus-like particles associated with porcine epidemic diarrhoea. Res Vet Sci **25:**255-6.

29. **Cheever, F. S., J. B. Daniels, and et al.** 1949. A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. J Exp Med **90:**181-210.

30. **Chen, C. J., K. Sugiyama, H. Kubo, C. Huang, and S. Makino.** 2004. Murine coronavirus nonstructural protein p28 arrests cell cycle in G0/G1 phase. J Virol **78:**10410-9.

31. **Child, S. J., M. K. Miller, and A. P. Geballe.** 1999. Translational control by an upstream open reading frame in the HER-2/neu transcript. J Biol Chem **274:**24335-41.

32. **Choi, K. S., P. Huang, and M. M. Lai.** 2002. Polypyrimidine-tract-binding protein affects transcription but not translation of mouse hepatitis virus RNA. Virology **303:**58-68.

33. **Choi, K. S., A. Mizutani, and M. M. Lai.** 2004. SYNCRIP, a member of the heterogeneous nuclear ribonucleoprotein family, is involved in mouse hepatitis virus RNA synthesis. J Virol **78:**13153-62.

34. **Chouljenko, V. N., X. Q. Lin, J. Storz, K. G. Kousoulas, and A. E. Gorbalenya.** 2001. Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia. J Gen Virol **82:**2927-33.

35. **Circella, E., A. Camarda, V. Martella, G. Bruni, A. Lavazza, and C. Buonavoglia.** 2007. Coronavirus associated with an enteric syndrome on a quail farm. Avian Pathol **36:**251-8.

36. **Decroly, E., I. Imbert, B. Coutard, M. Bouvet, B. Selisko, K. Alvarez, A. E. Gorbalenya, E. J. Snijder, and B. Canard.** 2008. Coronavirus nonstructural protein 16 is a cap-O binding enzyme possessing (nucloeside-2'O)-methyltransferase activity. J. Virol.

37. **Denison, M., and S. Perlman.** 1987. Identification of putative polymerase gene product in cells infected with murine coronavirus A59. Virology **157:**565-8.

38. **Denison, M. R., and S. Perlman.** 1986. Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. J Virol **60:**12-8.

39. **Denison, M. R., B. Yount, S. M. Brockway, R. L. Graham, A. C. Sims, X. Lu, and R. S. Baric.** 2004. Cleavage between replicase proteins p28 and p65 of mouse hepatitis virus is not required for virus replication. J Virol **78:**5957-65.

40. **Diakowski, W., M. Grzybek, and A. F. Sikorski.** 2006. Protein 4.1, a component of the erythrocyte membrane skeleton and its related homologue proteins forming the protein 4.1/FERM superfamily. Folia Histochem Cytobiol **44:**231-48.

41. **Donaldson, E. F., R. L. Graham, A. C. Sims, M. R. Denison, and R. S. Baric.** 2007. Analysis of murine hepatitis virus strain A59 temperature-sensitive mutant TS-LA6 suggests that nsp10 plays a critical role in polyprotein processing. J Virol **81:**7086-98.

42. **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:**6356-68.

43. **Dong, B. Q., W. Liu, X. H. Fan, D. Vijaykrishna, X. C. Tang, F. Gao, L. F. Li, G. J. Li, J. X. Zhang, L. Q. Yang, L. L. Poon, S. Y. Zhang, J. S. Peiris, G. J. Smith, H. Chen, and Y. Guan.** 2007. Detection of a novel and highly divergent coronavirus from asian leopard cats and Chinese ferret badgers in Southern China. J Virol **81:**6920-6.

44. **Dong, S., and S. C. Baker.** 1994. Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. Virology **204:**541-9.

45. **Dong, S., H. Q. Gao, and S. C. Baker.** 1995. Proteolytic processing of the MHV polymerase polyprotein. Identification of the P28 cleavage site and the adjacent protein, P65. Adv Exp Med Biol **380:**431-5.

46. **Doyle, L., Hutchings, L.M.** 1946. A Transmissible Gastroenteritis in Pigs. J Am Vet Med Assoc **108:**257-259.

47. **Eckerle, L. D., S. M. Brockway, S. M. Sperry, X. Lu, and M. R. Denison.** 2006. Effects of mutagenesis of murine hepatitis virus nsp1 and nsp14 on replication in culture. Adv Exp Med Biol **581:**55-60.

48. **Edgil, D., C. Polacek, and E. Harris.** 2006. Dengue virus utilizes a novel strategy for translation initiation when cap-dependent translation is inhibited. J Virol **80:**2976-86.

49. **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 U S A **101:**3792-6.

50. **Erles, K., C. Toomey, H. W. Brooks, and J. Brownlie.** 2003. Detection of a group 2 coronavirus in dogs with canine infectious respiratory disease. Virology **310:**216-23.

51. **Fields, B. N., D. M. Knipe, P. M. Howley, and D. E. Griffin.** 2001. Fields' virology. CD-Rom, 4th ed. Lippincott Williams & Wilkins, Philadelphia.

52. **Fouchier, R. A., N. G. Hartwig, T. M. Bestebroer, B. Niemeyer, J. C. de Jong, J. H. Simon, and A. D. Osterhaus.** 2004. A previously undescribed coronavirus associated with respiratory disease in humans. Proc Natl Acad Sci U S A **101:**6212-6.

53. **Gallie, D. R.** 1991. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev **5:**2108-16.

54. **Gamarnik, A. V., and R. Andino.** 1998. Switch from translation to RNA replication in a positive-stranded RNA virus. Genes Dev **12:**2293-304.

55. **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:**669-82.

56. **Gorbalenya, A. E., L. Enjuanes, J. Ziebuhr, and E. J. Snijder.** 2006. Nidovirales: evolving the largest RNA virus genome. Virus Res **117:**17-37.

57. **Gough, R. E., S. E. Drury, F. Culver, P. Britton, and D. Cavanagh.** 2006. Isolation of a coronavirus from a green-cheeked Amazon parrot (Amazon viridigenalis Cassin). Avian Pathol **35:**122-6.

58. **Graham, R. L., A. C. Sims, R. S. Baric, and M. R. Denison.** 2006. The nsp2 proteins of mouse hepatitis virus and SARS coronavirus are dispensable for viral replication. Adv Exp Med Biol **581:**67-72.

59. **Greig, A. S., D. Mitchell, A. H. Corner, G. L. Bannister, E. B. Meads, and R. J. Julian.** 1962. A Hemagglutinating Virus Producing Encephalomyelitis in Baby Pigs. Can J Comp Med Vet Sci **26:**49-56.

60. **Hall, T. M.** 2005. Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol **15:**367-73.

61. **Hamre, D., and J. J. Procknow.** 1966. A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med **121:**190-3.

62. **Hellen, C. U., G. W. Witherell, M. Schmid, S. H. Shin, T. V. Pestova, A. Gil, and E. Wimmer.** 1993. A cytoplasmic 57-kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. Proc Natl Acad Sci U S A **90:**7642-6.

63. **Hentze, M. W., Gebauer, F., Preiss, T.** 2007. cis-Regulatory Sequences and trans-acting Factors in Translational Control, p. 269-295. *In* M. Mathews, Sonenberg, N., Hershey, J. (ed.), Translational Control in Biology and Medicine. Cold Spring Harbor Laboratory Press, New York.

64. **Hilton, A., L. Mizzen, G. MacIntyre, S. Cheley, and R. Anderson.** 1986. Translational control in murine hepatitis virus infection. J Gen Virol **67 (Pt 5):**923-32.

65. **Hofmann, K., Stoffel, W.** 1993. TMbase - A database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler **374:**166.

66. **Hofmann, M. A., R. Y. Chang, S. Ku, and D. A. Brian.** 1993. Leader-mRNA junction sequences are unique for each subgenomic mRNA species in the bovine coronavirus and remain so throughout persistent infection. Virology **196:**163-71.

67. **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:**6911-21.

68. **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:**7567-78.

69. **Huang, J. C., S. L. Wright, and W. D. Shipley.** 1983. Isolation of coronavirus-like agent from horses suffering from acute equine diarrhoea syndrome. Vet Rec **113:**262-3.

70. **Huang, P., and M. M. 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:**5009-17.

71. **Huang, P., and M. M. Lai.** 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:**9110-6.

72. **Hughes, S. A., P. J. Bonilla, and S. R. Weiss.** 1995. Identification and analysis of MHV-A59 P28 cleavage site. Adv Exp Med Biol **380:**453-8.

73. **Hughes, S. A., P. J. Bonilla, and S. R. Weiss.** 1995. Identification of the murine coronavirus p28 cleavage site. J Virol **69:**809-13.

74. **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:**4933-42.

75. **Ivanov, K. A., T. Hertzig, M. Rozanov, S. Bayer, V. Thiel, A. E. Gorbalenya, and J. Ziebuhr.** 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. Proc Natl Acad Sci U S A **101:**12694-9.

76. **Jonassen, C. M., T. Kofstad, I. L. Larsen, A. Lovland, K. Handeland, A. Follestad, and A. Lillehaug.** 2005. Molecular identification and characterization of novel coronaviruses infecting graylag geese (Anser anser), feral pigeons (Columbia livia) and mallards (Anas platyrhynchos). J Gen Virol **86:**1597-607.

77. **Kahn, J. S.** 2006. The widening scope of coronaviruses. Curr Opin Pediatr **18:**42-7.

78. **Kamitani, W., K. Narayanan, C. Huang, K. Lokugamage, T. Ikegami, N. Ito, H. Kubo, and S. Makino.** 2006. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proc Natl Acad Sci U S A **103:**12885-90.

79. **Kim, C. S., S. K. Seol, O. K. Song, J. H. Park, and S. K. Jang.** 2007. An RNA-binding protein, hnRNP A1, and a scaffold protein, septin 6, facilitate hepatitis C virus replication. J Virol **81:**3852-65.

80. **Kolakofsky, D., and C. Weissmann.** 1971. Possible mechanism for transition of viral RNA from polysome to replication complex. Nat New Biol **231:**42-6.

81. **Kolakofsky, D., and C. Weissmann.** 1971. Q replicase as repressor of Q RNA-directed protein synthesis. Biochim Biophys Acta **246:**596-9.

82. **Kozak, M.** 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res **15:**8125-48.

83. **Kozak, M.** 1987. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. Mol Cell Biol **7:**3438-45.

84. **Kozak, M.** 2002. Pushing the limits of the scanning mechanism for initiation of translation. Gene **299:**1-34.

85. **Kozak, M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem **266:**19867-70.

86. **Krauss, S. W., J. R. Spence, S. Bahmanyar, A. I. M. Barth, M. M. Go, D. Czerwinski, and A. J. Meyer.** 2008. Downregulation of Protein 4.1R, a Mature Centriole Protein, Disrupts Centrosomes, Alters Cell Cycle Progression, and Perturbs Mitotic Spindles and Anaphase. Mol. Cell. Biol. **28:**2283-2294.

87. **Krishnan, R., R. Y. Chang, and D. A. Brian.** 1996. Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. Virology **218:**400-5.

88. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, and L. J. Anderson.** 2003. A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med **348:**1953-66.

89. **Lai, M. C., and K. V. Holmes.** 2001. Coronaviridae: The Viruses and Their Replication, p. 1163-1186. *In* D. M. Knipe, and P. M. Howley (ed.), Fields Virology, vol. 1. Lippincott Williams and Wilkins.

90. **Lai, M. M.** 1998. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. Virology **244:**1-12.

91. **Lapierre, J., G. Marsolais, P. Pilon, and J. P. Descoteaux.** 1980. Preliminary report on the observation of a coronavirus in the intestine of the laboratory rabbit. Can J Microbiol **26:**1204-8.

92. **Lapps, W., B. G. Hogue, and D. A. Brian.** 1987. Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. Virology **157:**47-57.

93. **Lau, S. K., P. C. Woo, K. S. Li, Y. Huang, H. W. Tsoi, B. H. Wong, S. S. Wong, S. Y. Leung, K. H. Chan, and K. Y. Yuen.** 2005. Severe acute respiratory syndrome

coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad Sci U S A **102:**14040-5.

94. **Law, A. H., D. C. Lee, B. K. Cheung, H. C. Yim, and A. S. Lau.** 2007. Role for nonstructural protein 1 of severe acute respiratory syndrome coronavirus in chemokine dysregulation. J Virol **81:**416-22.

95. **Letunic, I., R. R. Copley, B. Pils, S. Pinkert, J. Schultz, and P. Bork.** 2006. SMART 5: domains in the context of genomes and networks. Nucleic Acids Res **34:**D257-60.

96. **Li, H. P., P. Huang, S. Park, and M. M. Lai.** 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:**772-7.

97. **Li, H. P., X. Zhang, R. Duncan, L. Comai, and M. M. Lai.** 1997. Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription-regulatory region of mouse hepatitis virus RNA. Proc Natl Acad Sci U S A **94:**9544-9.

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

99. **Li, W., Z. Shi, M. Yu, W. Ren, C. Smith, J. H. Epstein, H. Wang, G. Crameri, Z. Hu, H. Zhang, J. Zhang, J. McEachern, H. Field, P. Daszak, B. T. Eaton, S. Zhang, and L. F. Wang.** 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science **310:**676-9.

100. **Liu, D. X., and S. C. Inglis.** 1992. Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. J Virol **66:**6143-54.

101. **Liu, D. X., H. Y. Xu, and K. P. Lim.** 1998. Regulation of mRNA 1 expression by the 5'-untranslated region (5'-UTR) of the coronavirus infectious bronchitis virus (IBV). Adv Exp Med Biol **440:**303-11.

102. **Liu, P., L. Li, J. J. Millership, H. Kang, J. L. Leibowitz, and D. P. Giedroc.** 2007. A U-turn motif-containing stem-loop in the coronavirus 5' untranslated region plays a functional role in replication. Rna **13:**763-80.

103. **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:**12105-13.

104. **Macejak, D. G., and P. Sarnow.** 1991. Internal initiation of translation mediated by the 5[prime] leader of a cellular mRNA. **353:**90-94.

105. **Marra, M. A., S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattra, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girn, O. L. Griffith, S. R. Leach, M. Mayo, H. McDonald, S. B. Montgomery, P. K. Pandoh, A. S. Petrescu, A. G. Robertson, J. E. Schein, A. Siddiqui, D. E. Smailus, J. M. Stott, G. S. Yang, F. Plummer, A. Andonov, H. Artsob,**

N. Bastien, K. Bernard, T. F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Stroher, G. A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R. C. Brunham, M. Krajden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper. 2003. The Genome sequence of the SARS-associated coronavirus. Science **300:**1399-404.

106. **Masters, P. S.** 2006. The molecular biology of coronaviruses. Adv Virus Res **66:**193-292.

107. **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:**911-40.

108. **Matthes, N., J. R. Mesters, B. Coutard, B. Canard, E. J. Snijder, R. Moll, and R. Hilgenfeld.** 2006. The non-structural protein Nsp10 of mouse hepatitis virus binds zinc ions and nucleic acids. FEBS Lett **580:**4143-9.

109. **Mehta, A., C. R. Trotta, and S. W. Peltz.** 2006. Derepression of the Her-2 uORF is mediated by a novel post-transcriptional control mechanism in cancer cells. Genes Dev **20:**939-53.

110. **Mihindukulasuriya, K. A., G. Wu, J. St Leger, R. W. Nordhausen, and D. Wang.** 2008. Identification of a novel coronavirus from a beluga whale by using a panviral microarray. J Virol **82:**5084-8.

111. **Minskaia, E., T. Hertzig, A. E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, and J. Ziebuhr.** 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc Natl Acad Sci U S A **103:**5108-13.

112. **Mostl, K.** 1990. Coronaviridae, pathogenetic and clinical aspects: an update. Comp Immunol Microbiol Infect Dis **13:**169-80.

113. **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.

114. **Nanda, S. K., and J. L. Leibowitz.** 2001. Mitochondrial aconitase binds to the 3'-UTR of mouse hepatitis virus RNA. Adv Exp Med Biol **494:**603-8.

115. **Narayanan, K., C. Huang, K. Lokugamage, W. Kamitani, T. Ikegami, C. T. Tseng, and S. Makino.** 2008. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. J Virol **82:**4471-9.

116. **Nelson, G. W., and S. A. Stohlman.** 1993. Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein. J Gen Virol **74 (Pt 9):**1975-9.

117. **Nelson, G. W., S. A. Stohlman, and S. M. Tahara.** 2000. High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. J Gen Virol **81:**181-8.

118. **Nishimura, T., M. Saito, T. Takano, A. Nomoto, M. Kohara, and K. Tsukiyama-Kohara.** 2008. Comparative aspects on the role of polypyrimidine tract-binding protein in internal initiation of hepatitis C virus and picornavirus RNAs. Comp Immunol Microbiol Infect Dis **31:**435-48.

119. **Nuttall, P. A., and K. A. Harrap.** 1982. Isolation of a coronavirus during studies on puffinosis, a disease of the Manx shearwater (Puffinus puffinus). Arch Virol **73:**1-13.

120. **O'Connor, J. B., and D. A. Brian.** 2000. Downstream ribosomal entry for translation of coronavirus TGEV gene 3b. Virology **269:**172-82.

121. **Parker, J. C., S. S. Cross, and W. P. Rowe.** 1970. Rat coronavirus (RCV): a prevalent, naturally occurring pneumotropic virus of rats. Arch Gesamte Virusforsch **31:**293-302.

122. **Pedersen, N. C., J. Ward, and W. L. Mengeling.** 1978. Antigenic relationship of the feline infections peritonitis virus to coronaviruses of other species. Arch Virol **58:**45-53.

123. **Peiris, J. S., S. T. Lai, L. L. Poon, Y. Guan, L. Y. Yam, W. Lim, J. Nicholls, W. K. Yee, W. W. Yan, M. T. Cheung, V. C. Cheng, K. H. Chan, D. N. Tsang, R. W. Yung, T. K. Ng, and K. Y. Yuen.** 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet **361:**1319-25.

124. **Pestova, T. V., C. U. Hellen, and E. Wimmer.** 1991. Translation of poliovirus RNA: role of an essential cis-acting oligopyrimidine element within the 5' nontranslated region and involvement of a cellular 57-kilodalton protein. J Virol **65:**6194-204.

125. **Ponnusamy, R., R. Moll, T. Weimar, J. R. Mesters, and R. Hilgenfeld.** 2008. Variable Oligomerization Modes in Coronavirus Non-structural Protein 9. J Mol Biol.

126. **Poon, L. L., D. K. Chu, K. H. Chan, O. K. Wong, T. M. Ellis, Y. H. Leung, S. K. Lau, P. C. Woo, K. Y. Suen, K. Y. Yuen, Y. Guan, and J. S. Peiris.** 2005. Identification of a novel coronavirus in bats. J Virol **79:**2001-9.

127. **Raaben, M., M. J. Groot Koerkamp, P. J. Rottier, and C. A. de Haan.** 2007. Mouse hepatitis coronavirus replication induces host translational shutoff and mRNA decay, with concomitant formation of stress granules and processing bodies. Cell Microbiol **9:**2218-29.

128. **Raman, S.** 2003. Stem-Loops III and IV in the 5' Untranslated Region are cis-Acting Elements in Bovine Boronavirus DI RNA Replication. Ph.D. Dissertation. University of Tennessee, Knoxville.

129. **Raman, S., P. Bouma, G. D. Williams, and D. A. Brian.** 2003. Stem-loop III in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J Virol **77:**6720-30.

130. **Raman, S., and D. A. Brian.** 2005. Stem-loop IV in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. J Virol **79:**12434-46.

131. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4:**406-25.

132. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. "Molecular Cloning: A Laboratory Manual", 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, NY.

133. **Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting.** 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A **95:**5857-64.

134. **Senanayake, S. D., and D. A. Brian.** 1997. Bovine coronavirus I protein synthesis follows ribosomal scanning on the bicistronic N mRNA. Virus Res **48:**101-5.

135. **Senanayake, S. D., and D. A. Brian.** 1999. Translation from the 5' untranslated region (UTR) of mRNA 1 is repressed, but that from the 5' UTR of mRNA 7 is stimulated in coronavirus-infected cells. J Virol **73:**8003-9.

136. **Sethna, P. B., and D. A. Brian.** 1997. Coronavirus genomic and subgenomic minus-strand RNAs copartition in membrane-protected replication complexes. J Virol **71:**7744-9.

137. **Sethna, P. B., S. L. Hung, and D. A. Brian.** 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. Proc Natl Acad Sci U S A **86:**5626-30.

138. **Shi, S. T., P. Huang, H. P. Li, and M. M. Lai.** 2000. Heterogeneous nuclear ribonucleoprotein A1 regulates RNA synthesis of a cytoplasmic virus. Embo J **19:**4701-11.

139. **Shi, S. T., P. Huang, H. P. Li, and M. M. Lai.** 2001. Regulation of mouse hepatitis virus RNA synthesis by heterogeneous nuclear ribonucleoprotein A1. Adv Exp Med Biol **494:**429-36.

140. **Silvera, D., A. V. Gamarnik, and R. Andino.** 1999. The N-terminal K homology domain of the poly(rC)-binding protein is a major determinant for binding to the poliovirus 5'-untranslated region and acts as an inhibitor of viral translation. J Biol Chem **274:**38163-70.

141. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Rozanov, W. J. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol **331:**991-1004.

142. **Spaan, W. J. M., Brian, D., Cavanagh, D., de Groot, R.J., Enjuanes, L., Gorbalenya, A.E., Holmes, K.V., Masters, P., Rottier, P., Taguchi, F. and Talbot, P.** 2005. Coronaviridae. *In* C. Fauquet, Mayo, MA, Maniloff, J, Desselberger, U, and Ball, LA

(ed.), Virus Taxonomy, Classification and Nomenclature of Viruses, 8th ICTV Report of the International Committee on Taxonomy of Viruses. Elsevier/Academic Press, pp. 1259.

143. **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:**5053-65.

144. **Spagnolo, J. F., and B. G. Hogue.** 2001. Requirement of the poly(A) tail in coronavirus genome replication. Adv Exp Med Biol **494:**467-74.

145. **Stair, E. L., M. B. Rhodes, R. G. White, and C. A. Mebus.** 1972. Neonatal calf diarrhea: purification and electron microscopy of a coronavirus-like agent. Am J Vet Res **33:**1147-56.

146. **Stephen F. Altschul, J. C. W., E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schäffer, Yi-Kuo Yu,.** 2005. Protein database searches using compositionally adjusted substitution matrices. FEBS Journal **272:**5101-5109.

147. **Stohlman, S. A., R. S. Baric, G. N. Nelson, L. H. Soe, L. M. Welter, and R. J. Deans.** 1988. Specific interaction between coronavirus leader RNA and nucleocapsid protein. J Virol **62:**4288-95.

148. **Tahara, S. M., T. A. Dietlin, C. C. Bergmann, G. W. Nelson, S. Kyuwa, R. P. Anthony, and S. A. Stohlman.** 1994. Coronavirus translational regulation: leader affects mRNA efficiency. Virology **202:**621-30.

149. **Tahara, S. M., T. A. Dietlin, G. W. Nelson, S. A. Stohlman, and D. J. Manno.** 1998. Mouse hepatitis virus nucleocapsid protein as a translational effector of viral mRNAs. Adv Exp Med Biol **440:**313-8.

150. **Tanguay, R. L., and D. R. Gallie.** 1996. Translational efficiency is regulated by the length of the 3' untranslated region. Mol Cell Biol **16:**146-56.

151. **Thiel, V., and S. G. Siddell.** 1994. Internal ribosome entry in the coding region of murine hepatitis virus mRNA 5. J Gen Virol **75 (Pt 11):**3041-6.

152. **Thomson, A. M., J. T. Rogers, C. E. Walker, J. M. Staton, and P. J. Leedman.** 1999. Optimized RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions. Biotechniques **27:**1032-9, 1042.

153. **Tompkins, W. A., A. M. Watrach, J. D. Schmale, R. M. Schultz, and J. A. Harris.** 1974. Cultural and antigenic properties of newly established cell strains derived from adenocarcinomas of the human colon and rectum. J Natl Cancer Inst **52:**1101-10.

154. **Tumlin, J. T., Pomeroy, B.S., Lindorfer, R.K.** 1957. Bluecomb disease of Turkeys IV. Demonstration of a filterable agent. J Am Vet Med Assoc **130:**360-365.

155. **Tyrrell, D. A., and M. L. Bynoe.** 1965. Cultivation of a Novel Type of Common-Cold Virus in Organ Cultures. Br Med J **1:**1467-70.

156. **van der Hoek, L., K. Pyrc, and B. Berkhout.** 2006. Human coronavirus NL63, a new respiratory virus. FEMS Microbiol Rev **30:**760-73.

157. **van der Hoek, L., K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. Berkhout, K. C. Wolthers, P. M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, and B. Berkhout.** 2004. Identification of a new human coronavirus. Nat Med **10:**368-73.

158. **van der Meer, Y., E. J. Snijder, J. C. Dobbe, S. Schleich, M. R. Denison, W. J. Spaan, and J. K. Locker.** 1999. Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. J Virol **73:**7641-57.

159. **van Hemert, M. J., S. H. van den Worm, K. Knoops, A. M. Mommaas, A. E. Gorbalenya, and E. J. Snijder.** 2008. SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. PLoS Pathog **4:**e1000054.

160. **Van Wynsberghe, P. M., H. R. Chen, and P. Ahlquist.** 2007. Nodavirus RNA replication protein a induces membrane association of genomic RNA. J Virol **81:**4633-44.

161. **Vijaykrishna, D., G. J. Smith, J. X. Zhang, J. S. Peiris, H. Chen, and Y. Guan.** 2007. Evolutionary insights into the ecology of coronaviruses. J Virol **81:**4012-20.

162. **Wang, J. T., and S. C. Chang.** 2004. Severe acute respiratory syndrome. Curr Opin Infect Dis **17:**143-8.

163. **Ward, J. M.** 1970. Morphogenesis of a virus in cats with experimental feline infectious peritonitis. Virology **41:**191-4.

164. **Wathelet, M. G., M. Orr, M. B. Frieman, and R. S. Baric.** 2007. Severe acute respiratory syndrome coronavirus evades antiviral signaling: role of nsp1 and rational design of an attenuated strain. J Virol **81:**11620-33.

165. **Weiss, S. R., and S. Navas-Martin.** 2005. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol Mol Biol Rev **69:**635-64.

166. **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:**8349-55.

167. **Woo, P. C., S. K. Lau, C. M. Chu, K. H. Chan, H. W. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai, W. K. Luk, L. L. Poon, S. S. Wong, Y. Guan, J. S. Peiris, and K. Y. Yuen.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol **79:**884-95.

168. **Woo, P. C., S. K. Lau, K. S. Li, R. W. Poon, B. H. Wong, H. W. Tsoi, B. C. Yip, Y. Huang, K. H. Chan, and K. Y. Yuen.** 2006. Molecular diversity of coronaviruses in bats. Virology **351:**180-7.

169. **Woo, P. C., M. Wang, S. K. Lau, H. Xu, R. W. Poon, R. Guo, B. H. Wong, K. Gao, H. W. Tsoi, Y. Huang, K. S. Li, C. S. 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:**1574-85.

170. **Wu, H. Y., and D. A. Brian.** 2007. 5'-proximal hot spot for an inducible positive-to-negative-strand template switch by coronavirus RNA-dependent RNA polymerase. J Virol **81:**3206-15.

171. **Wu, H. Y., A. Ozdarendeli, and D. A. Brian.** 2006. Bovine coronavirus 5'-proximal genomic acceptor hotspot for discontinuous transcription is 65 nucleotides wide. J Virol **80:**2183-93.

172. **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:**128-38.

173. **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:**2016-23.

174. **Zhang, X. M., W. Herbst, K. G. Kousoulas, and J. Storz.** 1994. Biological and genetic characterization of a hemagglutinating coronavirus isolated from a diarrhoeic child. J Med Virol **44:**152-61.

175. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J Gen Virol **81:**853-79.

176. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res **31:**3406-15.

177. **Zust, R., L. Cervantes-Barragan, T. Kuri, G. Blakqori, F. Weber, B. Ludewig, and V. Thiel.** 2007. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. PLoS Pathog **3:**e109.

Appendix I: SLV-VI

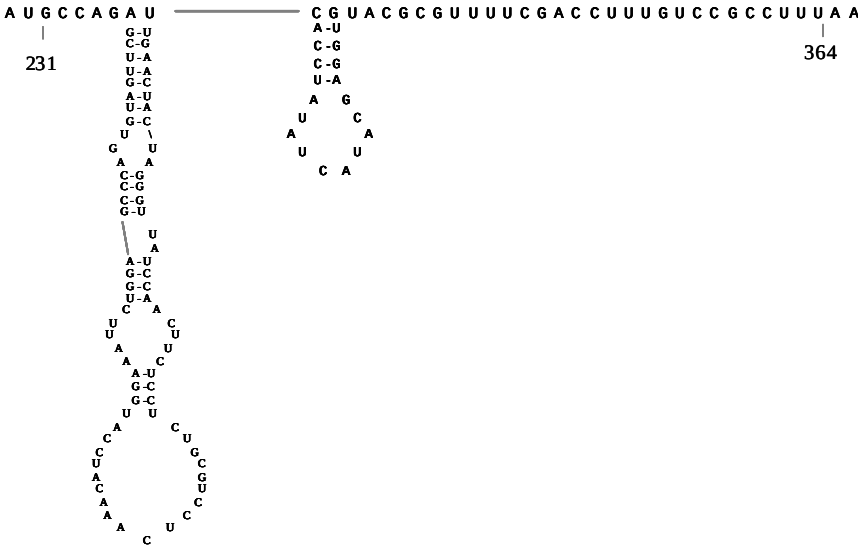
The Negative-Strand Counterparts of SLV and VI Bind Cellular Proteins and p28

Although it is not known if the negative strand counterparts to SLV and SL VI are required for replication, it is possible they are involved in RNA synthesis of the complementary strand. No viral proteins have been shown to bind negative strand *cis*-elements, but cellular proteins including PTB and hnRNP A1 have been shown to bind both positive and negative strand UTRs in MHV (70, 71, 96, 97). Although their exact function is not understood, they are thought to be involved in RNA synthesis of the complementary strand. Thus, cellular proteins that bind to the negative strand counterparts of these *cis*-acting replication elements could potentially be involved in RNA transcription. To determine if viral or cellular proteins bind to the negative strand regions of *cis*-acting SLV and SLVI gel shift assays were performed using BCoV infected and uninfected cellular lysates as described previously in Chapter II. The predicted secondary structure of the negative strand is shown in Figure A1-1A. To synthesize the negative strand of this region the previously described pSLV-VI was digested using the EcoR1 restriction site and transcribed using the SP6 promoter. Binding interactions were tested using gel shift assays.

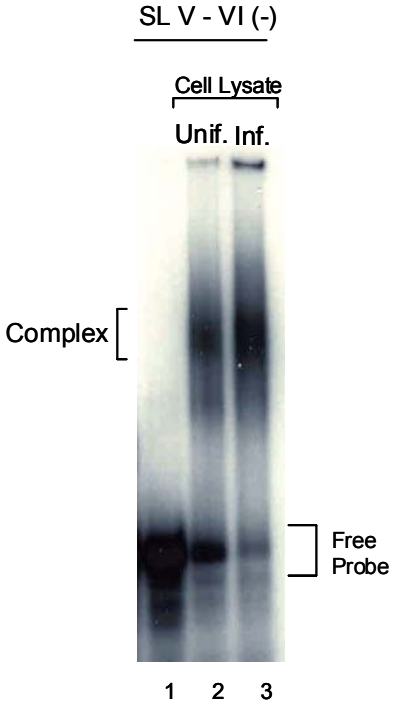
The Negative-Strand Counterparts of SLV and VI Bind Cellular Proteins

Gel shift identified one cellular complex that binds the complementary SLV and SLVI probe (Fig. A1-1B). The probe is completely bound when the infected lysate was used but there are no obvious indications that viral proteins are bound. This could be due to differences in the lysate quantitations or additional protein-protein interactions that do not bind directly to the RNA. In order to determine the number and approximate molecular weight(s) of the protein(s) in the complex, UV crosslinking was performed. UV crosslinking using the uninfected cell lysate revealed multiple proteins binding to the negative strand of SLV-VI (Fig. A1-1C). The same two cellular proteins of ~60kDa and 100kDa that bind the positive strand stem-loops also appear to bind the negative strand stem-loops. At least one other major protein of approximately 70kDa

A.



B.



C.

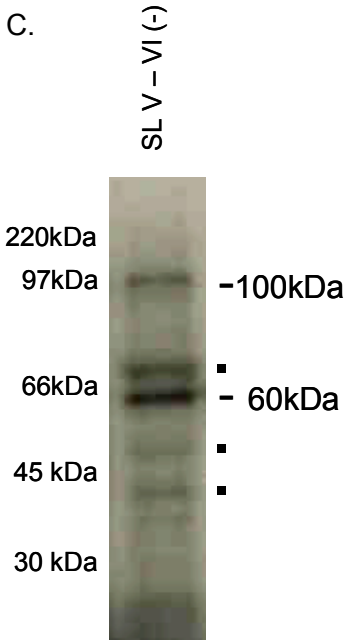


Figure A1-1. SLV-VI Negative-Strand. A. *Mfold* predicted secondary structure of SLV-VI(-). **B.** Gel shift assay indicates one cellular complex binds SLV-VI(-). C. UV crosslinking identifies five cellular proteins binding to SLV-VI(-). The ~60kDa and ~100kDa proteins also bind to the positive strand.

and several minor proteins (35-55kDa) also bind the negative strand. The identities of these proteins are currently unknown.

The Negative-Strand Counterparts of SLV and VI Bind p28

It is not known if the *cis*-acting requirement of stem-loop V and VI is in the positive or negative strand, and the negative strand region may have a role in virus replication such as RNA synthesis of the complementary strand.

The results show that p28 does bind the negative strand counterparts of these two stem-loops (Fig. A1-2A,left). The shifted band is decreased when a cold competitor RNA of the same sequence was added to the reaction, suggesting the interaction has both specificity and affinity for the SLV-VI (-) region. Although the significance of the negative strand stem-loops in this region is not known, there is support for the two stem-loops functioning together in the negative strand as well, as p28 does not bind the SL VI(-) or SLV(-) probes alone (Fig. A1-2A, right). In addition, the viral 2'-O-methyltransferase, nucleocapsid, and exonuclease do not bind the negative strand of this region (Fig. A1-2B).

Although the replication requirements for the negative strand counterparts of stem-loops V and VI are unknown, the results demonstrate this region binds both viral and cellular proteins, suggesting they may function in virus replication. As was shown for the positive strand, these predicted stem-loop structures in the negative strand may function as one element. In order to better understand the significance of these negative strand elements, more research is needed to determine the significance of p28 binding to this region as well as the identities of the cellular proteins.

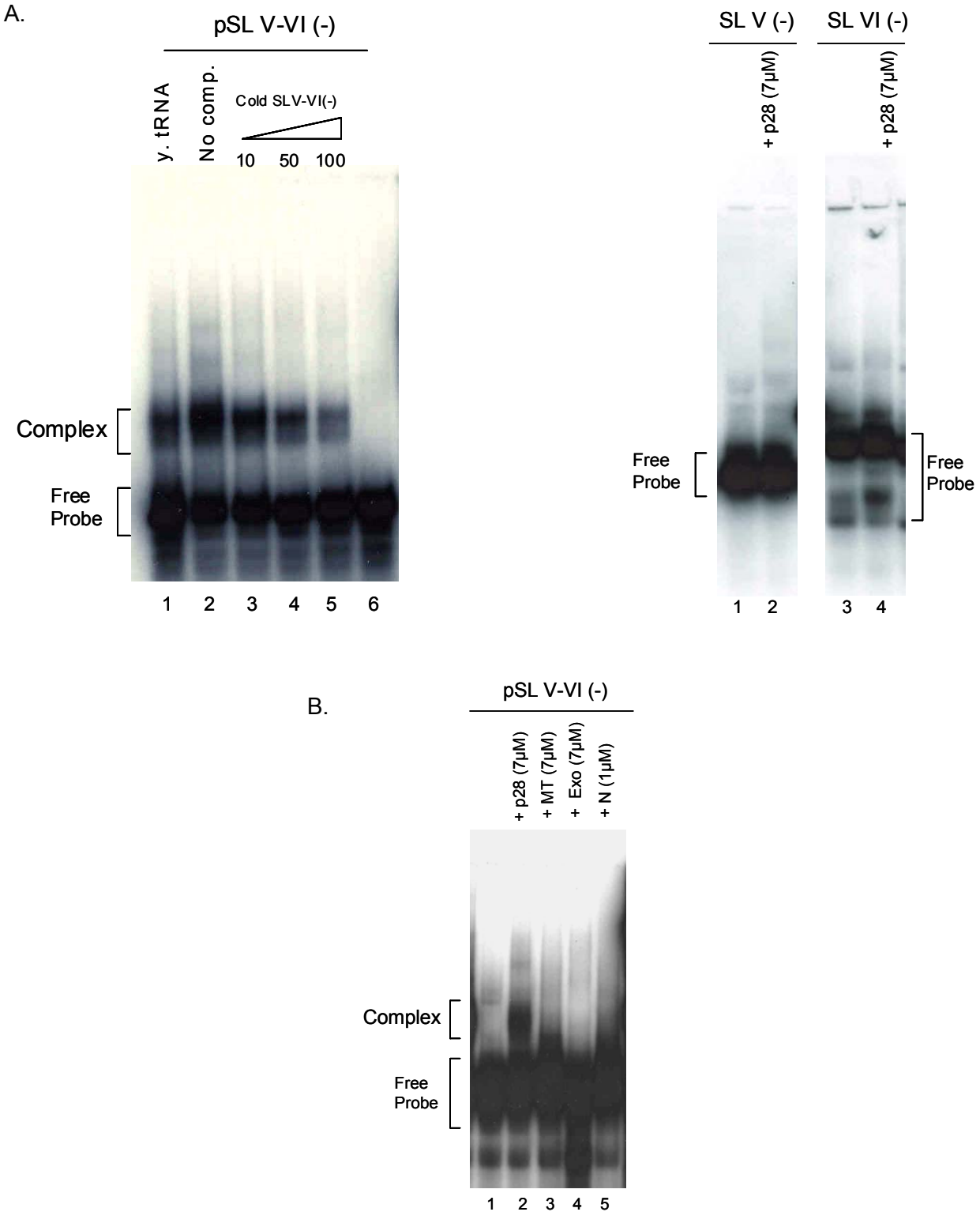


Figure A1-2. p28 Binds the Negative-Strand of SLV-VI. A. p28 specifically binds SLV-VI(-) (left), but not either stem-loop individually (right). B. The viral 2'-O-methyltransferase (MT), exonuclease (Exo), and nucleocapsid (N) proteins do not bind SLV-VI(-).

An ~60kDa Cellular Protein Binds SLVI

Three cellular complexes were found to bind SLV-VI by gel shift analysis as discussed in Chapter II. UV-crosslinking further resolved these complexes into two proteins of approximately 60kDa and 100kDa. To investigate protein complexes that bind each stem-loop independently, each was tested separately with BCoV infected and uninfected cellular lysates. In gel shift assays where probes representing stem-loop V (nts 231-316) and stem-loop VI (nts 309-343) individually were used, one cellular complex was found to bind each probe (Fig. A1-3A). UV crosslinking was used to determine how many factors comprised each complex and their approximate molecular weight. The results indicate that the ~60kDa protein found to bind the SLV-VI probe binds SLVI independently. The higher molecular weight protein that binds SL V and VI together, does not bind SLVI alone (Fig. A1-3B). It is not known if one of these proteins binds SLV only.

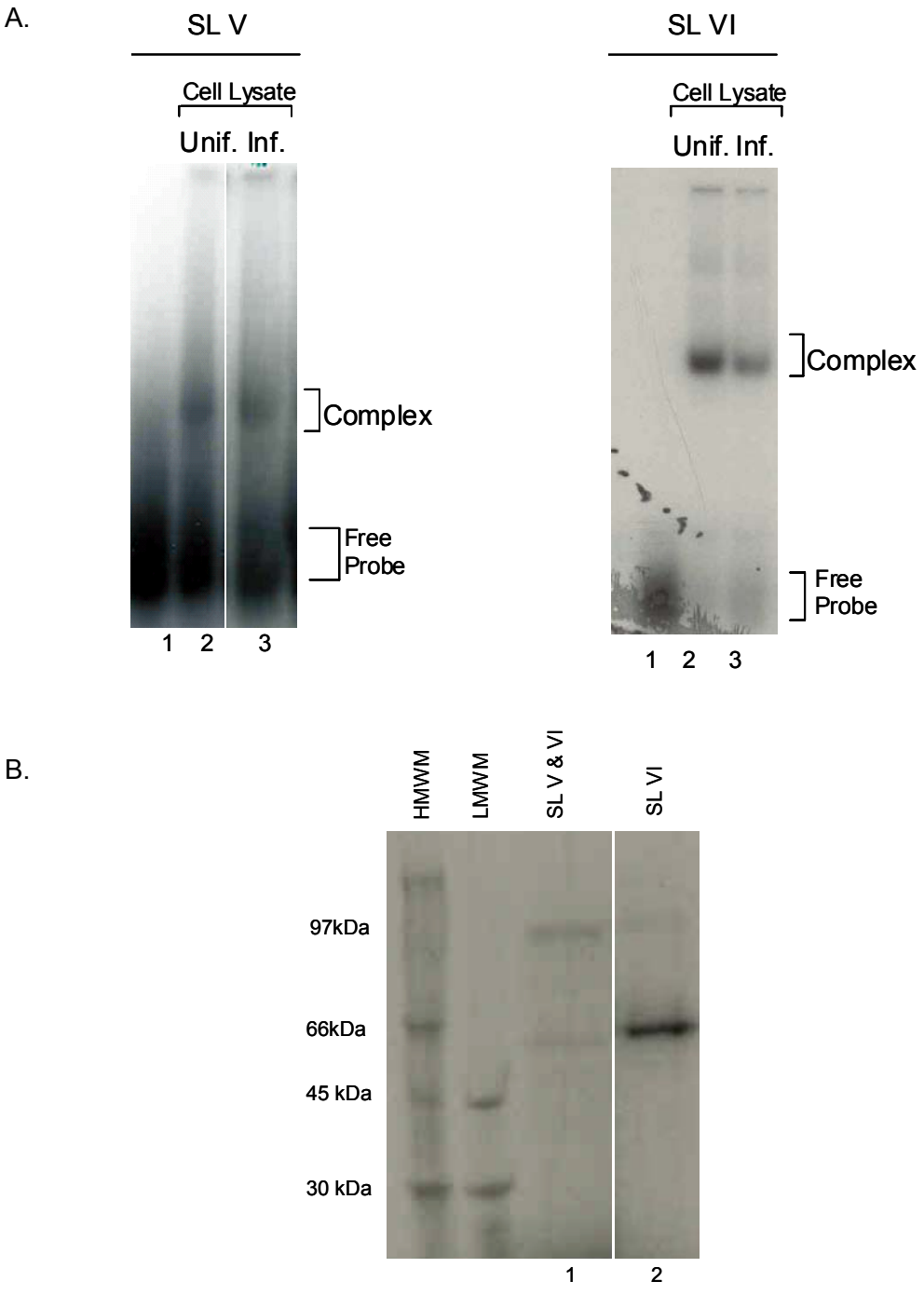


Figure A1-3. An ~60kDa Cellular Protein Binds SLVI. **A.** Gel shift assays show a single cellular complex binds SLV and SLVI individual probes. **B.** UV-crosslinking indicates SLVI binds only the ~60kDa protein.

p28 Nonspecifically Binds SLV-VI

As we were unable to detect viral proteins binding to the SLV-SLVI region with virus infected cellular lysates, we tested binding of *E.coli* expressed and purified recombinant viral proteins (Chapter II). The *cis* requirement for the p28 nucleotide sequence and separate requirement for the translation product made it tempting to speculate that p28 may bind the *cis*-acting stem-loops within its own coding region and function as a regulatory protein. Gel shift analysis was used to identify interactions between p28 and its own coding region. No binding of p28 to SLV-VI was found (Chapter II). To determine if p28 could interact with these stem-loops in a nonspecific manner, the experiments were repeated in the absence of competitor yeast tRNA (y.tRNA). The results indicate that even in the absence of the competitor RNA, p28 does not bind to stem-loop V or stem-loop VI (Fig. A1-4A). Interestingly, when the probe containing both stem-loop V and VI was tested in the absence of tRNA, some nonspecific, binding was observed (Fig. A1-4B), suggesting these two stem-loops which are close together may at least partially function together and form a protein binding site that can interact with p28. This nonspecific association may be the result of electrostatic interactions between positive residues in p28 and SLV-VI.

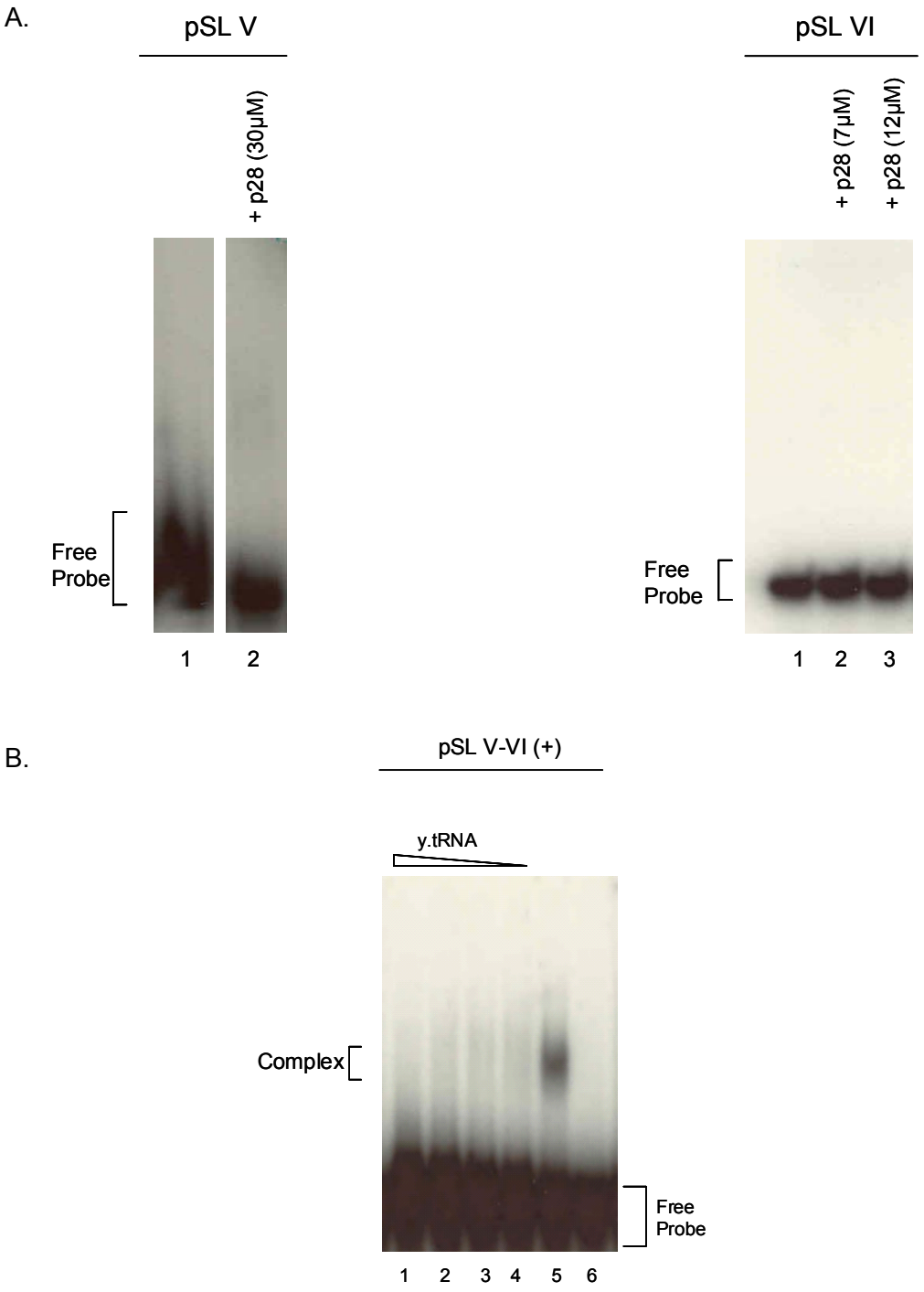


Figure A1-4. Nonspecific Binding of p28 to SLV-SLVI. A. Even in the absence of a nonspecific competitor, p28 does not bind to SLV or SLVI independently. **B.** Gel shift assay shows nonspecific binding of p28 to SLV-VI only in the absence of y.tRNA. The y.tRNA binds the RNA probe as the free probe is slightly shifted with increasing amounts of y.tRNA.

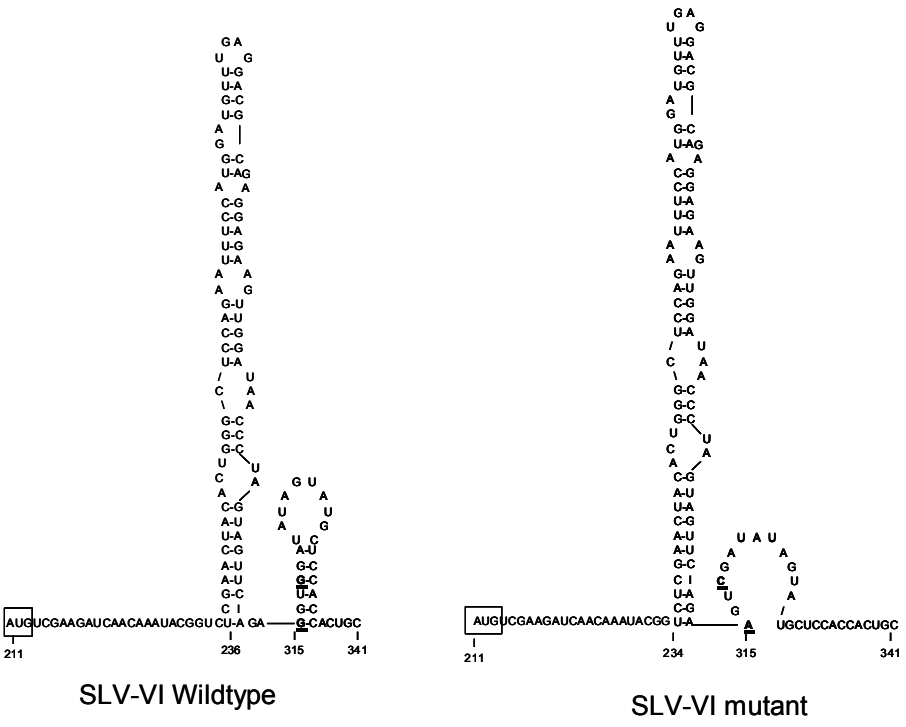
SLV-SLVI Mutant Binding Interactions

It is known that disruption of *cis*-acting stem-loop VI is detrimental to DI RNA replication (20). Mutation of the DI RNA to disrupt the base pairing of the stem of stem-loop VI at nucleotides 315 and 318 resulted in a loss of replication. Compensatory mutations that restored base pairing also restored replication, suggesting the structure of stem-loop VI is important for DI RNA replication (20). Therefore, a probe with the same mutations was used to determine if potential *trans*-acting interactions were modified. An inability of a protein to bind to the required stem-loop may explain the defect in replication. Gel shift assays were used to detect any differences in binding interactions between wildtype and mutant probes. The previously described SLV-VI probe (wildtype) was mutated to include base substitutions at nucleotide positions 315 and 318 in the stem of stem-loop VI (pSLV-SLVImut). The secondary structure of the mutant as predicted by *mfold* is shown in Figure A1-5A.

SLV-VI Mutant Binds Cellular Proteins

Incubation of the probe with the stem-loop VI mutant with cellular and virus infected lysates revealed only two cellular complexes compared to the three complexes in the wildtype (Fig. A1-5B; lanes 2 & 3 compared to lanes 5 & 6). UV crosslinking was used to determine how many factors comprised each complex and their approximate molecular weight. The results indicate that two cellular proteins (MW 60 & 100kDa) bind to the stem-loop V and VI region, including the stem-loop VI mutant (Fig. A1-5B). Although there is a difference in the number of complexes identified by gel shift analysis, it does not appear as though a unique protein-RNA interaction was disrupted with the mutant probe. It is possible that the three cellular complexes represent different molecular ratios of RNA to proteins and one of these was disrupted by the mutant. Alternatively, additional protein-protein interactions could exist between proteins that do not directly bind RNA that would alter the mobility of the complex but would not be identified by this assay.

A.



B.

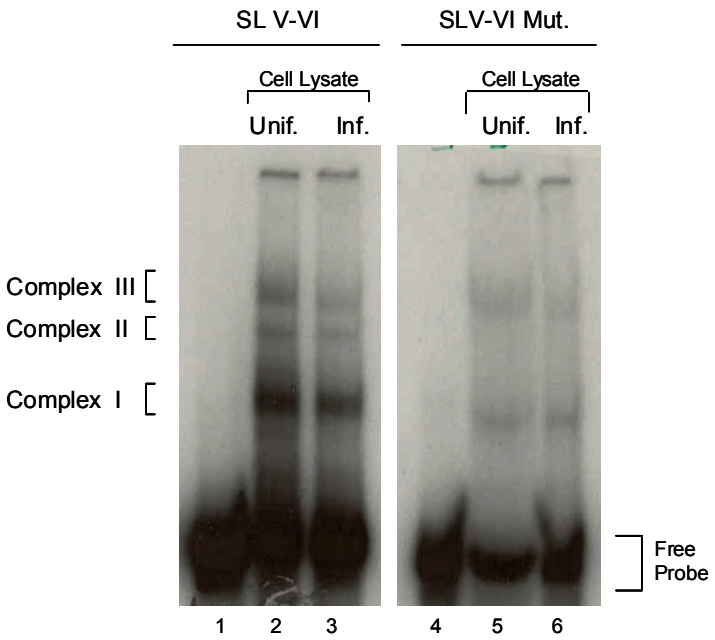


Figure A1-5. SLV-VI Mutant Binds Cellular Complexes. **A.** *Mfold* secondary structure predictions of wildtype and SLVI mutant stem-loops. **B.** Gel shift analysis reveals only two shifted cellular complexes in the SLV-VI mutant.

SLV-SLVI Mutant Binds p28

The results presented in Chapter II indicate that p28 does not bind to its own coding region with great affinity or specificity. Interestingly, when the probe containing both stem-loop V and VI was tested, some nonspecific, low affinity binding was observed (Fig. A1-4B,). To further examine the interaction between the two stem-loop structures and p28, pSLV-VImut. was used in gel shift assays. Incubation of the mutant probe with p28 showed similar results to what was observed with the wildtype probe. p28 binds the mutant probe weakly and nonspecifically (Fig. A1-6). Competition with increasing amounts of yeast tRNA outcompetes p28 binding. Collectively, it appears as though p28 may bind to the loop of SLVI and an unknown region of SLV, as the same binding pattern was observed with both wildtype and the mutated probe.

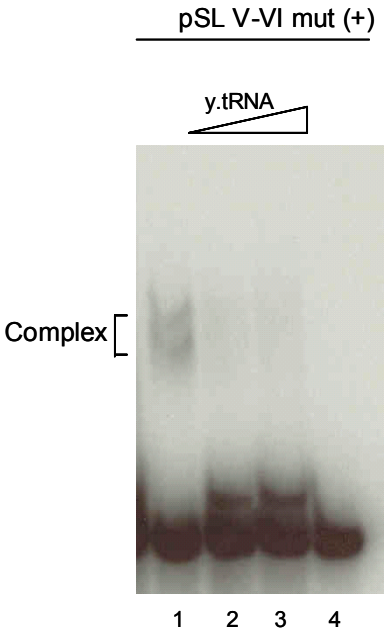


Figure A1-6. p28 Nonspecifically Binds SLV-VI Mutant. Although the mutations made to SLVI interfere with DI RNA replication, they do no interfere with nonspecific interactions with p28.

Electrophoretic Mobility Shift Assay (EMSA)/Gel Shift Assay

α^{32}P-UTP RNA probes were synthesized using linearized plasmid DNA and T7 RNA polymerase (Promega). Approximately 2.5µg of DNA was added to an in vitro transcription reaction following the manufacturer's protocol in a final volume of 50µl. Acetylated BSA (0.1µg/µl) was also included in the reactions. Transcription reactions were incubated at 37°C for one hour. RQ RNase free DNase was added (1U/µg DNA) and incubated an additional 30 minutes. 50µl of sequencing stop dye (80% Formamide, 1mM EDTA pH 8.0, 0.1% bromophenol blue, 0.1% xylene cyanol) was added and reactions were loaded onto 5-9% polyacrylamide/7%urea gel. Gels were exposed to Kodak XAR-5 film for 10minutes. Bands were excised from gel and eluted overnight with rotation at 4°C in elution buffer (0.5M ammonium acetate, 1mM EDTA). RNA was then ethanol precipitated and resuspended in water. Cerenkov counts were determined by scintillation counting.

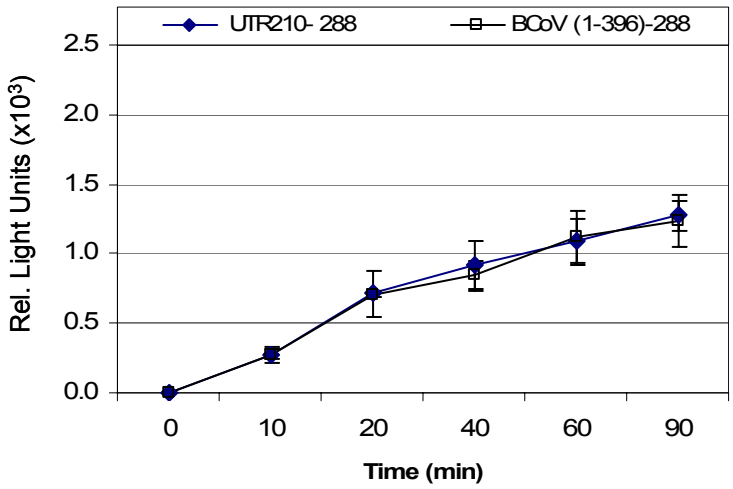
Gel shift assays were performed following the protocol in Silvera, et. al (140), with minor modifications. Recombinant *E.coli* expressed proteins, BCoV infected, or mock-infected cell lysates were incubated in binding buffer (5mM HEPES pH 7.9, 40mM KCl, 2mM MgCl2, 2mM DTT, 4% glycerol, 0.25% heparin, 20U RNasin (Promega), 1x EDTA-Free Complete protease inhibitors (Roche), 0.5ug/ul yeast tRNA) for 10 min. at 30°C. Approximately $4x10^4$cpm of each RNA probe was then added to the reaction and incubated an additional 10min. Following incubation, 2ul of 50% glycerol was added for gel loading and reactions were loaded onto a 5% nondenaturing polyacrylamide gel (prerun 100V, 20min) and run for approximately 4 hours. A Bio-Rad 16x16cm vertical electrophoresis system connected to a refrigerated circulating water bath (4°C) was used. Gels were dried using a Bio-Rad air dryer. Dried gels were exposed to Kodak XAR-5 film and results visualized by autoradiography.

Appendix II: Translation

cis-Acting Replication Elements Within the Nsp 1 Coding Region Do Not Affect Translation.

BCoV DI RNA replication has been shown to require the 5'-terminal portion of the coding region of the first nonstructural protein (nsp1) which includes two *cis*-acting stem-loop structures, SLV and SLVI. The exact roles these structures play in virus replication are unknown. Because the luciferase reporter constructs only include the UTR, it was possible an important translation regulatory element was missing. To determine if these stem-loops or nsp 1 sequence have a role in regulating translation, a construct with the required coding region sequence was included. This construct, **BCoV(1-396)-288** includes the entire 210nt 5' UTR and 186 nucleotides of nsp 1 sequence. The results show that the translation levels of **BCoV(1-396)-288** are the same as **UTR210-288**, indicating this *cis*-acting region did not impact translation (Fig. A2-1A). To eliminate the possibility that luciferase activity was altered due to the extra N-terminal amino acids, these results were confirmed with translation experiments using ^{35}S-Met *in vitro* translation. The resulting translation product is larger in size, but translation is equal to **UTR210-288** based on phosphorimager quantitation (Fig. A2-1B). These results are consistent with those presented in Figure 2-2C in which there was no change in translation of SLV mutant DI RNAs compared to wildtype.

A.



B.

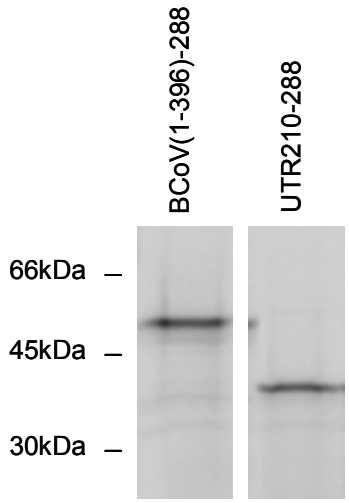


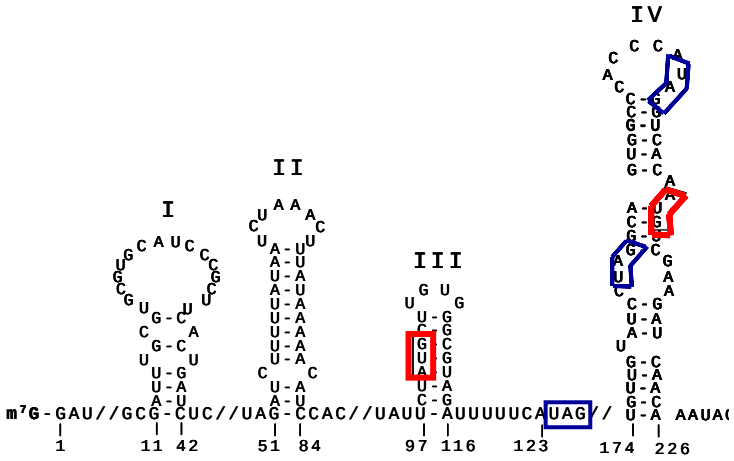
Figure A2-1. BCoV(1-396)-288 *in vitro* Translation. The partial p28 coding region present in the BCoV DI RNA does not appear to influence translation of an RLuc reporter construct. **A.** Translation levels of **BCoV(1-396)-288**, which includes *cis*-acting SLV-VI, are the same as **UTR210-288**. **B.** ^{35}S-Met. translation of **BCoV(1-396)-288** and **UTR210-288**.

uORF Mutants

As shown in Chapter III, Figure 3-4B, when the uAUG was mutated (**m.uAUG-288**), there is no change in translation and levels remain equal to that of the wildtype **UTR210-288**. This result was surprising since this uAUG is a highly conserved element and is found in all coronaviruses sequenced to date with only one exception (Brian, D.A., Gustin, K.M., unpublished data). To more fully understand the function or lack thereof of the uAUG, more mutants were made to test for reinitiation and usage of the uAUG.

Reinitiation of ribosomes after translation termination is believed to occur when there is sufficient distance between the uORF stop codon and the next start codon (31, 83). There must be appropriate timing for initiation factors to reassemble and reinitiate translation. A minimum of 50 nucleotides are needed for reinitiation to occur and approximately 80 are necessary for efficient reinitiation (31, 83). In BCoV, there would be sufficient distance for reinitiation, with over 80 nucleotides between the uORF stop codon (nt124) and the main ORF AUG at position 211. Two mutants were designed to test for reinitiation, **m.124STOP-288** and **m.124/184STOP-288** (Fig. A2-2). The uAUG is in-frame with the main ORF and contains three in-frame stop codons. Using the native sequence to our advantage, we were able to design mutants in which each stop codon was mutated to an amino acid codon, extending the uORF down to the next in-frame stop codon. Using this method we were able to extent the length of the uORF and decrease the distance to the 211 AUG, thus decreasing the chances for reinitiation. For **m.124STOP-288** the normal uORF stop codon (nt124) TAG was changed to TCG, increasing the length of the uORF to twenty-eight amino acids, and thus decreasing the distance from the stop codon (nt184) to the main ORF AUG at 211 to twenty-four nucleotides. Similarly, in **m.124/184 STOP-288** both the 124 and 184 stop codons were mutated (Fig. A2-2). Translation of this uORF would make a peptide of thirty-four amino acids and decrease the distance from the final stop codon at nucleotide 202 to the 211 AUG to only six nucleotides.

A.



AUG_____124UAG
AUG _____184UAG
AUG _____202UAG

Peptide	Nt. To AUG
8	84
28	24
34	6

B.

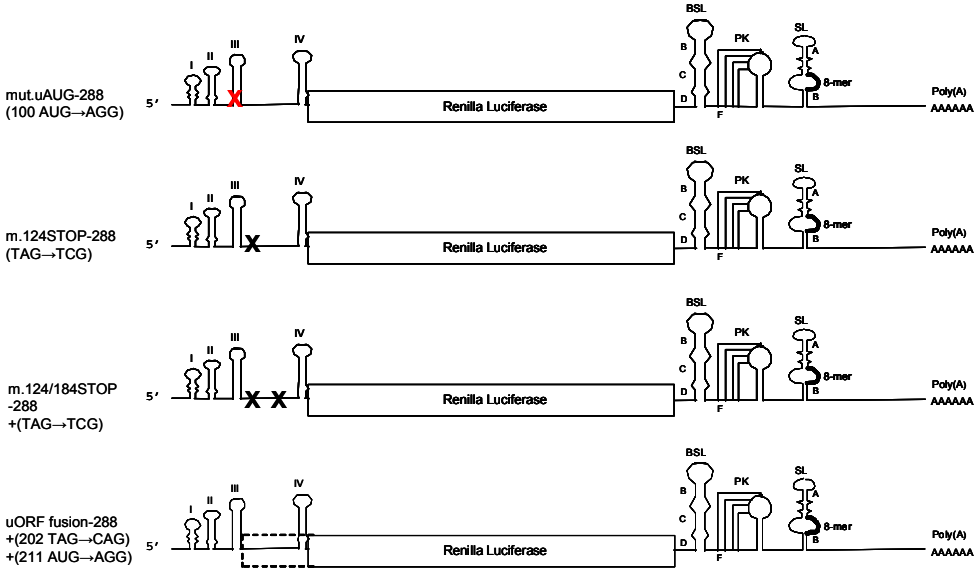
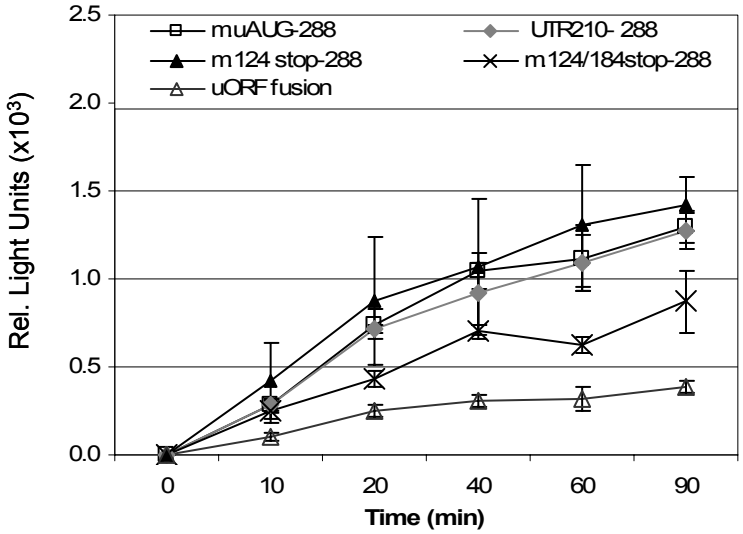


Figure A2-2. uORF Mutant Design. A. To test for uAUG usage and reinitiation, constructs were made by mutating in-frame stop codons, extending the length of the uORF and decreasing the distance between the termination codon and the 211AUG reinitiation site. **B.** Construct diagrams.

If the uAUG is used, translation of the main ORF should decrease because reinitiation should not be able to occur with only twenty-four and six nucleotides between the uORF stop and AUG (Fig. A2-2A). Consistent with the previous observation that mutation of the uAUG does not affect translation, **m.124STOP-288** translation levels are wildtype like indicating the uAUG is not being used and leaky scanning past the uAUG is most likely occurring. However, **m.124/184STOP-288** translation levels are significantly decreased at the 60 and 90 minute timepoints (Fig. A2-3A). In this case it is possible usage of the uAUG did attenuate translation, but it is also possible the mutations resulted in unintended changes that affected usage of the main AUG.

Figure A2-3B shows translation of these RNAs using ^{35}S-Met and SDS-PAGE. Based on phosphorimager quantitation, the major translation product in lanes 1, 2, 4, 5 are similar. Translation of **uORF fusion-288** resulted in one light upper band which is representative of initiation from the uAUG and a darker band which represents translation from an AUG within the RLuc coding region. **m.124STOP-288** and **m.124/184STOP-288** only have a single band, indicating translation from only the 211 AUG. From this experiment **m.124/184STOP** translation levels do not appear decreased as in the luciferase assay, indicating that perhaps the previously observed decrease in translation is not as significant. In addition, there are no lower bands in either **m.124STOP-288** or **m.124/184STOP-288** that would indicate reinitiation may have occurred at an AUG downstream within the coding region as with **uORF fusion-288**. Thus leaky scanning past the uAUG appears to occur most of the time.

A.



B.

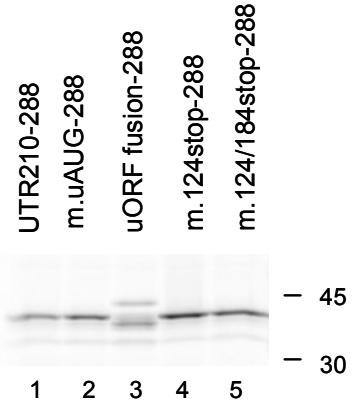


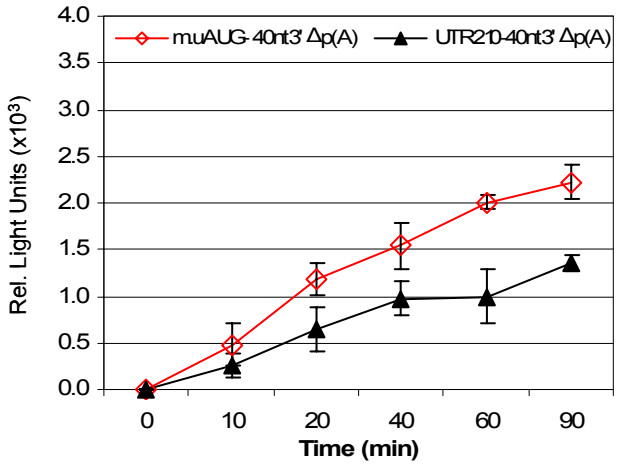
Figure A2-3. uORF Mutants. Results from *in vitro* translation of uAUG/uORF RLuc constructs. **A.** Translation levels of uORF mutant **m.124STOP-288** are similar to translation levels of **UTR210-288** and **m.uAUG-288**. **m.124/184STOP-288** translation levels are decreased compared to **m.124STOP-288**, **UTR210-288**, and **m.uAUG-288**. **B.** ^{35}S-Met. translation of uORF mutant constructs (90min.). Translation of **m.124STOP-288** and **m.124/184STOP-288** appear similar and initiation is from only the 211AUG.

The 5' UTR uAUG Attenuates Translation in the Absence of the Viral 288nt 3' UTR.

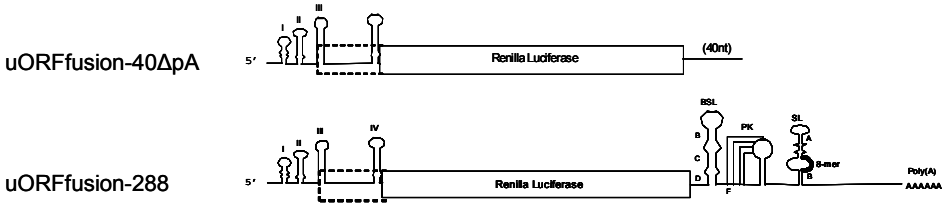
Translation of capped transcripts addressing the influence of a potential regulatory uAUG in the 5' UTR, (**m.uAUG-288**), did not show any apparent regulatory role and translation initiation from the uAUG occurred less than 25% of the time (**uORF fusion-288**) (Fig. 3-4C and 3-4D). However, translation of capped transcripts of **m.uAUG-40$_{\Delta p(A)}$** resulted in significantly increased translation over **UTR210-40$_{\Delta p(A)}$** at multiple timepoints (Fig. A2-4A). This result suggested that the uAUG in **UTR210-40$_{\Delta p(A)}$** was attenuating translation of the main ORF and that by removing the uAUG (**m.uAUG-40$_{\Delta p(A)}$**), translation was increased. Contradictory to the results shown in Figure 3-4, here the uAUG does appear to function in translation regulation of the main ORF, but only in the absence of the viral 3' UTR. In Chapter III, cap-dependent translation was shown to be repressed via a 5'-3' UTR mediated interaction. Therefore, we can hypothesize that in the absence of the viral 3' UTR, and thus the 5'-3' UTR interaction, the uAUG is used more and therefore regulates translation of the downstream ORF.

To more thoroughly examine the observation that the uAUG was regulating translation in the context of the non-viral 3' UTR, an additional construct was made. **uORF fusion-40$_{\Delta p(A)}$** was used to determine if usage of the uAUG was altered in the absence of the viral 3' UTR (Fig. A2-4B). For comparison, **uORF fusion-288**, which was shown to initiate translation from the uAUG ~20% of the time (Fig. 3-4), was used. Although this construct has a poly(A) tail, it has already been shown that the proposed 5'-3' UTR interaction is within the 5' UTR leader sequence and 3' UTR bulged stem-loop and adjacent pseudoknot region and is independent of the poly(A) tail. It is expected that translation levels of **uORF fusion-288** without a poly(A) tail would be significantly lower. Translation levels of **uORF fusion-40$_{\Delta p(A)}$** were significantly increased compared to **uORF-fusion-288** ($p=0.01$, 90min timept) (Fig. A2-4C). Although the increase was only 20% by the 90 minute timepoint, a definite change in translation is apparent at multiple timepoints, suggesting that in the absence of the viral 3' UTR, the uAUG is used

A.



B.



C.

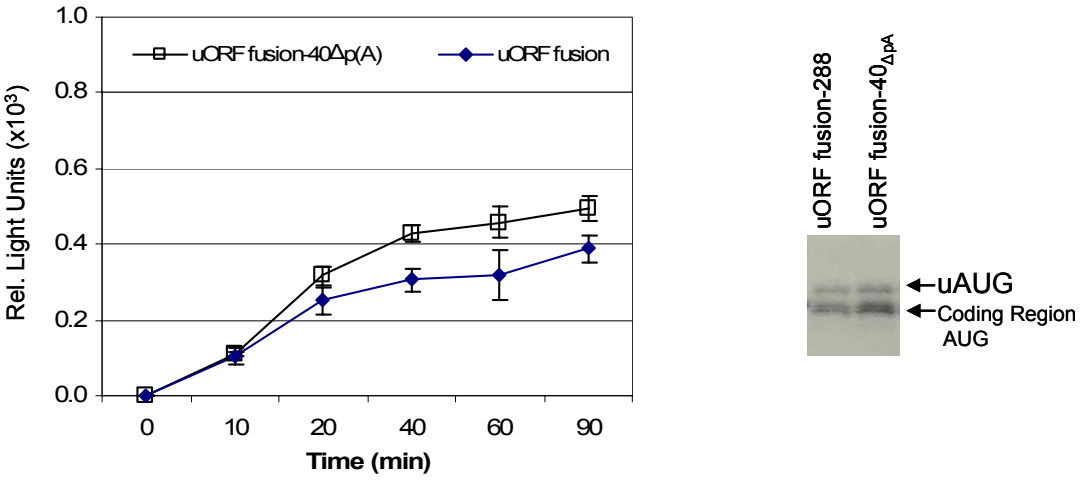


Figure A2-4. The 5' UTR uAUG Attenuates Translation in the Absence of the Viral 288nt 3' UTR. A. Translation levels of **m.uAUG-40$_{\Delta p(A)}$** are increased compared to **UTR210-40$_{\Delta p(A)}$**. **B.** Construct diagrams of **uORF fusion-40$_{\Delta p(A)}$** and **uORF fusion-288**. **C.** Left: Translation levels of **uORF fusion-40$_{\Delta p(A)}$** compared to **uORF fusion-288** (20% difference at 90min.). Right: ^{35}S-Met translation. Based on phosphorimager quantitation, translation from the uAUG in **uORF-40$_{\Delta p(A)}$** is 30% higher than **uORF fusion-288** (upper band). However, the ratio of translation products (upper band: lower band) indicate the uAUG is used ~23% of the time in both **uORF fusion-288** and **uORF-40$_{\Delta p(A)}$**.

more and therefore decreases translation of the main ORF. This data supports the hypothesis that in the absence of the viral 3' UTR, the uAUG is used more and therefore regulates translation of the downstream ORF.

As a secondary method of analysis, ^{35}S-Met translation of **uORF fusion-40$_{\Delta p(A)}$** was done to determine if the uAUG was being used more. Figure A2-4C shows ^{35}S-Met translation products at 90 minutes. There are two bands, a minor band representing translation from the uAUG and a major translation product that results from leaky scanning and initiation at a downstream coding region AUG. Quantitation of the upper bands, representing translation from the uAUG indicates that translation of **uORF fusion-40$_{\Delta p(A)}$** is increased 30% compared to **uORF fusion-288** (Fig. A2-4C). This indicates the uAUG in the context of the 40nt 3' UTR is being used 30% more than in the context of the viral 3' UTR. Therefore, in support of the hypothesis that the usage of the uAUG is affected, and that the uAUG is used more in the absence of the viral 3' UTR, ribosomes initiate translation at the uAUG in **UTR210-40$_{\Delta p(A)}$** ~30% more than in **UTR210-288**.

A comparison of the ratios of translation products from the uAUG and the downstream coding region AUG from **uORF fusion-40$_{\Delta p(A)}$** and **uORF fusion-288** indicate translation from the uAUG occurs ~23% of the time and leaky scanning accounts for over 70% of total translation in both cases (Fig. A2-4C). These results therefore suggest that the uAUG is not necessarily being used more in constructs with the 40nt 3' UTR, but perhaps that in the absence of the 3' UTR and thus the repressive 5'-3' UTR interaction, there are more initiation events overall. Subsequently, there are more ribosomes that initiate translation at the uAUG, resulting in an attenuation of translation of the main ORF. These results support a model for a 5'-3' UTR interaction and further suggest that the 5'-3' UTR interaction may repress cap-dependent translation through inefficient recruitment or binding of cap-binding initiation factors or ribosome subunits.

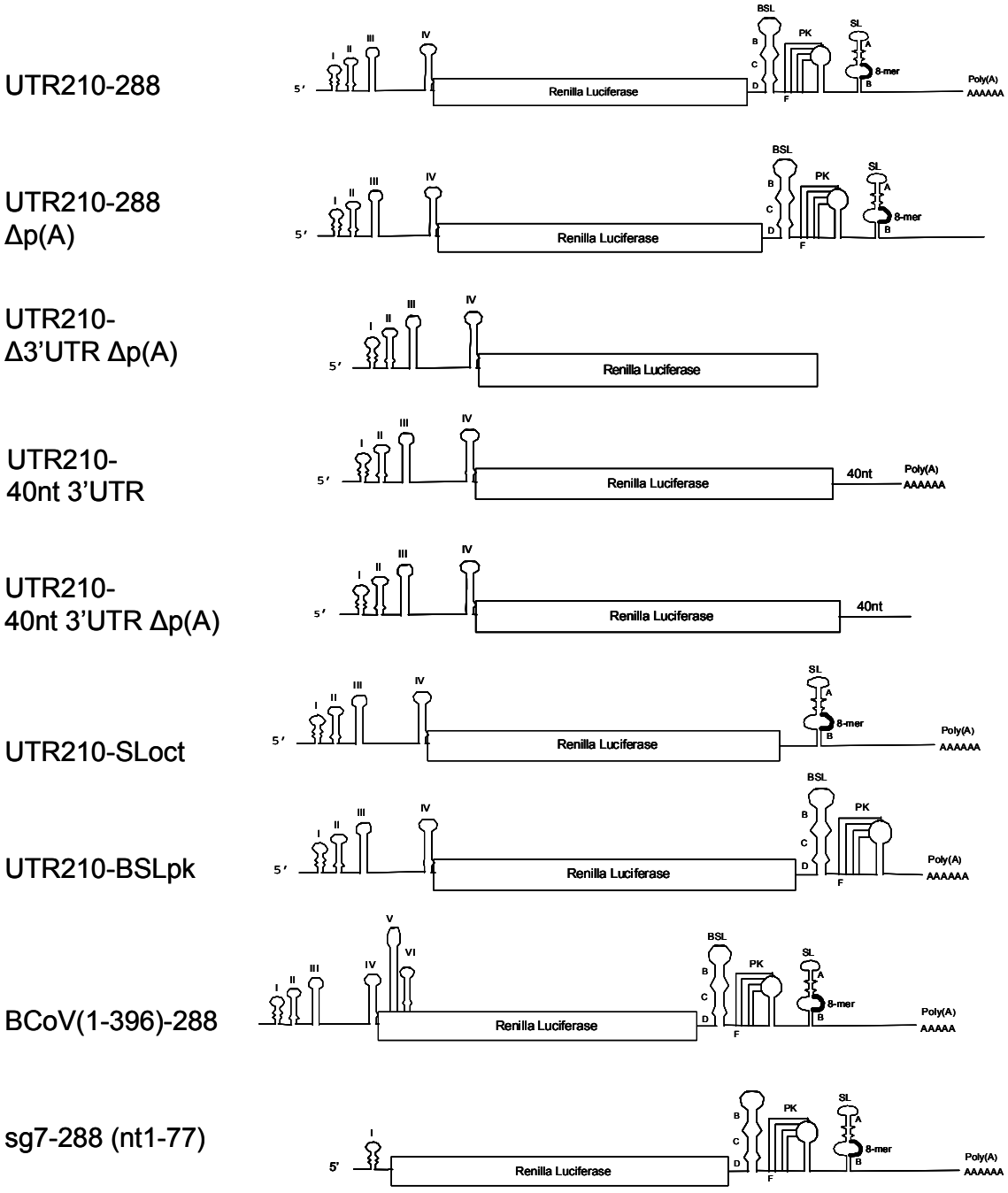


Figure A2-5. Renilla Luciferase Construct Diagrams. Construct names are given as 5' UTR sequence – 3' UTR sequence.

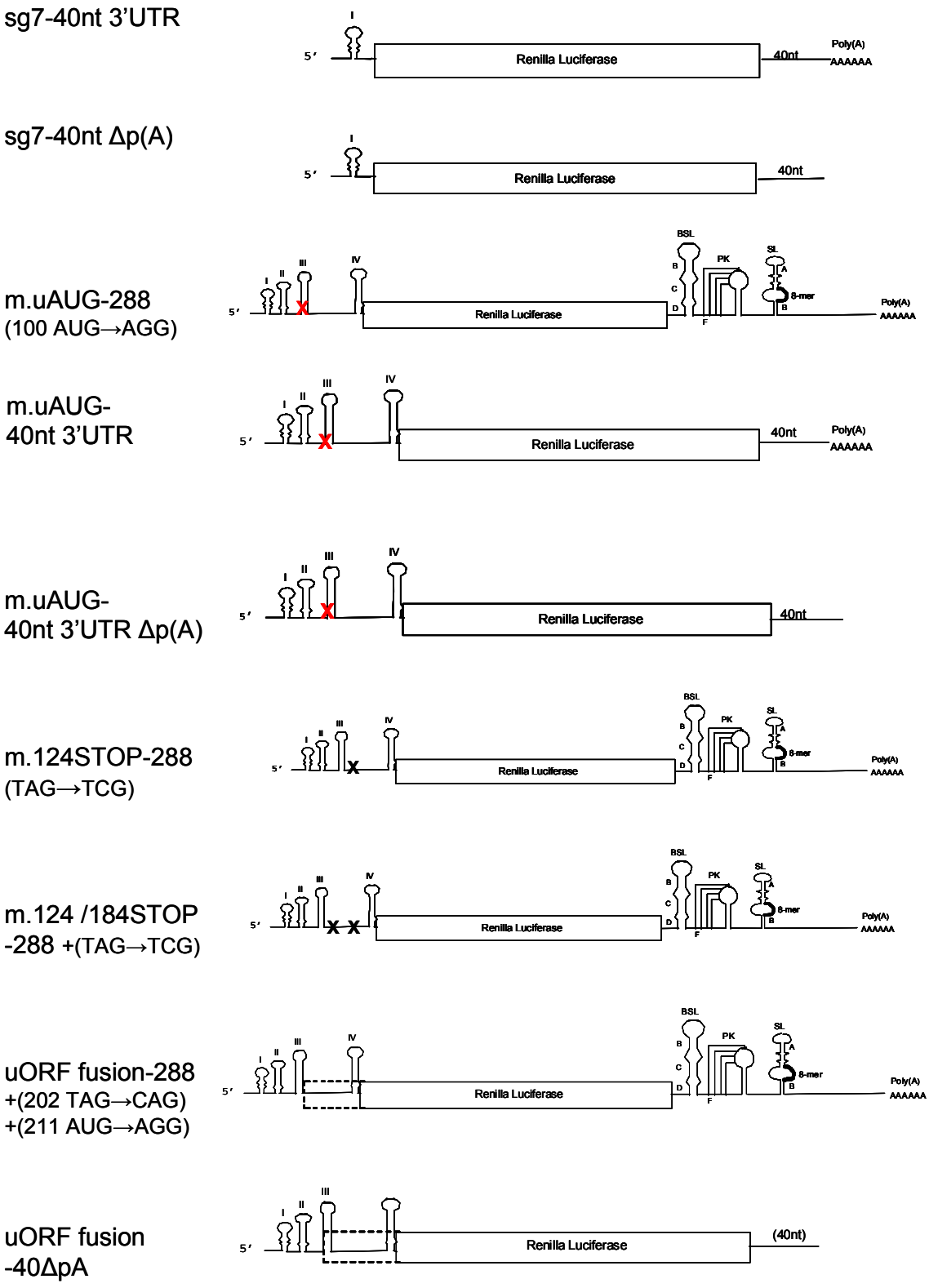
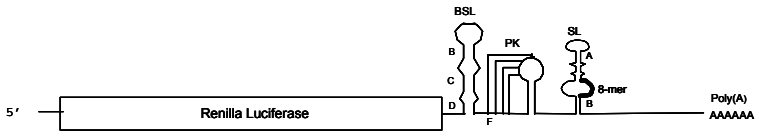
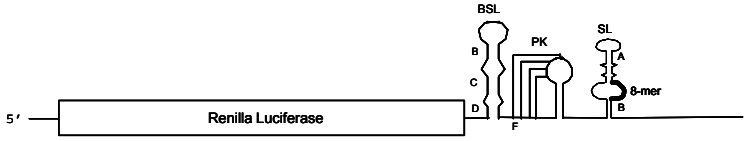


Figure A2-5 (continued). Renilla Luciferase Construct Diagrams.

5'OPT.-288
(20nt v.d.)



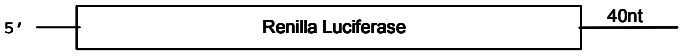
5'OPT.-288 Δp(A)



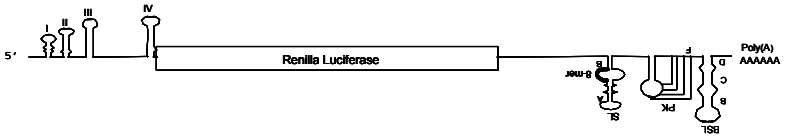
5'OPT. –Δ3'Δp(A)

5'OPT.-
40nt 3'UTR

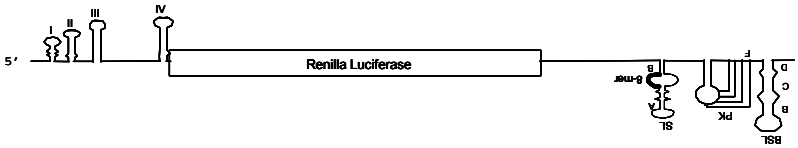
5'OPT.-40ntΔp(A)



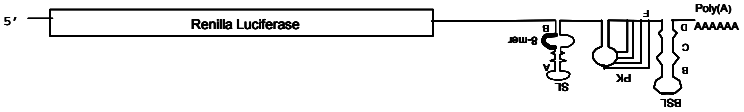
UTR210-FLIP



UTR210-FLIP
Δp(A)



5'OPT.-FLIP



5'OPT-FLIP Δp(A)

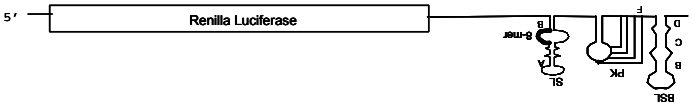


Figure A2-5 (continued). Renilla Luciferase Construct Diagrams.

Table A2-1. Renilla Luciferase Assay Results From *in vitro* Translation.

construct	BCoV UTR210		BCoV UTR210-Δp(A)		BCoV UTR210 Δ3'UTRΔp(A)		BCoV UTR210-40nt3'		BCoV UTR210-40nt3' Δp(A)		UTR210 FLIP		UTR210 FLIP-Δp(A)	
capped														
Time	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV
0	0.03	0.01	0.04	0.03	0.03	0.01	0.02	0.00	0.04	0.04	0.05	0.04	0.04	0.06
10	278.83	34.72	109.51	46.38	237.47	154.10	704.10	266.26	259.03	121.66	176.85	26.73	137.45	76.98
20	712.40	23.29	233.90	81.87	419.73	239.73	1423.33	526.28	648.13	243.66	544.03	190.94	381.82	140.01
40	915.90	174.49	310.57	83.32	458.50	201.74	2018.33	503.80	975.83	175.67	802.93	61.30	485.28	149.15
60	1093.47	161.07	351.00	78.23	551.67	254.67	2364.33	701.76	999.90	286.49	953.98	154.74	682.30	257.66
90	1272.00	103.01	427.33	28.35	515.83	163.94	2366.67	260.63	1351.33	86.03	988.73	120.94	711.72	192.95
uncapped														
0	0.04	0.02	0.03	0.01	0.04	0.04	0.02	0.00	0.06	0.02	0.14	0.05	0.06	0.03
10	235.27	99.76	110.03	41.20	121.26	81.48	359.45	116.18	158.27	80.65	154.57	51.88	180.95	60.89
20	561.83	235.16	258.40	66.97	262.53	118.29	731.60	169.31	401.53	54.28	460.93	72.40	475.58	63.45
40	848.93	162.82	331.10	140.76	389.65	93.51	937.73	334.87	523.07	72.28	620.70	112.63	515.45	55.95
60	1127.40	205.79	360.00	138.30	275.63	151.69	982.23	277.27	633.67	68.99	711.07	107.96	540.25	65.55
90	1263.67	127.42	428.60	122.09	414.90	78.16	1266.45	215.08	611.50	104.75	708.93	136.84	564.23	52.41

construct	5'OPT.		Δ5'UTR Δp(A)		Rluc 20 Δp(A)		Rluc 20 40nt3'		Rluc 20 40nt3'Δp(A)		RL20 FLIP		RL20 FLIP-Δp(A)	
capped														
Time	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV
0	0.03	0.00	0.05	0.03	0.03	0.03	0.04	0.02	0.06	0.01	0.07	0.06	0.02	0.02
10	300.30	171.33	186.03	71.73	163.97	25.59	268.37	81.70	146.91	74.46	148.05	22.98	385.80	211.64
20	751.67	256.03	305.40	58.19	327.80	82.85	671.70	169.62	365.30	172.07	454.58	82.25	645.53	38.79
40	1180.00	196.85	483.37	60.01	500.47	79.17	1090.17	121.97	688.77	237.65	653.80	73.50	780.47	147.90
60	1476.00	231.51	585.60	131.37	435.27	146.21	1173.67	85.15	868.87	15.73	818.10	76.71	829.10	245.14
90	1668.00	144.76	745.17	96.03	586.17	144.75	1276.00	139.53	1007.70	178.43	977.05	41.33	1022.07	100.03
uncapped														
0	0.04	0.02	0.03	0.00	0.05	0.06	0.06	0.01	0.03	0.03	0.06	0.05	0.03	0.03
10	153.33	21.85	80.41	34.15	182.85	141.96	151.60	41.30	146.97	16.70	57.35	18.57	86.75	57.62
20	384.17	40.22	157.42	54.91	295.00	172.18	352.25	57.91	393.53	49.20	178.30	60.82	194.09	89.25
40	576.83	61.02	247.23	86.77	388.33	233.66	544.50	34.37	502.33	85.01	269.47	113.08	262.50	107.15
60	817.37	47.08	281.67	93.79	434.47	202.07	610.40	104.93	582.37	100.36	310.27	140.07	309.37	115.33
90	1038.17	76.38	366.03	112.39	472.83	177.87	632.05	89.59	538.53	30.78	406.97	160.37	353.73	149.52

Table A2-1 (continued). Renilla Luciferase Assay Results From *in vitro* Translation.

construct	m.uAUG		m.uAUG-40nt3'		m.uAUG-40nt3' Δp(A)		m.124 stop		m.124/184stop		uORF fusion		uORF fusion-40Δp(A)	
capped														
Time	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV
0	0.06	0.01	0.03	0.01	0.03	0.02	0.04	0.04	0.04	0.03	0.03	0.03	0.05	0.05
10	279.13	32.89	381.93	103.49	478.10	224.34	420.03	218.92	249.37	63.96	105.16	23.25	110.10	5.50
20	743.37	82.51	1030.93	119.26	1187.00	170.24	876.53	361.92	432.67	45.24	250.80	34.15	316.00	22.27
40	1049.70	101.01	1444.33	51.83	1542.33	247.77	1068.93	386.03	700.17	35.89	305.88	31.60	428.10	23.39
60	1117.50	187.72	1781.67	139.09	2010.00	79.70	1301.33	347.93	621.50	44.16	319.15	65.57	457.83	42.64
90	1296.00	95.00	2218.00	300.13	2217.67	185.38	1425.00	153.16	869.77	174.35	388.28	34.48	493.87	34.90
uncapped														
0	0.03	0.01	0.02	0.00	0.04	0.01	0.04	0.01	0.04	0.02	0.04	0.03		
10	120.23	14.31	534.27	368.32	156.81	80.19	242.48	124.66	107.06	41.78	101.81	35.93		
20	419.90	57.51	1016.03	511.87	379.58	121.57	649.33	274.59	281.80	93.60	249.77	124.01	Not Tested	
40	885.97	167.15	1300.00	617.60	534.43	129.32	924.48	341.25	345.54	205.97	283.90	152.75		
60	1061.00	71.44	1206.33	271.77	607.40	96.55	870.75	143.58	459.80	94.81	319.43	193.72		
90	1337.33	54.17	1379.00	247.36	727.75	125.04	1050.48	167.82	586.33	156.87	369.03	136.14		

construct	sg7 Rluc		sg7 Rluc-40nt3' Δp(A)		sg7 Rluc-40nt3'		BCoV 1-396		UTR210-Oct.		UTR210-PK	
capped												
Time	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV	AVG	STDEV
0	0.05	0.03	0.02	0.01	0.03	0.00	0.02	0.00	0.23	0.08	0.17	0.07
10	169.46	66.01	247.13	66.43	346.00	70.48	272.63	63.43	155.60	28.96	81.42	14.25
20	483.44	90.69	425.13	37.53	833.07	118.35	706.67	164.47	463.07	55.98	276.93	19.20
40	786.08	160.28	588.63	112.51	1376.33	394.05	843.07	107.25	661.83	138.94	425.93	42.02
60	799.96	337.54	574.50	54.99	1819.33	803.69	1114.87	193.35	743.93	58.21	466.30	53.42
90	1105.86	169.81	747.80	84.49	1644.00	156.29	1236.67	184.80	815.40	87.60	542.07	86.20
uncapped												
0	0.04	0.01	0.02	0.00	0.03	0.01	0.03	0.01	0.16	0.08	0.09	0.01
10	200.00	24.45	76.24	25.25	347.50	57.69	84.57	14.74	93.29	15.76	99.41	22.14
20	483.73	140.55	160.10	45.44	472.61	376.69	210.80	36.03	341.87	71.15	323.17	27.42
40	839.70	44.26	183.40	25.03	751.80	149.82	319.90	27.38	481.50	23.19	469.67	17.50
60	1037.33	14.57	179.23	72.89	938.77	123.27	373.60	62.29	543.63	134.99	532.07	15.51
90	1200.73	187.11	222.40	48.51	1097.37	285.15	412.40	51.47	580.03	96.27	587.43	32.93

in vitro Translation Assay

in vitro Translation

Translation of RLuc RNAs were performed using rabbit reticulocyte lysate (RRL) (Promega), following the manufacturer's protocol. 100ng of RNA (quantitated by spectrophotometer and Ethidium Bromide formaldehyde gel electrophoresis) was added to reactions containing 40U of RNasin (Promega), 0.5µl each of 1mM amino acid mixtures (-)Met and (-)Cys and water to a total volume of 15µl. 35µl of RRL was added to each reaction. Reactions were incubated at 30°C for 90 minutes and 2.5µl aliquots were taken at the indicated timepoints, frozen in a dry ice/ethanol bath, and stored at -80°C. Independent translation reactions for each construct were performed a minimum of three times.

Luciferase Assay

50µl of Renilla luciferase assay reagent (Promega) was added to each 2.5µl timepoint aliquot and counted three times using a TD20/20 luminometer (luminometer settings: 2 second delay, 10 second integration). Results are presented as the averages of at least three independent experiments.

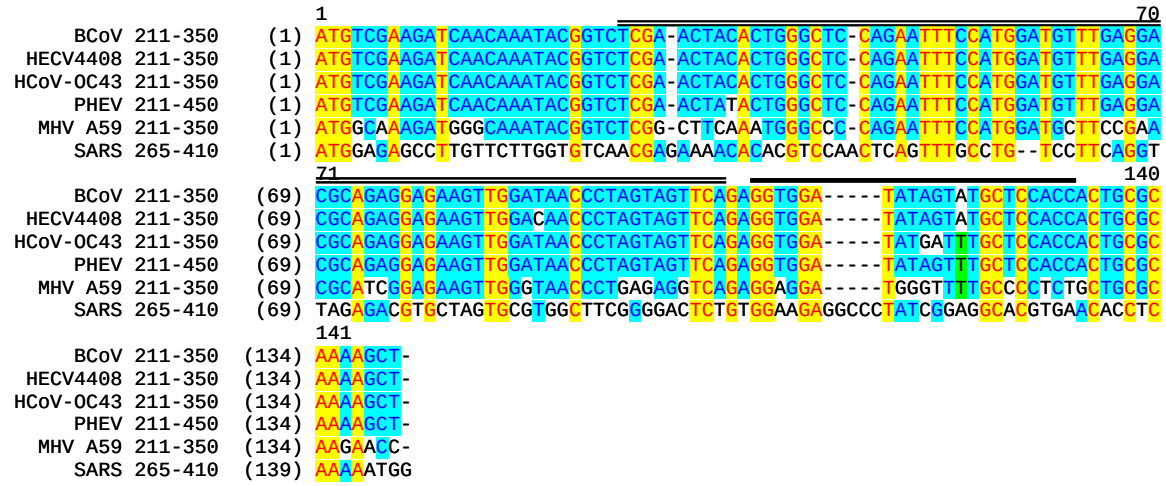
Radioactive *in vitro* translation

Translation reactions using ^{35}S-Met were performed using RRL (Promega), following the manufacturer's protocol. 500ng of RNA (quantitated by spectrophotometer and Ethidium Bromide formaldehyde gel electrophoresis) was added to reactions containing 40U of RNasin (Promega), 1.0µl of 1mM amino acid mixtures (-)Met, 2µl of ^{35}S-Met (20µCi) and water to a total volume of 15µl. 35µl of RRL was added to each reaction. Reactions were incubated at 30°C for 90 minutes and 5.0µl aliquots were taken after 90 minutes, frozen in a dry ice/ethanol bath, and stored at -80°C. SDS loading dye was added to each 5µl aliquot, heated at 100°C for 2 minutes and loaded onto a 12% Tris-glycine Ready gel (Biorad). Following SDS-PAGE, gels were fixed in 30% methanol/10% glacial acetic acid/ 3%glycerol for 1 hour, impregnated with

Enhance Fluor (Perkin Elmer) for 1 hour, precipitated in water for 30 minutes and then dried using a slab gel dryer. Results were visualized and quantitated using a phosphorimager.

Appendix III: Bioinformatics

A.



B.

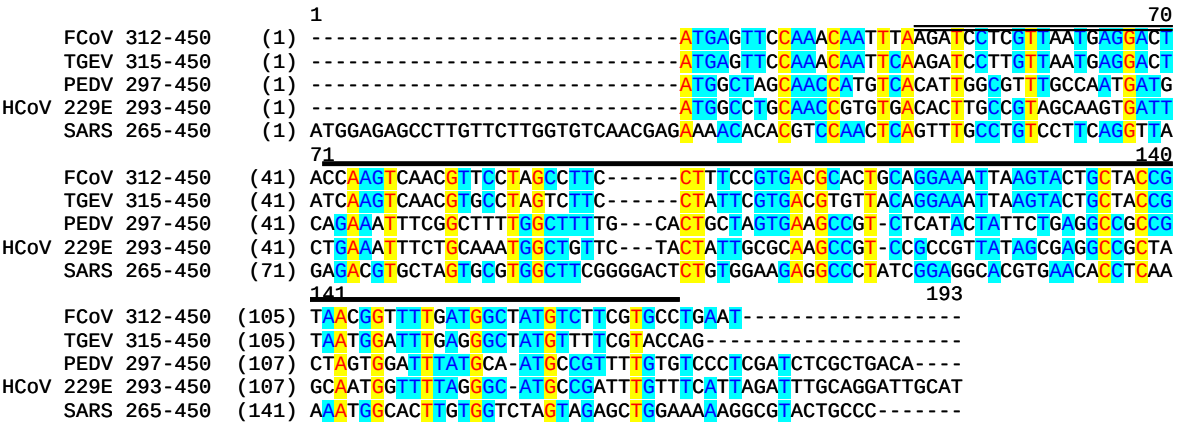


Figure A3-1. Nucleotide Sequence Alignment of Group 1 and Group 2 nsp 1 5' Terminal Region. A. Group 2 coronavirus sequence alignment. **B.** Group 1 coronavirus sequence alignment. SARS-CoV sequence is shown in both alignments. Approximate regions of secondary structures are indicated. Regions corresponding to SLV have double overline. Approximate regions corresponding to SLVI have bold overline.

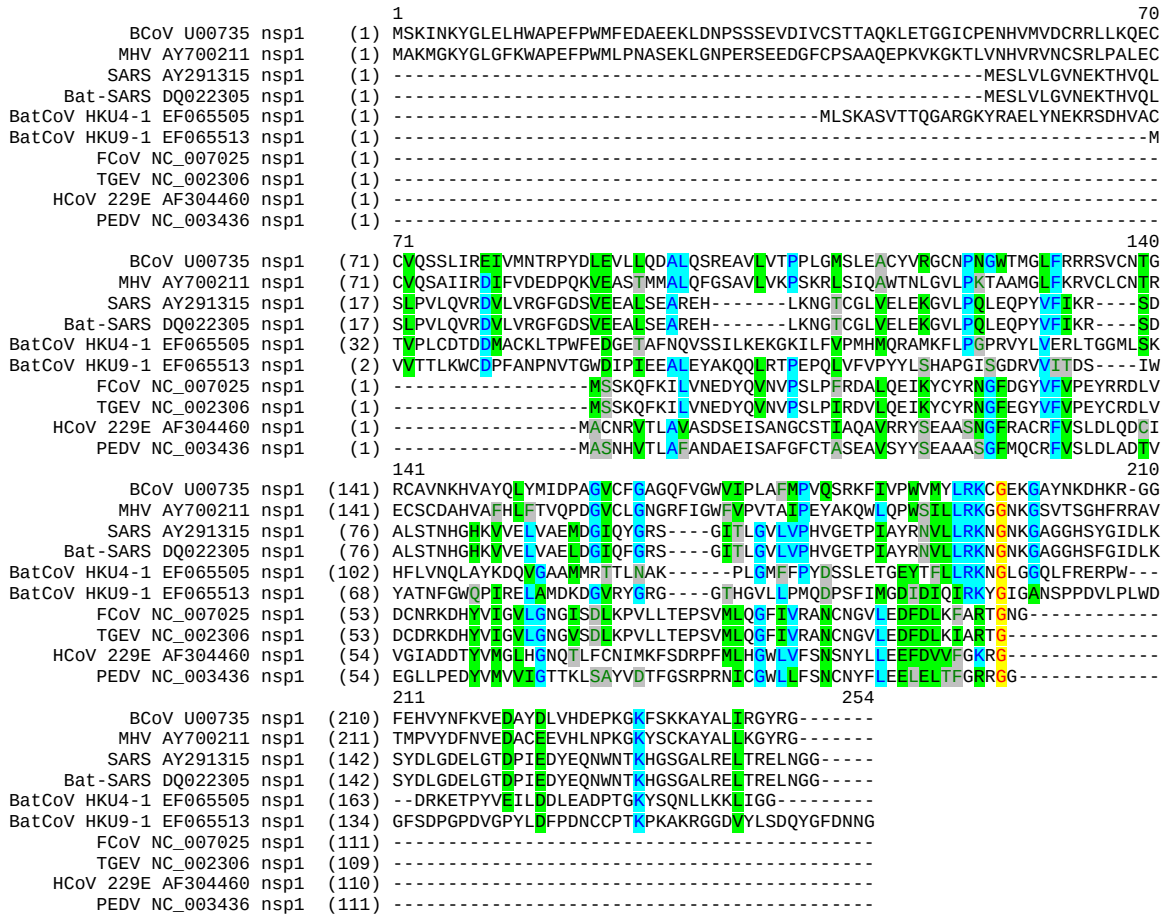


Figure A3-2. Group 1 and Group 2 nsp 1 Amino Acid Alignment.

BCoV nsp 1 Bioinformatics- PSI-BLAST Results

1. PSI-BLAST with BCoV p28 amino acids 37-76

The first non-CoV sequence was selected after psi-blast 2.

Physcomitrella patens subsp. *patens* (Genbank #EDQ51842): E-value: 5e-10

```
BCoV    37  DIVCSTTAQKLETGGICPENHVMVDCRRLLKQECCVQSSL  76
            D  C +  QK+     C   H   +C  L +Q+  +QS++
Sbjct  153  DTECPSPHQKMGQCMYCGGKHNTANCWYLPRQKTMIQSTV  192
```

Sequence used for SMART domain search

```
  1 MFVKESQPVT VAASLVRAKV WEECHYDQLL MVGTTMIPSI GNQVPMTNTM IESYPRMIAP
 61 SMGNPNSQGV VNATPLQAIP PVATYPSPMA YRQREVPMNA TINNSNEALL LNLTKKMEEL
121 AMHMTKDKEK RHKPSNMRPN VWCNNCKGQG HFDTECPSPH QKMGQCMYCG GKHNTANCWY
181 LPRQKTMIQS TVWDVNQVQT GNMGGWNGNR NNRVNRGFGN FQNPNYNQPY GTNMQNGITG
241 TLGGSDSRSI DKKSAKRQRI NSNIGEIDTS PGQQTNTKIC SIIGRSVTVT TQFDQYIGTS
301 INFTSTNITP IRWTIASVKG SNFQGAVTTK NRGSTNS
```

SMART Results

Name	Aa		*E-Value	Description
Pfam: Znf nanos	112	159	5.60e+00	Zinc Finger; RNA-binding
Pfam: DNAJ CXXCXGXG	111	186	9.70e+00	Zn dependent Cys rich domain
Pfam:zf-CXXC	135	166	9.00e+00	DNA binding zinc finger
**ZnF-C2HC	153	192	2.61e+02	Zinc finger; RNA/ssDNA binding
ZnF-AN1	156	183	9.28e+02	Zinc finger; protein binding
DNA Binding Motif	137	177	1.00e+03	DNA binding domain
eIF2B-5	55	172	3.25e+03	Contains putative C4 Zn finger

* E-value=score representing the probability that two sequences are matched by chance. A lower number indicates a lower probability.

** Domain identified using only amino acids 153-192

This hit identifies multiple types of Zinc fingers and DNA binding domains similar to those predicted to exist in BCoV.

2. PSI-BLAST with BCoV p28 amino acids 100-160

The first non- CoV sequence was selected after psi-blast 2.

Burkholderia pseudomallei (Genbank #ZP_02403227) E-value: 4e-10

```
BCoV  102  AVLVTPPLGMSLEACYVRGCNPNGWTMGLFRRRSVCNTGR  141
           A ++TPP    +   Y     P G   G   R + C +G+
Sbjct  61  ATVLTPPKRFDIGKWYNFAVMPGGKAEGPIYRGTFCISGK  100
```

Sequence used for SMART domain search

```
  1 MYSDERRRAS RNAARFHWRE PSHGASQSRC HLLASRSVGQ RSPVYLKRGD CIVYGQSIPG
 61 ATVLTPPKRF DIGKWYNFAV MPGGKAEGPI YRGTFCISGK QGDSGRVVLR SENKDACSSG
121 R
```

SMART Results

Name	Aa		*E-Value	Description
KAZAL type protease inhibitors	23	51	1.37e+03	Serine protease inhibitor

* E-value=score representing the probability that two sequences are matched by chance. A lower number indicates a lower probability.

This hit identifies a putative protein binding domain found in serine protease inhibitors. Although this particular domain was not predicted in BCoV, two different potential inhibitor domains were predicted.

3. PSI-BLAST with BCoV p28 amino acids 94-189

The first non- CoV sequence was selected after psi-blast 1

*Polaribacter dokdonensis (*Genbank #EAQ41748): E-value: 5e-21

```
BCoV  94    QDALQSREAVLVTPPLGMSLEACYVRGCNPNGWTMGLFRRRSVCNTGRCAVNKHVAYQLY  153
            QD  + + A ++     + L   +V   N   +    + R+ +       +N  V Y  Y
Sbjct  8    QDVKEGKLAAIICHFWIVGLVIGFVMNLNKKNYFTSFYLRQMIGMNLAQFLNGVVVYA-Y  66

BCoV  154   MIDPAGVCFGAGQFVGWVIPLAFMPVQSRKFIVPWV  189
            + D AG   G   FVGW+I L F  ++  + +VP+V
Sbjct  67    LGDTAGWIVGILLFVGWLISL-FGAIKGEEKLVPYV  101
```

Sequence used for SMART domain search

```
 1 MEVINNEQDV KEGKLAAIIC HFWIVGLVIG FVMNLNKKNY FTSFYLRQMI GMNLAQFLNG
61 VVVYAYLGDT AGWIVGILLF VGWLISLFGA IKGEEKLVPY VGEYFQDWFN QI
```

SMART Results

Name	Aa		*E-Value	Description
Transmembrane	72	91	-	Confidently predicted domain

* E-value=score representing the probability that two sequences are matched by chance. A lower number indicates a lower probability.

This hit confidently predicts a transmembrane domain within the region aligned with the sequence predicted to be a transmembrane domain in BCoV.

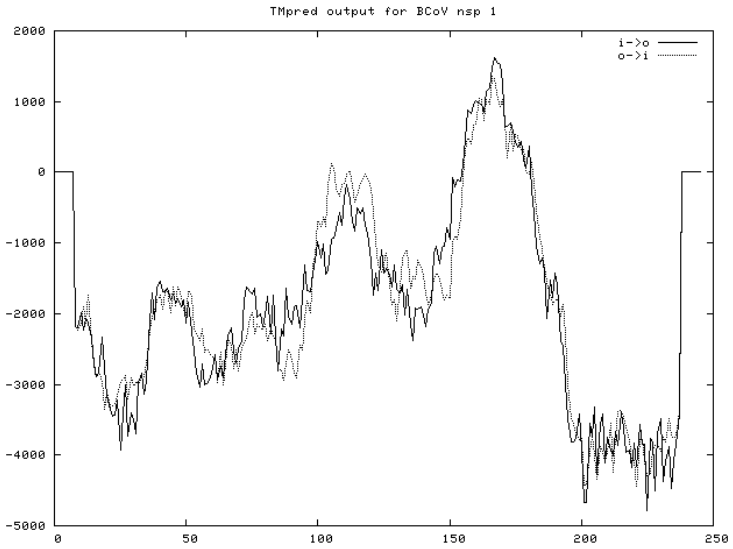
4. PSI-BLAST with BCoV p28 amino acids 90-246

No non-coronavirus sequences were identified.

TMpred Plot for BCoV and MHV p28 Transmembrane Domains

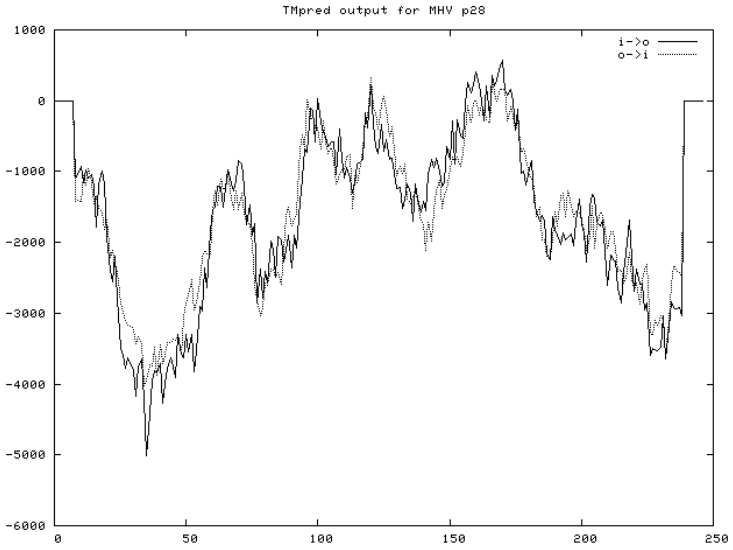
BCoV: STRONGLY preferred model: N-terminus inside

Amino Acid		Length	Score	Orientation
158	179	22	1625	i-o



MHV: STRONGLY preferred model: N-terminus inside

Amino Acid		Length	Score	Orientation
160	181	22	575	i-o



Vita

Kortney Gustin was born in Birmingham, Alabama. She received her Bachelor of Science Degree in Biology from the University of Montevallo in 2002. After working for approximately one year as a research technician at the University of Alabama at Birmingham School of Medicine, Department of Pediatrics, Division of Infectious Diseases, she entered graduate school at the University of Tennessee, Department of Microbiology, in 2003. She received her Doctorate Degree in Microbiology in 2008.