

To the Graduate Council:

I am submitting herewith a dissertation written by Ruey-Yi Chang entitled "Cis-Acting Sequence Requirements for Bovine Coronavirus Defective-Interfering RNA Replication." I have examined the final 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

David A. Brian, Major Professor

We have read this dissertation
and recommend its acceptance:

Bert Janssen
David Haehn

Grant Riggs
Beth C. Meltzer

Accepted for the council:

Lawrence Mink

Associate Vice Chancellor and
Dean of The Graduate School

***CIS*-ACTING SEQUENCE REQUIREMENTS FOR BOVINE
CORONAVIRUS DEFECTIVE-INTERFERING
RNA REPLICATION**

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

**Ruey-Yi Chang
December 1994**

**This dissertation is dedicated to my mother, Chiao-Mei Chiu
and my husband, Chipwu Cheng for their love and support
throughout.**

ACKNOWLEDGEMENTS

I wish to express the utmost appreciation to my major professor, Dr. David A. Brian, for his advice, guidance, support, encouragement, patience, and friendship for making this project an invaluable learning experience.

I also wish to thank my committee members: Dr. Beth Mullin, Dr. W. Stuart Riggsby, Dr. David Hacker, and Dr. Bert Lampson for their discussions, suggestions, and encouragement during the course of this study.

Finally, I wish to express my gratitude to my lab mates both past and present: Phiroze Sethna, Martin Hofmann, Manjiri Sethna, Susan Dombrowski, Rajesh Krishnan, Sylvie Rochat, Savithra Senanayake, Seulah Ku, Jennifer Black, and Gwyn Williams for their help, discussions, jokes and friendship. Thank all of them for making my foreign learning experience fun and the memories warm.

ABSTRACT

To study the *cis*-acting sequence requirements for coronavirus RNA replication, a 2.2 kilobase defective-interfering (DI) RNA of the bovine coronavirus was cloned, sequenced, and produced synthetically so that it could be analyzed as a replicon after transfection into helper virus-infected cells. This experimental approach was necessary since efforts have not yet proven successful for cloning the 30 kilobase single-stranded RNA coronavirus genome and synthesizing it as an infectious molecule. The DI RNA is comprised of the genomic 5' and 3' termini of 498 and 1635 nucleotides (nt), respectively. Within the DI RNA, part of the polymerase and all of the nucleocapsid protein genes are fused together as a single open reading frame.

Through a systematic mutagenesis analysis of the cloned reporter-containing DI RNA, named pDrep1, this study has established the following points regarding coronavirus RNA replication: (1) Transfected reporter containing mRNA, differing from pDrep1 by only 433 nt between the 5' terminal 65 nt leader and the start of the N open reading frame, does not undergo replication as hypothesized earlier by our laboratory. Therefore, coronavirus mRNA species probably do not amplify by a replication mechanism and at least part of the 5' signal for RNA replication resides in the 433 nt unique sequence. (2) The 65 nt terminal leader that is common to the genome, DI RNA, and all subgenomic mRNA species, although not sufficient as a 5' terminal *cis*-acting signal, is a necessary

part of the 5' terminal signal for RNA replication. This was demonstrated by terminal deletions of the leader for which it was shown that deletions of ≥ 13 nt, but not of ≤ 8 nt, were lethal for RNA replication. (3) When substitutions are made within the leader sequence, the 65 nt leader undergoes rapid reversion to the wild-type sequence during replication of pDrep1. This was demonstrated when a stem-loop of -8.0 kcal/mole between nucleotides 4 and 38, hypothesized to be an important part of the 5' cis-acting signal, was mutated to destroy the stem structure, and only wild-type leader was found in replicating molecules at 48 h posttransfection. The phenomenon of rapid leader reversion may reflect the mechanism of free leader synthesis and a requirement for leader priming of replication. That is, there may be two steps in RNA replication: (a) abortive leader synthesis, and (b) leader priming of replication. (4) Most deletion and insertion mutations made within the unique 433 nt 5' terminal region of pDrep1 are lethal to RNA replication. Thus, most of the 5' terminal 498 nt are probably part of the 5' terminal signal for RNA replication. (5) Except for the 3' terminal 11% of the open reading frame, there is a strict requirement for maintaining the open reading frame in the replication of pDrep1. This was demonstrated by the loss of replicating ability in out-of-frame mutations made at 7 sites throughout pDrep1. Differences in replication could not be explained by differences in mRNA stability since this was nearly the same for all constructs tested.

Thus, replication of coronavirus RNA, as determined by the study of DI RNA, requires an unusually complex behavior of the leader and a requirement for maintaining an open reading frame. With this study, the

coronavirus DI RNA has been established as a molecule with which to precisely determine many of the mechanisms involved in coronavirus plus and minus-strand RNA synthesis.

TABLE OF CONTENTS

CHAPTER	PAGE
I. LITERATURE REVIEW.....	1
BACKGROUND LEADING TO A STUDY OF BOVINE CORONAVIRUS DEFECTIVE-INTERFERING RNA REPLICATION.....	1
Introduction.....	1
Coronavirus genome structure and replication.....	2
Transcription of subgenomic RNAs.....	3
Defective-interfering (DI) RNA.....	8
Comparison between BCV and MHV.....	10
RATIONAL FOR THE STUDY OF BOVINE CORONAVIRUS DEFECTIVE-INTERFERING RNA.....	13
II. 5'-END CIS-ACTING SEQUENCE REQUIREMENTS FOR BOVINE CORONAVIRUS DEFECTIVE-INTERFERING RNA REPLICATION.....	16
INTRODUCTION.....	16
MATERIALS AND METHODS.....	18
Virus and cells.....	18
Synthetic oligonucleotides.....	18
Cloning and sequencing of the 5'-end of the BCV DI RNA.....	18
Construction of plasmids.....	21
<i>In vitro</i> mutagenesis of cloned reporter-containing DI RNA..	21
Assay for replication (Northern analysis).....	24
RNA stability test.....	26
Assay for leader reversion (RNase protection assay).....	26
Enzymes.....	27
Other methods.....	27
Nucleotide sequence accession number.....	27
RESULTS.....	28
The 5' <i>cis</i> -acting signal for RNA replication includes more than the leader.....	28
Deletions of the 5'-untranslated region downstream of the leader disrupt replication.....	34

Computer analysis of the genomic 5'-UTR predicts at least five stem-loops.....	36
5' terminal deletions of the leader disrupt a predicted intraleader stem-loop and also destroy replication ability..	38
Leader with stem-loop 1 disrupted by base substitutions is rapidly replaced by wild type leader.....	41
DISCUSSION.....	42
 III. TRANSLATIONAL REQUIREMENT IN <i>CIS</i> FOR BOVINE CORONAVIRUS DEFECTIVE-INTERFERING RNA REPLICATION.....	 54
INTRODUCTION.....	54
MATERIALS AND METHODS.....	57
Synthetic oligonucleotides.....	57
Construction of plasmids	57
Assay for replication (Northern analysis).....	61
<i>In vitro</i> translation.....	61
Measurement of mRNA stability.....	62
RESULTS.....	62
The single fusion protein within the BCV DI RNA is required for RNA replication.....	62
Out-of-frame mutations abolish replication while in-frame deletions or insertions do not.....	65
Truncating the ORF at various sites shows that almost the entire DI ORF is a <i>cis</i> -translationally required region for BCV DI RNA replication.....	68
Translational requirement in <i>cis</i> for RNA replication is not due to an increase in RNA stability conferred by translation.....	68
DISCUSSION.....	71

IV. FUTURE DIRECTION FOR USE OF THE CLONED DEFECTIVE- INTERFERING RNA IN THE ANALYSIS OF CORONAVIRUS RNA REPLICATION AND GENE EXPRESSION.....	78
TO DETERMINE THE SECONDARY STRUCTURE OF THE 5'- END OF THE BCV GENOME USING ENZYME PROBING...	78
TO LOCALIZE THE CROSS-OVER SITE DURING LEADER SWITCHING.....	79
TO CHARACTERIZE PROTEINS INVOLVED IN CORONAVIRUS RNA REPLICATION.....	80
TO IDENTIFY THE PACKAGING SIGNAL OF BCV.....	81
TO FURTHER EXAMINE THE <i>CIS</i> -TRANSLATIONAL REQUIREMENT FOR BCV DI RNA REPLICATION.....	82
TO EXPLORE THE POTENTIAL OF pDrep1 AS AN EXPRESSION VECTOR.....	82
REFERENCES.....	84
APPENDIX.....	93
DETAILED PROCEDURE FOR THE USE OF RNA BLOTTING IN THE ANALYSIS OF CORONAVIRUS RNA REPLICATION.....	94
DETAILED PROCEDURE FOR THE USE OF PCR SEQUENCING.....	104
LIST OF ABBREVIATIONS.....	107
VITA.....	108

LIST OF FIGURES

FIGURE	PAGE
1.1 Structure of BCV genomic and subgenomic RNAs.....	4
1.2 Leader-primed transcription model for generating subgenomic mRNAs in coronaviruses.....	5
1.3 Structure of BCV defective-interfering RNA.....	11
2.1 The aligned 5' ends of the bovine coronavirus DI RNA (which is the same as the BCV genome through base 498) and mouse hepatitis coronavirus genome.....	20
2.2 Constructs used for testing replication.....	22
2.3 RNA blotting evidence for replication and packaging of the reporter-containing BCV DI RNA (pDrep1).....	31
2.4 Stability of pDrep1 and pNrep2 RNA transcripts.....	33
2.5 Effects of 5' proximal internal mutations on the replication of pDrep1 RNA.....	35
2.6 Stem-loop structures within the BCV genomic 210 nt 5' UTR.....	37
2.7 Effects of 5' terminal leader deletions on the replication of pDrep1.....	39
2.8 Predicted stem-loop structures of the 5'-end in pDrep1, pDrep3, and the deletion mutants.....	40
2.9 Effects of substituted bases with stem-loop 1 on the replication pDrep1 RNA.....	43
2.10 Evidence of leader reversion for leaders with stem-loop 1 substitution mutations.....	44
2.11 Two models for leader reversion by the mechanism of template switching.....	48
3.1 Structure of BCV DI RNA and its fusion protein product.....	56
3.2 Effect of ORF mutations on the replication of pDrep1 RNA.....	64

3.3	SDS-polyacrylamide gel analysis of <i>in vitro</i> translation products of pDrep1 and its derivative mutants.....	67
3.4	Effect of premature stop codons on the replication of pDrep1 RNA.....	69
3.5	SDS-polyacrylamide gel analysis of <i>in vitro</i> translation products of pDrep1 truncated mutants.....	70
3.6	Stability of pDrep1, pDIN, and pT1i RNA transcripts.....	72
3.7	Model for the <i>cis</i> -acting function of the pDrep1 fusion protein.....	74

CHAPTER I. LITERATURE REVIEW

BACKGROUND LEADING TO A STUDY OF BOVINE CORONAVIRUS DEFECTIVE-INTERFERING RNA REPLICATION

Introduction.

Coronavirus virions are spherical, enveloped particles, ranging from 80 to 160 nm in diameter. They were named for their prominent surface projections that give them a "crowned" appearance, thus their Latin name, *corona* (Tyrrell, 1968). The *Coronaviridae* are one of the 15 families of animal RNA viruses (Cavanagh *et al.*, 1994). There are 16 members in the family (Sturman and Holmes, 1983) and many of them cause serious respiratory, gastrointestinal, and neurological diseases in domestic animal species of great economic importance that include cattle, pigs, chickens, and turkeys (Fenner *et al.*, 1993; Holmes, 1990). In humans, coronaviruses have been associated with common colds, diarrhea, and possibly multiple sclerosis (McIntosh, 1990; Murray *et al.*, 1992). It is now known that coronaviruses, in addition to having a unique structure, share a unique replication strategy among RNA virus families. However, the detailed mechanisms of coronavirus RNA replication are not yet clearly understood. The majority of studies on the molecular biology of coronavirus have been carried out on the mouse hepatitis virus (MHV), avian infectious bronchitis virus (IBV), transmissible gastroenteritis virus of swine (TGEV), bovine coronavirus (BCV), and to a lesser extent, the human coronavirus (HCV)

and feline infectious peritonitis virus (FIPV) (reviewed by Lai, 1990; Spaan *et al.*, 1988).

Coronavirus genome structure and replication.

Coronaviruses have a single-stranded, nonsegmented, positive-strand RNA (mRNA like) genome of approximately 30 kilobases (kb) in length that is polyadenylated at the 3' end and capped at the 5' end (Lai *et al.*, 1981; Lai, 1990). The genome is the largest known for any single-stranded animal RNA virus. Coronaviruses replicate in the cytoplasm of infected cells *via* a virus-encoded RNA-dependent RNA polymerase. The 5' two thirds of the genome (approximately 20 kilobases) is thought to encode the RNA-dependent RNA polymerase. The polymerase gene is translated as two polyproteins. The first is synthesized from ORF1a and undergoes extensive polyprotein processing apparently from the use of both viral and cellular proteases, but the details of this processing have not yet been completely determined. ORF1b is synthesized as a continuous polyprotein with ORF1a as a result of a -1 ribosomal frameshifting during translation of the polymerase gene (Brierley *et al.*, 1989). The -1 frameshift is considered to be an efficient one occurring in about 25-30% of the product, and is known to result from a pseudoknot in the RNA placed just downstream of the slippery sequence. A pseudoknot and frameshifting motif have been found both in the IBV and MHV genomes (Bredenbeek *et al.*, 1990) and is probably a mechanism used by all coronaviruses. The ORF1b product contains helicase and RNA polymerase motifs and may also undergo protease processing (Denison *et al.*, 1991).

During replication, a minus-strand RNA copy of the genome is produced after infection that is used as the template for genome replication and as the initial template for subgenomic mRNA synthesis. Replication of BCV RNA proceeds in a highly asymmetric manner, approximately 100 positive strands being produced for every negative strand (Hofmann *et al.*, 1990). Coronaviruses generate a 3' nested set of subgenomic mRNAs that are functionally monocistronic (only the 5' open reading frame of each message gets translated) but structurally polycistronic (except for the smallest and 3'-most mRNA). The structure of mRNAs relative to the genome is illustrated in Fig. 1.1 for BCV. Replication of BCV in the cytoplasm of the infected cells involves production of 8 species of subgenomic mRNA. The 3' one third of the BCV genome has so far been sequenced (Abraham *et al.*, 1990a; 1990b; Kienzle *et al.*, 1990; Lapps *et al.*, 1987).

Transcription of subgenomic RNAs.

The 5' terminal portion of the genomic and subgenomic messenger RNAs have a common sequence called "leader" of approximately 70 nt that is different in length depending on the coronavirus species. In BCV, the length of the leader is 65 nt. The leader sequence is only encoded at the very 5' end of the genome. It is not known how the leader is made and becomes attached to each mRNA. One popular hypothesis states that excess "free" leader is made and becomes fused to each mRNA by a leader-priming mechanism during transcription from a full-length minus-strand antigenome template (Fig. 1.2A) (Lai, 1990). That is, the leader sequence is

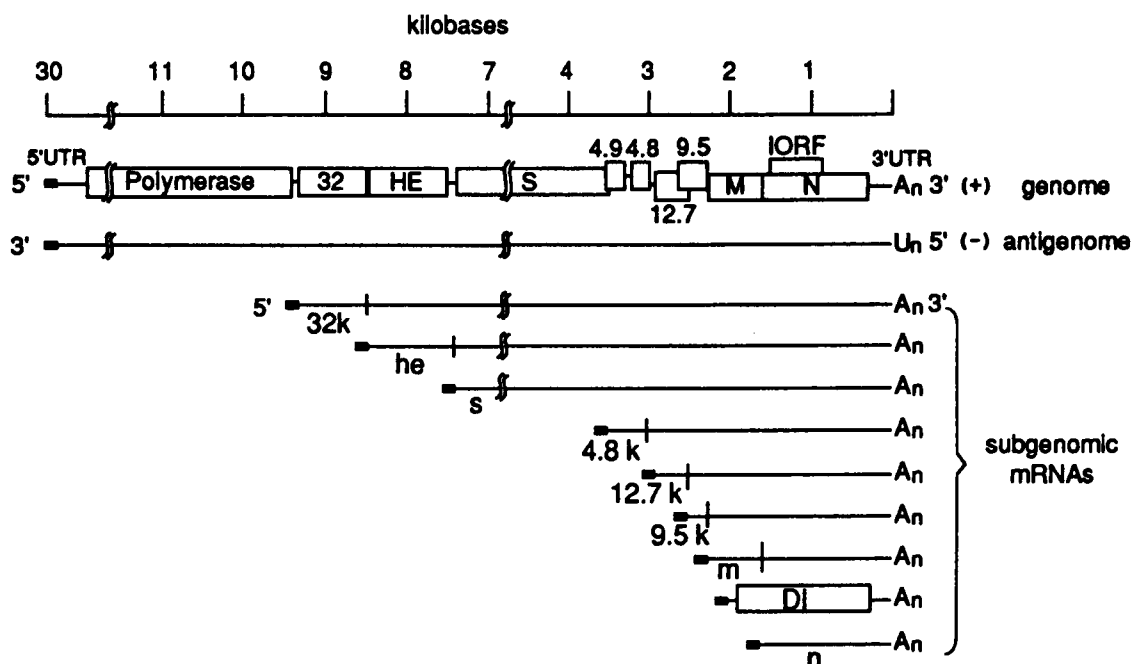


Fig. 1.1 Structure of BCV genomic and subgenomic RNAs. Each RNA contains a common leader sequence of 65 nucleotides, shown in black box at the 5' end, and poly(A) tail at the 3' end. The 5' terminal 498 nt and the 3' terminal 8955 nt have been cloned and sequenced. Open reading frames are indicated by open boxes. Abbreviations: HE, hemagglutinin esterase; S, spike; M, matrix; DI, defective-interfering; N, nucleocapsid; IORF, internal open reading frame.

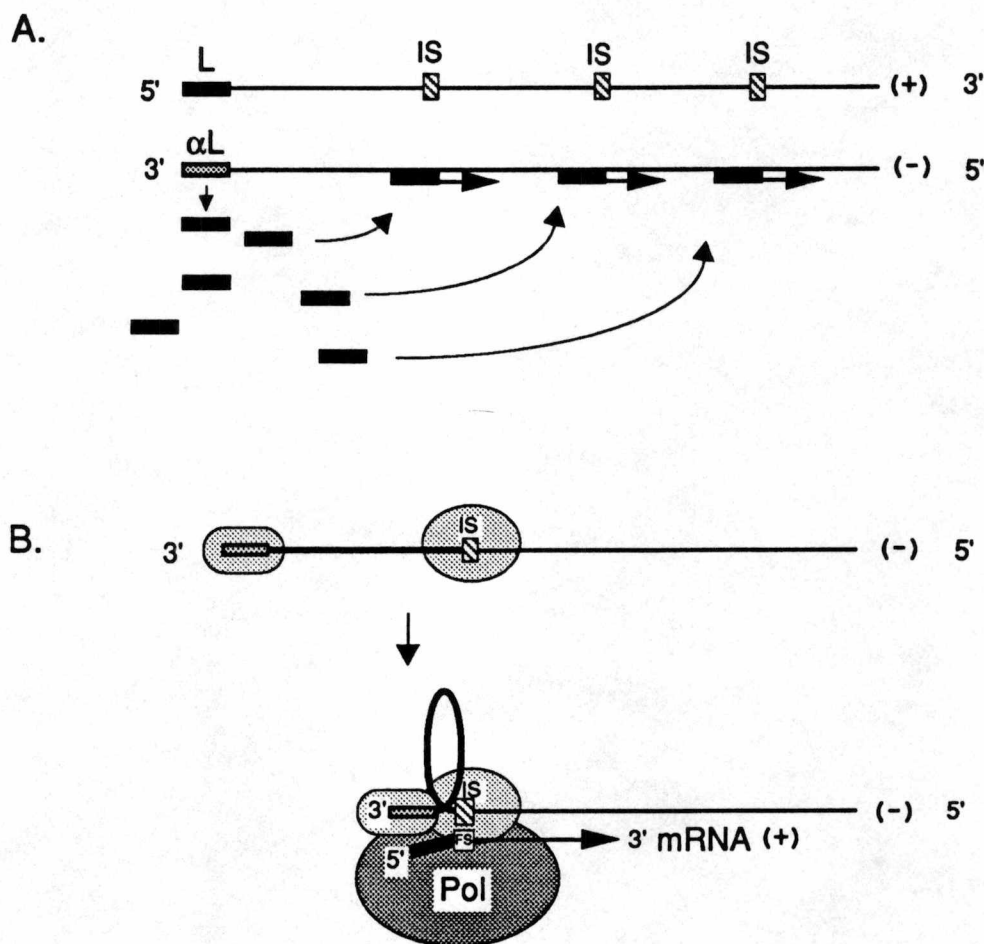


Fig. 1.2 Leader-primed transcription model for generating subgenomic mRNAs in coronaviruses. Models were redrawn from Lai, 1990 (A) and Zhang *et al.*, 1994 (B). A. Leader-primed transcription mechanism. Leader RNA (indicated as black boxes, L) is transcribed from the 3' end of the antigenome (minus strand), detaches, and with the RNA-dependent RNA polymerase, binds to different consensus intergenic sequences (denoted as hatched boxes, IS) on the antigenome. Leader priming then initiates transcription of the mRNAs resulting in a leader RNA being present on the 5' end of each subgenomic mRNA. B. A modified model of leader priming for subgenomic mRNA transcription. The small shaded ovals represent viral or cellular factors, and the large shaded oval denotes the viral RNA-dependent RNA polymerase (Pol). The thick lines between leader and IS represent the spacer sequence. The postulated transcriptional initiation complex contains an antileader and an anti-IS associated with viral or host factors. The *trans*-acting leader with the viral RNA-dependent RNA polymerase is fused within the IS region. FS indicates leader-mRNA fusion site.

transcribed by the RNA-dependent RNA polymerase from the 3'-end of the minus-strand in an abortive fashion to produce the free leader. Then, the leader, possibly in a complex with the RNA-dependent RNA polymerase, dissociates from the template and binds to the intergenic consensus sequence (AGA UUUG in BCV for example) between each open reading frame on the full-length minus-strand RNA. The RNA-dependent RNA polymerase copies each message from the full-length minus-strand toward the end of the template and generates the positive-sense 3' coterminal nested set of subgenomic RNAs (Fig. 1.1). Recently this model has been modified as a result of recent studies on MHV showing that leader-mRNA fusion in coronavirus transcription may not use direct RNA-RNA interaction between complementary sequences but rather RNA-protein and protein-protein interactions. A modified model of RNA transcription based on RNA-protein and protein-protein interactions is illustrated in Fig. 1.2B (van der Most *et al.*, 1994; Zhang *et al.*, 1994).

A second hypothesis to explain the origin of subgenomic mRNAs with a common leader states that the templates for the leader and mRNA become fused as minus-strand molecules, and that the minus-strand molecules then serve as templates for mRNA synthesis through the mechanism of mRNA replication (Sawicki and Sawicki, 1990). This idea was based on data first reported from our laboratory, in which it was shown that the subgenomic mRNAs, like full length genomic RNA, can be copied into minus-strands (Brian *et al.*, 1994; Hofmann *et al.*, 1990; Hofmann and Brian, 1991b; Sethna *et al.*, 1989). Minus-strand copies (anti-mRNAs) were found to exist for each and every subgenomic mRNA. In addition, plus-

and minus-strand RNAs were found to be components of double-stranded replicative forms (Sawicki and Sawicki, 1990). These results suggested that mRNAs might undergo replication and this notion was put forth as a hypothesis (Sethna *et al.*, 1989). Sawicki and Sawicki proposed that subgenomic minus strands containing the antileader are made by discontinuous RNA synthesis during minus-strand synthesis. This would require that the polymerase jump, either to a downstream site on its own template, or in *trans* to complete synthesis on another molecule.

Regardless of the mechanism used to generate the subgenomic mRNA templates, priming of transcription by a free leader during plus-strand synthesis or synthesis of subgenomic templates by a discontinuous action of the polymerase during minus-strand synthesis, amplification of subgenomic mRNAs was viewed as advantageous to the virus (Sethna *et al.*, 1989). mRNA replication as a mechanism of amplification would be a powerful mechanism by which a virus could increase the number of mRNA templates and thereby synthesize large amounts of proteins needed during an acute infection. As learned in this study, mRNAs probably do not undergo replication, but the subgenomic minus strands may still be contributing to mRNA amplification. This could happen by transcription from internal promoters on the subgenomic minus strands using the same mechanisms employed during transcription from the genomic minus strand (Brian *et al.*, 1994; Chapter II).

Do subgenomic mRNAs amplify *via* transcription by leader priming, mRNA replication, discontinuous transcription during negative-strand

RNA synthesis, or some other mechanism (e.g. *trans*-splicing)? Further studies in this area are needed to address these questions. Identification and characterization of the promoter sequences for plus-strand and minus-strand RNA synthesis are crucial for understanding coronavirus replication and transcription mechanisms.

Defective-interfering (DI) RNA.

Because the coronavirus genome is so large (30 kb), it is technically difficult at this time to cDNA clone the whole molecule in one piece in order to produce infectious genomic transcripts. However, it should be possible, based on precedents in other RNA virus systems (e.g. picornaviruses and togaviruses) to cDNA clone a defective-interfering genome that can replicate after transfecting of the RNA transcript into virus-infected cells and to characterize the *cis*-acting sequence elements required for replication of this molecule. Properties of DI particles and DI RNA as learned from the study of many viral systems include the following (reviewed by Holland, 1991; Schlesinger, 1988): (1) they often appear as a consequence of serial viral passaging at a high multiplicity of infection; (2) they contain only a part of the viral genome which, for plus strand RNA viruses, always include as a minimum, the termini of the genome; (3) helper virus is required for replication and encapsidation of the defective genome; (4) they replicate at the expense of the nondefective virus and interfere with the intracellular replication of the latter; and (5) interference is specific to the helper virus. In the absence of an infectious transcript generated from a cloned viral genome, the DI RNA is the next molecule of

choice to study replication since it retains the necessary *cis-acting* signals for RNA replication and its size is much smaller than genomic RNA facilitating cDNA cloning and other genetic mutational manipulation. In MHV, synthetic DI RNAs have proven to be a valuable tool for studies on various aspects of coronavirus replication, such as leader switching (Makino and Lai, 1989a), transcription initiation (Jeong and Makino, 1992; Makino and Lai, 1990; Makino *et al.*, 1991; Makino and Joo, 1993; van der Most *et al.*, 1994), and the identification of an encapsidation signal (Fosmire *et al.*, 1992; Makino *et al.*, 1990; van der Most *et al.*, 1991).

A defective subgenomic RNA species of approximately 2.2 kb [electrophoretically migrating between the mRNA species for the nucleocapsid (N) protein (1.7 kb) and the integral membrane (M) glycoprotein (2.5 kb)] was present in one of our cloned stocks of the Mebus strain of bovine coronavirus and appeared during a study of BCV persistent infection (Fig.1.1) (Hofmann *et al.*, 1990). It was this DI RNA that was chosen for use in this study. The cDNA clone of BCV DI RNA containing reporter sequence was cloned downstream of the phage T7 RNA polymerase promoter in a plasmid vector (Chapter II, Fig. 2.2). RNA transcripts of this construct were shown to undergo replication, persistence, and packaging (Chapter II). These results, therefore, show that the BCV DI RNA is a useful model for studying *cis*- or *trans*-acting elements required for coronavirus RNA replication.

Comparison between BCV and MHV.

BCV and MHV are closely related viruses based on their amino acid sequence identity. The identity of the BCV N protein with that of MHV is 70% (72% base identity). The BCV M protein shares 86% amino acid sequence identity with MHV (Lapps *et al.*, 1987). However, there are several differences between these two viruses. (1) Structure of DI: For MHV, there is one small DI RNA species called DIssF, which replicates at high efficiency when transfected into the helper virus-infected cells and is efficiently packaged into MHV virions. It has a mosaic structure consisting of sequences derived from five discontinuous regions of the parental virus genome. The first four regions (domain I-IV) from the 5' end are derived from four different regions of RNA-dependent RNA polymerase gene, and the last domain is derived from the 3' end of genomic RNA (Makino *et al.*, 1990). The DI RNA that we found in BCV is very similar to the DIssF of MHV with regard to its replicational competence and packaging efficiency. However, BCV DI RNA is a simple molecule resulting from a fusion between the genomic 5' 498 nt and 3' of 1635 nt (Fig. 1.3). (2) Packaging signal: For MHV, a signal required for packaging of the DI RNA into virions was found to map at 20,356 to 20,416 nucleotides from the 5' end of genomic RNA; thus, it is 1,381 to 1,441 nucleotides upstream from the 3' end of the polymerase gene (Makino *et al.*, 1990; Fosmire *et al.*, 1992). No equivalent stretch of nucleotides is found in the BCV DI RNA, yet it is packaged. Packaging is illustrated here by the fact that the DI RNA is found in purified virions. It is also perpetuated by serial

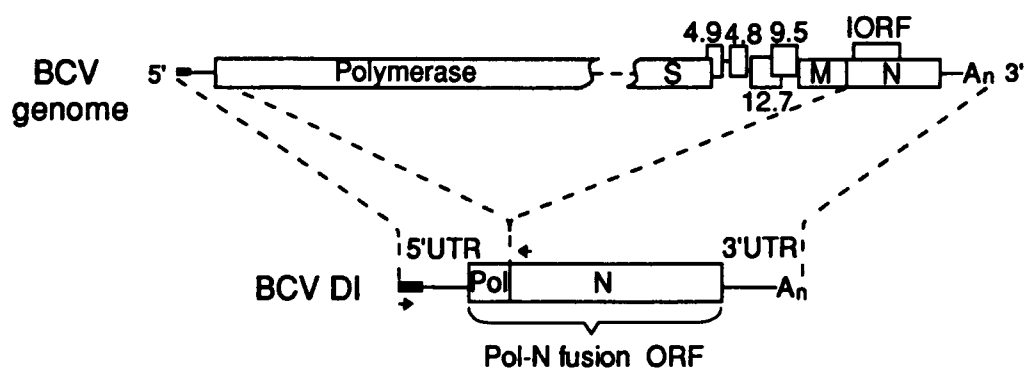


Fig. 1.3 Structure of BCV defective-interfering RNA. BCV DI RNA is a simple fused molecule between the genomic 5' 498 nt and 3' of 1635 nt. Two inverted arrows indicate the oligonucleotides that were used for reverse transcription and PCR amplification to clone 5' end of DI.

infections, indicating that it has been packaged. This information, along with an earlier demonstration that BCV mRNAs become packaged (Hofmann *et al.*, 1990), while MHV mRNAs are not shown to be packaged, suggests that the packaging signal for BCV may be significantly different from that of MHV, although they are closely related. Presence of subgenomic mRNAs in virions was also reported in the avian infectious bronchitis coronavirus (Zhao *et al.*, 1993) indicating that characteristics of MHV may not necessarily apply to the other coronaviruses. Interestingly, two closely related viruses having different packaging signals were also observed in alphaviruses. The subgenomic mRNA of Sinbis virus is not packaged (Weiss *et al.*, 1989), while the subgenomic RNA of Aura virus, a South American alphavirus related to Sinbis virus, is efficiently packaged into virions (Rümenapf *et al.*, 1994). (3) Intergenic consensus sequence: For MHV, within the 3' region of the genomic leader sequence there is a pentanucleotide sequence, UCUAA, that is repeated two to four times in different MHV strains (Makino and Lai, 1989b; Makino *et al.*, 1988). When leader-mRNA fusions are analyzed in MHV, each type of mRNA may differ in the number of UCUAA repeats resulting in heterogeneity at the site of fusion (Zhang and Lai, 1994; van der Most *et al.*, 1994). BCV, on the other hand, has only one intergenic consensus sequence, UCUAAAC, and leader-mRNA junction sequences are unique for each subgenomic mRNA (Hofmann *et al.*, 1993). The fusion event in BCV as compared to MHV therefore, is very stable. This feature may or may not directly involve replication but does concern the mechanism of subgenomic plus-strand RNA synthesis.

RATIONALE FOR THE STUDY OF BOVINE CORONAVIRUS

DEFECTIVE-INTERFERING RNA

There were two major reasons for cDNA cloning the BCV DI RNA and using it as a replicon for study. The first was to have a model replicon that could be used as a positive control in replication studies when testing our hypothesis that the coronavirus mRNA molecules can undergo replication. The second reason was, and remains, that it is technically difficult at this time to cDNA clone the entire 30 kb coronavirus genome in one piece in order to synthetically generate infectious molecules. The BCV DI RNA therefore is a model genome with which to study many of the mechanisms involved in coronavirus RNA replication.

To test the hypothesis that coronavirus mRNAs can undergo replication as a minigenome, a cDNA clone of the smallest message of subgenomic RNAs [nucleocapsid (N) mRNA for BCV] was made with the correct leader sequence and a poly(A) tail. In addition, a reporter sequence of 30 nt from the N gene of TGEV (this is a sequence that does not cross hybridize between the two viruses) was cloned into the N gene, in-frame with the N open reading frame, such that the plus-strand or minus-strand RNA could be identified by Northern analysis with a probe specific for the reporter sequence. The construct was named pNrep2. When plus-strand transcripts of pNrep2 were transfected into infected cells, no amplification of the input RNA was seen, indicating that viral mRNAs with reporter sequences were not replicating in infected cells. This result suggested that some other *cis*-acting sequence must be required for RNA replication.

Based on the precedent from other viruses (e.g., poliovirus, alphavirus, and MHV) that transcripts of cDNA clones of defective-interfering RNA can undergo replication, a BCV DI that was discovered by M. Hofmann in the study of BCV persistent infection (Hofmann *et al.*, 1990) was cloned and used for studying replication. The BCV DI RNA was found to be a simple fused molecule consisting of the 5' and 3' ends of the BCV genome (Fig. 1.3). The cDNA clone of BCV DI RNA containing a reporter sequence was named pDrep1. pDrep1 differs from pNrep2 by having 433 nt of continuous 5' proximal genomic sequence between the leader and the N ORF. It can undergo replication and become efficiently packaged. By comparing differences between pNrep2 which cannot replicate and pDrep1 which can replicate the localization of the *cis*-acting sequence requirements at the 5' end for replication ought to be facilitated. The features that distinguish pDrep1 from pNrep2 are: (1) a 5' untranslated region of 210 nt including the leader sequence (detailed studies shown in Chapter II), and (2) a fusion protein generated between the N-terminal portion of the polymerase gene and the entire N ORF (Fig. 1.3) (described in Chapter III).

Although a full-length cDNA clone of the BCV genome is not available, many aspects of coronavirus RNA genome replication should be obtainable from a study of the DI RNA. These include the potential for identifying (1) the *cis*-acting signals for minus-strand synthesis, (2) the *cis*-acting signals for plus strand synthesis, (3) the viral proteins that make up the polymerase complex, and (4) the cellular proteins that make up the polymerase complex. In addition, the cloned DI RNA might be useful as a replicon to study subgenomic transcription by engineering intergenic

sequences (internal promoters) into the replicon. Preliminary experiments by others in our laboratory (R. Krishnan) suggest that this might be possible. Preliminary studies have also shown that the BCV DI replicates in cells infected with human coronavirus OC43 (S. Ku, unpublished data), suggesting that information about replication of BCV may extend to the human coronavirus OC43. Characterization of the requirements for coronaviral replication should be useful in the design of an antiviral therapy for coronavirus diseases. It can be envisioned, for example, that antisense RNA or possibly ribozymes could be targeted to specific regulatory sequences.

CHAPTER II. 5'-END CIS-ACTING SEQUENCE REQUIREMENTS FOR BOVINE CORONAVIRUS DEFECTIVE-INTERFERING RNA REPLICATION.

INTRODUCTION

Coronaviruses possess a positive-strand polyadenylated RNA genome of approximately 30 kilobases that undergoes replication in the cytoplasm of infected cells (Lai, 1990). During virus replication, both full-length genomic and subgenomic RNAs are copied into minus-strands (Hofmann *et al.*, 1990; Sethna *et al.*, 1989). The subgenomic minus-strand molecules appear to be complete complements of the mRNAs except for a short oligo(U) of 12-20 nt at the 5' end of minus strand to complement the 3' terminal poly(A) tail of 100 to 300 nt (Hofmann and Brian, 1991b; Sethna *et al.*, 1991). The fact that the subgenomic minus-strands (anti-mRNAs) are components of metabolically active replicative intermediates (Sawicki and Sawicki, 1990) and that subgenomic mRNAs and anti-mRNAs amplify at a rate roughly proportional to their length (Hofmann *et al.*, 1990; Sethna *et al.*, 1989) led to the suggestion that the subgenomic mRNAs might serve as templates for synthesis of subgenomic minus strands that in turn serve as full-length templates for mRNA synthesis, i.e., amplification by the mechanism of replication (Hofmann *et al.*, 1990; Sethna *et al.*, 1989). Subgenomic mRNA amplification via replication, or possibly by transcription from internal promoters on the subgenomic minus strands (Brian *et al.*, 1994), would in

theory facilitate coronavirus replication since it would relieve demands on the low-abundance antigenome for mRNA production, a molecule already needed as a template for genome replication and for making the initial mRNA transcripts (Yokomori *et al.*, 1992).

The behavior of mRNAs as replicons would suggest that the 65 nt leader, encoded only at the genome 5' terminus but found as a common 5'-terminal sequence on all subgenomic mRNAs, suffices as a 5'-terminal *cis*-acting signal for replication. Assuming that a transfected mRNA molecule would become translocated to the replication complex, the sufficiency of the 5' leader as a promoter ought to be testable. However, when plus-strand transcripts of the nucleocapsid protein gene (the smallest mRNA) with a reporter sequence were transfected into infected cells, no clear evidence could be obtained that the mRNA underwent replication. Apparently an intact leader is not sufficient to promote replication of a subgenomic replicon. To extend the analysis and identify the sequence requirements for replication, a BCV DI RNA was cloned. Transcripts of the cloned defective-interfering RNA differ from the cloned N mRNA transcript by having an additional 433 nt derived from the 5' end of the genome and positioned between the 65 nt 5' leader and the N ORF start codon. These DI transcripts undergo replication and become packaged. The replicating construct was called pDrep1. Systematic deletions and mutations were made within pDrep1 to localize the *cis*-acting sequence requirements of coronavirus DI RNA replication.

MATERIALS AND METHODS

Virus and cells.

The Mebus strain of BCV was plaque purified three times and a stock at passage 5 was used as previously described (Hofmann *et al.*, 1990; Lapps *et al.*, 1987). The human rectal cell line, HRT-18, was used in all experiments. Cells were grown to monolayers in 35 mm² plastic dishes in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (Hyclone, Logan, UT.).

Synthetic oligonucleotides.

Oligonucleotides were synthesized by Oligo Etc. Inc. (Guilford, CT.). Oligonucleotides used in this study are listed in Table 2.1.

Cloning and sequencing of the 5'-end of the BCV DI RNA.

The 5' 557 nt of the BCV DI RNA was cDNA cloned using the N-binding primer [N(+), Table 2.1] for reverse transcription. The single-stranded cDNA then was amplified by PCR using the N(+) and the DI5'(-) primers. The amplified fragment was cloned into pCR1000 vector (Invitrogen, San Diego, CA.). The construct was named pCRDI (not shown). Sequencing of the cloned insert (Fig. 2.1) was done by dideoxynucleotidyl chain termination sequencing using the protocol supplied with the Sequenase sequencing kit (USB, Cleveland, OH.).

Table 2.1. Synthetic oligonucleotides used in this study

Oligonucleotide ^a	Binding site ^b	Sequence ^c
N(+)	<i>Xmn</i> I site-containing oligo which binds 34-39 nt in the N ORF	5' CCAGAACGATTTCCAAAGGACGCTCT 3'
TGEV#8(+)	nt 338-363 in the TGEV N ORF used as a reporter; binds to the plus strand	5' CATGGCACCATCCTTGGCAACCCAGA 3'
TGEV#7(-)	binds to the minus strand of reporter sequence	5' TCTGGGTTGCCAAGGATGGTGCCATG 3'
DI5'(-)	<i>Eco</i> RI and <i>Kpn</i> I site-containing antileader-binding primer	5' cggaaatcggiaccGATTGTGAGCGGATTTCGCTGCCG 3'
Ndel(+)	binds to the positive strand of nt 246-271 downstream from the DI ORF start codon	5' CCTCCAAATCATATGGACGCTGTATTTC 3'
DM(+)	binds to plus-strand nt 204-225	5' GTTGATCTTCGACAATGTGACC 3'
LM1(-)	binds to minus-strand of nt -7-21	5' cactataGATTCACCTCGCATTTGCGTGC 3'
LM2(-)	binds to minus-strand of nt 21-46	5' CGTGCATCCGCGGAGTGTGATCTCTTG 3'
pGEM3Zf(-)Ndel(-)	<i>Nde</i> I site-containing oligo	5' gagtgcaccataatgggtgt 3'
Sall(-)	binds to minus-strand of nt 196-219	5' CACCCATAGGTCGACAATGTGGAAG 3'
18S(+)	binds to plus-strand of human 18S rRNA base #595-620 (Torczynski <i>et al.</i> , 1985)	5' CTGCTGGCACCCAGACTTGCCCTCCAA 3'
ESK(+)	binds to the 22 nt of the 5' end of pDrep3	5' ggtaaccgagctgaattgccc 3'
leader(+) 344	binds to plus-strand nt 29-48	5' ACAAGAGATCAGTGAAGCCGG 3'

- a. Oligonucleotides bind either to plus-sense RNA as indicated by (+), or to minus-sense RNA as indicated by (-).
- b. Numbers (except for those describing the TGEV oligonucleotides) correspond to pDrep1 plus-strand sequence. TGEV numbers come from Kapke and Brian, 1986.
- c. The underlined sequence indicates a mutated region. The lower case sequence indicates that it is derived from the vector.

BCV Leader

```

BCV 5'  GATTGTGAGCGATTGCGTGCAT CCCGCTTCACT  GATCTCTT  GTTAGAT 53
MHV 5'  T- -AA--T---G---C---A-G-A---T-CT---CTAA-A-----GTA---TA-A 59

CTTTTATAA TCTAAAC TTTATAAAAACATCCACTCCCTGTAT TCTATGCTTGTGGGCG 112
TC-AA-C-----GG---T---CG-G-C---CC-----C 116

TAGATTTTTCATAGTGGTGTCTATATTCATTCTGCTGTTAACAGCTTTCAGCCAGGGAC 172
-G-TC--G-----C--A-AT-TG-GG--C-- --ACTTTC-T-C--T-----T--- 174

                                BCV POL ORF ->
                                M S K I N K Y 7
GTGTTGTATCCTAGGCAGTGGCCACCCATAGGT CACAATGTCGAAGATCAACAAATA 230
-----CCAT--GGC-C---CA-----TG-T---G-A-----GGG----- 234
                                - A - M G - - 7
                                MHV POL ORF ->

G L E L H W A P E F P W M F E D A E E K 27
CGGTCTCGAACTACACTGGGCTCCAGAATTCCATGGATGTTTGAGGACGCAGAGGAGAA 290
-----GCT-CA-A-----C-----C--CC-A-----TC----- 294
- - G F K - - - - - L P N - S - - 27

L D N P S S S E V D I V C S T T A Q K L 47
GTTGGATAACCTAGTAGTTCAGAGGTGGATATAGTATGCTCCACCACTGCGCAAAGCT 350
-----G-----GAG--G-----A---GGGT-T--C--T-TG-----G-A-C 354
- G - - E R - - E - G F - P S A - - E P 47

E T G G I C P E N H V M V D C R R L L K 67
GGAACAGGCGGAATTTGTCCTGAAAATCATGTGATGGTGGATTGTCGCCGACTTCTTAA 410
-A--GTTAAA-----AAAC-TTG-TT-----C---G-----A---G---CAGC 414
K V K - K T L V - - - R - - - S - - P A 67

Q E C C V Q S S L I R E I V M N T R P Y 87
ACAAGAGTGTGTGTGTCAGTCTAGCCTAATACGTGAAATTGTTATGAATACACGTCCATA 470
CTTG-----C-----T-----CGC-A---C-----T---T--G-TG-CGAGGA---CC- 474
L - - - - - A I - - D - F V D E D - Q 87

                                N ORF ->
D L E V L L Q E R
TGATTGGAGGTGCTACTTCAAGAAAGG ATG...
GA-GG-----CCTCGAC-ATGATGGCA
K V - A S T M M A

```

Fig. 2. 1 The aligned 5' ends of the bovine coronavirus DI RNA (which is the same as the BCV genome through base 498) and mouse hepatitis coronavirus genome. The BCV DI RNA nucleotide sequence and corresponding open reading frames are shown above, and those of MHV are shown below. The position of the 65 nt BCV leader is indicated. The consensus intergenic sequence and the start codons for the DI and N ORFs are boxed.

Construction of plasmids.

To construct pDrep1 (the reporter-containing DI RNA construct to be used in replication assays) (Fig. 2.2), the 78 nt *Bgl*II-*Xmn*I fragment of pNrep2 was replaced with the 499 nt *Bgl*III-*Xmn*I fragment from pCRDI.

pDrep3 was constructed by modifying the 5' end of pDrep1 (Fig. 2.2). (1) The DI5'(-) and NdeI(+) oligos were used to amplify the DI 5' end from virion RNA. A blunt-ended fragment was generated by PCR with *Pfu* DNA polymerase, and this was cloned into *Sma*I-linearized pGEM3Z to form pGDI5 (not shown). (2) The 2109 nt *Hpa*I-*Hind*III fragment from pDrep1 was used to replace the *Hpa*I-*Hind*III fragment of pGDI5, yielding pDrep2 (not shown). (3) The 15 nt *Kpn*I-*Kpn*I fragment from within the multiple cloning region of pDrep2 was removed and the resultant clone was named pDrep3 (Fig. 2.2).

In vitro mutagenesis of cloned reporter-containing DI RNA.

To prepare deletions of varying length between the leader and the N ORF of pDrep1, several approaches were taken. pD24-213, pD65-197, and pD130-170 were prepared by digesting pDrep1 at the unique *Hpa*I site, extending the digestion with *Bal* 31 exonuclease, blunt ending with mung bean nuclease and religating. Respectively, pD153-186 was prepared by removing the *Hpa*I-*Avr*II fragment, pD183-265 by removing the *Avr*II-*Xcm*I fragment, and pD181-471 and pD183-473 by removing the *Avr*II-*Nde*I fragment from pDrep1, blunt ending with mung bean nuclease and religating. pD201-209 and pD206-208 were prepared by cutting a modified

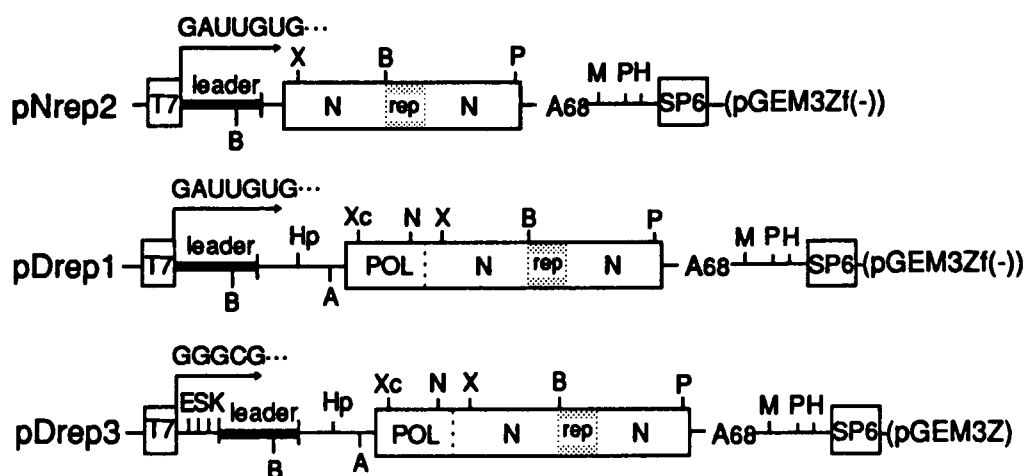


Fig. 2.2 Constructs used for testing RNA replication. Structure of plasmids pNrep2, pDrep1, and pDrep3 are shown. Abbreviations: A, *AvrII*; B, *BglII*; E, *EcoRI*; H, *HindIII*; Hp, *HpaI*; K, *KpnI*; M, *MluI*; N, *NdeI*; P, *PstI*; S, *SacI*; X, *XmnI*; Xc, *XcmI*; rep, 30 nt TGEV N-specific oligonucleotide reporter.

pDrep1 at a unique *SalI* site [created by PCR mutagenesis using a synthetic oligo *SalI*(-) with an inserted G after base 207], blunt ending with mung bean nuclease, and religating. pD469-470 was prepared by digesting pDrep1 at the unique *NdeI* site, blunt ending with mung bean nuclease, and religating.

To prepare a set of leader 5' deletions of varying length, pDrep3 was linearized with *EcoRI*, digested with *Bal* 31 exonuclease (0.2 u/ μ g DNA at 30°C), and fragments collected at 2 through 15 minutes were blunt-ended with mung bean nuclease, and digested with *HindIII* which cuts within the multiple cloning region downstream of the poly(A) tail. Gel purified fragments were ligated into pDrep3 that had been linearized with *EcoRI*, blunt-ended with mung bean nuclease and cut with *HindIII*. Deleted clones were screened by PCR and verified by sequencing. Clones with 5' terminal leader deletions of 4, 8, 13, 50 nt were used for replication studies.

To disrupt stem-loop 1 in pDrep1 transcripts, pLM1 (in which nucleotides 5 through 11 were replaced with CACTCGC) and pLM2 (in which nucleotides 30 through 37 were replaced with GCGAGTG) were generated. For this, a modification of the PCR-based gene splicing by overlap extension technique (Horton *et al.*, 1990) was used. For making pLM1, the gel purified 232 nt PCR product from a pDrep1-templated reaction using DM(+) and LM1(-) primers was used with the gel purified 738 nt product from a reaction using leader(+)344 and pGEM3Zf(-)NdeI(-) primers in an overlap PCR. In the overlap reaction, the 232 nt and 738 nt fragments were used together with DM(+) and pGEM3Zf(-)NdeI(-) primers

to form a product of 915 nt, from which the 660 nt *NaeI* vector-*HpaI* insert fragment was used to replace the analogous region in pDrep1, generating pLM1.

pLM2 was similarly generated, except that the overlap PCR used the 915 nt fragment from a reaction with pGEM3Zf(-)NdeI(-) and DM(+) primers, the the 460 nt fragment from a reaction with LM2(-) and the NdeI(+) primers, along with pGEM3Zf(-)NdeI(-) and NdeI(+) primers to generate a 1170 nt fragment. From this, the 660 nt *NaeI*-*HpaI* fragment was used to replace the analogous region in pDrep1 generating pLM2, or the analogous region in pLM1, generating pLMC, the compensatory double mutant.

Assay for replication (Northern analysis).

Transfections were done by the method of Felgner *et al.* (1987) essentially as described by Rice *et al.* (1989). Briefly, 5 µg of *MluI*-linearized DNA, blunt-ended with mung bean nuclease, was transcribed with 40 units of T7 RNA polymerase (Promega, Madison, WIS.) in a 100 µl reaction. Capping with ⁷mGpppG, when done, followed the procedure of the manufacturer (Promega). The transcription reaction mixture was treated with 5 units of RNase-free DNase (Promega) and the RNA was chromatographed through a Biospin 6 column (BioRad, Melville, N. Y.). Cells in 35 mm dishes at 50-80% confluency (approximately 1 x 10⁶ cells) were infected with BCV at a multiplicity of approximately 5 plaque forming units per cell by incubating cells with inoculum at 37°C for 1 h, refeeding with 2 ml of growth medium containing 10% fetal bovine serum, and

incubate at 37°C for 1 h. For transfection, each dish of cells was rinsed three times with Ca²⁺- and Mg²⁺-free phosphate-buffered saline (PBS) and treated for 10 min at 25°C with 200 µl of PBS containing 10 µg of lipofectin (Bethesda Research Laboratories, Grand Island, N. Y.) and 200 ng (unless otherwise noted) of transcript. Cells were rinsed with 2 ml of Dulbecco's Modified Eagle Medium and incubated at 37°C with 2 ml of medium containing 10% fetal bovine serum until cytoplasmic RNA extraction was done at the indicated times.

For serial passage of progeny virus, supernatant fluids were harvested at the indicated times and 500 µl was used directly to infect confluent cells in a 35 mm dish. Cytoplasmic RNA was extracted for analysis at 48h postinfection.

The yield of cytoplasmic RNA from one 35 mm dish was approximately 10 µg and 2.5 µg was used per well in a formaldehyde-agarose gel. Denaturing RNA electrophoresis, RNA blotting, and hybridization, were done as previously described (Hofmann *et al.*, 1990; Sethna *et al.*, 1989). 1 to 100 ng of transcripts was loaded per lane when used as marker, probe specific activities ranged from 1.5×10^6 to 4×10^6 cpm/pmole, and film exposure times ranged from 6 h to 1 week. Quantitation in Northern analyses was done with either the AMBIS Radioanalytic Imaging System (AMBIS, Inc., San Diego, CA.), or the imaging densitometer (Bio-Rad).

RNA stability test.

RNA stability testing was done exactly as described for the replication assay except that transcripts were transfected into uninfected cells. RNA was extracted every two hours after transfection and after electrophoresis was probed separately with end-labeled oligonucleotide specific to the reporter sequence and to 18S rRNA (Table 2.1). Quantitation was done simultaneously in order to normalize the amount of RNA in each sample.

Assay for leader reversion (RNase protection assay).

To generate minus-strand RNA probes of the pLM1 and pLMC leader sequences, pLM1 and pLMC were truncated to yield probes of 275 nt in length. For this, the *HpaI-HindIII* fragment [containing the entire pol-N ORF and poly(A) tail] was removed from each, the *HindIII* site blunt-ended by fill-in with Klenow enzyme, and the clones religated to form pLM1D and pLMCD. ³²P-labeled riboprobes were generated with SP6 polymerase after linearization of the plasmids with *PvuII* which cuts within the vector.

For the RNase protection assay, essentially the method of Vidal Kolakofsky (1989) was used. Approximately 60,000 cpm of riboprobe purified by electrophoresis in a denaturing polyacrylamide gel was coprecipitated in ethanol with 5 µg of cytoplasmic RNA or 100 ng of *in vitro* generated RNA transcripts. Dried RNA was dissolved in 9 µl of 0.1% N-lauroylsarcosine (Sigma, St. Louis, MO.) in water, heated for 1 min at 90°C, quick chilled, and made 0.1% N-lauroylsarcosine, 0.3 M NaCl, 20 mM Tris-HCl (pH 7.4), 2 mM EDTA (1X annealing buffer) by adding 1 µl 10X annealing buffer. After

annealing for 30 min at 70°C, 50 µl of 0.1% sodium dodecyl sulfate (SDS)-40 µg RNase A per ml in 1.25X annealing buffer was added. After 60 min at 30°C, 7 µl of 1% SDS and 2.1 µl of proteinase K (20 mg/ml) were added and the mix was incubated 30 min at 30°C, phenol-chloroform extracted after the addition of 5 µg tRNA, and ethanol precipitated. RNAs were dissolved in 10 µl of 80% formamide plus dyes, heated for 2 min at 90°C and separated on a 6% polyacrylamide sequencing gel at 5 µl per lane.

Enzymes.

Enzymes were purchased from New England Biolabs (Beverly, MA.), Promega Biotec (Madison, WIS.), and Stratagene (La Jolla, CA.).

Other methods.

Molecular biological manipulations were carried out using standard protocols (Sambrook *et al.*, 1989). All constructs were confirmed by sequencing through junction regions. RNA structural analyses were done with the aid of the Microgenie program (Beckman Instruments), which uses the RNA folding algorithm of Tinoco *et al.* (1973).

Nucleotide sequence accession number.

The GenBank/EMBL accession number is U00735.

RESULTS

The 5' cis-acting signal for RNA replication includes more than the leader.

To directly test the hypothesis that the leader is sufficient as a 5'-terminal signal for coronavirus RNA replication (Hofmann *et al.*, 1990; Sethna *et al.*, 1989), a clone of BCV N mRNA (Senanayake *et al.*, 1992) with a precise 5'-terminal sequence (Hofmann and Brian, 1991b; Hofmann *et al.*, 1993) on the 65 nt leader, a 3' poly(A) tail of 68 nt, and an in-frame 30 nt reporter sequence within the N ORF (construct pNrep2, Fig. 2.2) (Hofmann in Chang *et al.*, 1994) was used to prepare capped RNA transcripts *in vitro*. Transcripts were then tested for replication by transfection into infected cells. This system was based on the successful use of *trans*-acting RNA replication machinery in the analysis of replicating transfected subgenomic (DI) RNA molecules of picornaviruses (Hagino-Yamagishi and Nomoto, 1989), togaviruses (Levis *et al.*, 1986), and coronaviruses (Makino *et al.*, 1991; van der Most *et al.*, 1991). In our system, replication was monitored by Northern analysis with a strand-specific probe for the reporter sequence. When synthetic capped transcripts of the N mRNA construct (pNrep2) were transfected into infected cells and cytoplasmic RNA extracts were examined, no increase in the abundance of plus strands was observed and minus strands were not found (data not shown). This indicated that the N mRNA transcript alone could not replicate.

To determine what sequence in addition to the 5' leader might be required to promote RNA replication, the 5' 557 nt terminus of a DI RNA,

known from Northern blot experiments to contain the leader and at least a portion of the N mRNA (Hofmann *et al.*, 1990) was cloned and tested for its ability to restore replication to the reporter-containing mRNA (described below). Sequence analysis of the cloned 5' terminal 557 nt of the DI RNA demonstrated the first 498 nt to be identical to the 5' terminus of the BCV genome (Hofmann *et al.*, 1993) and base 499 to be the beginning of the N ORF (Senanayake *et al.*, 1992). Thus, the natural BCV DI RNA is apparently a simple recombinant between the genomic 5' terminal 498 nt and the N mRNA (Fig. 1.3). This DI RNA contains the 5' end sequence of the genome including the first 288 nt of the polymerase gene beginning at nt 211. Furthermore, the polymerase open reading frame is contiguous with the entire N open reading frame. A sequence comparison between the BCV and MHV JHM genomic 5' terminal 498 nt (Lee *et al.*, 1991) shows a nucleotide sequence identity of 64% and the putative polymerase beginning at the same relative position on the genome (Fig. 2.1). For the 96 amino acid portion of the polymerase proteins, there is a 48% amino acid sequence similarity between BCV and MHV which compares to a similarity of 60, 87, and 70% respectively for the S, M, and N proteins of the two viruses (Abraham *et al.*, 1990a; Lapps *et al.*, 1987).

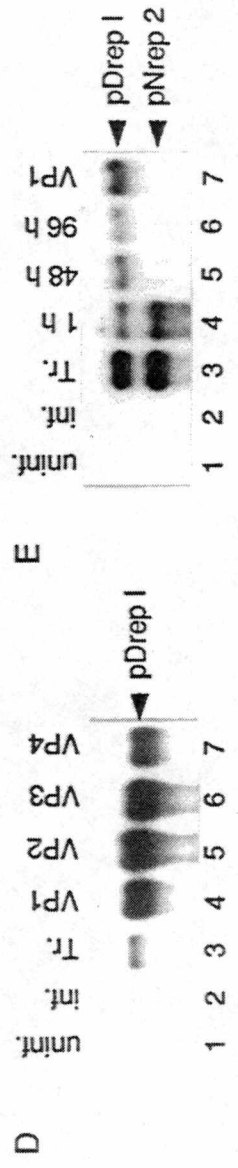
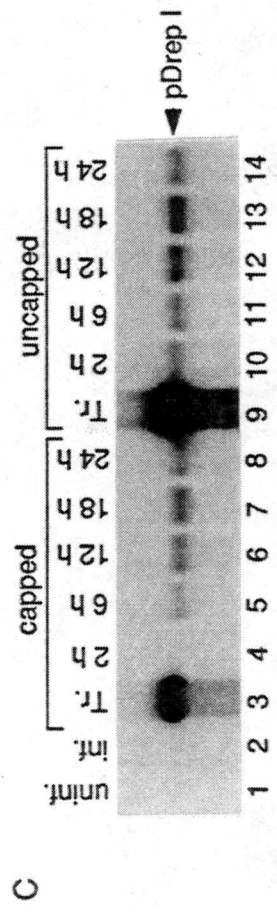
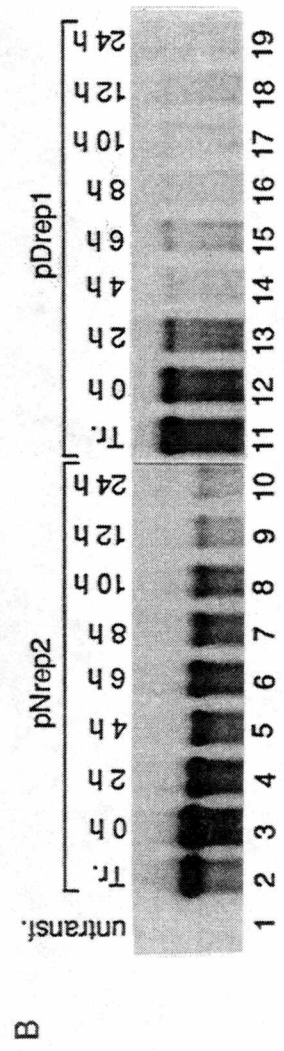
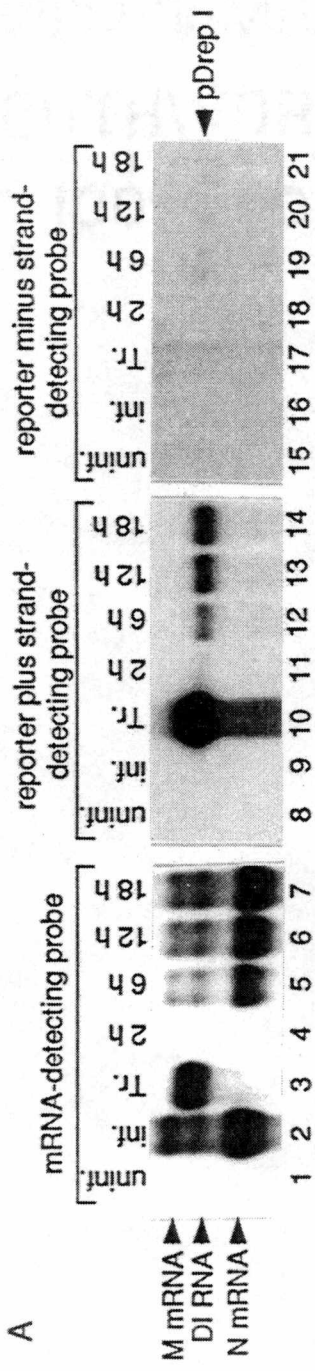
To test whether the DI RNA-specific sequence could impart an ability of N mRNA to replicate, the 5' terminal DI RNA sequence as a *Bgl*II-*Xmn*I fragment was used to replace the analogous reporter containing N mRNA construct (pNrep2) to form pDrep1 (Fig. 2.2). When synthetic uncapped transcripts from *Mlu*I-linearized pDrep1 were transfected into BCV-infected cells, replication was observed as demonstrated by an 11-fold

increase in abundance of plus-strand RNA through 18h posttransfection (Fig. 2.3A, lane 11-14), and by the appearance of minus-strand RNA that peaked at 6h posttransfection (Fig. 2.3A, lanes 18-21). Both the abundance of minus strands relative to plus strands and the kinetics of minus-strand accumulation and subsequent decay reflected patterns observed for genomic and subgenomic minus-strand RNAs in the BCV-infected cells (Hofmann *et al.*, 1990). To test whether successful pDrep1 replication under conditions in which there was no pNrep2 replication is due to a difference in RNA stabilities, equal amounts of RNA were transfected into uninfected cells and their decay rates were measured (Fig. 2.3). Amounts of RNA were normalized to the amounts of 18S rRNA present in each lane (Torczynski *et al.*, 1985). The result indicated that pDrep1 is more stable than pNrep2 (Fig. 2.4), suggesting that 5' 433 nt might involve protein binding and contribute to the stability of the molecule.

It was further demonstrated that transcripts used for transfection need not be capped since capped and uncapped transcripts of pDrep1 replicated equally well (Fig. 2.3C). Thus capping was not done in experiments described below. One notable feature of pDrep1 is that it underwent efficient packaging as evidenced by virion transmission for four serial passages (Fig. 2.3D).

Taken together, these experiments indicate that the leader sequence of 65 nt is not sufficient as a 5' *cis*-acting signal to promote replication of a transfected molecule, and that some or all of the additional 433 nt sequence between the leader and the start of the N gene in the DI RNA is required in

Fig. 2.3 RNA blotting evidence for replication and packaging of the reporter-containing BCV DI RNA (pDrep1). A. Evidence of pDrep1 replication. Denaturing RNA electrophoresis and RNA blotting were done as described in the Materials and Methods. Lanes 1-7, RNA was probed with a 5' end-labeled oligonucleotide that binds to plus strands of the N mRNA (N+ oligonucleotide), and all larger mRNAs in the nested set. Lane 1, RNA from uninfected cells. Lane 2, RNA from infected cells at 26 h postinfection. Lane 3, 10, and 17, pDrep1 transcript marker. Lane 4-7, cytoplasmic RNA from cells infected, then transfected. Times indicate hours posttransfection. Lanes 8-14, RNA probed with TGEV oligonucleotide #8 which detects the positive strand of the reporter sequence. Lanes 15-21, RNA probed with TGEV oligonucleotide #7 which detects the negative strand of the reporter sequence. B. Evidence of pNrep2 and pDrep1 stability in uninfected cells. Transcripts were transfected and cytoplasmic RNA was extracted at the times indicated. Tr., transcript marker. RNA extracts were analyzed to detect plus-strand reporter sequence. C. Evidence that both uncapped and capped transcripts of pDrep1 replicate. Tr., transcript marker. RNA extracts were analyzed to detect plus-strand reporter sequence. D. Evidence that pDrep1 is packaged. Virus produced by cells transfected with pDrep1 was used to infect new cells and cytoplasmic RNA from these (VP1) was analyzed at 49 h postinfection. Supernatant virus from this and subsequent infections was likewise passaged (VP2 through VP4). Tr., transcript marker. RNA extracts were analyzed to detect plus-strand transcript marker. RNA extracts were analyzed to detect plus-strand reporter sequence. E. Evidence that transfected mRNA transcripts (pNrep2) fail to replicate when cotransfected with DI RNA (pDrep1) transcripts. Tr., pDrep1 and pNrep2 transcript markers. For panels A and C, 2 μ g transcript was used for transfection. RNA extracts were analyzed to detect plus-strand reporter sequence.



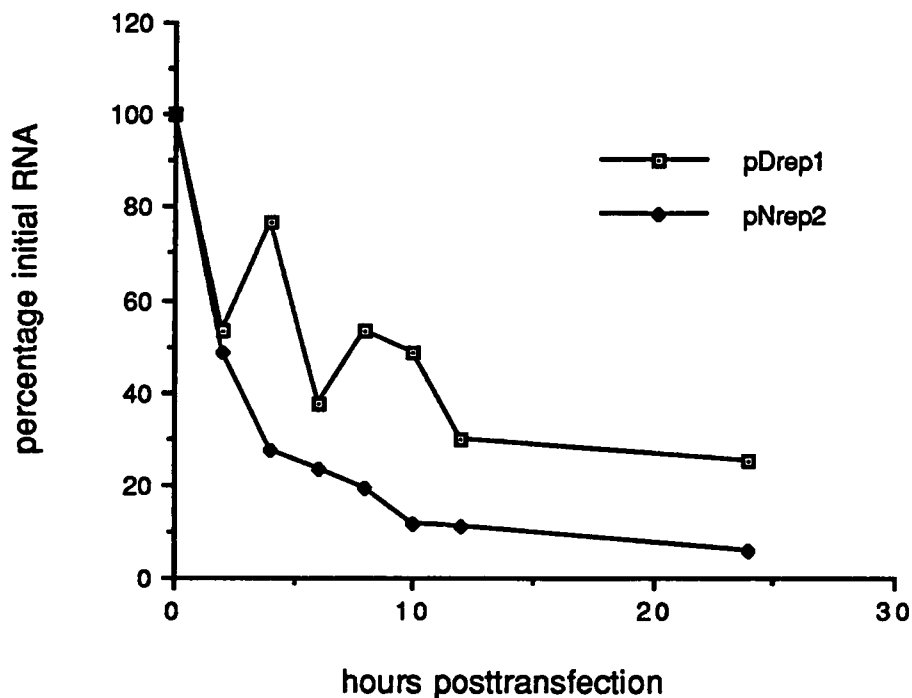


Fig. 2.4 Stability of pDrep1 and pNrep2 RNA transcripts. 200 ng of RNA transcripts were transfected into uninfected HRT cells. Cytoplasmic RNA was extracted at 0, 2, 4, 6, 8, 10, 12, and 24 hours posttransfection. RNA extracts were analyzed to detect plus-strand reporter sequence. Amounts of RNA were normalized to the amounts of 18S rRNA (Torczynski *et al.*, 1985) in each sample. Amounts of pDrep1 RNA and pNrep2 RNA were plotted as the percentage of the amount of that RNA present at 0 hr posttransfection.

cis for replication. To reexamine whether the N mRNA would replicate in the presence of replicating DI RNA, evidence of pNrep2 replication was sought after pNrep2 and pDrep1 were transfected together (Fig. 2.3E, lanes 4-7). Still no replication of pNrep2 was evident indicating again the insufficiency of the leader alone to provide the 5' *cis*-acting signal for replication of a transfected mRNA molecule.

It should be noted that although the amount of RNA taken up by cells, as determined by measurements at 1 or 2 h posttransfection, varied among experiments and was not entirely controllable (Fig. 2.3A to C and E) (see below), evidence of replication in all transfection experiments was based on both an abundant amount or increased amount of plus-strand RNA after long-term incubation (24, 48 or 96 h) and on the presence of plus-strand RNA at 48 h postinfection from passage 1 progeny virus (VP1).

Deletions of the 5'-untranslated region downstream of the leader disrupt replication.

To localize the minimum sequence requirement for RNA replication, systematic deletions were made within the 498 nt region using available restriction sites, site-specific mutagenesis, and *Bal* 31 exonuclease. To date, the 3' limit of the 5' terminal *cis*-acting signal for pDrep1 replication has not been identified. A limited number of deletion mutations located within the 210 nt 5' UTR and downstream of the 65 nt leader (Fig. 2.5) have all prevented replication with the exception of the pSM3 construct which has a one G insertion at base 207 suggesting that most of the entire 5' UTR may be a part of the signal required for RNA replication. To test whether

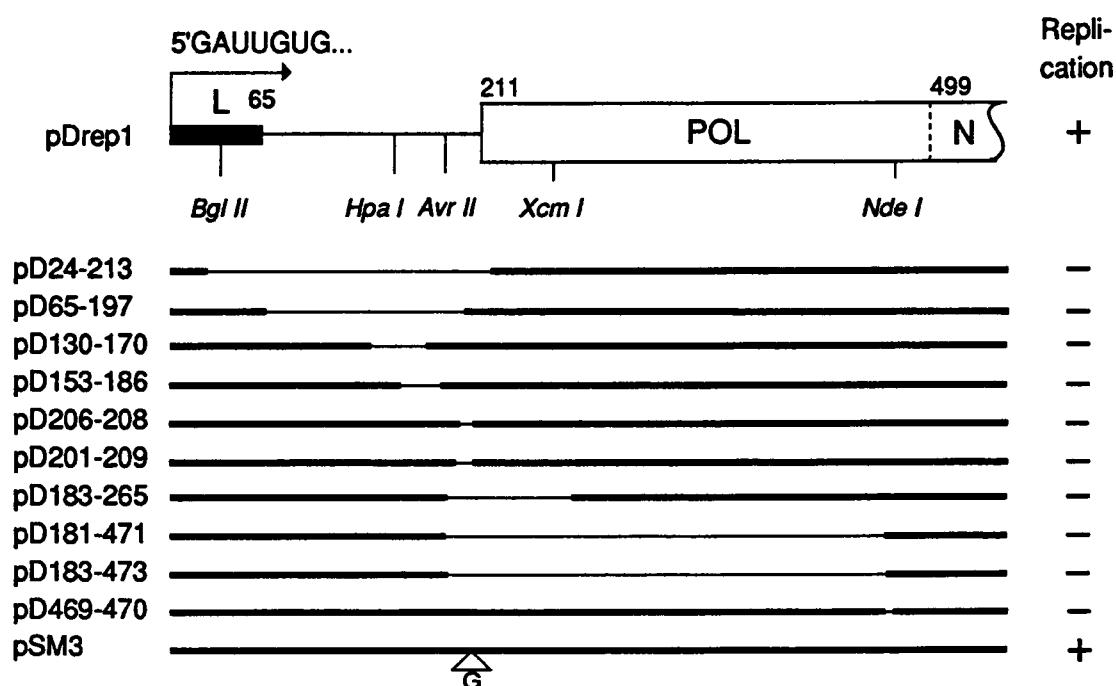


Fig. 2.5 Effects of 5' proximal internal mutations on the replication of pDrep1 RNA. Internal deletions of pDrep1 were prepared as described in Material and Methods. pSM3 has an inserted G after base 207, indicated by a triangle. Transcripts were tested for replication by RNA blotting analysis to detect plus-strand reporter sequence. Cytoplasmic RNA was analyzed at 1, 48, and 96h posttransfection, and at 48h postinfection for VP1.

the impairment in replication of these 5'-UTR deletion mutations is due to the effect on translation ability, equal amounts of pDrep1, pD206-208, pD201-209, and pD153-186 RNA transcripts, quantitated by a Northern assay, were used for *in vitro* translation. No significant differences among the molecules in translational efficiencies were observed (data not shown). In fact, pD153-186 showed slightly better translation than did pDrep1, suggesting that the loss of replication for this construct is not due to a loss in translational ability.

Computer analysis of the genomic 5'-UTR predicts at least five stem-loops.

Potential secondary structures within 5' untranslated region of BCV DI RNA were investigated with the Microgenie computer program (Beckman Instruments) which uses the RNA-folding algorithm of Tinoco *et al.* (1973). From this analysis, 13 possible stem-loop structures with free energies ranging from -4.4 kcal/mole to -14.2 kcal/mole were predicted within the first 210 nt sequence. Because of redundant usage of sequence in the secondary structures, however, only five unique stem-loops could be drawn. The most stable of the putative secondary structures are shown in Fig. 2.6. These include five stem-loops and two possible pseudoknots (structures in which nucleotides within a hairpin loop base pair to nucleotides outside the hairpin).

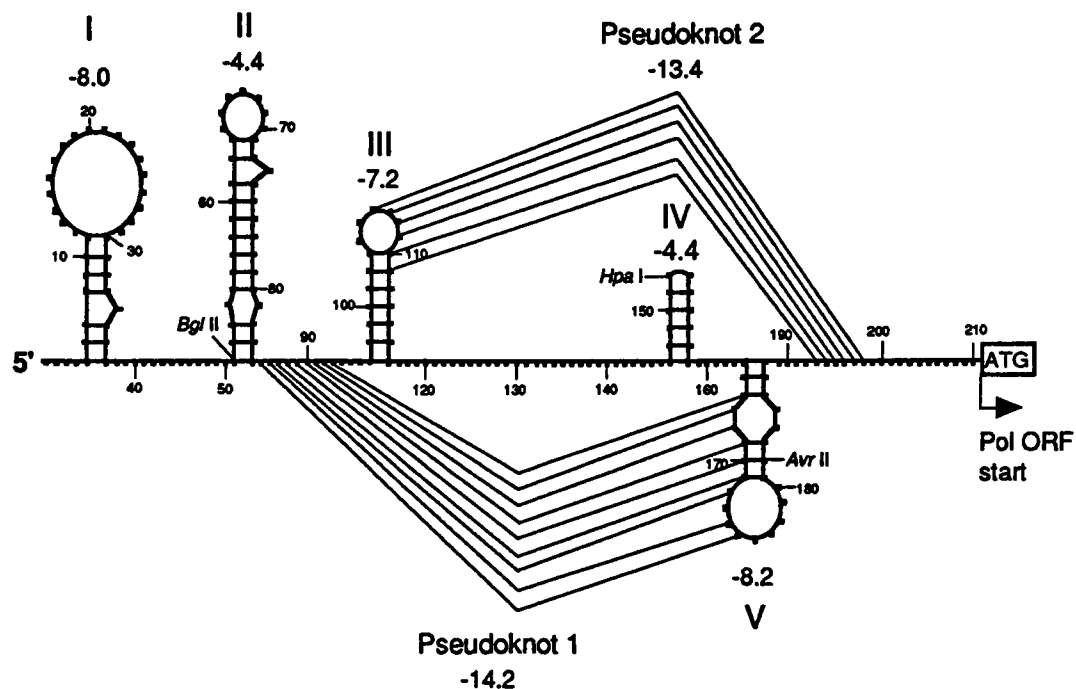


Fig. 2.6 Stem-loop structures within the BCV genomic 210 nt 5' UTR. The computer-predicted structures include five stem-loops and two pseudoknots. Each black spot on the sequence line indicates one nucleotide. The stability of each stem-loop is indicated by its free energy in kcal/mole.

5' terminal deletions of the leader disrupt a predicted intraleader stem-loop and also destroy replication ability.

To test whether the leader is required in *cis* for replication of the DI RNA, 5' truncation deletions were made and various deletion mutants were tested for replication. To generate a nested set of deletions with *Bal* 31, pDrep1 was modified to pDrep3 by inserting a 22 nt sequence containing *Eco*RI, *Sac*I and *Kpn*I restriction endonuclease sites between the T7 promoter and the first base of the leader. By using *Eco*RI-linearized DNA to prepare *Bal*31 digestions, clones containing deletions of 4, 8, and 50 nt were generated. However, it should be noted that for each of these, transcripts began with a 5' terminal gggcg sequence that was derived from the multiple cloning region of the vector. That is, the pD4 clone has nucleotides 1-4 of the leader (GAUU) replaced with the gggcg sequence, and so on (Fig. 2.7). To test whether the secondary structures play a role in viral replication, all of these mutants as well as pDrep1 were analyzed by the Microgenie computer program. There are two predicted secondary structures within the 5' end of pDrep1 with a free energy -8.0 Kcal/mole at stem-loop 1 and -4.4 Kcal/mole at stem-loop 2. Stem-loop 1 contains a bulging U at position 34. It should be noted that the predicted stem-loop 1 remains unchanged in pDrep3, but undergoes modifications to become a stem-loop of -9.4 Kcal/mole with a bulging U at original base position 6 for pD4, and one of -11.2 Kcal/mole with a bulging C at original base position 10 for pD8. For pD13 and pD50 the predicted stem loop disappears (Fig. 2.8). Whereas transcripts of pD4 and pD8 replicated almost as well as those from pDrep1. The removal of 13 or more nucleotides from the 5' end of the leader

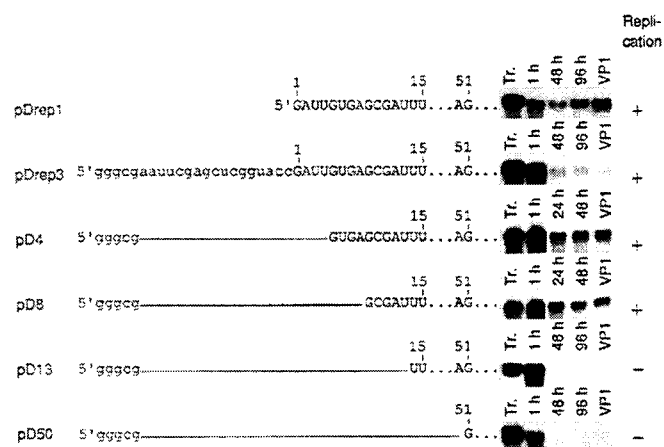


Fig. 2.7 Effects of 5' terminal leader deletions on the replication of pDrep1. Deletion sites and evidence of replication are shown. Lower case letters represent bases derived from the multiple cloning region of the vector and deleted bases are indicated by solid line. Tr., transcript marker. DI RNAs were identified by RNA blotting analysis with probe that detects plus-strand reporter sequence.

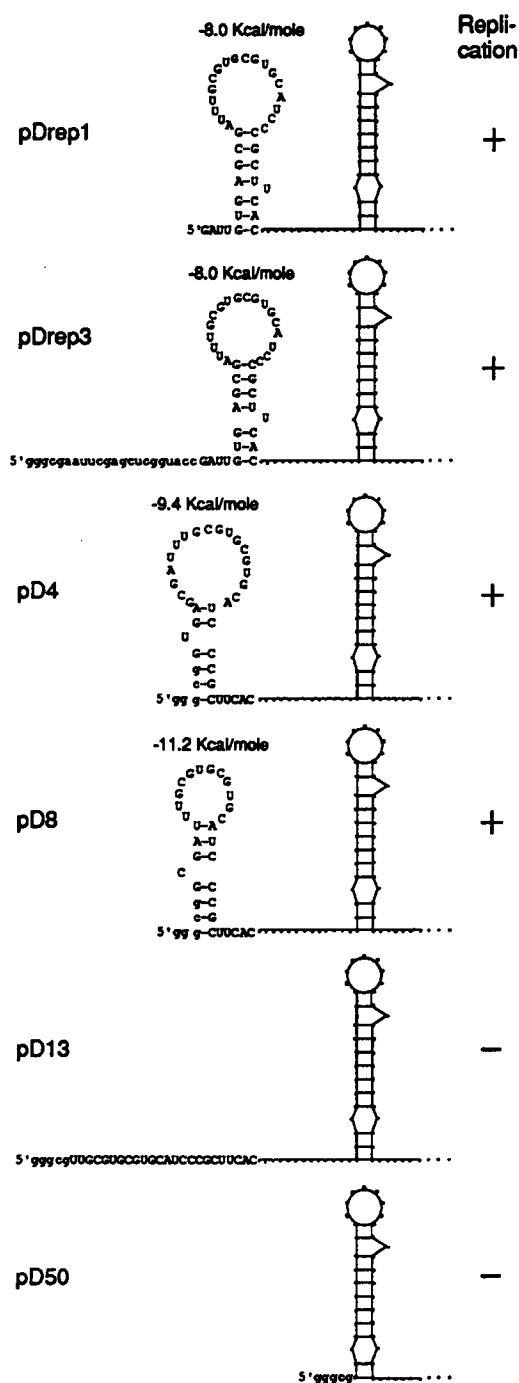


Fig. 2.8 Predicted stem-loop structures of the 5'-end in pDrep1, pDrep3, and the deletion mutants.

disrupted the first stem-loop structure and completely destroyed replicational competence (Fig. 2.7). Thus, although the 5' terminal portion of the leader up through 8 nt was dispensable, the tolerated limit for this truncation lies somewhere between nucleotides 8 and 13 from the 5' end. Whether the replicating molecules resulting from transfection with pD4 and pD8 are revertants with a wild-type leader (as described below for stem-loop mutants) was not tested.

Intriguingly, pDrep3 with a 5' extension of 22 nt also replicated but accumulated at a slower rate (Fig. 2.7). Whereas RNA blotting experiments using end-labeled oligonucleotide [Table 2.1, ESK(+)] specific for the 22 nt demonstrated the additional sequence to be missing in progeny DI RNA propagated from passaged virus (data not shown), 5'-terminal extensions were found in a small fraction of naturally occurring BCV mRNA (Hofmann *et al.*, 1993) and similar extensions have been shown to be tolerated in replicating molecules of other plus-strand RNA viruses (Ball and Li, 1993).

Leader with stem-loop 1 disrupted by base substitutions is rapidly replaced by wild type leader.

The correlation between the loss of replication and the loss of stem-loop 1 suggested that this stem-loop or its variants (Fig. 2.8) might be an important structural element in the 5' *cis*-acting signal for replication. To confirm that the first-stem loop structure within the leader is required for replication, three mutants were constructed using PCR site-directed mutagenesis. pLM1 and pLM2 are the base substitutions within each side

of the stem which generated mismatches and destroyed the stem region. pLMC has a double mutation within both sides of the stem which restored the stem-loop structure, however without a bulging U. Transcripts of these mutants were tested for replication. Surprisingly, the replication of pLM1 and pLM2 did not appear impaired relative to pDrep1 (Fig. 2.9, lane 3-14), although the compensatory mutant pLMC showed some impairment (Fig. 2.9, lane 15-18). All three mutants became packaged and replicated during passage 1 of progeny virus on fresh cells (Fig. 2.9, lanes 10, 14, 18).

To determine whether the leaders on replicating molecules resulting from transfection with mutants had retained their mutant genotype or had possibly reverted to the wild type, an RNase protection assay was carried out on each mutant. The results indicated that all three mutants had reverted to the wild type leader within 48 h of transfection and maintained this genotype in the progeny of packaged DI RNA (Fig. 2.10A, lanes 8-10, and 2.10B, lanes 9-10 and 13-14).

DISCUSSION

The hypothesis stated earlier that coronavirus subgenomic mRNAs behave as replicons (Hofmann *et al.*, 1990; Sethna *et al.*, 1989) were directly tested here using transcripts with a reporter sequence in the presence of viral polymerase and other factors. No evidence of mRNA replication from transfected molecules was found. The 5'-terminal 498 nt on the BCV DI RNA (including the 65 nt 5' terminal leader sequence) did impart the ability

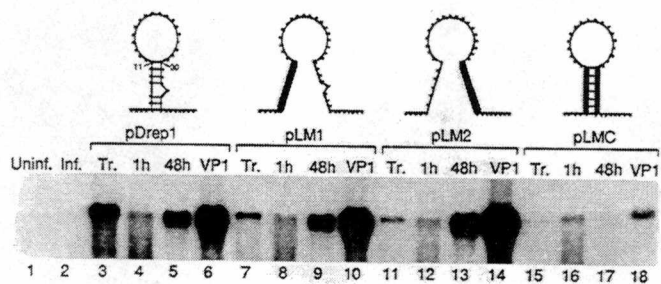


Fig. 2.9 Effects of substituted bases within stem-loop 1 on the replication of pDrep1 RNA. Substituted regions in pLM1, pLM2 and pLMC are shown with a heavy solid line. Tr., transcript marker. DI RNAs were identified in RNA blotting analysis with probe that detects plus-strand reporter sequences.

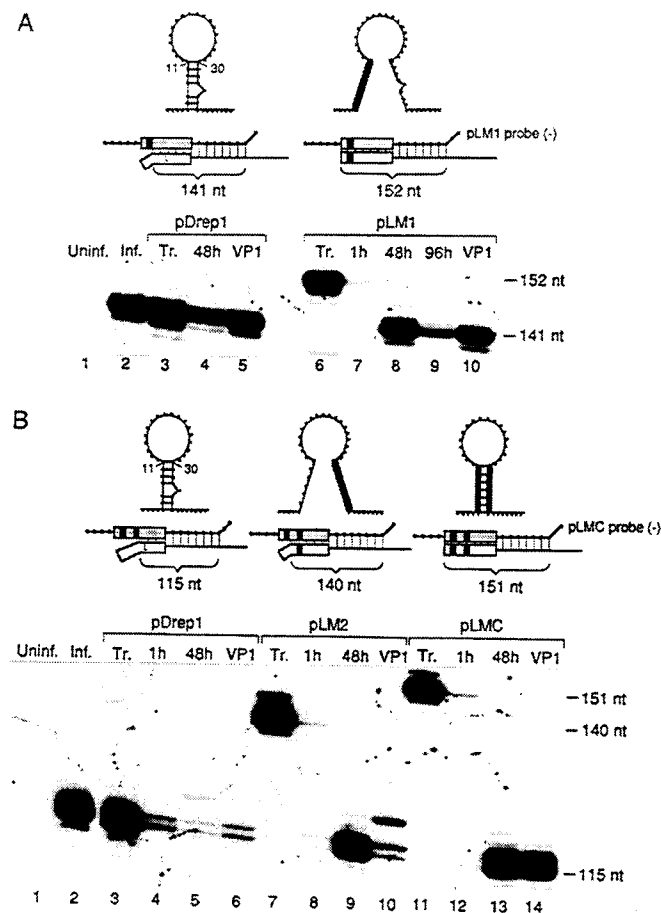


Fig. 2.10 Evidence of leader reversion for leaders with stem-loop 1 substitution mutations. Radiolabeled minus-strand riboprobes of pLM1 leader (panel A) or pLMC leader (panel B) were used in RNase protection assays with cytoplasmic RNA from infected and transfected cells. Tr., transcript marker. In panel A, the size of the expected protected fragment resulting from leader reversion to wild type is 141 nt. In panel B, the size of expected protected fragment resulting from leader reversion to wild type is 115 nt.

to replicate to the transfected reporter-containing N mRNA when it replaced the 5'-terminal 77 nt of the BCV N mRNA. The mechanistic contribution of the additional 433 nt in the DI RNA to the replication ability of the transfected molecule remains to be determined. Four possible functional elements that could be supplied by the additional sequence, thereby making the 5' terminal replication signal complete, are suggested by studies in other RNA replication systems.

(1) The minus-strand complement of the additional sequence may be providing a promoter or a portion of a promoter for plus-strand synthesis not supplied by the antileader alone. While promoters for some animal RNA viruses can be less than 20 nt in length (Levis *et al.*, 1986; Li and Palese, 1992), raising the possibility that the entire promoter for coronavirus plus-strand synthesis resides entirely within the 65 nt antileader, the full promoter may in fact be larger or even extend to discontinuous downstream regions. To date, all deletion mutations downstream of the leader destroy replication competence (Fig. 2.5). To map the 3' terminus of the replication signal or to define the full length of the promoter has not yet been completed. It is clear, however, that sequences mapping within discontinuous downstream regions of the polymerase gene and shown to be required for the replication of MHV-JHM DI RNAs (Kim *et al.*, 1993a; Lin and Lai, 1993), are not part of the BCV DI RNA replication signal. Moreover, such a sequence appears not to be a universal requirement for MHV DI RNA replication since a synthetic MHV DI RNA which shares a structure very similar to pDrep1 (Masters *et al.*, 1994) can replicate.

(2) The 5'-terminal plus-strand sequence extending downstream of the leader might be binding cellular or viral factors, protein or RNA, in a complex necessary for the initiation of plus-strand synthesis on the minus-strand template. Such functional complexes have been described for poliovirus (Andino *et al.*, 1993) and bromo mosaic virus (Pogue and Hall, 1992) RNA replication. It has also been shown for MHV that some cellular factors interact with the 5'-UTR regions (Furuya and Lai, 1994).

(3) The product of the DI ORF is required in *cis* for replication and will be described in more detail in Chapter III. To date, most out-of-frame mutations made at unique restriction endonuclease sites within the 2-kb pol-N ORF have destroyed replicating ability, suggesting there is a requirement in *cis* for the ORF product. Since the N ORF exists on the N mRNA, which does not replicate after transfection, the ORF requirement may reflect a qualitative need for the pol ORF product. The *cis*-acting requirement for coronavirus DI translation in replication remains a puzzle, since it seems not to be universal (de Groot *et al.*, 1992; Kim *et al.*, 1993b; Lin and Lai, 1993).

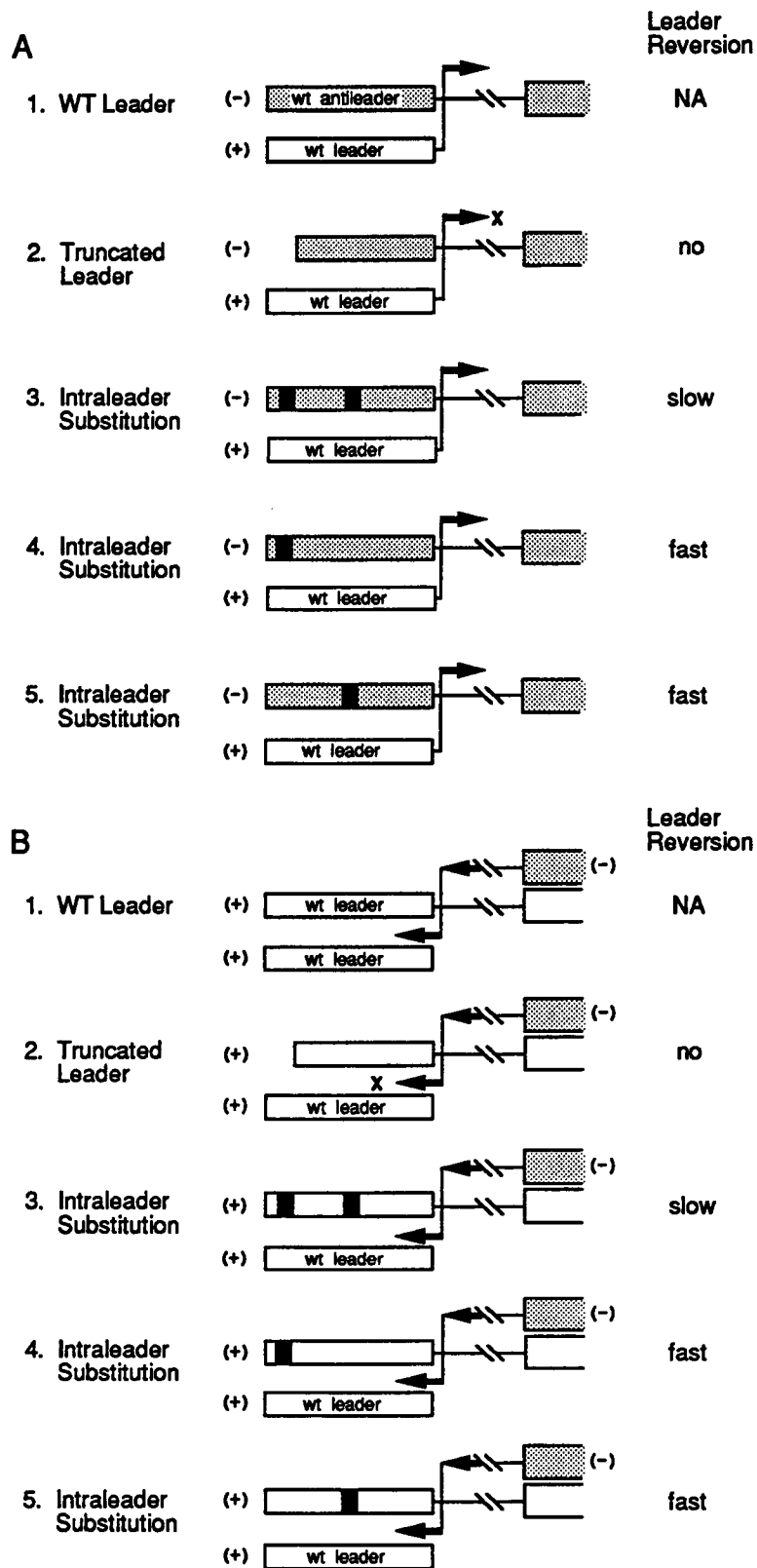
(4) The additional 433 nt may be providing a signaling function to carry the transfected molecule into a membrane-protected replication complex. It remains possible that coronavirus mRNAs do undergo replication during the virus multiplication cycle, but require a precise initiation event that can only occur within the replication apparatus. The transfected mRNA, therefore, may not be carried to the replication complex even when cotransfected with pDrep1 (Fig. 2.3E) if such a complex is in a physically

restricted compartment. Evidence for a membrane-protected replication complex in which roughly equimolar amounts of plus- and minus-strand counterparts exist (for both genome and subgenomic species) have been found (Sethna and Brian, unpublished data). The targeting function could conceivably reside within the unique part of DI RNA or possibly within the protein product of the truncated polymerase ORF on the replicating DI RNA.

Although the 65 nt leader is insufficient as a 5' terminal *cis*-acting signal for replication of the transfected mRNA molecule, it is a necessary part of the signal for DI RNA replication. Since the terminal deletions of 13 and 50 nt correlated with a loss in the predicted stem-loop 1 (-8 kcal/mole), it is postulated that this secondary structure is critical for the *cis*-acting replication signal. Surprisingly, base substitution mutations that disrupted and then restored stem-loop 1 appeared to have little or no effect on the replication of the DI RNA, presumably because of the rapid reversion back to wild type leader. This appears to have resulted from a high-frequency recombination event (Lai, 1992) that is reminiscent of the leader switching phenomenon described for MHV DI RNA (Makino and Lai, 1989a).

How does such a rapid and complete leader reversion occur mechanistically? Four mechanisms of RNA recombination have been described (Lai, 1992; Liao and Lai, 1992) that could be used to explain the leader reversion phenomenon. (1) Homologous recombination by template switching during positive-strand RNA synthesis (Fig. 2.11A). By this

Fig. 2.11 Two models for leader reversion by the mechanism of template switching. **A.** Model for template switching during plus-strand synthesis. During plus-strand synthesis, the wild type (wt) antileader on the helper virus genome, or possibly on antimRNAs of helper virus, would serve as template and polymerase switching would occur near the 5'-end of the antileader in a process of homologous recombination. If during RNA synthesis the plus-strand wt leader dissociates from its template, the putative "free" leader would have been produced. **B.** Model for template switching during minus-strand synthesis. During minus-strand synthesis, DI RNA would serve as template until the 3' terminus of the leader is reached by the polymerase. At that time the polymerase switches to wt leader on the genome, mRNA, or possibly a "free" leader in a process of homologous recombination. In both models A and B, the rate of reversion would be determined by the degree of homology between the two leader templates, and there would be a strict requirement for the 5' terminal sequence on the DI RNA leader. Leader and antileader sequences are depicted by closed rectangles. Solidly filled regions of leader or antileader represent regions of base substitutions. Regions of deletion (13 nt) and substitution are drawn to scale. Rates of leader reversion are deduced from the rate of DI RNA accumulation as compared to DI RNA with a wt leader (pDrep1).



mechanism (Lai, 1992), template switching by the polymerase would occur downstream of the leader on the minus-strand template at a time of polymerase pausing. It is pictured that the polymerase, complexed with nascent RNA and guided by sequence complementarity between the nascent RNA and template, switches templates and continues transcription. A variation of this mechanism has been formally proposed as the leader priming model of transcription for generation of coronavirus subgenomic mRNAs (Baker and Lai, 1990; Jeong and Makino, 1994; Makino *et al.*, 1986 and reviewed by Lai, 1990). In this model, template switching by a "free" polymerase-leader complex occurs at extremely high frequency and at more or less precise sites on the template. The leader priming mechanism has been used to explain 5' terminal leader switching between MHV genome and DI RNA (Makino and Lai, 1989a). (2) Homologous recombination by template switching during negative-strand synthesis (Fig. 2.11B). For leader reversion by this model, it is envisioned that the two positive-strand templates remain in close proximity and that template switching is precise and occurs near the UCUAAAC intergenic sequence. Leader reversion phenomenon could also be explained by (3) breakage and ligation (splicing) of the positive or (4) negative strand RNA.

Although none of these mechanisms has been rigorously ruled out, data described in this chapter are most consistent with a repair event utilizing homologous recombination by template switching during either plus- or minus- strand synthesis (Fig. 2.11). (1) In a system where it has been studied, the rate of homologous recombination by the postulated mechanism of template switching appears to be regulated by the degree of

sequence homology near the site of template switching (Lai, 1992; Zhang and Temin, 1994). In experiments described above, the most rapid accumulation of wild-type (reverted) leader occurred when intraleader nucleotide substitutions were made within a single region (pLM1 and pLM2), and a slower accumulation occurred when the two regions were mutated simultaneously (pLMC). The lower rate of accumulation presumably reflects a lower rate of leader reversion, although careful quantitative studies have not yet been done. (2) When the 5' terminal sequence of the leader is missing, no replication or reversion was observed. These results indicate that there is a strict requirement for terminal sequence in order for template switching to occur. Once a minimal structural requirement for leader is fulfilled for initiation of DI RNA replication, then the wild-type leader is strongly preferred and is selected for subsequent replication. Because of the rapid reversion of pLM1 and pLM2 in the absence of evidence for replicating mutated molecules, these results, however, leave unanswered the exact role played by the 5'-terminal portion of the leader in initiating replication.

It has been proposed that leader switching on the MHV DI RNA might arise from a leader-priming mechanism in which free leader anneals to the intergenic consensus sequence on the template just downstream of the genomic leader to prime replication (Baker and Lai, 1990; Liao and Lai, 1992; Makino and Lai, 1989). This is a variation of the leader-priming model for subgenomic transcription, in which the leader used for priming is provided in *trans* (Jeong and Makino, 1994; Kim *et al.*, 1993a; Makino *et al.*, 1986). If this is the mechanism of leader switching, one possible role for

the leader 5' terminus might be a necessary interaction with the leader-priming machinery. Further analysis of this must await a replicating DI RNA in which leader reversion (switching) does not occur. Such a DI RNA has been reported for MHV (Makino and Lai, 1989), and it is one in which a 9-nt sequence (UUUAUAAAC), occurring immediately downstream of the UCUAAC sequence, has been deleted. An identical sequence is found in an analogous region of the BCV DI RNA, but deletions made in this region abolished replication (data not shown). Therefore, leader reversion cannot be assessed with this mutant.

To date, deletions downstream of the leader sequence have all prevented replication with the exception of the pSM3 construct (Fig. 2.5). Interestingly, this mutant with one G insertion at base #207 resulting in generation of a unique *SalI* site, could still replicate and progeny viruses maintained the extra G at base #207. This result indicates that reversion does not happen with high frequency within this region. Deletion of three nucleotides (Fig. 2.5, pD206-208) at the 5' UTR abolished replication, suggesting that this region must contain part of the *cis*-acting signal for RNA replication.

Finally, perpetuation of a DI RNA may be aided in part by its ability to become packaged (Fosmire *et al.*, 1992; van der Most *et al.*, 1991). In this regard, the BCV DI RNA may differ mechanistically from the MHV DI RNAs, for which the packaging signal is described as being a portion of the 3' terminus of the polymerase gene that forms part of the DI RNA mosaic structure. This cannot be the case for the BCV DI RNA, or for the BCV

subgenomic mRNAs which also become efficiently packaged (Hofmann *et al.*, 1990), since neither possesses a sequence homologous to the identified packaging signal in MHV. The BCV DI packaging signal remains to be characterized.

CHAPTER III. TRANSLATIONAL REQUIREMENT IN *CIS* FOR BOVINE CORONAVIRUS DEFECTIVE-INTERFERING RNA REPLICATION

INTRODUCTION

Defective-interfering RNAs (DI RNAs) contain portions of the parental virus genome but are incapable of self-replication (reviewed by Roux *et al.*, 1991). For DI RNA replication, the viral specific RNA-dependent RNA polymerase must to be supplied in *trans* by the helper virus. However, DI RNAs contain the *cis*-acting signals required for RNA replication and sometimes encapsidation. All of the coronavirus defective interfering RNAs that has been analyzed so far contain an open reading frame from which DI RNA-specific protein(s) are translated, although the function of the DI-specific protein and the significance of the ORF are not known (Kim *et al.*, 1993b). For MHV DIssF, a 3.6 kilobases DI RNA, a single nucleotide deletion in the 5' region of the ORF resulted in a frame shift mutation and generated a shorter ORF that could still undergo replication (Kim *et al.*, 1993b). However, after one passage, the larger ORF was restored by RNA recombination during replication suggesting that DI RNAs containing a larger ORF have some advantage for accumulation (Kim *et al.*, 1993b). Experiments with another MHV DI, a 5.5 kilobase DI RNA, DI-a, also showed that truncation of the fusion ORF in its 3' terminal region was not lethal but led to reduced accumulation of DI RNA, and that the larger ORF

could be recovered after passage. This indicated that translation of the 3' end of MHV DI-a RNA is advantageous for replication or that the carboxyl end of the protein is required (de Groot *et al.*, 1992). Interestingly, the polypeptides encoded by DIssF and DI-a differ significantly in structure. Each has a different ORF because they are derived from different regions of the genome. These observations are difficult to fit into a model in which the products of DIssF and DI-a have identical functions.

BCV DI RNA, too, contains a single ORF (described in chapter II). In contrast to the DI RNAs described for MHV, the BCV DI RNA does not contain sequences derived from discontinuous internal regions of the genome. Polypeptides encoded by this DI RNA represent the N terminus of the RNA-dependent RNA polymerase (gene 1) fused in frame to the entire nucleocapsid (gene 7) as shown in Fig. 3.1A. The requirement for an open reading frame in the BCV DI RNA closely resembles that for the DI RNAs of poliovirus (Kuge *et al.*, 1986; Hagino-Yamagishi and Nomoto, 1989) and clover yellow mosaic virus (CYMV) (White *et al.*, 1992), both positive-strand RNA viruses. A sequence analysis of the RNA from DI particles derived from the Sabin strain of poliovirus revealed that deletions in these DI RNAs all occurred without changing the coding frame of the nucleotide sequence downstream from the deletion. This finding led Kuge *et al.* (1986) to propose that one or more of the viral nonstructural proteins is *cis*-acting, and, therefore, any genome in which the correct translation frame is not maintained would not survive (Kuge *et al.*, 1986; Hagino-Yamagishi and Nomoto, 1989). The structure of CYMV D RNAs is remarkably similar to that of the BCV DI RNA. It resulted from a simple fusion between the N

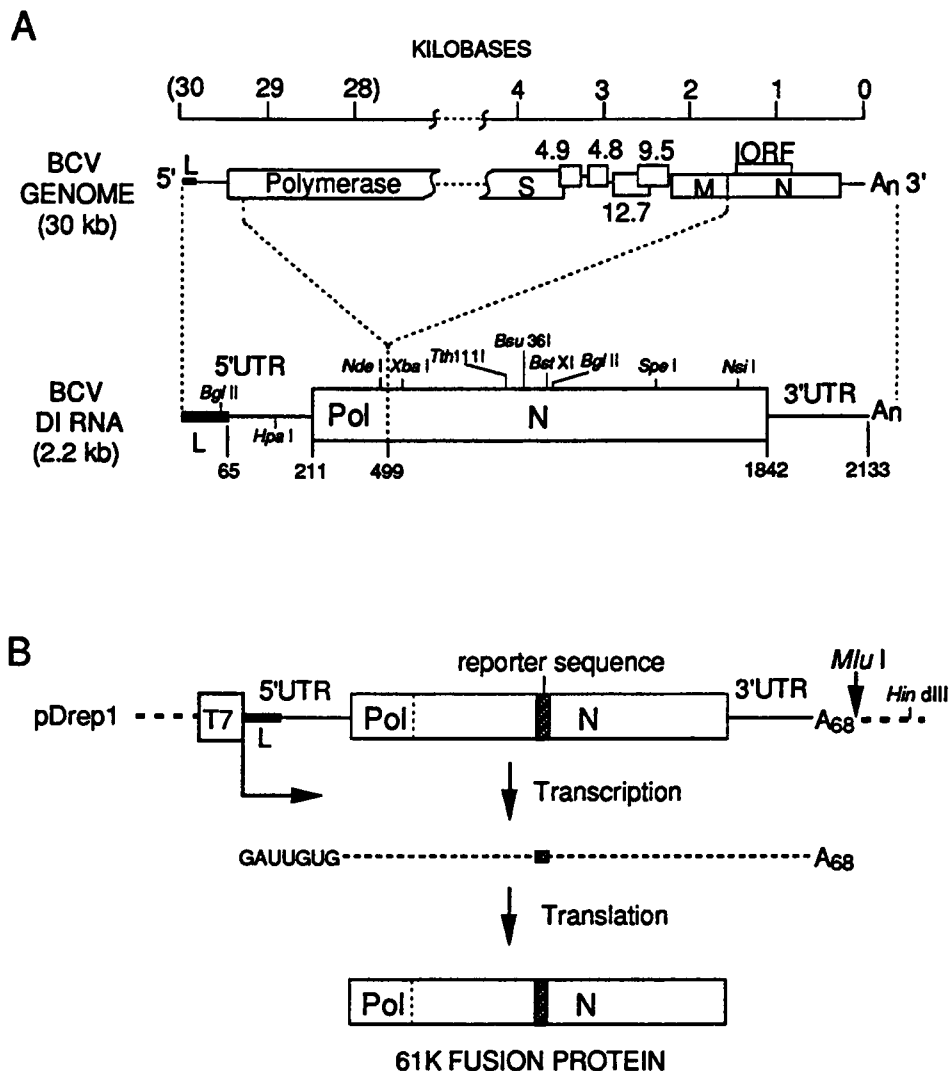


Fig. 3.1 Structure of BCV DI RNA and its fusion protein product. **A.** Structure of the BCV DI RNA relative to the BCV genome. Open bars denote open reading frames. The 5' and 3' untranslated regions are designated by a single line. Base 499 of the DI RNA is the first base in the N ORF which forms a contiguous ORF with the first 288 nt of the polymerase gene. Restriction endonuclease sites used in this study are shown. **B.** cDNA clone of BCV DI RNA was placed downstream of T7 promoter. 30 nt reporter sequence was inserted into *Bgl*II site. Transcripts of the pDrep1 construct contain the authentic 5' leader sequence and a 3' poly A tail. A 61-kD Pol-N fusion protein is produced by *in vitro* translation.

terminus of the 191-kD nonstructural protein (the putative replicase) gene and the C terminus of the 23-kD coat protein gene. Accumulation of the CYMV D RNA *in vivo* depends on the maintenance of its encoded fusion ORF (White *et al.*, 1992).

To test the idea that conservation of the fusion ORF within the BCV DI RNA is important for its replication, a full-length cDNA clone of the BCV DI RNA with a reporter sequence (pDrep1) was used for *in vivo* replication analysis (Fig. 3.1B). Mutations that included base substitutions and frame shifts were made throughout the ORF and the results of these changes on the replication of pDrep1 indicate that the ORF in the BCV DI RNA is required *in cis* for replication.

MATERIALS AND METHODS

Synthetic oligonucleotides.

Oligonucleotides used in this study are listed in Table 3.1.

Construction of plasmids.

To convert the start codon of the pol-ORF within pDrep1 to TTG, pDM1 was generated by using overlapping PCR mutagenesis (Horton *et al.*, 1990). The gel purified 219 nt PCR product from a pDrep1-templated reaction using leader(-) and DM(+) primers was used with the gel purified 790 nt product from a reaction using DM(-) and TGEV#8(+) primers in an overlap PCR. In the overlap, the 219 nt and 790 nt fragments were used together

Table 3.1 Synthetic oligonucleotides used in this study

Oligonucleotide ^a	Binding site ^b	Sequence ^c
leader(-)	binds to anti-sense leader nt 7-32	5' GAGCGATTGCGTGCATCCCGC 3'
DM(+)	binds to plus-strand nt 204-225	5' GTTGATCTTCGACAAATGTGACC 3'
DM(-)	binds to minus-strand of nt 204-225	5' GGTCACAATGTGGAAGATCAAC 3'
TGEV#8(+)	nt 338-363 in the TGEV N ORF used as a reporter; binds to the plus strand	5' CATGGCACCATCCTTGGCAACCCAGA 3'
TGEV#7(-)	binds to the minus strand of reporter sequence	5' TCTGGGTGCCCAAGGATGGTCCCATG 3'
NM(+)	binds to plus-strand nt 486-509	5' GTAAAAGACAACTTTCTTGAAGT 3'
NM(-)	binds to minus-strand of nt 486-509	5' ACTTCAAGAAAGGTTGCTCTTTTAC 3'
18S(+)	binds to plus-strand of human 18S rRNA base #595-620 (Torczynski <i>et al.</i> , 1985)	5' CTGCTGGCACCAAGACTTGCCCTCCAA 3'

- a. Oligonucleotides bind either to plus-sense RNA as indicated by (+), or to minus-sense RNA as indicated by (-).
- b. Numbers (except for those describing the TGEV oligonucleotides) correspond to pDrep1 plus-strand sequence. TGEV numbers come from Kapke and Brian, 1986.
- c. The underlined sequence indicates a mutated region.

with leader(-) and TGEV#8(+) primers to form a product of 987 nt, from which the 443 nt *HpaI-XbaI* insert fragment was used to replace the analogous region in pDrep1, generating pDM1.

pNM1, in which the start codon of the N ORF was converted to TTG, was similarly generated except that the overlap PCR used the 503 nt fragment from a reaction with the leader(-) and NM(+) primers, the 508 nt fragment from a reaction with NM(-) and TGEV#8(+) primers, and the leader(-) and TGEV#8(+) primers. The fragment generated was 987 nt. From this, the 443 nt *HpaI-XbaI* fragment was used to replace the analogous region in pDrep1 to generate pNM1. pNMTG, a construct with a one base deletion in the start codon of the N ORF which resulted in an out-of-frame mutation, was unexpectedly generated during PCR mutagenesis for pNM1 and this was tested for the effect of out-of-frame mutation of pol-N ORF in RNA replication. The out-of-frame mutation in the pol-N ORF of pNMTG caused pol to be in-frame with the I ORF, a second protein synthesized from within the N gene sequence (Senanayake *et al.*, 1992).

To create pDIN, a construct in which the C-terminal portion of the N ORF was put back in-frame with the I ORF (described in Results), a two step manipulation was done. (1) pD92, a construct identical to pDrep1 except that its 5' terminal 92 nt, and therefore also its leader with an intraleader *BglII* site, are deleted (not shown), was cut with *BglII* which cuts just upstream of the TGEV reporter (Fig. 3.1). The cut plasmid was then filled-in with klenow enzyme, which adds 4 nt and makes the site blunt, was religated to form pD92B (not shown). (2) The *XbaI-HindIII*

fragment from pD92B containing the 4 nt insertion was used to replace the analogous fragment in pNMTG, this generating pDIN. The fidelity of all new constructs was checked by sequencing the junctions at the ligation sites and *in vitro* translation of the RNA transcripts.

Constructs pD930-1047, pD930-1046, and pD930-1067, containing internal deletions of 118, 117, and 136 nt respectively, were generated by linearization of pDrep1 at the *Tth*111I site (position 925) followed by filling in with T4 DNA polymerase, and then digestion of the linearized fragment with *Bsu*36I (position 1045) and blunt ending with mung bean nuclease. The vector-containing large fragment was purified by agarose gel electrophoresis and religated. pD930-1046 had the desired 117 nt in-frame deletion. pD930-1047 (out of frame) and pD930-1067 (in frame) were generated as a result of overdigestion with mung bean nuclease and were tested for replication.

To introduce premature stop codons within the pDrep1 ORF, six unique restriction sites were used (described in Results). Linearized pDrep1 DNA was blunt-ended by using either T4 DNA polymerase, which fills in 5' overhangs and removes 3' protruding tails, or mung bean nuclease, which removes 5' overhangs. Modifying enzymes were selected depending on which modification would result in placement of the stop codon closest to the restriction site. Blunt-ended DNAs were religated and each mutated region was confirmed by sequencing. pN4d construct is a four base deletion of pDrep1 at the *Nsi*I site (nt 1662-1665) which resulted in the ORF stopping at base #1687. pN5d construct was selected during the

preparation of pN4d and it has a five base deletion at the *Nsi*I site (nt 1661-1665) causing the ORF to end at base #1748. pS4i construct has a four base insertion at the *Spe*I site resulting in the ORF stopping at base #1445. pB4d, pT1i, pX4i, and pNd2d with modifications at the *Bst*XI, *Tth*111I, *Xba*I, and *Nde*I sites resulted in each ORF stops at bases #1213, 947, 599, and 473, respectively.

Assay for replication (Northern analysis).

The replication of each mutant was analyzed by Northern hybridization as described in Chapter II.

In vitro translation.

In vitro translation of the transcripts was performed according to the manufacturer's protocol (Promega Biotec) in a 50 µl reaction volume using 100 units of RNasin, 1.17mM potassium acetate, 0.02 mM amino acid solution (without methionine), 1mCi/ml of ³⁵S methionine (ICN-1000 Ci/mmol), 1-2 µg transcripts, and 25 µl of wheat germ extract. The reactions were performed at 25°C for 60 minutes and stopped by the addition of an equal volume of 2X Laemmli standard treatment buffer (STB) (1X STB contains 0.0625 M Tris buffer pH 6.8, 2% SDS, 10% glycerol and 10 µg bromophenol blue). The products were analyzed on a 10% SDS-polyacrylamide gel according to the procedure of Laemmli (1970).

Measurement of mRNA stability.

RNA transcripts of pDrep1, pDIN, and pT1i were generated *in vitro* using T7 RNA polymerase and prepared for transfection as described (Chapter II). RNA transcripts were transfected into uninfected cells with lipofectin and cytoplasmic RNA was extracted as described in the Results. Northern analysis was performed using end-labeled oligonucleotide TGEV#8(+) and quantitation was performed as described in Chapter II. For quantitation of 18S rRNA, end-labeled probe 18S(+) was used on the same blot after total decay of the TGEV#8(+) probe. Quantitation was done with the Ambis Photoanalytic Imaging System.

RESULTS

The single fusion protein within the BCV DI RNA is required for RNA replication.

BCV DI RNA is a simple fusion between the N-terminal portion of the RNA-dependent RNA polymerase gene and the entire N ORF resulting in a single fused ORF (Fig. 3.1A). To test whether this ORF is necessary for RNA replication, site-specific point mutations were introduced in to pDrep1 (Fig. 3.1B) and each mutant was tested for replication under the same conditions that consistently gave positive results for pDrep1 replication (Chapter II). pDM1 contains a point mutation which converts the start codon ATG of the pol-N fusion protein to TTG so that the first available start

codon is that for the N ORF (Fig. 3.2). This single point mutation destroys the replication of the DI molecule suggesting that translation of this fusion protein is needed for BCV DI RNA replication. When the same mutation procedure was carried out independently on the N ORF, that is, the start codon for the N gene was mutated into TTG, the resulting mutant, pNM1, could still function as a replicon (Fig. 3.2). This was not unexpected, however, since the quality of the fused protein was nearly the same as in pDrep1. A construct called pNMTG (Fig. 3.2) in which the A in the start codon of the N gene was deleted, had a frame shift in the ORF such that the pol ORF became in frame with the I protein gene, an internal open reading frame within the N gene (Senanayake *et al.*, 1992). In pNMTG, no part of the N protein is predicted to be made and transcripts of this construct did not replicate. To test whether the replication of pNMTG could be rescued by restoring the continuity of the C-terminal portion of the N protein with the fusion ORF, a 4 base insertion at the *Bgl*III site in front of the TGEV reporter was made. This construct was named pDIN and resulted in the synthesis of a chimeric protein which is comprised of the N-terminal portion from pol ORF, the middle domain from I protein, and the C-terminal domain from N protein (Fig. 3.2). Despite replacement of the C-terminal domain of the N protein, however, pDIN still was unable to replicate. This result indicated that some quality of the pol-N fusion protein itself, rather than the action of translation through the C-terminal of N protein, is required for replication. Furthermore, it became apparent that since all of these constructs were transfected into HRT cells containing endogenous DI RNAs that there was no rescue of the replication ability of the mutants. These results suggested

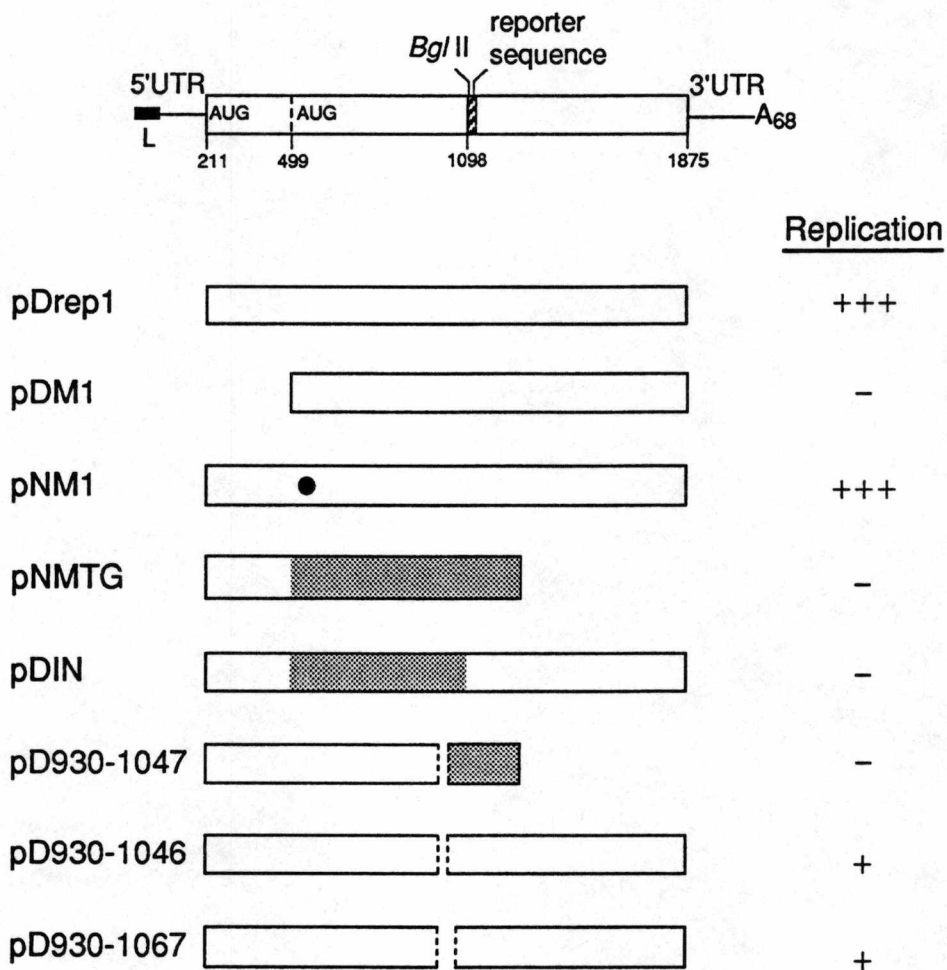


Fig. 3.2 Effect of ORF mutations on the replication of pDrep1 RNA. Translation products of each mutant are shown as open (maintaining the fused pol-N ORF) or shaded (the fused I ORF) boxes. The black circle indicates a point substitution. Internal deletions are shown as a space. Evidence of replication was based on increased amounts of plus-strand RNA after 24h posttransfection and on the presence of plus-strand RNA at 48h postinfection from passage 1 progeny virus. One plus sign (+) indicates that replicational efficiency was about 10-30% of pDrep1 level. Three plus (+++) indicates that the replicational efficiency was as good as pDrep1. Minus sign (-) denotes no plus-strand RNA could be detected after 24 h posttransfection.

that the fusion product must not act in *trans* to aid the replication of the transfected constructs. The striking difference in replication level among these mutants suggested that the fusion protein must play an important role in BCV DI RNA replication and is apparently needed in *cis* .

Out-of-frame mutations abolish replication while in-frame deletions or insertions do not.

Several in-frame mutations of pDrep1 have been made that still replicate. Insertion of the 30 nt reporter sequence that could be translated into 10 extra amino acids in pDrep1 was the first example of an in-frame insertion that could be tolerated. An in-frame intergenic sequence of 27 nt was cloned into pDrep1 at the *Bsu*36I site for the study of subgenomic transcription (Krishnan *et al.*, unpublished data) and this construct, named pDrepIS, could still function as a replicon. These results indicate that small in-frame changes can be made within pDrep1 and have little or no effect on replication. Three internal deletion mutations were also tested for effects on replication. pD930-1046 and pD930-1067 that contain in-frame deletions of 39 and 46 amino acids, respectively, still replicated but showed some impairment of replication (Fig. 3.2). On the other hand, pD930-1047 which contains an out-of-frame internal deletion resulting in termination of the ORF at base #1213, lost all ability to replicate. These results indicate that maintenance of the ORF is necessary for BCV DI RNA replication and that at least some changes in the quality of the ORF product can be tolerated when fulfilling the *cis*-acting function.

To make certain that all of the mutants maintained the predicted ORF, transcripts from the individual mutants were used as templates for *in vitro* translation in the presence of wheat germ extract. The product of pDrep1 translation was found in two forms, a pol-N fusion protein of 61 kD and an N protein of 50 kD (Fig. 3.3, lane 2). The N protein was produced from the pDM1 construct as predicted (Fig. 3.3, lane 3). However, for pDM1 two proteins appearing smaller than the pol-N fusion protein were also observed on the gel. They are probably translated from downstream start codons that map at bases #268 and #454. Interestingly, when mutations were made on the N start codon, translational efficiency of the pol-N fusion protein increased dramatically (Fig. 3.3, lanes 4 through 6 compared to lane 2) suggesting that translation of pol-N fusion protein might be down-regulated by the N protein product. The internal deletion mutants generated proteins of the predicted sizes (Fig. 3.3 lanes 7 through 9), indicating that there were no other unexpected mutations generated during the cloning procedures. The impairment of replication, therefore, is probably only the result of those modifications that we have described.

Taken together, all mutations which disrupted the ORF, such as base changes at the start codon or out-of-the frame mutations within the ORF, showed no evidence of replication, while in-frame deletions or insertions still maintained a certain level of replication, suggesting that the fusion protein or the ORF, or both encoded by the BCV DI RNA are absolutely required for replication.

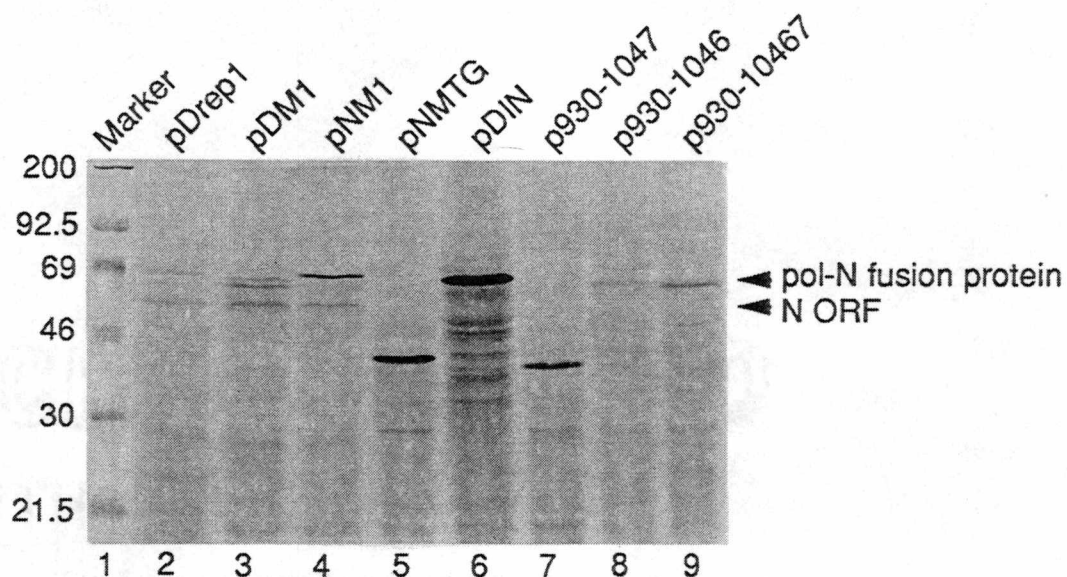


Fig. 3.3 SDS-polyacrylamide gel analysis of *in vitro* translation products of pDrep1 and its derivative mutants. Sequential transcription and translation of cloned cDNAs were carried out as described under Materials and Methods. Molecular weight markers are shown in lane 1, with sizes in kilodaltons indicated on the left. The positions of 61 kD pol-N fusion protein and 50 kD nucleocapsid (N) protein are shown on the right.

Truncating the ORF at various sites shows that almost the entire DI ORF is a *cis*-translationally required region for BCV DI RNA replication.

To localize the minimum ORF requirement within pDrep1 for replication, premature stop codons were introduced within the ORF by cutting at six separate unique restriction endonuclease sites, blunt-ending, religating, and testing transcripts of the resultant constructs for replication. Each blunt-ended site gave a shorter ORF than wild type pDrep1 ORF and only changed the sequence by less than 4 bases within the RNA transcript. Most of the mutants with a truncated DI RNA ORF were disrupted in their replication (Fig. 3.4). Only the construct that was blunt-ended at the *Nsi*I site of pDrep1, resulting in a termination of the ORF at a site 62 aa upstream from the carboxyl terminus, still replicated. However, by comparison with wild type pDrep1, the replication of this construct was much less efficient. Translation results for each of the mutants is shown in Fig. 3.5. Truncated proteins of the predicted size were observed for each mutant. These experiments indicated that almost the entire DI ORF is required in *cis* for BCV DI RNA replication.

Translational requirement in *cis* for RNA replication is not due to an increase in RNA stability conferred by translation.

One possible mechanism for the observed dependence of viral RNA replication on translation in *cis* is that the absence of translating ribosomes beyond a certain position in the coding region might result in rapid mRNA decay (Sachs, 1993; Vancanneyt *et al.*, 1990). To test whether the translational requirement in *cis* is due to differences in RNA stabilities,

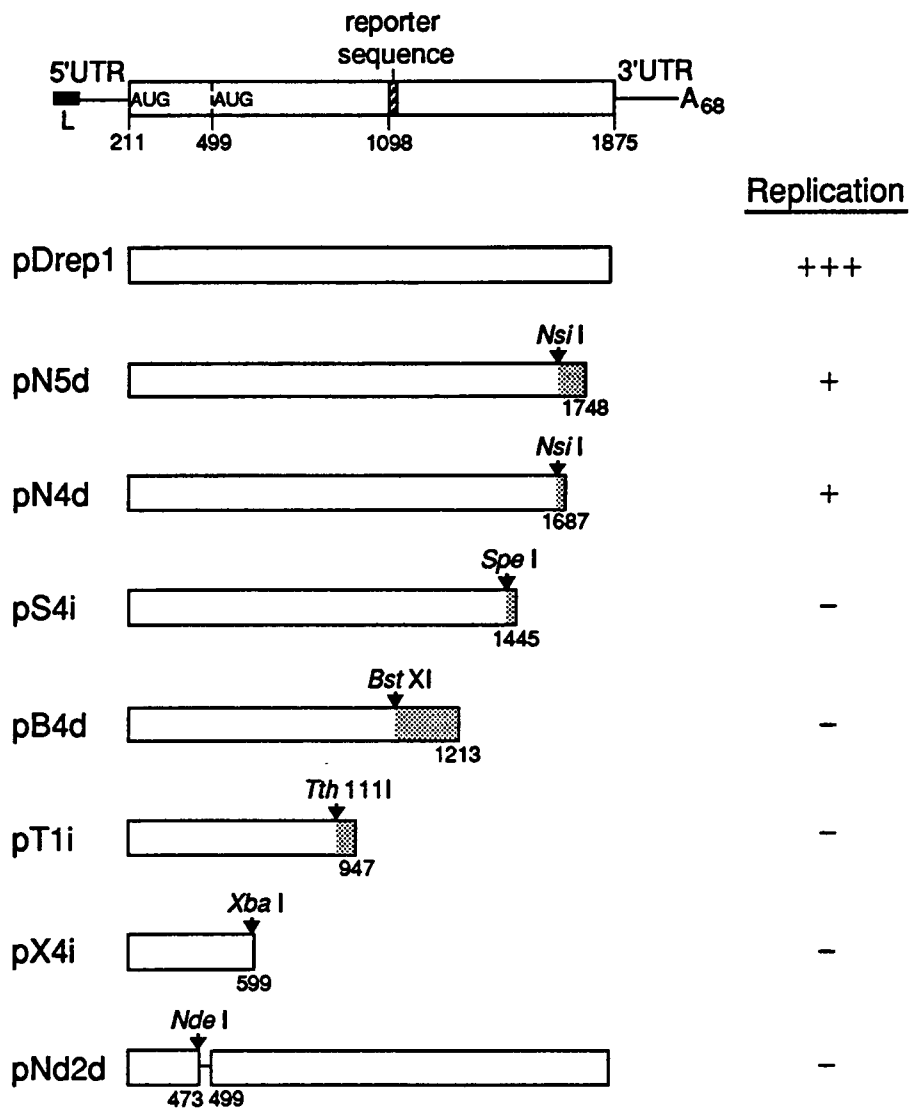


Fig. 3.4 Effect of premature stop codons on the replication of pDrep1 RNA. Translation products of each mutant are shown as open (maintaining the fused pol-N ORF) or shaded (representing non-N sequence) boxes. pN5d was generated by a 5 nt deletion, pN4d by a 4 nt deletion, pT1i by a 1 nt insertion, pX4i by a 4 nt insertion, and pNd2d by a 2 nt deletion. Replication efficiency was measured as described in Fig. 3.2.

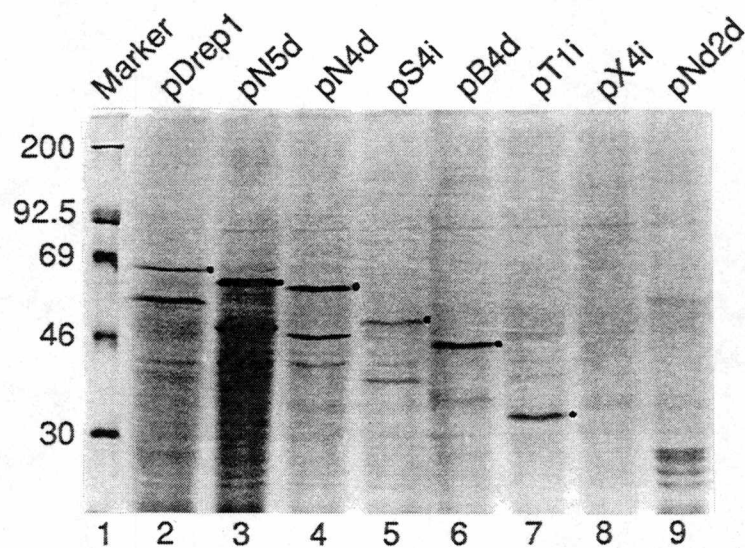


Fig. 3.5 SDS-polyacrylamide gel analysis of *in vitro* translation products of pDrep1 truncated mutants. Molecular weight markers are shown in lane 1. The position of the expected product is indicated by a dot. For lanes 8 and 9, the expected product are less than 15 kDa and were therefore not retained on the gel.

pDrep1, pDIN, and pT1i RNA transcripts were tested for stability after transfection into uninfected cells. The decrease in the amount of transfected RNAs was shown by Northern and quantified by the Ambis Photoanalytic Imager. pDrep1 and pDIN encode ORFs of the same size but encode different amino acids in the central region of the ORF (Fig. 3.2). pT1i encodes a truncated protein (Fig. 3.4). All of these RNA transcripts have similar decay rates (Fig. 3.6) although pDrep1 shows slightly more stability than pDIN and pT1i, indicating that the requirement for translation in *cis* for BCV DI RNA replication does not result from an increase in RNA stability.

DISCUSSION

To date, all mutations which truncate the 1665 nt pol-N fusion ORF of pDrep1 within its 5' proximal 1477 nt (5' proximal 89%) region abolish replication of the RNA. Small changes as a result of in-frame insertions or deletions yield a construct still showing some degree of replication function. These experiments indicate that the fusion ORF encoded by the BCV DI RNA plays an important role for RNA synthesis and is needed in *cis*. The requirement of a fusion ORF indicates that either the encoded product, or the process of translation, or both are essential for BCV DI RNA replication. The results of experiments described in this study suggest that both are required for RNA replication. Information suggesting that the quality of the protein product is important comes from a study of mutant pDIN. The pDIN construct which encodes a protein of the same size as

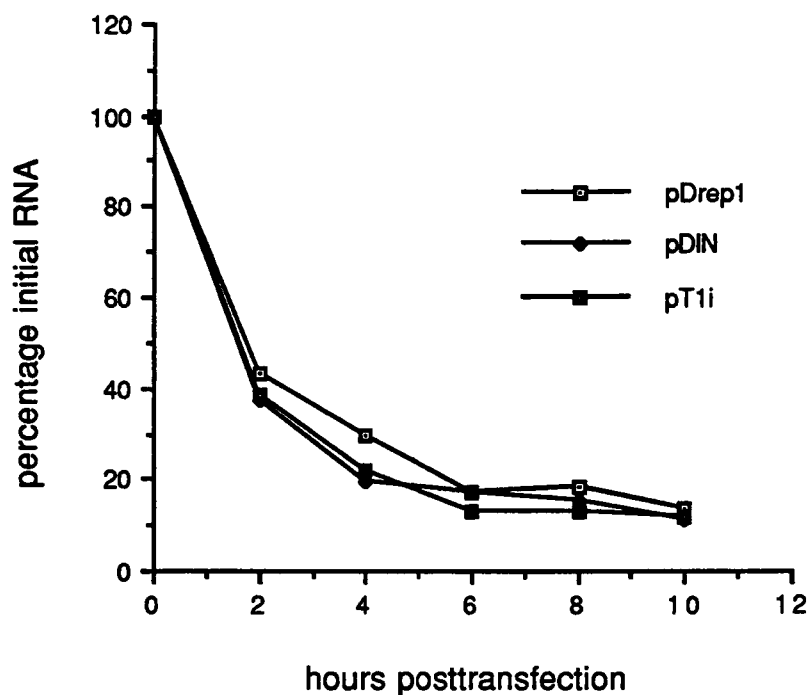


Fig. 3.6 Stability of pDrep1, pDIN and pT1i RNA transcripts. 200 ng of RNA transcripts were transfected into uninfected HRT cells. Cytoplasmic RNA was extracted at 0, 2, 4, 6, 8, and 10 hours posttransfection and analyzed to detect plus-strand reporter sequence. Amounts of RNA were normalized to the amount of 18S rRNA in each sample. Amounts of each RNA were plotted as the percentage of the amount of RNA present at 0 hr posttransfection.

pDrep1 but which has a different coding sequence in the central domain (Fig. 3.2) showed no evidence of replication, suggesting that the encoded product in addition to the act of translation is required for BCV DI RNA replication.

The results of ORF truncation experiments suggest that the process of translation itself is important for RNA replication. The translational requirement in *cis* for viral replication could be explained by the action of the associated ribosomes in the translation process. For example, the binding of ribosomes could unfold the secondary structure of the template and thereby facilitate polymerase binding for minus-strand synthesis. It could be that ribosomes are involved in binding of cofactors required for replication. Ribosomes, on the other hand, may be merely contributing to the stability of the mRNA molecule in eukaryotic cells. Destabilization of incompletely translated RNAs supplied evidence for a correlation between mRNA decay and ribosome loading (Sachs, 1993; Vancanneyt *et al.*, 1990). However, this explanation could not be applied here because different ORFs were tested and shown no significantly difference in RNA stability (Fig.3.6).

We propose that one possible mechanism might be that binding of ribosomes to the RNA molecules and subsequent translation of the entire fusion ORF is required for targeting the RNA molecule to the membranous replication complex (Fig. 3.7). A possible mechanism for this is that the nucleocapsid protein binds to RNA, specifically with the leader (Stohlman *et al.*, 1988), and also interacts with the pol portion of the fusion protein. The nucleocapsid domain of the pol-N fusion protein is possibly required for

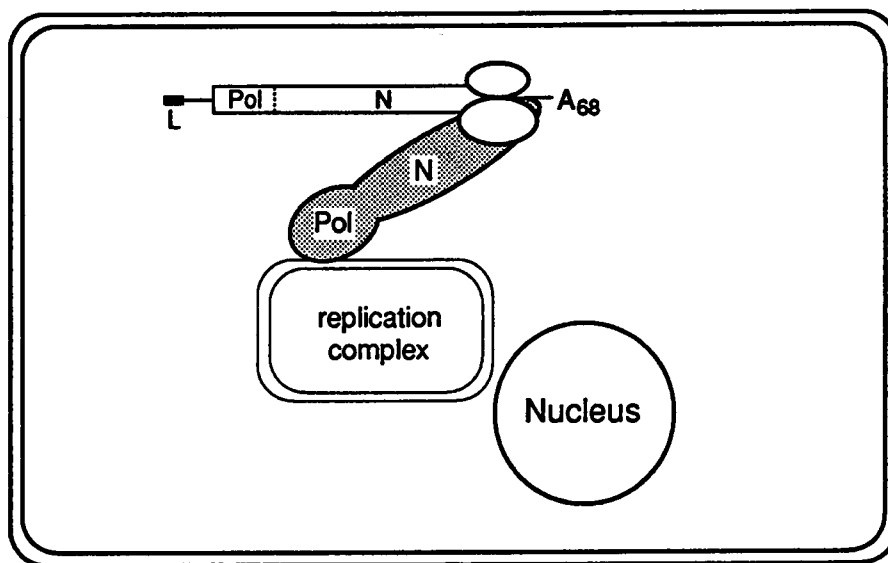


Fig. 3.7 Model for the *cis*-acting function of the pDrep1 fusion protein. pDrep1 RNA with an intact leader and poly A tail is shown. The fusion ORF is indicated as an open box. Ribosomes are depicted as two ovals. The fusion protein product is shown as a shaded structure. The replication complex is illustrated as a membrane-protected compartment. The model predicts that the pol portion of the pol-N fusion protein is made and then targets the protein-pDrep1 RNA complex to the membranous replication complex. Once inside the replication complex, pDrep1 can undergo replication.

binding of the template, and the pol domain is possibly responsible for targeting pDrep1 to the replication complex. Some other host or viral factors might also be involved in targeting to the membrane since hydrophobicity analysis within all of the pol-N ORF did not reveal a transmembrane domain. *Cis*-acting sequences for the BCV DI RNA replication may reflect a requirement in the genome for replication. However, a *cis*-translational requirement may not necessarily take the same form as found in the DI RNA since a fusion protein does not exist in the genome. The polymerase protein, of course, does exist as a natural viral protein and may provide the targeting function for the genome. Thus, the replication of pDrep1 may be useful in determining the targeting function of the polymerase protein.

The requirement in *cis* for an RNA replicon to be translated in order to replicate is not universal. Most DI RNAs of influenza virus arise by a single internal deletion, and only few of them have DI-specific polypeptides. By mixed infection of different DI RNAs of influenza viruses, it has been shown that those DIs producing detectable fusion polypeptides did not have a selective advantage (Akkina *et al.*, 1984). Maintenance or disruption of the open reading frame for protein α of flock house virus, a positive-strand RNA virus of insects, did not affect the replication ability of the defective RNA2 molecules (Ball and Li, 1993). In plant viruses such as tomato bushy stunt virus (TBSV) or turnip crinkle virus (TCV) DI RNAs, there is no evidence for DI-specific translation products, yet these DI RNAs replicate efficiently (Hillman *et al.*, 1987; Li *et al.*, 1989). However, evidence suggesting a requirement for translation in *cis* has also been reported for

other positive-strand RNA viruses: (1) Mutations in the 2B nonstructural protein of poliovirus are noncomplementable in *trans* (Johnson and Sarnow, 1991). Recent studies show that the CTR (*cis*-translational required) region of poliovirus genome maps between the 2A and 3C nonstructural viral genes (Novak and Kirkegaard, 1994). (2) The largest ORF of the turnip yellow mosaic virus RNA genome encodes a 206 kD protein that is cleaved to yield N-terminal 150 kD and C-terminal 70 kD proteins. These two proteins are required in *cis* to form a complex with the 3' end of the RNA for replication (Weiland and Dreher, 1993). (3) Naturally occurring defective RNAs derived from the potexvirus clover yellow mosaic virus contain a single fusion ORF which is required in *cis* for replication (White *et al.*, 1992). (4) Cowpea mosaic virus B-RNA encodes proteins required for viral RNA replication. Amplification of B-RNA itself requires translation in *cis* (Bokhoven *et al.*, 1993). Common features among these examples of *cis*-translational requirement are: (1) they are all plus-sense RNA viruses and replicate in the cytoplasm, and (2) the product of the *cis*-required protein all contains part of the RNA-dependent RNA polymerase which is required for viral RNA synthesis, suggesting that translation is potentially coupled to replication. For plus-strand RNA viruses, templates are subject to both translation and replication. These two events must be closely associated and highly regulated.

Why should an RNA virus require translation in *cis*? *Cis*-translational requirement could be a substantial selective pressure evolved for late proofreading (Novak and Kirkegaard, 1994). The RNA-dependent RNA polymerase does not have proof reading activity, and mutations are

known to happen frequently during viral replication. The protein required in *cis* is related to the function of the RNA-dependent RNA polymerase which is crucial for viability and could not afford to be mutated. Whatever the mutation is that results in a shorter ORF could not survive since function of these proteins is important for viral replication. This form of late proofreading may help many plus-sense RNA viruses to be less lethal by coupling replication and translation because those mutants without functional proteins being translated properly could not survive. Although the actual mechanism for the *cis*-translational requirement is not known, it appears that many plus-strand RNA viruses may have evolved this strategy to overcome the lack of proof reading mechanism by the RNA-dependent RNA polymerase.

CHAPTER IV. FUTURE DIRECTION FOR USE OF THE CLONED DEFECTIVE-INTERFERING RNA IN THE ANALYSIS OF CORONAVIRUS RNA REPLICATION AND GENE EXPRESSION

In this dissertation, *cis*-acting sequence requirements for the replication of bovine coronavirus were examined by using a cloned defective-interfering RNA. As a result of this study, the bovine coronavirus DI RNA containing a reporter sequence, pDrep1, has been established as a molecule with which to study precisely many of the mechanisms involved in RNA replication and subgenomic mRNA transcription. Further interesting questions can now be asked regarding the unique genome organization, replication and transcription strategies of coronaviruses.

TO DETERMINE THE SECONDARY STRUCTURE OF THE 5'-END OF THE BCV GENOME USING ENZYME PROBING

The computer predicted secondary structure (Tinoco algorithm) of the 5'-end of the genome reveals two stem-loops within the leader. Truncation deletion analysis of the 5'-end shows that the first stem-loop is important for replication. However, base substitutions to destroy the stem-loop or compensatory changes to restore the stem-loop structure did not abolish replicational competence and the mutated leader was reverted into wild type. The role of the first stem-loop structure remains undetermined. To

reexamine the secondary structure of the leader, enzyme probing within this region needs to be carried out. Secondary structure could be determined using specific ribonucleases to digest the RNA molecule. For example, cobra venom ribonuclease V1 cuts double-stranded regions and ribonucleases T1, T2, and A specifically cleave single-stranded regions. The digested products can be analyzed on a sequencing gel. The secondary structure of the 5'-end can be determined.

TO LOCALIZE THE CROSS-OVER SITE DURING LEADER SWITCHING

Mutations made upstream of the leader consensus UCUAAAC sequence appear to be lost by leader reversion into wild type leader. Most mutations made downstream of the consensus sequence but within the 5'-UTR abolish replicational competence. We have learned that a four base insertion at the *AvrII* site (R. Krishnan, unpublished data) and a one base insertion at base #207 (Chapter II) within the 5' UTR in pDrep1 each allows the RNA to be maintained as a replicon without reversion to wild type sequence. This suggests that the high rate of RNA recombination found in the MHV genome may not apply to BCV. Leader switching appears to be an important feature for coronavirus RNA replication, and this appears to be a special case of high frequency homologous site-specific recombination common to both BCV and MHV, and probably all coronaviruses. Experiments described in Chapter II showed that the leader can be provided in *trans* during replication, suggesting that a leader priming of RNA replication may exist. The exact cross-over site needs to be

determined. This should help to determine the sequence determinants of the reversion event. We predict the cross-over site for leader switching is probably located downstream of the consensus UCUAAAC sequence. This consensus sequence is within the loop region of the second stem-loop structure (Fig. 2.6). Whether this stem-loop plays a role in leader switching needs to be determined. A systematic site-directed mutagenesis destroying the second stem loop and the effect on leader switching should be carried out.

TO CHARACTERIZE PROTEINS INVOLVED IN CORONAVIRUS RNA REPLICATION

Experiments to measure the binding of cellular or viral proteins will be important for establishing the composition of the complexes necessary for initiating plus-strand synthesis on minus-strand RNA templates and minus-strand synthesis on plus-strand RNA templates. Identification of these cellular or viral factors will be essential for understanding the replication mechanism of coronaviruses. Gel retardation assays and ultraviolet (UV) crosslinking assays can be carried out with RNA sequences from the 5'-UTR and the 3'-UTR regions of the BCV DI RNA and proteins from infected or uninfected cells.

TO IDENTIFY THE PACKAGING SIGNAL OF BCV

The packaging signal of MHV has been identified (Makino *et al.*, 1990; Fosmire *et al.*, 1992; van der Most *et al.*, 1991) with a specific stem-loop region at the 3' end of the polymerase gene. The MHV DI failed to be packaged after site-directed mutagenesis destroyed the secondary structure. Maintaining the putative stem-loop did not affect the packaging efficiency. Two pieces of evidence show that the location of BCV packaging signal is different from that for MHV. (1) For MHV, a signal required for packaging of the DI RNA into virions was found to map at 20,356 to 20,416 nucleotides from the 5' end of genomic RNA (within 1,381 to 1,441 nucleotides upstream from the 3' end of the polymerase gene) (Fosmire *et al.*, 1992). No equivalent stretch of nucleotides is found in the BCV DI RNA, yet it is packaged. (2) Packaging of mRNAs has been reported for BCV (Hofmann *et al.*, 1990). The N and M proteins are found abundantly within BCV virions, while subgenomic RNAs are not packaged in MHV. This result indicates that the BCV encapsidation signal could be located within the leader sequence or within the 3' end of the genome, which is present in each subgenomic RNA. The actual packaging signal of the BCV RNA needs to be characterized.

TO FURTHER EXAMINE THE C/S-TRANSLATIONAL REQUIREMENT FOR BCV DI RNA REPLICATION

Experiments described in Chapter III showed that the translation of the DI open reading frame is required for BCV DI RNA replication, suggesting that translation is coupled to RNA replication. The reason of this coupling is not known. Termination of translation theoretically would stop RNA replication if these two events are associated. A testable experiment could be performed by putting translation inhibitors and measuring RNA accumulation.

Does pol-N fusion protein bind to the BCV DI RNA? This fusion protein can be cloned into an *E. coli* expression vector to produce large amounts of protein. The gel shift and UV crosslinking assays could be carried out to test for this protein-RNA interaction.

TO EXPLORE THE POTENTIAL OF pDrep1 AS AN EXPRESSION VECTOR

The fact that pDrep1 can replicate with a foreign sequence, potentially express the foreign sequence, and become packaged into virions, opens up the exciting possibility that the bovine coronavirus could be used as an expression vector with immediate medical usefulness. Many questions are therefore open for investigation: (1) Can a foreign epitope be expressed in infected cells transfected with pDrep1 or infected via a pDrep1-carrying

virus? (2) Can the foreign epitope be used to stimulate mucosal immunity or other desired response? (3) Can a cytokine gene (~600 nt) be expressed from the packaged pDrep1 and therefore modulate the infection or immune response? These and many other questions can now be explored with this characterized coronavirus replicon.

REFERENCES

REFERENCES

- Abraham, S., T. E. Kienzle, W. Lapps, and D. Brian. 1990a. Deduced sequence of the bovine coronavirus spike protein and identification of the internal proteolytic cleavage site. *Virology* 176: 296-301.
- Abraham, S., T. E. Kienzle, W. Lapps, and D. Brian. 1990b. Sequence and expression analysis of potential nonstructural proteins of 4.9, 4.8, 12.7, and 9.5 kilodaltons encoded between the spike and membrane protein genes of the bovine coronavirus. *Virology* 177:488-495.
- Akkina, R. K., T. M. Chambers, and D. P. Nayak. 1984. Expression of defective-interfering influenza virus-specific transcripts and polypeptides in infected cell. *J. Virol.* 51: 395-403.
- Andino, R., G. E. Rieckhof, and D. Baltimore. 1990. A functional ribonucleo-protein complex forms around the 5' end of poliovirus RNA. *Cell* 63: 369-380.
- Baker, S. C., and M. M. C. Lai. 1990. An *in vitro* system for the leader-primed transcription of coronavirus mRNAs. *EMBO J.* 9: 4173-4179.
- Ball, L. A., and Y. Li. 1993. *cis*-Acting requirements for the replication of flock house virus RNA2. *J. Virol.* 67: 3544-3551.
- Baric, R. S., S. A. Stohlman, M. K. Razavi, and M. M. C. Lai. 1985. Characterization of leader-related small RNAs in coronavirus-infected cell: further evidence for leader-primed mechanism of transcription. *Virus Res.* 3: 19-33.
- Bokhoven, H. V., O. L. Gall, D. Kasteel, J. Verver, J. Wellink, and A. V. Kammen. 1993. *Cis*- and *trans*-acting elements in cowpea mosaic virus RNA replication. *Virology* 195: 377-386.
- Bredenbeek, P. J., C. J. Pachuk, A. F. H. Noten, J. Charité, W. Luytjes, S. R. Weiss and W. J. M. Spaan. 1990. The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucl. Acids Res.* 18: 1825-1832.
- Brian D. A., R.-Y. Chang, P. B. Sethna, and M. A. Hofmann. 1994. Role of subgenomic minus-strand RNA in coronavirus replication. *Archives of Virology [Suppl]* 9: 173-180.

- Brierley, I., P. Digard, and S. C. Inglis. 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* 57: 537-547.
- Brown, T. D. K., and M. E. G. Bournsnel. 1984. A leader sequence is present on mRNA of avian infectious bronchitis virus. *J. Gen. Virol.* 65: 1437-1442.
- Cavanagh, D., D. A. Brian, M. Brinton, L. Enjuanes, K. V. Holmes, M. C. Horzinek, M. M. C. Lai, H. Laude, P. G. W. Plagemann, S. Siddell, W. J. M. Spaan, F. Taguchi, and P. J. Talbot. 1994. Revision of the taxonomy of the coronavirus, torovirus and arterivirus genera. *Arch. Virol.* 135: 227-237.
- 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: 8223-8231.
- Chastain, M., and I. Tinoco. 1991. Structural elements in RNA. *Prog. Nucl. Acid. Res. and Mol. Biol.* 41: 131-177.
- Collins, P. L., M. A. Mink, and D. S. Stec. 1991. Rescue of synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. *Proc. Natl. Acad. Sci. USA* 88: 9663-9667.
- de Groot, R. J., R. G. van der Most, and W. J. M. Spaan. 1992. The fitness of defective interfering murine coronavirus DI-a and its derivatives is decreased by nonsense and frameshift mutations. *J. Virol.* 66: 5898-5905.
- Denison, M. R., P. W. Zoltick, J. L. Leibowitz, C. J. Pachuk, and S. R. Weiss. 1991. Identification of polypeptides encoded in open reading frame 1b of the putative polymerase gene of the murine coronavirus mouse hepatitis virus A59. *J. Virol.* 65: 3076-3982.
- Fenner, F. J., E. P. J. Gibbs, F. A. Murphy, R. Rott, M. J. Studdert, and D. O. White. 1993. Coronaviridae. In *Veterinary Virology*. pp457-469. Academic Press, Inc. San Diego, Ca.
- Felgner, P. L., T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielsen. 1987. Lipofectin: A highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. U. S. A.* 84: 7413-7417.
- Fosmire, J. A., K. Hwang, and S. Makino. 1992. Identification and characterization of a coronavirus packaging signal. *J. Virol.* 66: 3522-3530.

- Furuya, T., and M. M. C. Lai. 1993. Three different cellular proteins bind to complementary sites on the 5'-end-positive and 3'-end-negative strands of mouse hepatitis virus RNA. *J. Virol.* 67: 7215-7222.
- Hagino-Yamagishi, K., and A. Nomoto. 1989. *In vitro* construction of poliovirus defective interfering particles. *J. Virol.* 12: 5386-5392.
- Hillman, B. I., J. C. Carrington, and T. J. Morris. 1987. A defective interfering RNA that contains a mosaic of a plant virus genome. *Cell* 51: 427-433.
- Hofmann, M. A., and D. A. Brian. 1991a. Sequencing PCR DNA amplified directly from a bacterial colony. *Biotechniques* 11: 30-31.
- Hofmann, M. A., and D. A. Brian. 1991b. The 5-prime end of coronavirus minus-strand RNAs contain a short poly(U) tract. *J. Virol.* 65: 6331-6333.
- 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: 163-171.
- Hofmann, M. A., S. D. Senanayake, and D. A. Brian. 1993. A translation attenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. *Proc. Natl. Acad. Sci. USA* 90: 11733-11737.
- Hofmann, M. A., P. B. Sethna, and D. A. Brian. 1990. Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. *J. Virol.* 64: 4108-4114.
- Holland, J. J. 1991. Defective viral genomes. *In*: B. N. Fields and D. M. Knipe (ed.), *Fundamental virology*. pp.151-165. Raven Press, New York.
- Holmes, K. V. 1990. Coronaviridae and their replication. *In*: B. N. Fields (ed.), *Virology* pp. 841-856. Raven Press, New York.
- Horton, R. M., Z. Cai, S. N. Ho, and L. R. Pease. 1990. Gene splicing by overlap extension: Taylor made genes using the polymerase chain reaction. *Biotechniques*. 8: 528-535.
- Jeong, Y. S., and S. Makino. 1992. Mechanism of coronavirus transcription: duration of primary transcription initiation activity and effects of subgenomic RNA replication. *J. Virol.* 66: 3339-3346.
- Jeong, Y. S., and S. Makino. 1994. Evidence for coronavirus discontinuous transcription. *J. Virol.* 68: 2615-2623.

- Johnson, K. L., and P. Sarnow. 1991. Three poliovirus 2B mutants exhibit noncomplementable defects in viral RNA amplification and display dosage-dependent dominance over wild-type poliovirus. *J. Virol.* 65: 4341-4349.
- Joo, M., and S. Makino. 1992. Mutagenic analysis of the coronavirus intergenic consensus sequence. *J. Virol.* 66: 6330-6337.
- Kapke, P. A., and D. A. Brian. 1986. Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. *Virology* 151: 41-49.
- Kienzle, T. E., S. Abraham, B. G. Hogue, and D. A. Brian. 1990. Structure and orientation of expressed bovine coronavirus hemagglutinin-esterase protein. *J. Virol.* 64: 1834-1838.
- Kim, Y.-N., Y. S. Jeong, and S. Makino. 1993a. Analysis of *cis*-acting sequences essential for coronavirus defective interfering RNA replication. *Virology* 197: 53-63.
- Kim, Y.-N., M. M. C. Lai, and S. Makino. 1993b. Generation and selection of coronavirus defective interfering RNA with large open reading frame by RNA recombination and possible editing. *Virology* 194: 244-253.
- Kuge, S., I. Saito, and A. Nomoto. 1986. Primary structure of poliovirus defective- interfering particle genomes and possible generation mechanisms of particles. *J. Mol. Biol.* 192: 473-487.
- Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Aca. Sci. USA* 82: 488-492.
- Lai, M. M. C. 1990. Coronavirus: organization, replication and expression of genome. *Annu. Rev. Microbiol.* 44: 303-333.
- Lai, M. M. C. 1992. RNA recombination in animal and plant viruses. *Microbiol. Rev.* 56: 61-79.
- Lai, M. M. C., R. S. Baric, P. R. Brayton, and S. A. Stohlman. 1984. Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc. Natl. Acad. Sci. USA* 81: 3626-3630.
- Lai, M. M. C., P. R. Brayton, R. C. Armen, C. D. Patton, C. Pugh, and S. A. Stohlman. 1981. Mouse hepatitis virus A59: messenger RNA structure and genetic localization of the sequence divergence from the hepatotropic strain MHV 3. *J. Virol.* 39: 823-834.
- Lai, M. M. C. and S. A. Stohlman. 1981. Comparative analysis of RNA genomes of mouse hepatitis viruses. *J. Virol.* 38: 661-670.

- Lapps, W., B. G. Hogue, and D. A. Brian. 1987. Sequence analysis of bovine coronavirus nucleocapsid and matrix protein genes. *Virology* 157: 47-57.
- Lee, H.-J., C.-K. Shieh, A. E. Gorbalenya, E. V. Eugene, N. La Monica, J. Tuler, A. Bagdzhadzhyan, and M. M. C. Lai. 1991. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* 180: 567-582.
- Levis, R., B. G. Weiss, M. Tsiang, H. Huang, and S. Schlesinger. 1986. Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. *Cell* 44: 137-145.
- Li, X.-H., L. A. Heaton, T. J. Morris, and A. E. Simon. 1989. Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated *de novo*. *Proc. Natl. Acad. Sci. USA*. 86: 9173-9177.
- Li, X., and P. Palese. 1992. Mutational analysis of the promoter required for influenza virus virion RNA synthesis. *J. Virol.* 66: 4331-4338.
- Liao, C.-L., and M. M. C. Lai. 1992. RNA recombination in a coronavirus: recombination between viral genomic RNA and transfected RNA fragments. *J. Virol.* 66: 6117-6124.
- Liao, C.-L., and M. M. C. Lai. 1994. Requirement of the 5'-end genomic sequence as an upstream *cis*-acting element for coronavirus subgenomic mRNA transcription. *J. Virol.* 68:4727-4737.
- Lin, Y.-J., and M. M. C. Lai. 1993. Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinuous sequence for replication. *J. Virol.* 67: 6110-6118.
- Makino, S., and M. Joo. 1993. Effect of intergenic consensus sequence flanking sequences on coronavirus transcription. *J. Virol.* 67: 3304-3311.
- Makino, S., M. Joo, and J. K. Makino. 1991. A system for study of coronavirus mRNA synthesis: a regulated, expressed subgenomic defective interfering RNA results from intergenic site insertion. *J. Virol.* 65: 6031-6041.
- Makino, S., and M. M. C. Lai. 1989a. High-frequency leader sequence switching during coronavirus defective interfering RNA replication. *J. Virol.* 63: 5285-5292.
- Makino, S., and M. M. C. Lai. 1989b. Evolution of the 5' end of genomic RNA of murine coronaviruses during passages *in vitro*. *Virology* 169: 227-232.
- Makino, S., and M. M. C. Lai. 1990. Studies of coronavirus DI RNA replication using *in vitro* constructed DI cDNA clones. *Adv. Exp. Med. Biol.* 276: 341-347.

- Makino, S., L. H. Soe, C.-K. Shieh, and M. M. C. Lai. 1988. Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. *J. Virol.* 62: 3870-3873.
- Makino, S., S. Stohlman, and M. M. C. Lai. 1986. Leader sequence of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. *Proc. Natl. Aca. Sci. USA* 83: 4204-4208.
- Makino, S., K. Yokomori, and M. M. C. Lai. 1990. Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible RNA-packaging signal. *J. Virol.* 64: 6045-6053.
- Masters, P., C. A. Koetzner, C. A. Kerr, and Y. Heo. 1994. Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in coronavirus mouse hepatitis virus. *J. Virol.* 68: 328-337.
- McIntosh, K. 1990. Coronaviruses. In: B. N. Fields (ed.), *Virology* pp. 857-864. Raven Press, New York.
- Murray, R. S., G. Y. Cai, K. Hoel, J. Y. Zhang, K. F. Soike, and G. F. Cabirac. 1992. Coronavirus infects and causes demyelination in primate central nervous system. *Virology* 188: 247-284.
- Nesters, H. G. M., and J. H. Strauss. 1990. Defined mutations in the 5' nontranslated sequence of Sindbis virus RNA. *J. Virol.* 64: 4162-4168.
- Novak, J. E., and K. Kirkegaard. 1994. Coupling between genome translation and replication in an RNA virus. *Gene & Development* 8: 1726-1737.
- Pachuk, C. J., P. J. Bredenbeek, P. W. Zoltick, W. J. M. Spaan, and S. R. Weiss. 1989. Molecular cloning of the gene encoding the putative polymerase of mouse hepatitis coronavirus strain A59. *Virology* 171: 141-148.
- Pogue, G. P., and T. C. Hall. 1992. The requirement for a 5' stem-loop structure in brome mosaic virus replication supports a new model for viral positive-strand RNA initiation. *J. Virol.* 66: 674-684.
- Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers. 1989. Transcription of Yellow Fever RNA from full-length cDNA templates produced by *in vitro* ligation. *New Biol.* 1:285-296.
- Roux, L., A. E. Simon, and J. J. Holland. 1991. Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo*. *Adv. Virus Res.* 40: 181-211.
- Rümenapf, T., E. G. Strauss, and J. H. Strauss. 1994. Subgenomic mRNA of Aura alphavirus is packaged into virions. *J. Virol.* 68: 56-62.

- Sachs, A. B. 1993. Messenger RNA degradation in eukaryotes. *Cell*. 74: 413-421.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
- Sawicki, S. G., and D. L. Sawicki. 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates functional mRNA synthesis. *J. Virol.* 64: 1050-1056.
- Schlesinger, S. 1988. The generation and amplification of defective interfering RNAs. *In*: E. Domingo, J. J. Holland, and P. Ahlquist (ed). *RNA genetics: Vol II* pp.167-185. CRC Press, Boca Raton, Florida.
- Senanayake, S. D., M. A. Hofmann, J. L. Maki, and D. A. Brian. 1992. The nucleocapsid protein gene of the bovine coronavirus is bicistronic. *J. Virol.* 66: 5277-5283.
- Sethna, P. B., M. A. Hofmann, and D. A. Brian 1991. Minus-strand copies of replication coronavirus mRNAs contain antileaders. *J. Virol.* 65: 320-325.
- Sethna, P. B., S.-L. Hung, and D. A. Brian. 1989. Coronavirus subgenomic minus-strand RNA and the potential for mRNA replicons. *Proc. Natl. Acad. Sci., USA* 86: 5626-5630.
- Spaan, W., D. Cavanagh, and M. C. Horzinek. 1988. Coronaviruses: Structure and genome expression. *J. Gen. Virol.* 69: 2939-2952.
- Spaan, W. J. M., P. J. M. Rottier, M. C. Horzinek, and B. A. M. van der Zeijst. 1981. Isolation and identification of virus-specific mRNAs in cells infected with mouse hepatitis virus (MHV-A59). *Virology* 108: 424-434.
- 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: 4288-4295.
- Sturman, L. S., and K. V. Holmes. 1983. The molecular biology of coronaviruses. *Adv. Virus. Res.* 28: 35-112.
- Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla. 1973. Improved estimation of secondary structure in ribonucleic acids. *Nature New Biology* 246: 40-41.
- Torczynski, R. M., M. Fuke, and A. P. Bollon. 1985. Cloning and sequencing of a human 18S ribosomal RNA gene. *DNA* 4: 283-291.
- Tyrrell, D., D. Alexander, J. Almeida, C. Cunningham, D. Hamre, M. Hofstad, L. Malluci, K. McIntosh. 1968. Coronaviruses. *Nature* 220:650.

- van der Most, R. G., P. J. Bredenbeek, and W. J. M. Spaan. 1991. A domain at the 3' end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. *J. Virol.* 65: 3219-3226.
- van der Most, R. G., R. Groot, and W. J. M. Spaan. 1994. Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: a study of coronavirus transcription initiation. *J. Virol.* 68: 3656-3666.
- Vancanneyt, G., S. Rosahl, and L. Willmitzer. 1990. Translatability of a plant-mRNA strongly influences its accumulation in transgenic plants. *Nucl. Acids Res.* 18: 2917-2921.
- Vidal, S., and D. Kolakofsky. 1989. Modified model for the switch from Sendai virus transcription to replication. *J. Virol.* 63: 1951-1958.
- Weiland, J. J., and T. W. Dreher. 1993. *cis*-Preferential replication of the turnip yellow mosaic virus RNA genome. *Proc. Natl. Acad. Sci. USA* 90: 6095-6099.
- Weiss, B., H. Nitschko, I. Ghattas, W. R., and S. Schlesinger. 1989. Evidence for specificity in the encapsidation of Sinbis RNAs. *J. Virol.* 63: 5310-5318.
- White, K. A., J. B. Bancroft, and G. A. Mackie. 1992. Coding capacity determines *in vivo* accumulation of a defective RNA of clover yellow mosaic virus. *J. Virol.* 66: 3069-3076.
- Yokomori, K., L. R. Banner, and M. M. C. Lai. 1992. Coronavirus mRNA transcription: UV light transcriptional mapping studies suggest an early requirement for a genomic-length template. *J. Virol.* 66: 4671-4678.
- Zhang, X., and M. M. C. Lai. 1994. Unusual heterogeneity of leader-mRNA fusion in a murine coronavirus: implications for the mechanism of RNA transcription and recombination. *J. Virol.* 68: 6626-6633.
- Zhang, X., C.-L. Liao, and M. M. C. Lai. 1994. Coronavirus leader RNA regulates and initiates subgenomic mRNA transcription both in *trans* and in *cis*. *J. Virol.* 68: 4738-4746.
- Zhang, J., and H. M. Temin. 1994. Retrovirus recombination depends on the length of sequence identity and is not error prone. *J. Virol.* 68: 2409-2414.
- Zhao, X., and D. Cavanagh. 1993. Presence of subgenomic mRNAs in virions of Coronavirus IBV. *Virology* 196: 172-178.

APPENDIX

DETAILED PROCEDURE FOR THE USE OF RNA BLOTTING IN THE ANALYSIS OF CORONAVIRUS RNA REPLICATION

This protocol is a refinement of the one developed by Sethna (Sethna *et al.*, 1989) in which it was shown for the first time that coronavirus subgenomic minus-strands are present in infected cells. A key feature of this approach is that end-labeled oligodeoxynucleotides are used for the probe in the probing procedure. This approach avoids any possibility of a mixed probe population that might arise from biologically-generated radiolabeled probes such as those made in the riboprobe system. With this procedure the relative number of plus- and minus-strand species can be estimated from knowledge of the specific activity of our radiolabeled probe.

TRANSFECTION OF *IN VITRO* TRANSCRIPTS AND RNA EXTRACTION

Linearizing plasmid DNA.

1. Add as following components:

H ₂ O:	77.0 μ l
NEB buffer 3:	10.0 μ l
DNA (5-10 μ g):	10.0 μ l
<u>Mlu I (10 u/μl):</u>	<u>3.0 μl (3 u/μg DNA)</u>
	100.0 μ l
2. Incubate at 37°C for 3 hr (or O/N), and check linearization by running on a 1% agarose mini-gel.

Blunt ending DNA with mung bean nuclease.

1. Heat inactivate restriction enzyme and add 1 μ l of 100 mM ZnSO₄ (final concentration is 1mM) or extract DNA with phenol/chloroform and

precipitate with EtOH. Spin down DNA, dry the pellet, and then, add 10 μ l of 10X mung bean nuclease buffer. Add H₂O up to 100 μ l.

2. Add mung bean nuclease (0.1-0.2 u/ μ g DNA)
 - a. Incubate at 30°C for 30 min
 - b. Extract with phenol/chloroform once.
 - c. Extract with chloroform once.
 - d. Precipitate with EtOH on dry ice for 15 min
 - e. Spin down, wash once with 80% EtOH, and dry
 - f. Resuspend in H₂O to give a concentration of 1 μ g/ μ l

In vitro transcription.

1. Add as following components:

	<u>uncapped RNA</u>	<u>capped RNA</u>
H ₂ O	35.5 μ l	25.5 μ l
5X transcription buffer	20.0 μ l	20.0 μ l
100 mM DTT	10.0 μ l	10.0 μ l
acetylated BSA (2 mg/ml)	5.0 μ l	5.0 μ l
RNasin (40 u/ μ l)	2.5 μ l	2.5 μ l
rNTP mix (2.5 mM of each rA, U, C, GTP) (replace 2.5 mM rGTP with 0.25 mM GTP if capping)	20.0 μ l	20.0 μ l
5 mM ⁷ mGpppG	-	10.0 μ l
linearized DNA (5 μ g)	5.0 μ l	5.0 μ l
T7 RNA polymerase (40 u)	2.0 μ l	2.0 μ l
	100.0 μ l	100.0 μ l

Note: The mixture should be kept at room temperature during the addition of each successive component, since DNA can precipitate in the presence of spermidine if kept at 4°C. (Add DNA last, right before adding enzyme).

2. Incubate at 37°C for 1 hr.
3. Add 1 μ l of T7 RNA polymerase, and incubate at 37°C for another 1 hr.
4. Add 5 μ l (5u) of RQ RNase-free DNase (1u/ μ l), and incubate at 37°C for 30 min.
5. Extract with phenol/chloroform once.
6. Extract with chloroform once.

7. Pass through Biospin 6 column (equilibrated with 1 ml RNase-free H₂O for three times).
8. Take 5 µl of RNA, add 495 µl of H₂O, and quantitate by using spectrophotometer.
9. Aliquot into small amount (20-30 µl/tube), and store at -20°C.
10. Run 500 ng of RNA on a fresh 1% agarose gel.

Transfection.

1. Cells grown in 35 mm dish. Trypsinize well so that cells are not clumped. Cells should be young (12-18 hrs old). Incubate dishes in 37°C CO₂ incubator for overnight to give 50~80 % confluent.
2. Infect with 200 µl of BCV stock for 1 hr at 37°C, and shake every 10 min (leave one plate for uninfected control).
3. Suck off virus, and add 2 ml of DMEM with 10% FBS (leave one plate for BCV-infected-untransfected control).
4. Prepare RNA-lipofectin mixture as following components for transfection, mix by vortexing gently, and keep on ice for 10 min.

1X PBS:	to make volume to 200 µl
Lipofectin (10 µg)	10 µl/plate
<u>RNA</u>	<u>200 ng/plate</u>
	200 µl /plate

Note: use polystyrene tubes, not polypropylene tubes or eppendorf tubes (lipofectin may react with them).

5. Plates for transfection:
 - a. Wash three times with 1X PBS prior to transfection.
 - b. Add 200 µl RNA-Lipofectin mix to each plate.
 - c. Transfect for 10 min at room temperature, and shake occasionally.
 - d. Wash with serum-free DMEM once.
 - e. Refeed 2 ml of DMEM with 10% FBS, and incubate at 37°C.
 - f. Extract RNA at each designated time point.

RNA extraction.

1. Wash cells with 1X EBSS (ice-cold) three times.
2. Add 0.5 ml of lysis buffer.
3. Scrape cells off by using broad end of a blue tip, and transfer cell lysate into an eppendorf tube.
4. Vortex for 10 sec, and then spin at 7g for 4 min at 4°C.
5. Take out supernatant into a new tube containing 0.1 ml 10% SDS and add 2 µl of proteinase K (20 mg/ml).
6. Incubate at 37°C for at least 30 min.
7. Add 1/4V of 7.5M NH₄OAc (less than 150 µl).
8. Extract with phenol/chloroform once.
9. Add 3V of EtOH and store -20°C.

NORTHERN BLOTTING

Prepare denaturing gel (volume for maxi gel apparatus, 300 ml).

1. Rinse gel apparatus with distilled water. Clean gel cast and comb with 95 % EtOH and let dry.
2. Place gel apparatus in a fume exhaust hood and level it with a bubble till exactly horizontal.
3. Dissolve 3 grams of agarose in 216.6 ml H₂O.
4. Heat in microwave, and let it boil until the agarose is completely melted.
5. Cool down a little bit, and then add 30 ml of 10X MOPS buffer.
6. After reaching a temperature of about 60°C, add 53.4 ml of formaldehyde in the hood. Shake the flask to make sure formaldehyde mixed well.
7. Pour the agarose into the gel cast, and remove air bubbles with a clean pipette tip.
8. Cover the apparatus with lid, and let gel solidify.

9. When gel has become solid, fill in 1X MOPS (prepared in millipore-filtered water) to cover the gel. Remove the combs carefully, remove tapes, and place the gel cast back into the apparatus. Fill about 2.5 liters of 1X MOPS into the gel apparatus until buffer is about 2 mm below gel surface.

Preparing RNA samples and running on the denaturing gel.

1. Take approximately 2.5 µg of cytoplasm RNAs (1/4 of RNAs from 35-mm petri dishes, about 500 µl) per well (5 µg of RNA for 2 wells).
2. Pellet RNAs at 13000 rpm for 12 min at 4°C.
3. Discard EtOH, and wash pellet once by adding 500 µl of 80% EtOH.
4. Pellet samples again at 13000 rpm for 12 min at 4°C
5. Remove EtOH with pipette. Dry pellets in speedvac for 5 min.
6. Prepare loading mix as following components per sample:

RNase-free H ₂ O (-20°C):	8.2 µl
10X MOPS (-20°C):	3.7 µl
formaldehyde 37% (RT):	6.6 µl
<u>formamide 100% (-20°C):</u>	<u>18.5 µl</u>
	37.0 µl

Note: after removing formamide from the tube, and add N₂ gas into tube before refreezing to prevent ionization.

7. Heat samples for 5 min at 65°C after vortexing and quickly spinning down.
8. Place the samples on ice for 5 min to chill.
9. Add 8 µl of loading dye per sample, vortex, and spin down quickly.
10. Load 22 µl per well. Load twice for two panels if needed.
11. Run the gel immediately after finishing with loading. Set at constant voltage (80V for midi gel, and 140V for maxi gel). Current should be about 60 mA.
12. Run the gel for about 3.5 to 4 hours, until the bromo phenol blue dye has migrated about 10 to 12 cm.

Transferring RNAs onto a membrane.

1. Cut a piece of hybond-N+ nucleic acid transfer membrane (Amersham, Arlington Heights, IL.) to fit exactly in the plastic window of the vacuum blotting apparatus (VacuGene XL vacuum blotting system, Pharmacia LKB Biotechnology, Piscataway, NJ).
2. Prewet the membrane in deionized water.
3. Place the membrane in the plastic window. Smooth it with a rubber roller, and make sure that there is no bubble.
4. Place the agarose gel on the top of the membrane. Smooth again with a rubber roller.
5. Seal wells and edges of the gel with melted agarose to prevent leakage.
6. Turn on the vacuum pump and set pressure at 55 mbar.
7. Make sure the gel is tightly sucked. Then, pour 20X SSC carefully on the center of the gel till the whole gel is covered with buffer.
8. Blot for at least 4 hours.

Fix RNAs on the membrane by UV cross-linking.

1. Remove 20X SSC buffer from the apparatus and turn off the vacuum pump.
2. Mark position of wells and discard the gel.
3. Place a piece of 3M paper on a used X-ray film and wet the paper with water. Smooth it with rubber roller.
4. Transfer the blotted hybond-N+ membrane on the top of the wet 3M paper.
5. Place the X-ray film with the membrane and 3M paper into the Stratalinker UV crosslinker (Stratagene, La Jolla, Ca.). Set at auto cross link mode ($1.2 \times 10^5 \mu$ joules for 30 seconds).
6. Take out the membrane and let air dry. The membrane is ready for hybridization.

HYBRIDIZATION

(USING BELLCO AUTOBLOT HYBRIDIZATION OVEN)

Prehybridization.

1. Cut a piece of mesh (supplied by Bellco Glass Inc., Vineland, N. J.) to the size of the membrane.
2. Pre-wet the mesh and the hybridization membrane in a suitable tray containing 2X SSC.
3. Ensuring that the hybridization membrane exactly overlays the mesh, roll both up into a tight roll.
(If more than one membranes is to be hybridized, simply add further meshes and membranes as required before rolling. It is important that each membrane is separated from any other by a piece of mesh. Up to five 20 x 20 membranes can be hybridized in a single hybridization bottle.)
4. Transfer the roll to a hybridization bottle and add 30 ml of 2X SSC (15 ml for a small bottle).
5. Slowly unwind the membrane and mesh around the inside of the bottle by gently rolling the bottle. No air bubbles should be visible between the membrane and the bottle. If bubbles are present, the membrane should be removed and rerolled. The procedure should then be repeated more gently.
Note: The membrane has to be at least overlapped a little bit in order to be in fixed place during rotation in the oven.
6. Once the mesh and membrane is in place in the bottle, 2X SSC can be discarded and replaced with hybridization buffer (10 ml for a large bottle or 5 ml for a small bottle. Higher volumes will be required if there is more than one membrane in the bottle).
7. Put the bottle in the hybridization oven at 55°C for at least 1 hour.
Note: The rotisserie should be balanced in the same manner as a centrifuge. When place the bottle into the oven, the inner part of the overlapping membrane should be in the same direction of the fluid flow (rule of the lower lip).

End label an oligonucleotide as a probe.

1. Mix the following components:

oligonucleotide (10 μ M):	2 μ l
H ₂ O:	13 μ l
10X kinase buffer:	3 μ l
γ^{32} P-ATP:	10 μ l
<u>T4 polynucleotide kinase (10u/μl):</u>	<u>2 μl</u>
	30 μ l

2. Incubate for 30 min at 37°C.
3. Stop reaction by adding 70 μ l of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).
4. Purify the end-labelled primer by passing through a BioSpin6 column (Bio-Rad).
5. Take out 1 μ l of probe and add 9 μ l of H₂O for counting. Ideally you should have at least 1×10^7 cpm Cerenkov.

Hybridization.

1. Denature the purified probe by heating at 90°C for 5 min. Then, store on ice.
2. Remove the bottle from the hybridization oven.
3. Hold the bottle upright and unscrew the bottle.
4. Carefully pipette the probe directly into the hybridization buffer. Be careful to avoid injecting the probe directly onto the membrane as this will result in hot spot.
5. Replace the cap and gently agitate the bottle to ensure an even distribution of the probe in the hybridization solution.
6. Place the bottle back in hybridization oven and leave it to hybridize at 55°C for overnight (more than 12 hours).

Washing the membrane.

1. Remove the bottle from the hybridization oven. Turn off the temperature switch of the oven and leave the door open so the temperature will drop down fast to room temperature.

2. Pour off the hybridization fluid and store in -20°C.
3. Fill the bottle with 50 ml of wash solution (2X SSC and 0,1% SDS).
4. Replace the bottle in the oven (temperature should be below 30°C) and leave for 10 min.
5. Discard wash solution.
6. Repeat steps 3-5 for two to three times. Meanwhile, prewarm another bottle of wash solution to 55°C.
7. Remove the bottle from the hybridization oven. Turn on the temperature switch and let the temperature back to 55°C for high stringency wash.
8. Fill the bottle with the 55°C wash solution and place back to the oven at 55°C for 30 min.
9. Pour off wash buffer. Rinse once with wash solution.
10. Remove the mesh and the membrane from the bottle. Quickly dip in water and remove excess water by placing on a piece of 3M paper. Do not let dry if further wash is needed.
11. Wrap the membrane with Saran wrap and autoradiography.

REAGENTS AND SOLUTIONS

10X mung bean nuclease buffer.

<u>components</u>	<u>concentration</u>
NaOAc (pH 5.0)	500 mM
NaCl	300 mM
ZnSO ₄	10 mM

10X Dulbecco's Phosphate-Buffered Saline (10X-PBS).

<u>components</u>	<u>concentration</u>
KCl	2.0 g/l
KH ₂ PO ₄	2.0 g/l
NaCl	80.0 g/l
Na ₂ HPO ₄ •7H ₂ O	21.6 g/l

Note: The pH of 10X is 6.7 to 6.9. The pH of 1X should be checked after 1:10 dilution to ensure that it is 7.1 ± 0.1 . Adjust pH by using sterile 1N NaOH or 1N HCl where appropriate.

Lysis buffer for RNA extraction.

<u>components</u>	<u>for 500 ml</u>	<u>concentration</u>
1M Tris-HCl (pH 7.0 at 37°C)	5.0 ml	10 mM
5M NaCl	1.0 ml	10 mM
1M MgCl ₂	2.5 ml	5 mM
NP40	5.0 ml	1 %
H ₂ O	481.5 ml	

Mix well, aliquot into small amount and store at -20°C.

20X SSC.

<u>components</u>	<u>for 20 liters</u>	<u>concentration</u>
NaCl	3506 g	3.0 M
Sodium Citrate	1764 g	0.3 M

Adjust pH with 1N HCl to pH 7.0.

Prehybridization and hybridization solution.

<u>components</u>	<u>for 500 ml</u>	<u>concentration</u>
20X SSC	125 ml	6X SSC
0.5M phosphate buffer, pH6.5	50 ml	50 mM
50X Denhardt's solution	50 ml	5X
10% SDS	50 ml	1%
ST DNA (10 mg/ml)	5 ml	100 µg/ml
yeast tRNA (5 mg/ml)	5 ml	50 µg/ml

1. ST DNA and tRNA must be denatured at 100°C for 10 min before adding to the prehybridization solution.
2. Mix well and store at -20°C.

DETAILED PROCEDURE FOR THE USE OF PCR SEQUENCING

This procedure is a modification of the protocol of Hofmann and Brian (1991a).

PCR for generating ss DNA template.

1. Mix the following components:

33.2 μ l H₂O

4.0 μ l 10X PCR buffer (Promega)

0.8 μ l dNTP mix (1.25 mM each) (final concentration of 20 μ M each)

1.0 μ l primer 1 (10 μ M), not used as sequencing primer

1.0 μ l primer 2 (0.1 μ M), used in sequencing reaction

40.0 μ l

2. Add a few cells from an isolated *E.coli* colony into the reaction mix with a toothpick.
3. Boil for 15 min at 95°C, and then chill on ice.
4. Add 10 μ l of diluted Taq DNA polymerase (Promega, 1.25 units per reaction) in 1X PCR buffer. Mix, and cover with 2 drops of mineral oil.
5. PCR: 40 cycles: 94°C (30 sec), 50°C (1 min), and 72°C (2-3 min).

Labeling reaction for sequencing primer (for 10 sequencing reactions).

1. Mix: 1 μ l primer (10 μ M)

11 μ l H₂O

2 μ l 10X kinase buffer

5 μ l α^{32} P-ATP

1 μ l (10u) T4 polynucleotide kinase

20 μ l

2. Incubate for 30 min at 37°C. Add 30 μ l of H₂O to increase the volume to 50 μ l.
3. Purify end-labeled primer by passing through a BioSpin6 column (Bio-Rad) (equilibrated with 1 ml of H₂O for three times).

4. Estimate volume of end-labeled oligonucleotide after purification.
Calculate volume to be taken for 1 pmole per reaction (about 5 μ l).

Sequencing reaction, using end-labeled primer.

1. Prepare sequencing mix as following components for one reaction:
 - 1.2 μ l 10X PCR buffer
 - 5.0 μ l (or corresponding to 1 pmole) labeled primer
 - 5.6 μ l H₂O
 - 0.2 μ l Taq DNA polymerase (1u)
 - 12.0 μ l
 - a. Spin down PCR tubes for 1 min at 6000 rpm from above PCR reactions.
 - b. Take 20 μ l of ssDNA (of 50 μ l) from PCR reaction into a new tube.
 - c. Add 12 μ l of sequencing mix into each tube and mix by pipetting 10 times.
2. Immediately distribute 7.5 μ l of this mix into G/A/T/C tubes, which contain 2.5 μ l of the respective termination mix*. Mix by pipetting up and down for five times.
3. Incubate the termination reactions at 72°C for 5 min.
4. Add 4 μ l of formamide stop buffer, and store at -20°C.
5. Heat denature samples at 90°C for 2 min and analyze on a 6% polyacrylamide gel.

* Termination mix:

G: 250 μ M ddGTP:	1.25 μ l ddGTP (10 mM)
	5.00 μ l 10X PCR buffer
	0.80 μ l dNTP (1.25 mM each)
	<u>42.95 μl H₂O</u>
	50.00 μ l

A: 1.28 mM ddATP:	6.4 μ l ddATP (10 mM)
	5.0 μ l 10X PCR buffer
	0.8 μ l dNTP (1.25 mM each)
	<u>37.8 μl H₂O</u>
	50.0 μ l

T: 1.92 mM ddTTP: 9.6 μ l ddTTP (10 mM)
5.0 μ l 10X PCR buffer
0.8 μ l dNTP (1.25 mM each)
34.6 μ l H₂O
50.0 μ l

C: 640 μ M ddCTP: 3.2 μ l ddCTP (10 mM)
5.0 μ l 10X PCR buffer
0.8 μ l dNTP (1.25 mM each)
41.0 μ l H₂O
50.0 μ l

LIST OF ABBREVIATIONS

BCV	bovine coronavirus
MHV	mouse hepatitis virus
TGEV	porcine transmissible gastroenteritis virus
IBV	avian infectious bronchitis virus
HCV	human coronavirus
DI RNA	defective-interfering RNA
UTR	untranslated region

VITA

Ruey-Yi Chang was born in Chungli, Taiwan, R. O. C. on November 1, 1964. She received a Bachelor of Science degree in Botany from National Chung-Hsing University, Taiwan, in 1987. She received a Master of Science degree in Microbiology from the same university in 1989. From 1989 to 1990 she worked as a research technician in the Agricultural Biotechnology Laboratories at National Chung-Hsing University. In the fall of 1990, she entered the graduate program in the Department of Microbiology at the University of Tennessee, Knoxville, and received her Ph.D. degree in December 1994.