INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

REPLICATION AND RECOMBINATION OF THE *RED CLOVER NECROTIC MOSAIC VIRUS*

by

Ziming Weng

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PLANT PATHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2002

UMI Number: 3060946

UMI Microform 3060946

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by ZIMING WENG

entitled REPLICATION AND RECOMBINATION OF THE RED CLOVER NECROTIC MOSAIC VIRUS

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of DOCTOR OF PHILOSOPHY

	8/1/02 Date
	8/12/02 Date
Christina Kennedy	8/12/02 Date
	Date
	Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director	8/1/02 Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interest of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Zimin Wang

ACKNOWLEDGEMENTS

First of all, I would like to thank Dr. Zhongguo Xiong, my major professor, for providing me the opportunity to pursue Ph'D degree, for teaching me, supporting me, and encouraging me during the whole course of this program. I would also like to thank Dr. Merritt Nelson, our former department head and my former committee member, for his guidance and support since my MS program. I would like to thank Dr. Christina Kennedy and Dr. Hans VanEtten for their helps during this work and for their critical reading and valuable suggestions of my dissertation. I would also like to acknowledge Dr. Don Bourque and Dr. Jennifer Hall to have served on my committee.

Additionally, I would like to thank Dr. Leland Pierson for his help in writing my preliminary proposal and Dr. Martha Hawes for her emotional support whenever I have trouble. I would also like to thank members in Dr. Xiong's Lab, Rodolfo Acosta-Leal, Rick Langham, and Marie Solofoharivelo for their cooperation and friendship.

Finally, I would like to thank my parents and my husband Yinming for their love, support, and understanding.

DEDICATION

This dissertation is dedicated to my parents who love me, educate me, and support me all the time, and to my little daughter Bailey who is the present from the God and was born during this program.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE OF ABBREVIATIONS

AlMV: Alfalfa mosaic virus
BMV: Brome mosaic virus
CaMV: Cauliflower mosaic virus
CCMV: Cowpea chlorotic mottle virus
CMV: Cucumber mosaic virus
CNV: Cucumber necrotic virus
CP: coat protein or capsid protein
CRSV: Carnation ringspot virus
CyRSV: Cymbidium ringspot virus
DdDp: DNA-dependent DNA polymerase
DdRp: DNA-dependent RNA polymerase
DI RNA: defective interfering RNA
dsRNA: double-stranded RNA
EF-Ts: elongation factor-Ts
EF-Tu: elongation factor-Tu
eIF-3: eukaryotic translation initiation factor 3
ER: endoplasmic reticulum
GFP: green fluorescent protein
HCV: Hepatitis C virus
HF-I: host factor I
HIV: Human immunodeficiency virus
MCMV: Maize chlorotic mottle virus
MP: movement protein
ORF: open reading frame
PABP: poly A-binding protein
PCBP: poly C-binding protein
PPV: Plum pox virus
PVY: Potato virus Y
RCNMV: Red clover necrotic mosaic virus
RdDp: RNA-dependent DNA polymerase
RdRp: RNA-dependent RNA polymerase
RNP: ribonucleoprotein complex
RT: reverse transcriptase
RT-PCR: reverse transcription-polymerase chain reaction
satRNA: satellite RNA
SCNMV: Sweet clover necrotic mosaic virus
SDS-PAGE: SDS-polyacrylamide gel electrophoresis
SEL: size exclusion limit
sgRNA: subgenomic RNA
SLS: stem-loop structure
ssRNA: single-stranded RNA
TBSV: Tomato bushy stunt virus

TCV: Turnip crinkle virus
TLS: tRNA-like structure
TMV: Tobacco mosaic virus
TNTase: terminal nucleotidyl transferase
TYMV: Turnip yellow mosaic virus
UTR: un-translated region
VPg: genome-linked virus protein
ZYMV: Zucchini yellow mosaic virus

ABSTRACT

In this study, *Red clover necrotic mosaic virus* (RCNMV) was used to better understand the functions of replication proteins and to identify the terminal promoter element involved in viral replication. RCNMV genome contains two positive-sense, single-stranded RNAs. RNA-1 encodes two proteins essential for viral replication: p27 and p88. p88 is a fusion protein containing p27 at its N terminus and RNA dependent RNA polymerase motifs at its C-terminal domain. The function of p27 is not known. In this work, studies of RNA-1 chimerical clones between a highly infectious clone and a poorly infectious clone and subsequent mutagenesis demonstrated that the N-terminal 14 amino acids of p27 and p88 were required for efficient RNA replication. Sequence analysis indicated that it is possibly involved in membrane interaction.

Another important aspect of viral replication is template recognition by the replicase at the 3' promoter. The 3'-29 nucleotides of both RCNMV RNA-1 and RNA-2 can be predicted to form an identical stem-loop structure (SLS). Mutational analysis of the SLS indicated that both the structure and the loop sequence were required for viral replication. Within the 5-nt loop region, three discontinuous nucleotides were identified as critical nucleotides for RNA-replicase interaction. The functional groups in these key nucleotides involved in replicase recognition are predicted.

The 3' promoter element of RCNMV not only affects viral RNA replication but also influences transgenic recombination. RCNMV RNA-2 encodes a movement protein (MP) that is required for viral cell-to-cell movement and systemic infection. Transgenic *Nicotiana benthamiana* plants expressing different versions of MP mRNA neither

resisted RCNMV nor complemented RNA-1 infection. However, systemic infection was observed in transgenic lines expressing 5' truncated MP mRNA when only RNA-1 was inoculated. Further analysis showed that the infection was resulted from nonhomologous RNA recombination events between infecting RNA-1 and MP transgene mRNA. A replicase-mediated template switch model of the transgenic recombination was proposed. The presence of the 3' promoter element in the transgene mRNA thus was a major factor determining transgenic recombination frequencies. As predicted from the model, transgene mRNA lacking the 3' promoter element would not be a good donor RNA for transgenic recombination. Consequently, no transgenic recombination was detected in transgenic plants expressing the 3' truncated MP mRNA upon inoculation with RCNMV RNA-1.

CHAPTER 1.

Current Research Status of the Molecular Biology of *Red Clover Necrotic Mosaic Virus* and Other Related Viruses

Red clover necrotic mosaic virus (RCNMV) is a member of *Dianthovirus* genus (Murphy *et al.*, 1995). It is characterized by a bipartite genome composed of positive-sense, single-stranded RNA-1 of 4 kb and RNA-2 of 1.5 kb. Most physical and biological properties of this virus were studied during the 70's and 80's (reviewed by Hiruki, 1987). Since the elucidation of the complete nucleotide sequence and genome organization of RNA-1 and RNA-2 in late 80's (Lommel el al., 1988; Xiong and Lommel, 1989), and the availability of infectious full-length cDNA clones (Xiong and Lommel, 1991), RCNMV has been developed as a model system to study at a molecular level viral movement and replication in the last decade. A review of the current knowledge about RCNMV as well as viral RNA replication and recombination is presented.

1.1. RED CLOVER NECROTIC MOSAIC VIRUS

1.1.1. Physical and biological characters

RCNMV has a 28-32 nm icosahedral virion composed of 180 copies of the 37 kd capsid protein. Each virion contains both genomic RNA-1 and -2 (Murphy *et al.*, 1995). Since the first report that RCNMV infected red clover (*Trifolium pratense*) in

Czechoslovakia in 1969, different isolates have been reported from Czechoslovakia, Canada, England, Australia, Sweden, New Zealand, Ireland, and Poland (Hiruki, 1987). In nature, it was found in red clover, sweet clover, white clover, and alfalfa. Although it is not known how the virus is transmitted naturally, RCNMV can be transmitted mechanically by sap inoculation in laboratories. The host range of RCNMV includes *Nicotiana benthamiana*, *N. clevelandii*, cowpea, and red kidney bean. The infection typically causes chlorotic and necrotic spots on the infected leaves, necrotic mosaics and deformation on the young systemic leaves, and plant stunting (Hiruki, 1987).

1.1.2. Viurs infection cycle (Figure 1.1)

Plant virus can infect host plant and single cell infection can result in systemic infection of whole plant. As shown in figure 1.1, after entering into plant cell, virus disassemblies to remove the coat protein. Early genes such as genes encoding for replication proteins are expressed from the exposed viral RNA. These viral replication proteins and host proteins form a replicase complex that recognizes and catalyze the viral genomic RNA replication. A single genomic RNA can be amplified to millions in a single infected cell by genome replication. The progeny viral RNA can be either directly moved to neighboring cells by the function of viral movement protein expressed from viral genomic RNA as late gene product, or encapsidated by coat protein into virions that are moved to adjacent cells. Eventually, the virus can move long distantly to whole plant and cause systemic infection.

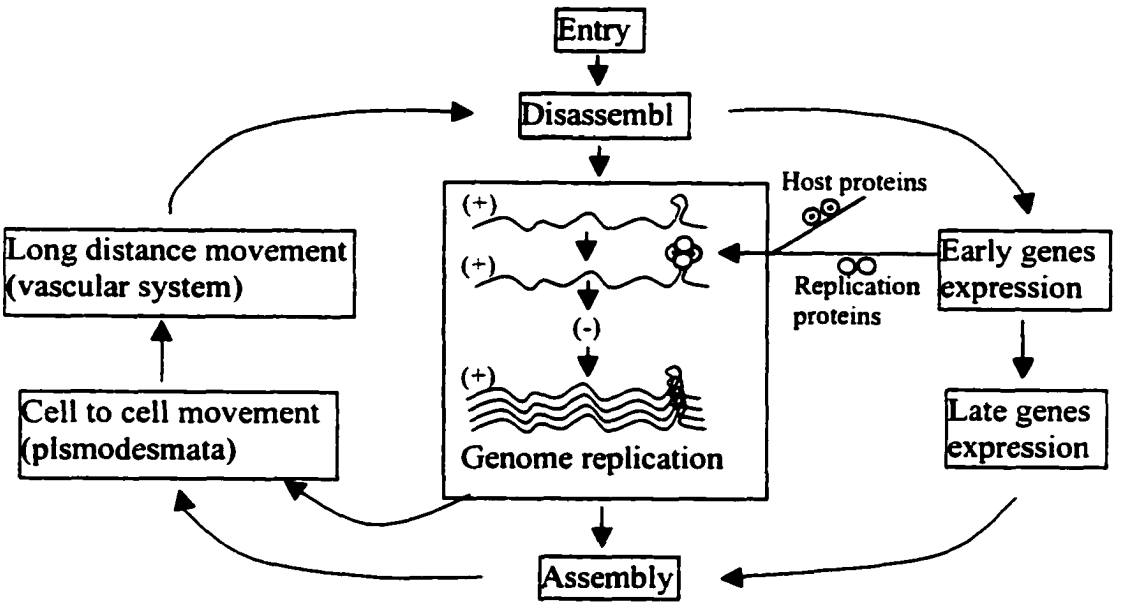


Figure 1.1 Diagram of virus infection cycle. In order to infect plant, virus must be able to enter into host cell and remove the coat protein. The exposed viral genomic RNA is replicated by the replicase complex composed of host proteins and viral replication proteins. The amplified viral progeny RNA is either directly moved to neighboring cells or encapsidated by coat proteins and then moved to neighboring cells via plasmodesmata. Eventually, the virus is moved systemically to whole plant via vascular system.

1.1.2. Genome organization (Figure 1.2)

RCNMV contains two positive-sense, single-stranded genomic RNAs with 5' m-^{7}GpppA cap (Xiong and Lommel, 1989; Xiong *et al.*, 1993a). RNA-1 of 3889 nucleotides in length contains three large open reading frames (ORFs) encoding for a 27 kd polypeptide (p27), a 57 kd polypeptide (p57), and a 37 kd capsid protein (CP) from the 5' end to the 3' end (Figure 1.2). In addition, a p27-p57 fusion protein, p88 RNA dependent RNA polymerase (RdRp), is produced (Xiong *et al.*, 1993b). RNA-2 contains a single ORF encoding for a 35 kd movement protein (MP) (Lommel *et al.*, 1988).

RCNMV is most closely related to other two dianthoviruses, *Sweet clover necrotic mosaic virus* (SCNMV) (*Ge et al.*, 1993) and *Carnation ringspot virus* (CRSV) (Ryabov *et al.*, 1994). In addition to dianthoviruses, p88 RNA RdRp of RCNMV shows significant amino acid similarity to homologues from *Carmovirus* (Guilley *et al.*, 1985; Weng and Xiong, 1997; Carrington *et al.*, 1989), *Tombusvirus*, *Luteovirus* (Miller *et al.*, 1988), *Necrovirus*, and *Maize chlorotic mottle virus* (MCMV) (Nutter *et al.*, 1989). Phylogeny of RdRp of positive-stranded RNA viruses reveals three supergroups (Koonin, 1991). RCNMV RdRp was placed in supergroup II composed of *Carmovirus*, *Tombusvirus*, *Dianthovirus*, *Pestivirus*, *Flavirus*, a subset of *Luteovirus* (BYDV), and single-stranded RNA bacteriophages. The amino acid sequence of 37 kD CP of RCNMV also shares high similarity to those of other small icosahedral RNA viruses. Multiple alignments of the shell (S) domains of the CP of small icosahedral RNA plant viruses generated three groups (Dolja and Koonin, 1991). RCNMV CP is among the group I containing

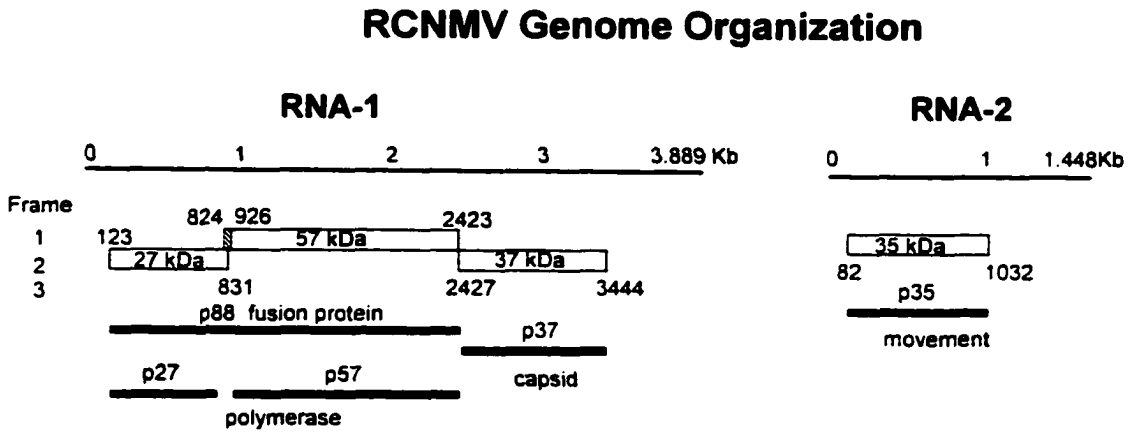


Figure 1.2 Schematic representation of RCNMV genome organization. RNA-1 and RNA-2 are depicted as solid lines with ORFs identified as open boxes. Numbers above or below ORFs identify the positions of ORF initiation and termination codons. The ribosomal frameshifting region in the RNA-1 is marked as a hatched box. Predicted and observed protein products are represented by black boxes. Known functions of RCNMV proteins are described below each protein.

Carmovirus, *Tombusvirus*, and *Dianthovirus*. Amino acid sequence of 35 kD MP is highly conserved within *Dianthovirus*, but is less homologous to other viruses. Interestingly, RCNMV MP exhibited a considerable degree of homology with the MP from *Furovirus*, a group of rod shaped, fungus-transmitted viruses. Phylogenic analysis of plant RNA and DNA viruses identify two families of MP (Koonin *el al.*, 1991). RCNMV MP belongs to family II consisting of *Bromovirus*, *Cucumovirus*, *Alfamovirus*, *Dianthoviruses*, and *Tobacco streak virus* (TSV).

In addition to the coding region, RNA-1 and -2 contain two terminal noncoding regions (Xiong and Lommel, 1989; Lommel *et al.*, 1988). The 5' terminal noncoding regions of RNA-1 (122 bases) and RNA-2 (79 bases) form no potential secondary structure, and share no homology with each other except the first six nucleotides (ACAAAC) at the 5' termini. Secondary structure analysis of the 3' noncoding region of RNA-1 (445-nucleotide) and RNA-2 (417 nucleotide) revealed non-tRNA like secondary structures. Alignment of this region identified a homologous region composed of the 3' terminal 29 nucleotides which is predicted to form nearly identical, highly stable stem-loops and is required for minus strand RNA synthesis (Turner and Buck, 1999) (details see 1.1.5 section).

1.1.3. Gene expression strategies

Multigenic RNA plant viruses have developed several different mechanisms to translate their polycistronic genome. These mechanisms include: (1) genome segmentation, (2) terminator readthrough, (3) ribosomal frameshifting, (4) subgenomic

RNA synthesis, (5) proteolytic processing, and (6) overlapping open reading frames (Xiong, 1988). RCNMV employs genome segmentation strategy to express the MP from the monocistronic RNA-2, ribosomal frameshifting to control the production of p88 fusion protein of RNA polymerase from polycistronic RNA-1, and subgenomic RNA synthesis to translate the CP gene located at the 3' terminus of RNA-1.

Ribosomal frameshifting

This mechanism was first reported to express the *gag-pol* fusion protein in retroviruses (Jacks and Varmus, 1985; Varmus, 1988). It is also involved in the expression of viral RNA polymerase in plant viruses, such as *Barley yellow dwarf virus* (BYDV) (Brault and Miller, 1992) and *Potato leaf roll virus* (PLRV) (Prufer *et al.*, 1992). Two RNA sequence elements are required for the -1 ribosomal frameshifting to occur (Jack *et al.*, 1988a). One element is the specific 7-nucleotide sequence (heptanucleotide) where the ribosomal shifting occurs. Different heptanucleotides can be generalized as sequences X XXY YYZ (triplets represent codons from the original frame; X = A, G, or U; Y = A or U; Z = A, C, or U). The second element is an RNA stem-loop structure downstream of the heptanucleotide. A slippage model was proposed in which the stable downstream secondary structure slows or stalls ribosome migration, resulting in increased probability of tRNA slippage at the shifty hepatanucleotide (Jacks *et al.*, 1988a).

In RCNMV, the existence of p88 as a p27-p57 fusion protein was suggested by its immunoprecipitation with antibodies against either p27 or p57 (Xiong *et al.*, 1993b). *In*

vitro and *in vivo* studies of mutant clones confirmed that this p88 RNA polymerase was indeed expressed through a -1 ribosomal frameshifting event (Xiong *et al.*, 1993b; Kim and Lommel, 1994). In accordance with the -1 slippage model (Jacks *et al.*, 1988a), a shifty heptanucleotide (G GAU UUU) was identified in RNA-1 immediately 5' to the amber termination codon of the p27 ORF (Xiong *et al.*, 1993b). Replacing RCNMV shifty heptanucleotide with those from a number of animal retroviruses and RNA plant viruses was able to support RCNMV frameshifting *in vitro* (Kim and Lommel, 1994). However, only a limited number of the heterologous shifty heptanucleotides were functional in plant cells. It was suggested that specific shifty tRNA populations in the cell facilitate viral -1 ribosomal frameshifting. In addition to the heptanucleotide, a stable stem-loop secondary structure downstream the heptanucleotide sequence was predicted (Xiong *et al.*, 1993b). Further mutation analysis of other sequence features adjacent to the heptanucleotide indicated that p27 stop codon and p57 start codon are not required for frameshifting (Kim and Lommel, 1994).

Subgenomic RNA (sgRNA) synthesis

Although accumulated to a very low level, the 1.5 kb sgRNA was detected in cowpea protoplasts inoculated with RNA-1 only (Osman and Buck, 1987). A 1.5 kb double-stranded (ds) RNA homologous to the 3' terminus of RNA-1, in addition to 4 kb and 1.5 kb dsRNAs of genomic RNA-1 and RNA-2, was detected consistently from RCNMV infected plant tissue (Osman and Buck, 1990). The 5' terminus of the sgRNA was mapped by RNA sequencing of the purified dsRNA to an adenosine residue

corresponding to RNA-1 nucleotide 2365 (Zavriev *et al.*, 1996). The CP sgRNA was 1525 nucleotides in length with a 62-nucleotide 5' leader sequence upstream of the CP ORF (Zavriev *et al.*, 1996). The 62-nucleotide 5' leader sequence contains a 14-nucleotide promoter sequence homologous to the RNA-1 5' terminus and sequences highly similar to the 5' leader translational enhancer element Ω identified in *Tobacco mosaic virus* (TMV) (Gallie *et al.*, 1997, 1998; Zavriev *et al.*, 1996).

The synthesis of CP sgRNA from RNA-1 required RNA-2 sequence (Vaewhongs and Lommel, 1995). RNA-1 alone can produce local lesion in the inoculated leaves of transgenic *N. benthamiana* plants expressing RCNMV MP. However, CP synthesis is observed only when RNA-1 is co-inoculated with wild-type RNA-2 or mutant RNA-2 that is incapable of MP synthesis. A 34-nucleotide trans-activator in RNA-2 MP coding region at nucleotide positions 756-789 is required to transactivate the synthesis of CP sgRNA. This trans-activator is predicted to form a stem-loop structure and functions *via* base-pairing between its 8-nt loop sequence with the promoter of CP subgenomic RNA (Sit *et al.*, 1998).

1.1.4. Viral movement

There are two types of viral movement in host plants. While the cell-to-cell movement through plasmodesmata (plasma membrane-lined cylindrical pores that traverse the cell wall) allows viruses to migrate from originally infected cell to the adjacent cells, long-distance movement through vascular system permits viruses to move

rapidly from inoculated leaves to uninoculated leaves and eventually throughout the entire plant.

The role of RCNMV CP in viral cell-to-cell and systemic movement has been studied by mutational analysis (Xiong *et al.*, 1993a). All CP mutants including the deletion of the entire CP gene do not affect RNA replication and symptom production in the inoculated leaves, suggesting that it is not required for cell-to-cell movement. Mutations deleting up to 16 amino acids of CP N-terminus affect neither CP production, virion assembly nor wild-type infection, whereas internal and C-terminal deletions prevent CP expression, virion formation and wild-type systemic infection, suggesting a positive relation between CP expression, virion formation and long distance movement. Moreover, when inoculated onto the MP-expressing transgenic plants, RNA-1 is able to move from cell-to-cell at a rate similar to wild-type virus, but cannot move from leaf-to-leaf (Vaewhongs and Lommel, 1995). Neither CP nor virions was detected in the inoculated leaves. However, when RNA-1 was coinoculated with RNA-2 mutants, which do not express a functional MP but still contains the 34-nucleotide trans-activator for CP sgRNA synthesis (Sit and Lommel, 1998), CP and virions were detected and a systemic infection were resulted (Vaewhongs and Lommel, 1995). Taken together, these data demonstrated that CP production and virion formation are not required for cell-to-cell movement but required for long-distance movement of RCNMV. It was hypothesized that virions protect the viral genome from nuclease activity present in the vascular fluid or/and that interaction between virion-forming CP and a host factor is required for RCNMV to enter the vascular system (Cookmeyer *et al.*, 1994).

RCNMV RNA-1 alone is able to replicate in single plant cells, but is restricted to the originally infected cell, indicating that p35 MP expressed from RNA-2 is required for cell-to-cell movement. Wild-type MP of RCNMV is able to bind to single-stranded nucleic acids (RNA or DNA) *in vitro* in a cooperative and sequence non-specific manner (Osman *et al.*, 1992). To identify the functions that are associated with RCNMV MP, 12 MP mutants were generated by alanine-scanning site-directed mutagenesis, in which clusters of two or three charged residues dispersed throughout the protein were changed to alanines (Giesman-Cookmeyer and Lommel, 1993). All 12 mutants were asssayed for RNA binding *in vitro* by gel shift assay and for their ability to potentiate RCNMV cell-to-cell movement *in vivo* by infectivity analysis. Mutant proteins 204, 242, and 280 bound RNA *in vitro* less efficiently than wild type but still facilitated cell-to-cell movement and systemic infection. Mutants 27-31 and 305 bound RNA at only 20% of the wild-type level and lost cooperative binding, but still were able to move from cell-to-cell and leaf-to-leaf. Mutant 278 has wild-type level and cooperative *in vitro* RNA binding properties, but was incapable of potentiating *in vivo* cell-to-cell movement and systemic infection. These results suggested three functional domains associated with RCNMV MP: an RNA binding domain, a cooperative RNA binding domain, and a third domain required for cell-to-cell movement *in vivo*. These data also indicated that only limited RNA binding is necessary and that cooperative binding is not required for cell-to-cell movement.

By using the combination of fluorescently labeled macromolecules and electron microscopy, Wang *et al.* (1998) provided direct evidences to show that RCNMV MP can

potentiate the cell-to-cell trafficking of RCNMV RNAs through plasmodesmata in plants. Plasmodesmata normally have a size exclusion limit (SEL) of 0.8 to 1 kd (2.5 nm in diameter), which restricts the transport to only small molecules such as metabolites, irons, and hormones (Fujiwara *et al.*, 1993). The fluorescently labeled 35 kd RCNMV MP, when directly microinjected into mesophyll cells of a cowpea leaf, moved rapidly from cell-to-cell (Fujiwara *et al.*, 1993). When this MP was coinjected into mesophyll cells of a cowpea leaf, a 9.4 kd fluorescein-conjugated dextran (F-dextran) spread into the neighboring and more distant cells (Fujiwara *et al.*, 1993). The data suggested that RCNMV MP interacts with plasmodesmata to increase the SEL to 10-fold higher than the normal size in a manner analogous to that of TMV MP expressed in transgenic tobacco plants (Wolf *et al.*, 1989). This MP was able to potentiate the cell-to-cell trafficking of fluorescently labeled RCNMV RNAs, but not the single-stranded or double-stranded DNA derived from infectious cDNA clones of RCNMV. The ability to differentially traffic RNA and DNA indicated that trafficking of nucleic acids through plasmadesmata mediated by RCNMV MP is a selective process. Twelve alanine-scanning mutants were tested for the capability of inducing the plasmodesmatal SEL increase and potentiating cell-to-cell trafficking of RCNMV RNAs. All of the mutants that failed to increase SEL did not support cell-to-cell trafficking of RCNMV RNAs *in vitro* and virus infection in host plant, whereas all of the mutants that elicited an increase in SEL also potentiated RCNMV cell-to-cell movement except mutant 280. Mutant 280, which did not potentiate RNA cell-to-cell trafficking and viral infection, did potentiate an increase of plasmodesmal SEL. These data demonstrated that an increase in the plasmodesmal SEL

is necessary, but not sufficient, for RNA cell-to-cell trafficking and that RCNMV moves from cell to cell as ribonucleoprotein complex (RNA-MP). Interestingly, electron microscopic studies showed that although the RCNMV MP bound to and potentiated trafficking of RCNMV RNAs, it did not unfold and extend these molecules *in vitro*, a phenotype observed for TMV MP. Unfolding of RCNMV RNAs is perhaps not necessary for their cell-to-cell movement *via* plasmodesmata. Alternatively, if unfolding of RCNMV RNAs is a prerequisite for plasmodesmal trafficking, it may involve the participation of host factors.

Because cell-to-cell movement is a prerequisite for systemic infection, it has been difficult to determine what role(s), if any, viral MPs may play in long-distance transport that are distinct from those associated with cell-to-cell movement. Plasmodesmata connections between different types of plant cells (tissues) are not necessary equivalent and, consequently, the requirements for cell-to-cell movement may not be common across all cell types (Ding *et al.*, 1992; Ghoshroy *et al.*, 1997; Gilbertson and Lucas, 1996). Wang *et al.* (1998) hypothesized that viral MPs possessed separate functional motifs essential for cell-to-cell movement of the viral genome across specific tissue barriers. In *N. benthamiana*, all 6 alanine scanning mutants (27-31, 204, 242, 301, 305, and double mutant 27-31/305), which were capable of potentiating cell-to-cell trafficking of RCNMV RNAs and inducing viral infection in the inoculated leaves, can also cause systemic infection (Giesman-Cookmeyer and Lommel, 1993; Fujiwara *et al.*, 1993). However, these same mutants produced more interesting phenotypes in *N. edwardsonii* and cowpea. In *N. edwardsonii*, all 6 MP mutants can move cell-to-cell and cause

infections on inoculated leaves, but mutants 27-31, 204, and 27-31/305 did not induce systemic infection and mutant 305 induced systemic infections of only 25% of the inoculated plants. In cowpea plants mutants 27-31, 305, and 27-31/305 did not induce systemic infection and mutant 204 caused systemic infection of only 13% of the inoculated plants. These data demonstrated that mutations in the RCNMV MP could restrict viral movement to the inoculated leaves in two systemic hosts, *N. edwardsonii* and cowpea. Cellular distribution of RCNMV within the inoculated leaves and the upper leaves of the inoculated *N. edwardsonii* and cowpea were examined under the light and electron microscope. Immunolocalization using a polyclonal antibody against the RCNMV CP revealed one cellular boundary at which the RCNMV MP functions to facilitate entry into the phloem long-distance transport system. This boundary is located at the interfaces between the bundle sheath and phloem parenchyma cells and the companion cell-sieve element (CC-SE) complex (Wang *et al.*, 1998). Interestingly, in a local infection host plant of RCNMV, *N. tabacum*, the cell-to-cell movement of RCNMV was blocked at this same intercellular boundary. Together, these data demonstrated that RCNMV MP played a role in long-distance transport in a host-specific manner, which was distinct from its role in cell-to-cell movement. Since RCNMV CP was also required for systemic infection (Xiong *et al.*, 1993a), Wang *et al.* (1998) speculated that RCNMV CP might act at the same cellular location to facilitate long-distance transport. They further hypothesized that RCNMV MP and CP function in a host-specific manner to potentiate the trafficking of the infectious agent from either the bundle sheath or phloem

parenchyma cells into the CC-SE complex, and that further entry of RCNMV into the phloem translocation stream may require additional host and /or viral factors.

1.1.5. Viral replication

RCNMV RNA-1 encodes three functional proteins of p27 and p88 polymerase subunits, and p37 CP. RNA-2 encodes a single p35 MP. Genetic evidence suggested that p27 and p88 are the only two viral proteins needed for RNA replication, although expression of all four proteins is required for wild type infection in systemic host plants (Cookmeyer *et al.*, 1994). RNA-1 alone is able to replicate in protoplasts, indicating that MP encoded in RNA-2 is dispensable for RNA replication (Osman and Buck, 1987; Paje-Manalo and Lommel, 1989). A CP deletion mutant can cause infection in the inoculated leaves of *N. benthamiana* when co-inoculated with wild type RNA-2, suggesting that CP is also not required for RNA replication (Xiong *et al.*, 1993a). Changing the initiation codon of p57 ORF to an aspartate codon precluded p57 synthesis *in vitro*, but did not affect the infectivity (Kim and Lommel, 1994), suggesting that p57 was not required for wild type virus infection and that p57 ORF might only be needed for the production of p88. Although the mutations that abolish the p27 termination codon did not affect the synthesis of p88 fusion protein, their infectivity were lost (Kim and Lommel, 1994), demonstrating that p27 protein, in addition to p88, is required for RNA synthesis.

An RNA-dependent RNA polymerase (RdRp) complex was isolated from RCNMV infected *N. clevelandii* plants and characterized (Bates *et al.*, 1995). It was membrane-associated and template-bound. RdRp preparations contained viral p27 and p88 proteins

along with several minor proteins of possible host origin. The polymerase became template-dependent and template-specific after endogenous RNA was removed by micrococcal nuclease treatment. It was able to utilize RCNMV RNAs as templates, but not RNAs of three viruses in different taxonomic groups, *Cucumber mosaic cucumovirus* (CMV), *Tomato bushy stunt tombusvirus* (TBSV), and *Tomato mosaic tobamovirus* (TMV). The products of the *in vitro* RNA polymerase reactions were full-length double-stranded RNAs corresponding to RNA-1 and -2, indicating that the RNA polymerase preparation lacks a factor(s), possibly a helicase that is required for the synthesis of the single-stranded, positive-sense, progeny RNA. A helicase domain was absent in the RdRP of RCNMV and other viruses in the supergroup II (Koonin, 1991; Koonin and Dolja, 1993). Alternatively, the purified RNA polymerase might contain an inhibitor (s) for the synthesis of positive-stranded RNA. Moreover, RNA synthesis by RdRp *in vitro* was completely inhibited by antibodies to a peptide containing the GDD motif that is conserved in all RdRps, whereas the activity of the template-bound polymerase was not affected by these antibodies. It was hypothesized that binding of these antibodies to the GDD region may interfere with the binding of the template to the catalytically active site of the polymerase (Bates *et al.*, 1995).

The roles of the 3' and 5' untranslated region (UTR) in RNA replication were recently reported (Turner and Buck, 1999). Since all the proteins required for RNA replication are encoded in RNA-1, mutations were incorporated in RNA-2 without affecting the production of replication proteins. The mutant RNA-2 was co-inoculated with wild type RNA-1 into *N. clevelandii* protoplasts. The replication and accumulation

of RNA-2 were detected by northern blot or by reveres transcription PCR (RT-PCR). Like many other positive RNA viruses, the 3' noncoding regions of RCNMV RNA-1 and -2 are predicted to form a cluster of stem-loop secondary structures. The 3' 29 nucleotide of both RNA-1 and -2 is predicted to form a stable and nearly identical 3'-terminal stem-loop structure (SLS). Deletion analysis showed that most of the 3' UTR including this 3'-SLS was needed for the efficient synthesis of negative-strand RNA-2. The roles of the sequence and structure of the 3'-SLS was investigated further. Mutants that disrupted stem base-pairing or altered the 3 nt. loop sequence were unable to support RNA-2 replication and accumulation to a detectable level, whereas stem compensatory mutant that reconstituted stem base-pairing restored the production of RNA-2 to the wild type level. These results suggested that the secondary stem-loop structure as well as the three-nucleotide loop sequence was important for the synthesis of minus strand RNA-2, but the primary sequence at the stem was not critical. The 5' UTR of RNA-1 and RNA-2 is not predicted to form any secondary structure and does not share any significant homology except the 5' ACAAAC. Deletions in the 5' noncoding region of RNA-2 showed that 5'ACAAAC sequences as well as sequences across the whole region were required for synthesis of the positive strand but not for production of the negative strand.

1.2 REPLICATION OF POSITIVE-STRANDED RNA VIRUSES

Over 90% of plant viruses and many important animal viruses have positive RNA genomes (Murphy *et al.*, 1995). Viral genome replication is the most fundamental and critical step in the virus infection. The enzyme catalyzing virus replication is replicase

complex composed of viral proteins and host proteins. Genome replication has two stages: 1) the replicase complex recognizes the template elements of genomic positive-stranded RNA and initiates the synthesis of minus-stranded RNA from the 3' end of the template; and 2) replicase complex synthesizes the progeny positive-stranded RNA (and subgenomic RNA if applicable) using minus-stranded RNA as template. (Buck, 1996).

1.2.1 Replication proteins

RNA-dependent RNA polymerases (RdRp)

RdRp is the catalytic subunit of the replication complex. All sequenced positive-stranded RNA viruses encode a putative RdRp that is required for viral RNA replication and has been biochemically shown to have the RNA polymerase activity (Buck, 1996). Phylogenetic alignments of the amino acid sequences of RdRps of all groups of positive-stranded RNA viruses revealed eight conserved motifs with the most conserved Gly-Asp-Asp (GDD) domain located at motif VI (Koonin, 1991; Koonin and Dolja, 1993). The alignment delineates RdRps into three supergroups: Group I contains RdRps in picornaviruses, nodavirses, comoviruses, nepoviruses, potyviruses, bromoviruses, sobemoviruses, and a subset of luteoviruses; group II is composed of RdRps in carmoviruses, tombusviruses, dianthoviruses, another subset of luteoviruse (BYDV), pestiviruses, *Hepatitis C virus* (HCV), flaviviruses, and single-stranded RNA bacteriophages; group III includes RdRps in tobamoviruses, bobraviruses, hordeiviruses, tricornaviruses, *Beet yellows virus*, alphaviruses, rubiviruses, furoviruses, *Hepatitis E virus*, potexviruses, carlaviruses, bymoviruses, and *Apple chlorotic leaf spot virus*

(Koonin, 1991). The importance of the GDD motif and some other conserved motifs was confirmed by mutagenesis in different viral RdRps (Buck, 1996). Amino acid sequence alignments of RdRps to other three types of polymerases, RNA-dependent DNA polymerases (RdDps) or reverse transcriptases (RTs), DNA-dependent RNA polymerases (DdRps), and DNA-dependent DNA polymerase (DdDps), suggested the presence of the motifs IV and VI counterparts in all type of polymerases, which led to the speculation of similar protein folding (Heringa and Argos, 1994).

Crystallographic studies of the three-dimensional structures of the Klenow fragment of *Escherichaia coli* DNA polymerase I (a DdDp) (Ollis *et al.*, 1985), *Human immunodeficiency virus* (HIV) type I RT (a RdDp) (Koehlstaedt *et al.*, 1992), phage T7 RNA polymerase (a DdRp) (Sousa *et al.*, 1993), and RNA polymerases (RdRp) of *Poliovirus*, HCV, and bacteriophage Φ6 (Bressanelli *et al.*, 1999; Butcher *et al.*, 2001; Hansen *et al.*, 1997; Hong *et al.*, 2001; Lesburg *et al.*, 1999) further demonstrated that all types of polymerases have a similar right hand configuration composed of three subdomains, "fingers", "thumb", and "palm" (Figure 1.3). Although the structure of the fingers and thumb subdomains of *Poliovirus* polymerase differed from those of other types of polymerases, the palm subdomain was very similar to that of other polymerases. The palm domain contained a core catalytic site consisting of four amino acid sequence motifs described for RNA dependent polymerase (Poch *et al.*, 1989; Xiong and Eickbush, 1990). In addition to the three subdomains, *Poliovirus* RdRp contained a unique structure element composed of N-terminal residues that interacted with the thumb subdomain and extended to the catalytic site in the palm subdomain.

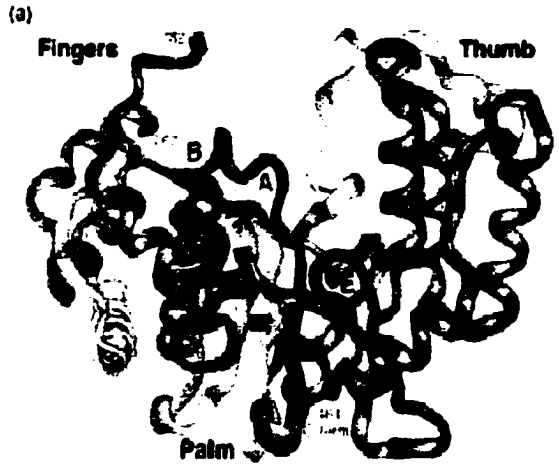


Figure 1.3 Three-dimensional structure of poliovirus $3D^{pol}$ protein. It composed of three subdomains of "fingers", "thumb", and "palm". The palm subdomain contains four motifs of "A" in red, "B" in green, "C" in yellow, and "E" in blue (From Hansen *et al.*, 1997).

By using the structure of *Poliovirus* polymerase as a guide, the secondary structures of the RdRps from six different viral families of *Bromoviridae*, *Tobamoviridae*, *Tombusviridae*, *Leviviridae*, *Flaviridae*, and *Picornaviridae* were predicted (O'Reilly and Kao, 1998). The RdRp of all six families shared a similar order of secondary structure elements. All except the leviviruses contained a similar unique N-terminal region as observed in the *Poliovirus* polymerase. Based on the phenotypes of mutations in several RNA polymerases, the functions of the motifs and subdomains were predicted: motif A for magnesium binding, motif B for the discrimination of ribose versus deoxyribose, motif C for magnesium binding, motif D completed palm core structure with unknown function, motif E for hydrophobic interaction with thumb, the fingers subdomain determined RNA templates, the thumb subdomain may be involved in template binding, the N-terminal unique region for catalytic activity and oligomerization of polymerase (O'Reilly and Kao, 1998).

In addition to the polymerization activity, *Poliovirus* polymerase carries other activities, including RNA binding (Pata *et al*, 1995; Beckman and Kirkegaard, 1998), NTP binding (Richards *et al*, 1995; Richards and Ehrenfeld, 1997), and RNA duplexes unwinding (Cho *et al.*, 1993).

RNA Helicases

Most positive-stranded RNA viruses, including alpha-like, picorna-like, sobemo-like, corona-like and flavi-like supergroups, contain putative helicases (Buck, 1996). The ORFs encoding the helicases are required for virus replication, such as in BMV (Kroner *et al.*, 1990). Several functions are proposed for helicases: unwinding dsRNA replication

intermediate to make the strands available for further replication; removing template RNA secondary structure to facilitate the initiation of minus-stranded RNA synthesis from the 3' end of positive-stranded RNA template and to facilitate the process of RNA polymerase through internal secondary structure. Based on amino acid sequence alignment, RNA helicases of positive-stranded RNA viruses were placed into three superfamilies (Koonin and Dolja, 1993). Two ATP-binding motifs are highly conserved among all three superfamilies (Buck, 1996). The motif II in superfamily 2 contains sequence DEXH, a variant of DEAD. Some of the DEAD box proteins, such as the cellular RNA helicase eIF-4A (Buck, 1996), have been shown to have activities of RNA-dependent ATPase, RNA binding, and RNA helicase. These activities have also been shown for a number of DEXH-containing helicases in superfamily 2, such as the CI protein of *Plum pox virus* (PPV) (Fernandez *et al.*, 1995). While no helicase activity has been demonstrated for the putative helicases in superfamilies 1 and 3, mutation studies of the conserved amino acids in the motif I (A) or motif II (B) demonstrated the essential role of the putative helicases in virus replication, such as the helicase-like domain of the *Turnip yellow mosaic virus* 140 kd protein (Weiland and Dreher, 1993), 58 kd protein of *Cowpea mosaic virus* (Peters *et al.*, 1994), and 2C protein of *Poliovirus* (Mirzayan and Wimmer, 1992; Teterina *et al.*, 1992). Interestingly, some viruses with genomes smaller than 6 kb, including Carmo-like and Sobemo-like superfamilies, lack the putative helicase (Buck, 1996). Several possibilities have been proposed to explain the helicase-lacking phenomenon (Buck, 1996): (1) the NTP-binding motifs may be present, but have diverged too much to be easily recognized by primary sequence alignment, such as the

variants of motif A (I) and motif B (II) in a cellular NTP-binding protein actin; (2) the RdRp may have the duplex unwinding activity, as shown for poliovirus RdRp; (3) unwinding may be accomplished by a helix-destabilizing protein which utilizes the energy by stoichiometric binding to single-stranded nucleic acid in the absence of NTP hydrolysis, such as the cooperatively binding of poliovirus $3D^{pol}$ protein to single-stranded RNA; (4) a cellular helicase may be recruited by the virus to unwind the duplex.

Capping and methylation enzymes

The presence of a 5' cap structure is characteristic of most eukaryotic mRNAs and mRNAs of DNA viruses, retro- and pararetroviruses, negative-stranded RNA viruses, some double-stranded RNA viruses, and some positive-stranded RNA viruses. The 5' cap structure is M^7GpppN, where N is usually A or G. Capping of the cellular mRNAs and those of most DNA viruses, retro, and pararetroviruses occurs in the nucleus using cellular capping enzymes. Capping of the genomic RNAs of positive- and negative-stranded RNA viruses occurs in the cytoplasm using virus-encoded enzymes (Murphy *et al.*, 1995). The capping reaction consists of several reactions:

(1) pppNpN' $\rightarrow$ ppNpN' + p_i by an RNA triphosphatase;

(2) guanylytransferase + GTP $\rightarrow$ guanylyltransferase-GMP + pp_i ;

(3) guanylyltransferase-GMP + ppNpN' $\rightarrow$ GpppNpN' + guanylyltransferase;

(4) GpppNpN' + S-adenosylmethionine $\rightarrow$ M^7GpppNpN' + S-adenosylhomocysteine

Transgenic tobacco plants expressing antisense RNA for S-adenosylhomocysteine hydrolase, which controls the cellular S-adenosylhomocysteine / S-adenosylmethionine ratio, were resistant to infection by viruses with 5' capped genomic RNAs, such as tobacco mosaic virus, potato virus X, and cucumber mosaic virus, but much less resistant to infection by viruses containing 5'-VPg genome, such as potato virus Y (Masuta *et al.*, 1995). BMV-encoded replication protein 1a, with capping domain in N-terminus and helicase-like domain in C-terminus, was biochemically shown to catalyze the capping reaction (Ahola and Ahlquist, 1999). Ahola *et al.* (2000) provided direct evidences for the important function of viral RNA capping. Single amino acid mutation at the predicted capping active site of BMV 1a caused very low levels of negative strand accumulation and no detectable subgenomic RNA in yeast expressing 1a and 2a. However, deletion of XRN1 gene encoding the major exonuclease that degrades uncapped mRNAs suppressed the capping mutants, allowing synthesis and accumulation of large amounts of uncapped RNA 3 and subgenomic RNA.

Genome-linked virus proteins (VPgs)

The 5'-termini of genomic RNAs of positive-stranded RNA viruses in the picorna-like and sobemo-like superfamilies are covalently linked by a phosphodiester linkage to VPg, a virus-encoded protein. Mutation analysis has indicated that the poliovirus VPg (3B) consisting of 22 amino acids is essential for RNA replication (Xiang *et al.*, 1995a). VPg was found covalently linked to the 5' ends of newly synthesized negative and positive strands of poliovirus (Wimmer *et al.*, 1993). VPg-pUpU was present in infected

cells and could be synthesized by purified replicase complex and elongated to longer VPg-RNA molecule (Buck, 1996). It was proposed that VPg-pUpU might function as a primer for poliovirus RNA synthesis.

Host factors

The replicase complex that catalyzes the positive-strand RNA replication is composed of virus-encoded protein(s) and host-encoded protein(s). The host proteins in the replicase complex of phage Qβ, a positive-stranded RNA virus infecting *E. coli* bacteria, was identified as 30S ribosomal protein S1, protein synthesis elongation factors EF-Tu and EF-Ts, and host factor I (HF-I) (Blumenthal and Carmichael, 1979). HF-I is required only for minus-strand RNA synthesis.

The host proteins involved in eukaryotic positive-stranded viral RNA replication are not well known. Various approaches had been employed to study the host proteins. One of these approaches utilized partial purification of soluble and template-dependent replicase complex from infected hosts to search for the co-purified host proteins. Purified BMV replicase preparations capable of minus-strand RNA synthesis contained two host proteins of 160 kd and 45 kd, in addition two BMV-encoded proteins, 1a (helicase) of 115 kd and 2a (polymerase) of 100 kd, (Wuadt and Jaspars, 1990). The 45-kd host protein was identified as the barley analogue of the p41 subunit of the wheat germ eukaryotic translation initiation factor 3 (eIF-3), or a closely related protein (Quadt *et al.*, 1993). It reacted specifically with antiserum against wheat germ eIF-3, was able to bind with high affinity and specificity to 2a in the way similar to p41 subunit of wheat germ

eIF-3. eIF-3 as an integral part of the BMV replicase was further demonstrated by a 3-fold stimulation of minus-strand RNA synthesis resulted from the addition of wheat germ eIF-3 or its p41 subunit. The purified replicase preparation of RCNMV from infected *Nicotiana clevelandii* contained virus-encoded proteins of p27 and p88 polymerase protein, and several host proteins (Bates *et al.*, 1995). One of the co-purified host proteins has a similar size (45 kD) to the p41 subunit of eIF-3 co-purified with BMV replicase preparation.

Another approach to study host proteins was to search for those that bind specifically to termini of viral RNAs, which act as promoters for (-) and (+) RNA synthesis. Specific binding of host proteins to viral terminal sequences has been reported in diverse viruses including BMV (Duggal *et al.*, 1994; Duggal and Hall, 1995), RCNMV (Hayes *et al.*, 1994b), and *Poliovirus* (Najita and Sarnow, 1990). A cellular poly(C)-binding protein (PCBP) and the poliovirus protein $3CD^{pro}$ were found to bind to and form the ribonucleoprotein complex (RNP) with the 5' cloverleaf structure (Andino *et al.*, 1990a, 1993; Gamarnik and Andino, 2000). PCBP was first found to bind to and greatly increase the stability of human α-globin mRNA (Kiledjian *et al.*, 1995). Recent evidence suggested that PCBP was involved in controlling the expression of numerous cellular and viral RNAs (Ostareck-Lederer *et al.*, 1998). A poly(A) binding cellular protein (PABP) and $3CD^{pro}$ was also found to bind to the 3' end poly(A) and upstream stem-loop structure (Andino *et al.*, 1990a,b, 1993; Harris *et al.*, 1994). The formation of these ribonucleoprotein complexs in the poliovirus termini was essential for virus replication.

A new approach was to use the powerful yeast genetic system. BMV was demonstrated to replicate in yeast *Saccharomyces cerevisiae*, indicating that all cellular factors essential for BMV replication were present in yeast (Janda and Ahlquist, 1993). By mutating yeast cells followed by multi-step selection and screening process, Ishikawa *et al.* (1997) isolated a series of yeast mutants that inhibited BMV RNA replication and characterized them into multiple complementation groups. Noueiry *et al.* (2000) identified a mutation in yeast gene DED1 that inhibited BMV RNA replication but not yeast growth. DED1 encodes a DEAD (Asp-Glu-Ala-Asp)-box RNA helicase required for translational initiation of yeast mRNAs. Inhibition of BMV RNA replication by the mutant DED1 allele (ded 1-18) resulted from specific inhibition of BMV polymerase 2a expression. Lee *et al.* (2001) showed that BMV RNA replication was inhibited by a mutation in OLE1, a yeast gene encoding delta9 fatty acid desaturase, an integral ER membrane protein and the first enzyme in unsaturated fatty acid synthesis. OLE1 deletion and medium supplementation revealed that BMV RNA replication required unsaturated fatty acids, not the Ole1 protein.

1.2.2 Interaction between replicase proteins and assembly of replicase complex

Replication proteins interact with each other and with RNA template to form a functional replication complex in association with different cellular membranes. such as endoplasmic reticulum membrane in BMV, vacuolar membrane in AlMV, chloroplast outer membrane in TYMV, and nuclear membrane in potyviruses (Buck, 1996).

The polymerase-polymerase interaction and oligomerization were indicated for poliovirus polymerase from several lines of evidence. The RNA binding and template utilization by the purified polymerase exhibited a high degree of cooperativity (Pata *et al.*, 1995; Beckman and Kirkegaard, 1998), suggesting that direct polymerase-polymerase interaction might be important for polymerase function. The crystal structure of poliovirus polymerase revealed an extensive polymerase-polymerase interaction *via* two interfaces (Hansen *et al.*, 1997). Interface I mediated the interaction between the front of the thumb subdomain (C-terminal) of one molecule and the back of the palm subdomain of the adjacent molecule. Interface II involved in the interaction between the N-terminal residues 12-37 (the unique region) of one molecule and the thumb subdomain of another molecule. This polymerase-polymerase interaction was also demonstrated *in vitro* by chemical cross-linking study (Pata *et al.*, 1995) and in the yeast two-hybrid system (Hope *et al.*, 1997; Xiang *et al.*, 1998), and by mutagenesis of the specific amino acids at each of the two interfaces identified in the crystal structure (Hobson *et al.*, 2001). The results of these studies demonstrated that both interfaces of the polymerase-polymerase interaction observed in crystal structure were indeed functionally important. While interface I was required for efficient substrate RNA binding, interface II was crucial for forming catalytic sites.

In addition to polymerase-polymerase interaction, polymerase also interacted with other viral proteins. In poliovirus, 3D polymerase was able to bind 3AB, an intracellular membrane-associated protein, *in vitro* and in yeast two-hybrid system (Molla, *et al.*, 1994; Debra *et al.*, 1997; Xiang *et al.*, 1998). Detergent-solubilized 3AB had been shown

to stimulate both the polymerase activity of purified 3D polymerase and the proteolytic activity of 3CD (Lama *et al.*, 1994; Paul *et al.*, 1994; Plotch and Palant, 1995; Richards and Ehrenfeld, 1998). 3D mutants that were specifically defective in interaction with 3AB but not other viral and host proteins were defective in RNA synthesis (Debra *et al.*, 1997). Similarly, in BMV, 2a protein containing a polymerase-like domain was able to form a complex with 1a protein containing an N-terminal methyltransferase-like domain and a C-terminal helicase-like domain *in vitro* (Kao *et al.*, 1992) and *in vivo* (O'Reilly *et al.*, 1995). The 1a-2a interaction observed *in vitro* was shown to be essential for BMV RNA replication *in vivo*. A three-amino-acid insertion in 1a that blocked BMV RNA replication *in vivo* also blocked *in vitro* interaction with 2a, while another two-amino-acid insertion that rendered the 1a protein temperature sensitive for RNA replication interacted with 2a *in vitro* at 24°C but not at 32°C. The domains required for the direct 1a-2a interaction were mapped to the C-terminal helicase-like domain of 1a and 115 amino acid N-terminal segment preceding the polymerase-like domain of 2a (Kao and Ahlquist, 1992). These two identified domains are necessary and sufficient for 1a-2a interaction *in vitro* and in yeast (O'Reilly *et al.*, 1995). Requirement of the large size of the 1a helicase-like domain suggested that high order structures might be necessary for the interaction with 2a. Consistent with this hypothesis, the 1a helicase-like domain was found to be protease-resistant. Mutations in this region that render this region protease-sensitive were also defective for RNA replication *in vivo* (O'Reilly *et al.*, 1995). The BMV 1a-2a interaction was in agreement with the observation that 1a and 2a proteins co-localized on the endoplasmic reticulum (ER) at sites of viral RNA synthesis in plant cells

and yeast cells (Restrepo-Hartwig and Ahlquist, 1996, 1999). While 1a could independently localize on ER in the absence of 2a and viral RNA template, 2a is directed to ER by 1a *via* interaction with 1a (Chen and Ahlquist, 2000).

Polymerase was capable of interacting with host proteins in addition to viral proteins. Using the yeast two-hybrid system to screen a HeLa cDNA expression library, a human protein Sam68 (Src-associated in mitosis, 68 kDa) was found to strongly interact with poliovirus 3D polymerase (McBride *et al.*, 1996). Upon poliovirus infection, Sam68 relocated from the nucleus to the cytoplasm where poliovirus replication occurred.

In addition to the replication proteins, the assembly of BMV replicase complex in yeast cell required viral RNA template (Quadt *et al.*, 1995). The functional BMV replicase complex can be extracted from yeast expressing 1a and 2a replication proteins and containing RNA-3, but not from yeast that had normal amount of 1a and 2a expression but lacked RNA-3. Deletion of RNA-3 revealed that both the 3' and the intercistronic noncoding sequences, which contained the *cis*-acting elements required for RNA-3 replication *in vivo*, was required for replicase activity. These results suggested that assembly of active replicase required not only viral proteins but also viral RNA to recruit essential host factors into the replicase complex and that sequences at both the 3'-terminal initiation site and distant internal sites of RNA-3 templates might participate in BMV replicase assembly and initiation of minus-strand RNA synthesis.

1.2.3 Template recognition and initiation: cis-acting elements required for RNA replication

Synthesis of viral RNAs is always initiated from the 3' ends, which possess essential *cis*-acting sequences including promoters for viral RNA synthesis and sequences for the assembly of replicase complexes. Based on the fact that viral RNA replication required replication proteins in addition to the RNA elements, approaches that alter viral RNA sequences without any changes to the production of replication proteins have been used to study the *cis*-acting replication elements.

1.2.3.1 3'-terminal sequences of positive strand RNAs

The 3' termini of positive strand viral RNAs which contain the promoter elements are able to form various structures, including tRNA-like structures, stem-loop structures, and polyadenylated (poly(A)) tail.

3' tRNA-like structures (TLS)

TLS have been found at the 3' temini of various positive-stranded RNA viruses such as BMV, CMV, and tobacco mosaic virus (TMV). Deletion studies showed that the TLS were essential for RNA replication (Buck 1996). Exchanges of 3' tRNA-like termini between different viruses had been carried out to determine the template specificity of replicase complexes. BMV RNA-1 and RNA-2 were able to amplify BMV RNA-3 containing a 3'-terminal region of CMV RNA-3 (Ishikawa *et al.*, 1991a), but not BMV RNA-3 containing a 3'-terminal region of TMV RNA-3, (Rao and Grantham, 1994),

indicating the greater structural similarity of the 3' end between BMV and CMV RNAs than BMV and TMV RNAs. These data suggested that 3'-terminal sequences were important determinants of template specificity and that replicase complexes were able to recognize some common structural elements in related viruses.

Structural elements required for viral RNA replication have been well characterized in BMV TLS, which is formed by approximately 150-nucleotides at the 3' end of each of BMV genomic RNAs (Dreher and Hall, 1988). BMV TLS can be subdivided into A, B1, B2, B3, C and D regions (Felden *et al.*, 1994). Mutation analysis idenitfied several regions important for minus-strand RNA synthesis, They are (1) two C residues adjacent to the 3' terminus (+1 and +2); (2) 3' pseudoknot at region A; (3) region B1; (4) region C (Dreher and Hall, 1988). Competitions between templates and purified replicase preparations showed that only mutations in the bulge and hairpin loops of stem C decreased the ability of the TLS to interact with RdRp complex (Chapman and Kao, 1998). Mutations that disrupted the pseudoknot and stem B1 as well as mutations in the +1 (C) and +2 (C) apparently had no effect on RdRp binding. Furthermore, in the absence of the rest of the BMV TLS, region C is sufficient to interact with RdRp complex. Addition of the initiation sequence ACCA3' to stem C created a minimal, aritificial RNA promoter capable of directing RNA synthesis in an *in vitro* transcription system. Synthesis from this minimal minus-strand template required the stem-loop structure as well as the sequences in the bulged and hairpin loops. While all 4 nucleotides in the bulged loop contributed to recognition by replicase complex, only A1 in the hairpin loop was most critical for template-replicase interaction.

After template recognition and binding, replicase complex would start the initiation of minus strand synthesis. Initiation of BMV minus-strand synthesis was shown to begin at the penultimate cytosine residue (CCA3') (Miller *et al.*, 1986). By using an *in vitro* transcription system, Kao and Sun (1996) found that low GTP concentrations had a more detrimental effect on RNA synthesis than the limited amounts of the other nucleotides. Approximately 15-fold higher concentration of GTP than CTP and UTP was required. Only at limited GTP concentration, minus-strand RNA synthesis was stimulated by the inclusion of guanosine monophosphate (GMP) or specific oligoribonucleotides. Furthermore, guanylyl-3',5'-guanosine (GpG) could be incorporated into minus-strand RNA and increased the rate of minus-strand RNA synthesis. In the presence of oligoribonucleotide GpG, k_m for GTP decreased form 50 uM to approximately 3 uM while the k_m for other nucleotides remained unaffected (~3uM). These results suggested that BMV RdRp may possesses two distinct nucleotide binding sites, one was used in elongation and bound to all four NTPs and another recognized only the initiating nucleotide GTP, however, with some flexibility.

3' stem-loop structure (SLS)

The 3' ends of the genomic RNAs of carmo-like and sobemo-like viruses and some alpha-like viruses are capable of forming simple SLS (Buck, 1996). Among carmo-like viruses, *Turnip crinkle virus* (TCV) is best studied. TCV is a positive-strand virus with a single RNA genome of 4054 bases. The virus naturally associates with sat-RNAs and DI RNAs that range in size from 194 to 356 bases. Sat-RNAs and DI RNAs do not encode

any functional proteins. Their replication and accumulation in plants requires co-inoculation of TCV genomic RNA. By using an *in vitro* transcription system, it was shown that the promoter for the synthesis of minus-strand sat-RNA C (356 base) was completely contained within the 3'-terminal 29 nt of the positive strand (Song and Simon, 1995). Structural probing revealed the presence of hairpin structure in this region. Mutational analysis indicated that the primary sequences of the loop and the lower part of the stem were not important for replication, however, the primary sequence of the upper stem and the maintenance of the stem were required for replication. The ability of these sat-RNA C mutants to accumulate *in vivo* was further analyzed (Stupina and Simon, 1997; Carpenter and Simon, 1998). The *in vivo* results confirmed that the loop sequence was not important for replication, however, transcripts containing compensatory exchanges in the upper stem region that had limited activity *in vitro* were biologically active *in vivo*, indicating that positioning of specific bases in the stem was not required to form an active promoter for minus strand synthesis. The 3'-terminal sequences of TCV genomic RNA and sat-RNA C are 90% identical and the 3' genomic RNA could be folded into a similar hairpin structure. In addition to the 3'-terminal SLS, mutagenesis showed that upstream motifs in the 3' nontranslated region of TCV RNA were also important for RNA replication *in vivo* (Carpenter *et al.*, 1995).

The DI RNA of *Cymbidium ringspot virus* (CyRSV) retained a block of 102 nucleotides from the 3' terminus of the genomic RNA (Havelda *et al.*, 1995). Mutation analysis showed that the 3' 77 nucleotides were required for replication. This 3' 77 nt of CyRSV DI RNA were predicted to form a structure composed of three SLSs and two

non-basepaired region. Mutagenesis suggested that the maintainance of the SLSs was essential for replication (Havelda and Burgyan, 1995).

The 3' 29 nt of RCNMV RNA-1 and RNA-2 are nearly identical and can be predicted to form a stable SLS. Deletion analysis shown that this 3' terminal 29 nt region was required for replication in *N. clevelandii* protoplasts (Turner and Buck, 1999). Substitution mutations shown that the SLS as well as 3 nt loop sequence were important for replication, whereas the primary sequence of the stem region was not important. In addition to the 3' terminal SLS, deletion mutations suggested that upstream sequences in the 3' untranslated region were also important for RNA replication.

3' poly(A) tail

The 3' ends of all picorna-like viruses and some alpha-like viruses are polyadenylated. The poly(A) tail is present not only to protect the RNA from 3' exonuclease degradation but also to provide an essential *cis*-acting sequence for RNA replication (Buck, 1996). A comparison of poliovirus RNA transcripts with $poly(A)_{12}$ and $poly(A)_{80}$ tails showed that a short poly(A) tail had no significant effect on translation but dramatically inhibited minus strand synthesis (Barton *et al.*, 1996). In addition to poly(A) tail, the upstream sequences and structure were also important for replication. The conserved sequences and structures in the 3'-untranslated region in poliovirus RNA were required for efficient negative strand synthesis (Rohll *et al.*, 1995; Sarnow *et al.*, 1986; Wang *et al.*, 1999). In cowpea mosaic virus, the 3' 151 nt of RNA M contains all the 3'-terminal *cis*-acting elements required for RNA replication (Roll *et al.*, 1993). The 3' 65 nt

upstream of the poly(A) in both RNA B and RNA M have a high degree of sequence similarity and can be folded into a SLS, containing four A residues from the poly(A). Mutation analysis suggested the importance of both the poly(A) and the SLS for replication. Another upstream SLS within the 3'-terminal 151 nt of RNA M was also required for replication, although the sequences of this region have little homology between RNA M and RNA B. Exchanging of the 3'-terminal 200 nt of RNA M by the 3'-terminal 500 nt of RNA B had only a minimum effect on replication (van Bokhoven *et al.*, 1993).

1.2.3.2 5'-terminal sequences of positive strand and 3'-terminal sequences of negative strand

The 5' sequences and structures were frequently conserved among the segmented genomic RNAs. Since the positive strand RNA was synthesized from the 3' end of the negative strand, it was thought that the 5' end of genomic RNA, *via* its complementary sequence at the 3' of negative strand, was involved in the synthesis of progeny positive strand RNA.

On the contrary, the 5' termini of RNA viruses have been shown to affect the synthesis of progeny negative strand RNA. The 5'-terminal 90 nucleotides of poliovirus RNA was capable of folding into a cloverleaf structure. This region was able to form a similar structure at the 3' end of negative strand (Andino *et al.*, 1990, 1993). Mutations which disrupted the base-pairing of the stems of the cloverleaf in positive strand but not the negative strand by using G-U pairs resulted in no RNA replication in host cells. In

contrast, mutations that disrupted the cloverleaf structure in the 3' end of negative strand but maintained the structure at the 5' end of positive strand did not affect the replication. Both uncleaved viral 3CD and host factor PCBP could bind to the 5' cloverleaf structure at loops d and b, respectively, to form a functional ribonucleoprotein (RNP) complex *in vitro*. The formation of this 5' RNP was essential for RNA replication *in vivo*. Mutations that disrupted RNP complex formation, either within the cloverleaf RNA or within 3CD, eliminated RNA synthesis (Andino *et al.*, 1990a, b). It was proposed that the initiation complex consisting of the 5' cloverleaf positive RNA, 3CD and PCBP catalyzed *in trans* the initiation of the nascent positive strand RNA synthesis.

Most recently, Barton *et al.* (2001) provided additional evidence to show that the 5' cloverleaf in poliovirus RNA was required for negative strand synthesis. Mutations in the 5' cloverleaf usually decreased the stability of poliovirus RNA and dramatically inhibited negative strand synthesis *in vitro*. Adding a 5' 7-methyl guanosine cap fully restored the stability of the mutant RNAs, indicating that the 5' cloverleaf provided protection to the uncapped poliovirus RNA. However, the 5' cap did not restore the ability of the mutant RNAs to serve as templates for negative strand synthesis. These data indicated that the 5' cloverleaf structure was a *cis*-acting replication element that was required for both stability and negative-strand RNA synthesis. Based on these data and previous results, a model was proposed in which the 5' and 3' ends of poliovirus RNA interacted with viral and host proteins to form a circular RNP complex that coordinately regulated poliovirus RNA stability, translation, and initiation of negative strand synthesis. Interactions between the proteins bound to the two termini formed a protein bridge and helped to form

a circular RNP complex. For example, 3CD bound to both termini formed a homodimer, PCBP bound to the 5' cloverleaf interacted with PABP bound to the 3' poly(A) tail (Wang *et al.*, 1999; Wang and Kiledjian, 2000).

A similar requirement of the 5' SLS of BMV RNAs was shown for viral positive strand synthesis. Both the 5'-positive and 3'-negative termini of RNAs from BMV and related viruses can form a SLS with the ICR2-like motif (internal control regions of tRNA gene promoters) located at the loop region (Pogue *et al.*, 1990, 1992; Pogue and Hall, 1992). Mutation analysis showed that the ICR2-like motif as well as the SLS at the 5'-positive strand but not at the 3'-negative strand were essential for replication (Pogue *et al.*, 1990, 1992; Pogue and Hall, 1992). It was suggested that the ICR2-like motifs in BMV RNAs might act analogously to ICR2 sequences in tRNA genes in binding a protein factor that is required for formation of the replicase complex. These data resemble the requirement in poliovirus of a 90-nt cloverleaf structure present at the 5' terminus of positive-strand RNA for viral replication. Although replicase complex can separate internal double-stranded regions by a helicase function, a single-stranded terminus is mandatory for strand initiation (Ahlquist *et al.*, 1984). The unexpected requirement of a 5' positive-sense SLS for synthesis of positive-strand RNA leaded to the propose of a model in which binding of the host factor to the ICR2-like motif elicit folding of the 5'-terminal stem-loop as a hinge, thereby releasing a single-stranded 3' negative-strand terminus for replication recognition and initiation (Pogue and Hall, 1992).

The 5' terminus of BMV genomic RNA (or the 3' terminus of negative strand RNA) also plays important role in the synthesis of positive strand RNA. In an *in vitro*

replication system containing the purified replicase complex, initiation of RNA synthesis from the 3' end of negative-strand RNAs required one non-template nucleotide 3' of the initiation nucleotide(Sivakumaran and Kao, 1999; Sivakumaran *et al.*, 1999). Mutational analysis and template competition assay of the minus-strand template revealed that the -1 non-template nucleotide along with the +1 cytidylate and +2 adenylate was important for plus-strand synthesis *via* stable interaction with BMV RdRp. Although RdRp of BMV was able to bind to CMV minus strand whose 3' sequence (3'GCA) was identical to the BMV 3' sequence, it was not able to direct efficient CMV RNA synthesis, suggesting that in addition to the 3' nucleotides, upstream sequences were required for efficient RNA synthesis. By deletion analysis, the 3' terminal 27 nt was identified as the minimal promoter and template for positive strand synthesis *in vitro*. This 27 nt RNA can form a stem-loop secondary structure determined by nuclear magnetic resonance, however, this secondary structure can be significantly altered and still retained the ability to direct RNA synthesis. In contrast, position-specific changes in the RNA sequence resulted in reduced RNA synthesis by affecting replicase recognition (positions -1, +1, and +2) and modulating the polymerization process (U to G substitution at +3 to +5; nucleotides at +17 to +24). These data indicated that the 3' primary sequences rather than the secondary structure was important for BMV negative strand synthesis.

1.2.3.3 Internal sequences

The *cis*-acting signals required for RNA replication have generally been located to the ends of virus genomic RNAs. However, internal *cis*-acting elements, either in

intercistronic or in coding regions have been reported to be required for efficient RNA replication in some RNA viruses, including bacteriophage Qβ, poliovirus, BMV, PVX (Duggal *et al.*, 1994; Buck, 1996). The function of internal *cis*-acting elements is not well known, but roles in RNA folding, host factor and RdRp binding, and replicase assembly have been proposed (Buck, 1996; Klovins *et al*, 1998; Klovins and Duin, 1999; Kim and Hemenway, 1999).

At the initiation of viral RNA replication, bacteriophage Qβ replicase complex binds genomic RNA *via* subunit S1 at two internal regions, the S-site and the M-site (Meyer *et al.*, 1981). The S-site, which overlaps the start of the coat protein gene, was not essential for the transcription reaction but was required for replicase to compete with ribosomes for the template by blocking the translation initiation (Weber *et al*, 1972). The M-site, around nucleotides 2750, was essential for replication and its deletion strongly reduced replication (Schuppli *et al.*, 1998). Two long-range interaction *via* base-pairing (2981-2988 and 4049-4056, 2972-2979 and the loop of 3'-terminal hairpin U1) in Qβ RNA that bridged the thousand nucleotides between the M-site and the 3' end were identified (Klovins *et al*, 1998; Klovins and Duin, 1999). While mismatch mutations abolished RNA replication, compensatory mutation restored the replication.

Similar long-distance RNA-RNA interactions between the conserved octanucleotide sequences located upstream of two subgenomic RNAs and the 5' end of the genome RNA were found in PVX (Kim and Hemenway, 1999). Mutations in the 5' end resulted in lowering genomic RNA and sgRNA levels. However, compensatory changes in the octanucleotide elements restored levels of gRNA and sgRNA, suggesting that the

identified long-distance interactions were important for the accumulation of plus-strand RNA (Kim and Hemenway, 1999).

A sequence of 150 nt at the intercistronic region of BMV RNA-3, which separated the ORFs of movement protein and coat protein, was required *in cis* for efficient RNA replication (French and Ahlquist, 1987). This region contained ICR2-like motif. Removal of the ICR2-like motif dramatically decreased RNA replication, indicating that this motif contributed to the function of the intercistronic region in RNA-3 replication. In yeast, synthesis of BMV negative strand RNA-3 required this intercistronic sequence, as well as the 3'-terminal sequence (Quadt *et al.*, 1995). *In vitro* synthesis of BMV negative-strand RNA-3 by an isolated RdRp preparation did not require the intercistronic sequence, suggesting that this sequence was required for the assembly of the replication complex. This intergenic region also mediated the 1a protein-induced stability of RNA-3 in yeast (Janda and Ahlquist, 1998; Sullivan and Ahlquist, 1999). The replication of RNA-3 in yeast required both replication proteins 1a and 2a. 1a protein expressed in yeast dramatically increased the accumulation of RNA-3 *via* stabilization of RNA-3. However, the increased RNA-3 accumulation did not increase RNA-3 translation. This intergenic region was sufficient to confer 1a-induced stability on a heterologous β-globin RNA. Deletions within this intergenic region inhibited 1a-induced stability as well as decreased RNA-3 replication. Deletion of ICR2-like motif in this region dramatically reduced the ability of RNAs to respond to 1a, suggesting that IRC2-like motif play a major role in 1a-induced stabilization. These results suggested that 1a-induced RNA-3 stabilization *via* direct or indirect 1a-IRC2 interaction may reflect processes involved in recruiting RNA-3

templates into replication as well as removing them from the interfering pathways of translation and degradation.

Internal sequences important for viral RNA replication were also studied in poliovirus. By searching the poliovirus genome for local secondary structural elements, Goodfellow *et al.* (2000) identified a 61-nt stem-loop in the 2C protein-coding region. While mutants disrupting the structure in the positive strand without affecting the 2C protein eliminated the synthesis of negative strands, the revertant viruses were recurred, suggesting that the integrity of the structure in the positive sense was critical for minus strand synthesis. Furthermore, the addition of the second copy of this 61-nt element at another location to the mutant replicons bearing lethal mutations in the native structure restored its replication. This *cis*-acting element was also required for the synthesis of VPg-pUpU *in vitro* (Paul *et al.*, 2000; Rieder *et al.*, 2000).

1.2.3.4 Subgnomic RNA promoters

In addition to genomic RNA(s), many RNA viruses transcribe subgenomic RNA (sgRNA) to express the gene(s) located at the 3' end. sgRNA is syntheized by replicase complex from the negative strand genomic RNA. The subgenomic promoters have been well studied in a number of plant RNA viruses.

The promoter for BMV coat protein subgenomic RNA is located in a region extending between 74 and 95 nt upstream and 16 nt downstream of the subgenomic initiation site (Marsh *et al.*, 1988; French and Ahlquist, 1988). It consists of three functional domains: the core promoter and two upstream enhancers. The core promoter

includes the transcription initiation site, 20 nt upstream sequence, and 15 nt downstream sequences. Oligoribonucleotides that contained the core promoter and the short regions of the (-)-strand RNA-3, were sufficient to direct accurate *in vitro* synthesis of (+)-strand sgRNA by the purified BMV replicase (Adkins *et al*, 1997). Further nucleotide substitution revealed that nucleotides at -17, -14, -13, and -11 relative to the subgenomic initiation site were critical and must be maintained for RNA synthesis from the subgenomic core promoter (Siegel *et al.*, 1997). Template competition assay demonstrated that at least one of these nucleotides, -17, was contacted by BMV replicase. The different phenotypes resulting from nucleotide substitution at these four positions allowed predictions of the base-specific functional groups required for replicase recognition. Substitution of G at -17 with A, U and C reduced the promoter activity to 1.1%, 13% and 1.6%, of wild type level respectively, suggesting that the C6 keto and the N1 amine were the functional groups. Replacing -14 A with G, U or C resulted in 1.2%, 2.8% or 2.1% of RNA synthesis, indicating that multiple groups were involved in replicase recognition. Changes of -13 C to G, A or U decreased the promoter activity to 1.7%, 1.8% or 1.4% of wild type level, suggesting that C4 amine was the functional group. Substitution of -11 G to A, U or C caused reduction of RNA synthesis to 8.8%, 7.1% or 1.1%, revealing both C6 keto and N7 imine as the functional groups.

The roles of these predicted functional groups in interacting with replicase and RNA synthesis were examined by using chemically synthesized RNAs containing base analogs at each of the four positions (Siegel *et al.*, 1998). Base analog 2-aminopurine riboside removed the C6 keto from G, 7-deaza-2'-deoxyguanosine removed the N7 imine from

deoxyG, 2'-O-TBDSi-3'-O-phosphoraidites of purine riboside removed the C6 amine group from A, pyrimidine-2-one riboside removed the C4 amine group from C. In theory, replacing one nucleotide at four essential positions with the base analog that removed the predicted functional group will reduce RNA binding by RdRp and RNA synthesis to background level if the removed group was the only functional group or to certain percentage of wild type if more than one functional group were involved. Indeed, all predicted functional groups in four essential positions were shown to be required for RdRp recognition and RNA synthesis. These results suggested the sequence-specific recognition of RNA by the BMV RdRp, which was analogous to the recognition of DNA promoter by DdRp. Moreover, nucleotides at these four key positions were highly conserved in subgenomic promoters of other members of the alpha-like superfamily, including closely related *Cowpea chlorotic mottle virus* (CCMV), in which all except the -11 nt were conserved (Allison *et al.*, 1988). Both BMV and CCMV RdRps were found to recognize the BMV core subgenomic promoter in the same manner, requiring specific functional groups at positions -17, -14, -13, and -11 relative to the subgenomic initiation site (Adkins and Kao, 1998). In addition to these four nucleotides CCMV subgenomic RNA synthesis by both RdRps required four additional nucleotides at -20, -16, -15, and -10. Of these eight key nucleotides, only nucleotides at -20, -11, and -10 differed between the CCMV and BMV core subgenomic promoters. By using a set of CCMV/BMV hybrids, it was shown that nucleotides -20 and -11 functioned together to promote subgenomic RNA synthesis. It was suggested that the -20 nt in CCMV triggers recognition of the -11 nt and contributed to the difference in RdRp recognition of the

CCMV and BMV core subgenomic promoters and that RNA can induced RdRps to alter the mode of promoter recognition.

In RCNMV, one subgenomic RNA coding for coat protein was synthesized internally from the minus-strand RNA-1 (Zavriev *et al.*, 1996). Thirteen of the first fourteen nucleotides of the coat protein subgenomic RNA were identical to those at the 5' terminus of genomic RNA-1. A SLS was predicted in the region around the transcriptional initiation site of the sgRNA (-53 to +27, or 2312-2391). Mutation analysis suggested that both the core 14-nucleotide element and the SLS were required for sgRNA synthesis. In addition, a 34-nucleotide sequence (756-789) in RNA-2 was identified as a transactivator for sgRNA synthesis from RNA-1 (Tim *et al.*, 1998). This 34-nt element was predicted to form a SLS in which the 8-nt loop sequence was complementary to an 8-nt sequence in RNA-1 located 2 nt upstream of the initiation site of sgRNA (-3 to -10). RNA-1 mutations that disrupted this 8-nt base pairing abolished the sgRNA synthesis and compensatory mutations restored the sgRNA synthesis. It was proposed that direct binding of the transactivator element in RNA-2 to the 8-nt element in RNA-1, possibly mediated by interactions with protein factors, caused the premature termination of the minus-strand RNA-1 synthesis. This prematurely terminated minus-strand RNA containing the highly conserved 5' sequences was then used as template for sgRNA synthesis.

1.3 VIRAL RNA RECOMBINATION

RNA recombination is a common feature and has two functions in RNA viruses (Simon and Bujarski, 1994). It is used to repair, *via* recombining the mutant genome with the wild type genome, the high-rate errors that is caused by the nucleotide misincorporation of RdRp during viral replication. It is also thought to be a major driving force in the evolution of RNA viruses.

RNA recombination can occur between the segmented genomic RNAs which is thought to lead to the generation of monopartite virus from multipartite virus or within the same RNA molecule resulting in the generation of defective interfering (DI) RNAs (Roux *et al.*, 1991). It can also occur between viruses. The crossover between two different viruses could result in a new virus with some genes derived from one parental virus and other genes from the second parental virus (gene modulation). Recombination between the infecting viral RNA and the host cellular RNA could lead to the acquisition of cellular genes by RNA viruses, such as the acquisition of an ubiquitin coding sequence in bovine diarrhea virus genome (Meyers *et al.*, 1991), a portion of 28S rRNA sequences in influenza virus (Khatchkian *et al.*, 1989), a tRNA sequence in Sindbis virus RNA, $tRNA_{Asp}$ of *Escherichia coli* in Qβ genome (Munishkin *et al.*, 1988), and a sequence homologous to an exon of tobacco chloroplast open reading frame 196 in *Potato leafroll virus* (Mayo and Jolly, 1991). Recently the recombination between the infecting viral RNA and the transgene transcripts from the virus-resistant transgenic plants were reported from several groups using different systems, including *Cowpea chlorotic mottle virus* (CCMV, Greene and Allison, 1994), *Tomato bushy stunt virus* (TBSV, Borja, *et al.*,

1999), TMV (Adair and Kearney, 2000), and *Plum pox virus* (PPV) (Varrelmann *et al.*, 2000).

1.3.1 Homologous recombination versus non-homologous recombination

RNA recombination is generally classified into two categories: homologous and non-homologous recombination (Lai, 1992). The crossover in homologous recombination occurs between two closely related RNA molecules at sites precisely matched between two parental RNAs (precise homologous recombination) or imprecisely matched (aberrant homologous recombination). The precise homologous recombination plays important roles in correcting the RNA replication errors. The imprecise crossover in aberrant homologous recombination causes either the sequence duplication or sequence deletion in the recombinant RNA.

In contrast, the crossover in non-homologous recombination occurs between two unrelated RNA molecules. Sequence phylogenetic comparisons and genomic organization of RNA viruses suggest that the nonhomologous recombination has been a major driving force in RNA virus evolution (Dolja and Carrington, 1992; Zaccomer *et al.*, 1995).

1.3.2 Transgenic recombination

Several cases of recombination between viral RNA and viral transgene mRNA expressed in transgenic plants have been reported. These studies are summarized below.

The first reported case of transgenic recombination is CCMV, a member of *Bromovirus* genus. CCMV contains three genomic RNAs, the monocistromic RNA-1 and RNA-2 encode replication proteins and the dicistronic RNA-3 encodes the movement protein and the capsid protein (CP). CP protein is dispensable for viral replication in the inoculated cell, but is required for systemic infection. The transgenic plants expressing the 3' two-thirds of the CCMV CP gene and the 3'-untranslated region (3' UTR) of RNA-3 were constructed and inoculated with a CCMV deletion mutant lacking the 3' one-third of the CP gene (Greene and Allison, 1994). The systemic infections occur only if RNA recombination occurs and restores a functional CP gene. Four of 125 (3%) inoculated transgenic plants became systemically infected. Sequence analysis of the recombinant viral RNA revealed that RNA recombination occurs *via* different events of single aberrant homologous recombination between the transgenic mRNA and the infecting viral RNA at the central one-third common region. This initial report of transgenic recombination was followed by several other cases as discussed below.

TBSV, a tombusvirus, has a single genomic RNA that encodes p33 and p92 replication proteins, p41 CP, and p19 and p22 movement protein. While the wild type virus can cause lethal systemic infection, the CP mutants can establish systemic infection with mild symptoms. The transformed *Nicotiana benthamiana* plants expressing the CP gene of TBSV without 3'-UTR were inoculated with mutant TBSV having the 46 nucleotide within the 5' of CP gene deleted (Borja, *et al.*, 1999). The lethal systemic symptom was used as marker to identify the RNA recombination that reconstitutes the functional CP gene. A total of 32 of 200 inoculated transgenic plants (16%) representing

different lines developed the lethal symptom. The recombinant virus was recovered and the sequence analysis of the vial RNA confirmed that the wild type CP gene was regenerated *via* the double homologous recombination events. The effect of the length of homologous region on the recovery of recombination was characterized. Whereas the 4% (46 nt) CP gene deletion resulted in the average 16% recombination, the 56% CP deletion at the corresponding region decrease the recombination frequency to 4%, indicating a positive relationship between the recombination frequency and the length of the common sequence shared by the transgene mRNA and the challenging viral RNA. The effect of the extent of the homologous region was also tested by inoculation of the TBSV CP transgenic plants with the CP defective *Cucumber necrosis virus* (CNV). CNV is a relative of TBSV; its CP can be functionally replaced by the TBSV CP. No nonhomologous recombination between the defective CNV CP and the TBSV CP transgene was recovered in 160 plants tested, confirming the positive relation between the recombination frequency and the homologous extent around the crossover site.

TMV is the type member of *Tobamovirus* genus and contains a single-stranded RNA genome encoding foru proteins: p126 and p183 replication proteins, MP and CP. CP is required for efficient systemic infection. A TMV vector (30B.GFP) carrying a green fluorescent protein (GFP) gene under the control of TMV subgenomic promoter can infect *N. benthamiana* plant in a similar way as wild type TMV. *N. benthamiana* plants was transformed with a 3' portion of 30B.GFP. The transgene included a part of the replicase gene, MP gene, GFP gene, and CP gene, but lacked the 3'-UTR (Adair and Kearney, 2000). When the transformed plants were inoculated with 30B.GFP.ΔCP, a CP

defective mutant of 30B.GFP with a 168 nt deletion in the central region of CP gene, 28 out of 44 infected plants showed efficient systemic infection, indicating regeneration of functional CP gene *via* recombination in these plants. However, recombinant RNA was detected in only 14 (32%) infected plants by RT-PCR using a downstream primer specific for the 3'-UTR of TMV and an upstream primer specific for the deleted CP region. Among these 14 plants, no virions were detected, indicating that the recombinant RNA was generated *via* aberrant homologous recombination. In contrast, when transgenic plants were challenged with a mutant of 30B.GFP (30B.KO.GFP) which had the wild type CP gene but GFP gene knocked out by substitution of the AUG start codon, fluorescent patches which would occur if recombination occurred and restored the start codon of GFP gene were not detected in 120 infected plants. Although viral RNA was detected in all tested plants, no recombinant RNA that restores the GFP AUG start codon was detected by RT-PCR which is capable of detecting 1 recombinant in a background of 10,000 parental. These results suggested that the recombination under no selective pressure was much lower than that under strong selective pressure.

In PPV, a potyvirus, viral assembly by CP is required for cell-to-cell movement and systemic infection in *N. benthamiana* plants. While a CP assembly mutant PPV was restricted in the originally infected cells, CP chimeric PPV with CP core region exchanged with that of either *Zucchini yellow mosaic virus* (ZYMV) or *Potato virus Y* (PVY) was able to form virions and systemically infect *N. benthamiana* plants. In both cases, under either the strong selection pressure for cell-to-cell movement or the moderate selection pressure due to the selection advantage of wild type sequence, wild

type CP sequence was recovered via recombination between the mutant virus and the transgene CP containing the complete 3' UTR (Varrelmann *et al.*, 2000). In contrast, no recombinant molecules were recovered from transgenic lines expressing CP gene with 3' UTR truncated. These results indicated that transgenic recombination could occur under either strong or moderated selection pressures once the complete 3' UTR containing the replication element was present in the transgene.

1.3.3 Mechanism of RNA recombination

Genetic RNA recombination is a process that joins together two non-contiguous genomic RNA molecules (for reviews see, Lai, 1992). Since its first discovery in early 60's (Ledinko, 1963), RNA recombination has been found and studied in different RNA viruses. Two models of RNA recombination were most commonly accepted: the replicative template-switching (copy-choice) mechanism and the nonreplicative RNA breakage-rejoining mechanism (for reviews see Alejska *et al.*, 2001; Chetverin, 1999; Figlerowicz and Bujarski, 1998; Lai, 1992; Nagy and Simon, 1997). The template-switching mechanism proposes that RNA recombination occurs during RNA synthesis when the replicase complex pauses and leaves the first template (donor strand), and resumes synthesis on a second template (acceptor strand). In contrast, the breakage-rejoining mechanism suggests that the recombinant RNAs are formed by RNA breakage and joining the resulting fragments from different molecules. Currently, there are evidences supporting both models.

1.3.3.1 Replicative template-switching

The template-switching mechanism was first proposed by Kirkegaard and Baltimore in 1986 to explain the homologous recombination in poliovirus (Kirkegaard and Baltimore, 1986). Since then, this mechanism has been commonly accepted and used to explain both homologous and nonhomologous recombinations in different systems including well studied viruses, such as BMV, TCV, and TBSV. Both RNA components (donor strand, nascent strand, and acceptor strand) and RdRp complex are involved in template-switching recombination. The effects of RNA and RdRp on the recombination frequency and distribution of crossover sites were well studied in BMV and TCV. Based on these studies, two template-switching models, the heteroduplex-mediated template switching model and homology-mediated template switching model were proposed (Nagy and Simon, 1997; Figlerowicz and Bujarski, 1998).

Heteroduplex-mediated template switching model

In this model, nonhomologous recombination is mediated by heteroduplex (base-pairing) formed between the donor and acceptor strands down stream the crossover site. It proposes that RdRp complex pauses at donor strand due to heteroduplex and then occasionally switches template into and resumes transcription on receptor strand brought into proximity by the heteroduplex. In BMV, inserting 66 nt sequence complementary to RNA-1 into RNA-3 resulted in the production of RNA-1/RNA-3 recombinants (Nagy and Bujarski, 1993). However, insertion of the same region in the sense orientation did not induce any detectable recombinant. Moreover, the crossover sites clustered at left side

of the heteroduplex region. In addition, longer heteroduplex regions supported recombination at a higher frequency than shorter ones among the tested range (20-140 nt). Introduction of mismatch mutations at the left side within the deteroduplex region resulted in the right-ward shift of crossover sites (toward more stable portions of the heteroduplex). A role for similar heteroduplexes formed from intramolecular base-pairing in promoting RdRp pausing and template switching was also indicated in the formation of DI RNAs in *Tombusvirus* (White and Morris, 1995) and *Bromovirus* (Pogany *et al.*, 1995). In addition, stable internal hairpin structures, which may also promote RdRp pausing, are found to constitute recombination hotspots in *Tombusvirus* (White and Morris, 1995). Recently, by using an *in vitro* RNA recombination system, Kim and Kao reported that RdRp complexes from BMV and other three viruses can produce easily detectable recombination products and that NTP availability affected the frequency of template switch (Kim and Kao, 2001).

Homology-mediated template switching model

In contrast to heteroduplex mediated recombination, homology mediated recombination is supported by the base-pairing between the nascent RNA and acceptor strand at the homologous region between donor and acceptor strands. In BMV, insertion of a 60 nt sequence derived from RNA-2 in direct orientation into RNA-3 resulted in efficient homologous (precise and aberrant) recombination between wild type RNA-2 and mutant RNA-3 at the common region (Nagy and Bujarski, 1995). Deletion analysis revealed that 15 or longer sequence identity between RNA-2 and RNA-3 can support

efficient RNA-2/RNA-3 homologous recombination. Common regions of five to nine nt support reduced levels of homologous recombination. No homologous recombination was detected when the sequence of identity was 4 or less. Mismatching mutation analysis indicated that the presence of four or more mismatch nucleotides in the common region dramatically reduced the frequency of crossovers and also shifted the location of crossovers towards non-mutagenized portion.

In addition to sequence identity, the A/U and G/C contents also play very important roles in determining the frequency and precision of the homologous recombination. Sequence analysis of the recombinant RNAs revealed that the crossover sites were clustered in two recombination hot-spots of short A/U-rich sequences (Nagy and Bujarski, 1995). Increasing the number of the A/U-rich nucleotides in both donor and acceptor strands increased recombination frequency significantly (Nagy and Bujarski, 1996). In addition to the A/U-rich region, the upstream 23 nt is G/C-rich region. Insertion of artificial A/U-rich sequences at the 3' of this G/C-rich region can also increase the recombination frequency and shift the crossover site to the A/U-rich region (Nagy and Bujarski, 1997). In contrast, insertion of the same A/U-rich sequences at the 5' of this G/C-rich region neither increased the recombination frequency nor shifted the crossover sites. Deletion of this G/C-rich region in either or both of donor and acceptor RNA supports no or very low level of recombination in the presence of A/U-rich sequence, suggesting that this G/C-rich region is required and AU-rich sequences alone in both parental RNAs are not sufficient for recombination to occur. However, in the presence of AU-rich sequences, the G/C-rich region can be replaced in both parental RNAs with

other G/C-rich or -average sequences and support recombination at a level higher than that of AU-rich sequences alone, indicating that the G/C content but not the particular sequence in this region was important for recombination (Nagy and Bujarski, 1998). Based on these data, a model was proposed: the 5' GC-rich sequences (as the "core" region) and the downstream AU-rich sequences (as the "enhancer" region) play different roles in homologous recombination. During (+)-strand RNA synthesis, the AU-rich sequences in the donor RNA may cause replicase pausing and/or facilitate the release of the 3' end of the nascent plus strand due to the weak A-U base-pairing. The GC-rich sequence in the upstream portion of the nascent plus strand may facilitate its landing on the complementary region (landing site) of the acceptor RNA *via* strong G-C base-pairing. The AU-rich sequences in the acceptor RNA may further prompt the opening of the double-stranded replication intermediate *via* local "bubble structure" formation (Nagy and Bujarski, 1997; Figlerowicz and Bujarski, 1998).

Consistent with the proposed model, studies of the sequences flanking the recombination hot-spots revealed that insertion of 20-30 nt GC-rich sequences at downstream positions either in both receptor and donor strands or in receptor strand only dramatically reduced the recombination frequency (Nagy and Bujarski, 1998). It indicated that the downstream GC-rich sequences acted as homologous recombination silencers on the upstream hot-spots. According to this model, the downstream GC-rich sequences in the acceptor strand interfere with the formation of recombination intermediates via inhibiting the formation of the bubble structure in double stranded replication intermediate of acceptor RNA at the AT-rich hot spots, or affect the re-

initiation of the recombinant RNA synthesis by the landing replicase on the acceptor RNA due to the strong G-C base-pairing which is difficult to unwind during replicase landing or reinitiation.

TCV is another well-studied system. The sequences and structures required for recombination between TCV-associated satellite RNAs (sat-RNA C and sat-RNA D) had been determined. Plant inoculated with TCV genomic RNA, sat-RNA D, and sat-RNA C containing lethal deletion in the 5' region resulted in the generation of recombinant sat-RNAs comprising sat-RNA D at the 5' end and a portion of sat-RNA C at the 3' end via aberrant homologous recombination (Cascone *et al.*, 1990, 1993). Analysis of the crossover sites revealed a recombination hot spot located at a position 13 bases from the 3' end of the plus-strand sat-RNA D and one of five consecutive bases in sat-RNA C beginning at position 175. Further analysis of the sequence and structure around the junctions revealed two elements: a short homologous region (upstream of crossover sites) between two parental RNAs and a stable hairpin (named motif1-hairpin). Motif1-hairpin shares sequence similarity with mapped TCV promoter for positive strand synthesis (Nagy and Simon, 1997). Mutations disrupting the hairpin eliminated recombination, while compensatory mutations restored recombination (Cascone *et al.*, 1993). A model was proposed to explain the recombination observed in TCV. The (+) nascent RNA transcribed from the (-) sat-RNA D template basepaired with (-) sat-RNA C immediately downstream the motif1-hairpin. The promoter-like motif1-hairpin had the ability to recruit RdRp to the site. RdRp then reinitiated RNA synthesis using the 3' of the (+) nascent sat-RNA D as primer and the receptor RNA ((-) sat-RNA C) as template. Similar

features were also found in the recombinant RNAs between sat-RNA D and genomic RNA *in vivo* (Zhang *et al.*, 1991; Carpenter *et. al.*, 1995). An *in vitro* system that mimics the strand switch and primer extension (reinitiating of RNA synthesis on receptor strand) was developed (Nagy *et al.*, 1998b). This system contains purified TCV RdRp preparations and hybrid RNAs that resembled putative recombination intermediates (containing 28 nucleotides from sat-RNA D plus strand sequence extending from position -13 from the 3' end, joined by a six-base artificial sequence (5'-GAAUUC-3') to sat-RNA C minus strand sequence that includes 33 nucleotides 3' of motif-hairpin, the motif1-hairpin, and 18 nucleotides 5' of the hairpin. The 3'-terminal extension by RdRp on these hybrid RNAs was analogous to reinition/primer extension proposed to occur *in vivo* by the TCV RdRp. Mutation and competition studies shown that efficient generation of the 3'-terminal extension products depended on two key factors: 1) the hairpin structure in the receptor RNA *via* binding RdRp, and 2) a short base-paired region formed between the nascent RNA and the acceptor RNA by bringing them together and placing the 3' end of the nascent strand at a proper position to RdRp. Further mutation analysis using this system defined the optimal sequences and structures in 3'-terminal extension reactions *in vitro*: wild-type like structure of motif1-hairpin, longer than 10 bp of priming stem, 12-36 nt pyrimidine-rich single-stranded spacer between motif1-hairpin and priming stem, and U-rich single-stranded 5' flanking region of motif1-hairpin (Nagy and Simon, 1998a, b).

The role of RdRp complex in template switching

The roles of replicase in RNA recombination were clearly demonstrated by mutagenesis of BMV replication proteins 1a and 2a (Nagy *et al.*, 1995; Figlerowicz *et al.*, 1997, 1998). Viable mutations within the helicase-like domain of 1a protein resulted in an increase in the frequencies of recombination as well as a shift of the crossover sites into the energetically less stable portion of the heteroduplex formed between two parental RNAs (Nagy *et al.*, 1995). In contrast to the 1a mutations, a viable 2a mutation within the central RdRp core domain did not support heteroduplex-mediated non-homologous recombination (Figlerowicz *et al.*, 1997). While the overall frequency of homologous recombination remained unchanged, the frequency of imprecise (aberrant) recombination was increased (doubled) and the crossover sites were shifted towards 5' regions as compared to wild type. Mutations were also introduced into the N terminus of 2a that was known involved in 1a-2a interaction (Figlerowicz *et al.*, 1997). Two of these mutants retained near wild type 1a-2a interaction and were infectious. However, the non-homologous recombination frequency in both mutants was dramatically reduced. While one mutant also decreased the frequency of homologous recombination with increased occurrence of imprecise homologous crossovers, another one maintained wild type level of homologous recombination but shifted the crossover sites toward the 3' (Figlerowicz *et al.*, 1998). Taking together, these data suggested that some regions in the RdRp complex are required for recombination but not for RNA replication. Different sites on 2a apparently participated in the homologous and non-homologous recombination.

Consequently, mechanisms underneath both types of recombination are probably different.

1.3.3.2 Nonreplicative breakage-rejoining

In contrast to replicative template switching mechanism, nonreplicative breakage-rejoining model proposes that recombinant RNA resulted from parental strand breakage followed by ligation of the resulting fragments from different molecules. While the results from most systems supported template-switching model, studies using an *in vitro* system of bacteriophage Qβ provided the first direct evidence for breakage-rejoining mechanism (Chetverin *et al.*, 1997; Chetverin, 1999). This system contained the recombination substrates, 5'- and 3'- fragments of a broken phage Qβ satellite RQ135-1 RNA of 124 nt long, and pure Qβ replicase under physiological condition similar to those in cells (Chetverin *et al.*, 1997). RQ135-1 RNA can be replicated by Qβ replicase efficiently, but neither the 5'- nor the 3'-fragments alone can be replicated. In addition, homologous sequences were added to the broken ends to facilitate the homology-based replicase switch. The resulting replicable recombinant RNAs were then amplified into RNA colonies on Qβ- and rNTP-containing agarose. The colony number reflects the number of recombinant molecules generated in the system. Despite the homology of the foreign sequences, cloning and sequencing of the recombinants shown that all of them were nonhomologous, and that all contained the complete 5'-fragment joining to 3'-fragment at the common region. The sequence features at the junctions suggested the transesterification mechanism of the attacking of the 3'-hydroxyl of the 5'-fragment at

the phosphoester bonds within the 3'-fragment. Accordingly, removal of the 3'-hydroxyl of the 5'-fragment but not the 3'-fragment eliminated recombination. Restoration of the 3'-hydroxyl restores original recombination level.

In addition to the replicase-assisted recombination describe above, self-recombination in the absence of replicase was also observed. Self-recombination was characterized differently from replicase-assisted recombination in three ways. It recombined the parental RNAs at the internal region, occured at much lower frequency, and was not affected by the elimination of 3'-hydroxyl (Chetverina *et al.*, 1999), indicating that the 3'-hydroxyl did not participate in the self-ligation. Theses data suggested a different mechanism for self-recombination in which 2'-hydroxyls attacked intramolecularlly at the phosphoester bonds resulting in the formation of intermediates of the 2',3'-cyclic phosphate and the 5'-hydroxyl termini, these intermediates were then cross-ligated to produce the final recombinant RNAs.

Gmyl *et al.* (1999) provided evidences to show the existence of nonreplicative RNA recombination in poliovirus *in vivo*. Viable viruses were recovered from cells cotransfected with different pairs of genomic RNA fragments (5' partners and 3' partners) that are nontranslatable and nonreplicatable. The 5' partners contained truncated 5'UTR that had all of its essential replicative and translational *cis* elements and a part of the spacer separating the IRES and the polyprotein initiator AUG but lack the viral polyprotein coding region. The 3' partners preserved the polyprotein coding region, 3' replicative elements and poly(A) tail, however, contain lethal mutations in the essential *cis* elements of the 5'UTR. Whereas none of the 5' or 3' partners alone can cause any

infection, cotransfection of different pairs of the partners resulted in the recover of viable recombinant genome. Analysis of the crossover sites revealed some "hot spots" with ribozyme like structure. Surprisingly, periodate oxidation of the 3' end of the 5' partners increased rather than decreased the nonhomologous recombination frequency, and caused the entire 5' partner incorporated into a significant proportion of the recombinant genomes. Treatment with aniline to remove the modified nucleoside and generate a 3'-phosphate terminus did not markedly affect the recombination properties of the oxidized 5' partners. However, subseqent dephosphorylation by phosphatase to remove the 3'-phosphate resulted in restored the recombinational frequency and features similar to original nonoxidized fragment. Moreover, direct introduction of the 3'-phosphate to the 5' partner by pCp ligation had similar results as those of oxidized 5' partners. Base on these results, a mechanism was proposed in which the intermediates were formed by 2'-hydroxyls attacking intramolecularly at the phosphosester bonds and the subsequent cross-ligation between the 2',3'-cyclic phosphate and 5'-hydroxyle termini produced the recombinant RNAs.

1.4 OBJECTIVES

The goals of my Ph'D program are: 1) to generate highly infectious cDNA clones of RCNMV RNA-1 and RNA-2, 2) to determine the role of the N-terminus of p27 and p88 replication proteins in RCNMV RNA replication, 3) to characterize the role of the 3'-terminal stem-loop structure in RCNMV RNA replication and how the 3'-SLS is recognized by replicase complex, and 4) to study the recombination between infecting

RCNMV RNA-1 and transgenically expressed RNA-2 and determine the factors affecting the transgenic recombination in RCNMV.

As the results of this program, we have generated highly infectious cDNA clones of RCNMV RNA-1 and RNA-2 by using RT-PCR. One of the newly generated clone, pRC1IG69, is as infectious as wt RNA and about 100 times more infectious than pRC1IG previously generated. The comparison between these two clones and the predicted transmembrane helices and glycosylation site suggested that the N-terminal 14 amino acids of p27 and/or p88 replication proteins were required for viral RNA replication possibly via membrane targeting. The 3'-SLS was recognized by replicase complex. The secondary structure and the loop sequence of the 3'-SLS was required for viral RNA replication. Three key nucleotides recognized by replicase complex were identified and the functional groups were predicted. This 3'-SLS recognized by replicase complex was important for transgenic recombination in RCNMV. Recombination between infecting RNA-1 and transgenically expressed 5'-terminally truncated RNA-2 was demonstrated under experimental conditions. A replicase-mediated template-switching model was proposed. Factors including the 3'-SLS, the transgene mRNA level that could dramatically affect transgenic recombination were identified, which may help to generate safer, virus-resistant transgenic plants.

REFERENCES

Adair, T.L., and Kearney, C. M. 2000. Recombination between a 3-kilobase tobacco mosaic virus transgene and a homologous viral construct in the restoration of viral and nonviral genes. Arch Virol145:1867-83.

Adkins, S. and Kao, C. C. 1998. Subgenomic RNA promoters dictate the mode of recognition by bromoviral RNA-dependent RNA polymerases. Virology. 252:1-8.

Adkins, S., Siegel, R. W., Sun, J. H., and Kao, C. C. 1997. Minimal templates directing accurat initiation of subgenomic RNA synthesis *in vitro* by the brome mosaic virus RNA-dependent RNA polymerase. RNA 3:634-647.

Adkins, S., Stawicki, S. S., Faurote, G., Siegel, R. W., and Kao, C. C. 1998. Mechanisitic analysis of RNA synthesis by RNA-dependent RNA polymerase from two promoters reveals similarities to DNA-dependent RNA polymerase. RNA 4:455-470.

Ahola, T. and Ahlquist, P. 1999. Putative RNA capping activities encoded by brome mosaic virus: methylation and covalent binding of guanylate by replicase protein 1a. J. Virol. 73:10061-10069.

Ahola, T., den Boon, J. AL., and Ahlquist, P. 2000. Helicase and capping enzyme active site mutations in brome mosaic virus protein 1a cause defects in template recruitment, negative-strand RNA synthesis, and viral RNA capping. J. Virol. 74:8803-8811.

Alejska, M., Kurzyniska-Kokorniak, A., Broda, M., Kierzek, R., and Figlerowicz, M. 2001. How RNA viruses exchange their genetic material. Acta Biochim Pol. 48:391-407.

Andino, R., Rieckhof, G. E., Achacoso, P. L., and Baltimore, D. 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. EMBO 12:3587-3598.

Andino, R., Rieckhof, G. E., and Baltimore, D. 1990a. A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. Cell 63:369-380.

Andino, R., Rieckhof, G. E., Trono, D., and Baltimore, D. 1990b. Substitutions in the protease (3C) gene of poliovirus can suppress a mutation in the 5' noncoding region. J. Virol. 64:607-612.

Bates, H. J., Farjah, M., Osman, T. A., and Buck, K. W. 1995. Isolation and characterization of an RNA-dependent RNA polymerase from Nicotiana clevelandii plants infected with red clover necrotic mosaic virus. J. Gen. Virol. 76:1483-1491.

Beckman, M.T., and Kirkegaard, K. 1998. Site size of cooperative single-stranded RNA binding by poliovirus RNA-dependent RNA polymerase. J Biol Chem 273:6724-30.

Behrens, S. E., Tomei, L., De Francesco, R. 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. EMBO J. 15:12-22.

Borja, M., Rubio, T., Scholthof, H. B., and Jackson, A. O. 1999. Restoration of wild-type virus by double recombination of tombusvirus mutants with a host transgene. Mol. Plant Microbe. Interact. 12:153-162.

Brault, V., and Miller, W. A. 1992. Translational frameshifting mediated by a viral sequence in plant cells. Proc Natl Acad Sci U S A. 89:2262-2266.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R., and Rey, F. A. 1999. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Proc Natl Acad Sci U S A. 96:13034-13039.

Buck, K. W. 1996. Comparison of the replication of positive-stranded RNA virus of plants and animals. Adv. Virus Res. 47:159-251.

Butcher, S., Grimes, J., Makeyev, E., Bamford, D., and Stuart, D. 2001. A mechanism for initiating RNA-dependent RNA polymerization. Nature 410, 235-240.

Carpenter, C. D. and Simon, A. E. 1994. Recombination between plus and minus strands of turnip crinkle virus. Virology 201:419-423.

Carpenter, C. D. and Simon, A. E. 1998. Analysis of sequences and predicted structures required for viral satellite RNA accumulation by *in vivo* genetic selection. Nucleic Acids Res. 26:2426-2432.

Carpenter, C. D., Oh, J. W., Zhang, C., and Simon, A. E. 1995. Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. J. Mol. Biol. 245:608-622.

Carrington, J. C., Heaton, L. A., Zuidema, D., Hillman, B. I., and Morris, T. J. 1989. The genome structure of turnip crinkle virus. Virology. 170:219-226.

Cascone, P. J., Carpenter, C. D., Li, X. H., and Simon, A. E. 1990. Recombination between satellite RNAs of turnip crinkle virus. EMBO 9:1709-1715.

Cascone, P. J., Haydar, T. F., and Simon, A. E. 1993. Sequences and structures required for recombination between virus-associated RNAs. Science 260:801-805.

Chapman, M. R. and Kao, C. C. 1999. A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. J Mol Biol. 286:709-20.

Chapman, M. R., Rao, A. L., and Kao, C. C. 1998. Sequences 5' of the conserved tRNA-like promoter modulate the initiation of minus-strand synthesis by the brome mosaic virus RNA-dependent RNA polymerase. Virology 252:458-467.

Chen, J. and Ahlquist, P. 2000. Brome mosaic virus polymerase-like protein 2a is directed to the endoplasmic reticulum by helicase-like viral protein 1a. J. Virol. 74:4310-4318.

Chen, M. H., Roossinck, M. J., and Kao, C. C. 2000. Efficient and specific initiation of subgenomic RNA synthesis by cucumber mosaic virus replicase *in vitro* requires an upstream RNA stem-loop. J. Virol. 74:11201-11209.

Chetverin, A. B. 1997. Recombination in bacteriophage Qβ and its satellite RNAs: the *in vivo* and *in vitro* studies. Seminars in Virology 8:121-129.

Chetverin, A. B. 1999. The puzzle of RNA recombination. FEBS Letters 460:1-5.

Chetverin, A. B., Chetverina, H. V., Demidenko, A. A., and Ugarov, V. I. 1997. Nonhomologous RNA recombination in a cell-free system: evidence for a transesterification mechanism guided by secondary structure. Cell 88:503-513.

Chetverina, H. V., Demidenko, A. A., and Ugarov, V. I., and Chetverin, A. B. 1999. Spontaneous rearrangements in RNA sequences. FEBS Letters 450:89-94.

Cho, M. W., Richards, O. C., Dmitrieva, T. M., Agol, V., and Ehrenfeld, E. 1993. RNA duplex unwinding activity of poliovirus RNA-dependent RNA polymerase 3Dpol. J Virol. 67:3010-3018.

Dies, J., Ishikawa, M., Kaido, M., and Ahlquist, P. 2000. Identification and characterization of a host protein required for efficient template selection in viral RNA replication. Proc. Natl. Acad. Sci. USA. 97:3913-3918.

Ding, B., Haudenshield, J. S., Hull, R. J., wolf, S., Beachy, R. N., and Lucas, W. J. 1992. Secondary plasmodesmata are specific sites of localization of the tobacco mosaic virus movement protein in transgenic tobacco plants. Plant Cell 4:915-928.

Dolja, V. V. and Carrington, J. C. 1992. Evolution of positive-strand RNA viruses. Semin. Virol. 3:315-326.

Dolja, V. V. and Koonin, E. V. 1991. Phylogeny of capsid proteins of small icosahedral RNA plant viruses. J. Gen. Virol. 72:1481-1486.

Dreher, T. W., and Hall, T. C. 1988. Mutational analysis of the sequence and structural requirements in brome mosaic virus RNA for minus strand promoter activity. J. Mol. Biol. 201:31-40.

Duggal, R., Lahser, F., C., and Hall, T. C. 1994. cis-Acting sequences in th replication of plant viruses with plus-sense RNA genomes. Annu. Rev. Phytopathol. 32:287-309.

Eggen, R., Verver, J., Wellink, J., Pleij, K., Van Kammen, A., and Goldbach, R. 1989. Analysis of sequences involved in cowpea mosaic virus RNA replication using site-specific mutants. Virology 173:456-64.

Felden, B., Florentz, C., Giege, R., and Westhof, E. 1994. Solution structure of the 3'-end of brome mosaic virus genomic RNAs. Conformational mimicry with canonical tRNAs. J Mol Biol. 235:508-531.

Figlerowicz, M. and Bujarski, J. J. 1998. RNA recombination in brome mosaic virus, a model plus strand RNA virus. Acta Biochimica Polonica 45:847-868.

Figlerowicz, M., Nagy, P. D., and Bujarski, J. J. 1997. A mutation in the putative RNA polymerase gene inhibits nonhomologous, but not homologous, genetic recombination in an RNA virus. Proc. Natl. Acad. Sci. USA. 94:2073-2078.

Figlerowicz, M., Nagy, P. D., Tang, N., Kao, C. C., and Bujarski, J. J. 1998. Mutations in the N terminus of the brome mosaic virus polymerase affect genetic RNA-RNA recombination. Journal of Virology 72:9192-9200.

Francki, R. I. B., Fauquet, C. M., Knudson, D. L., and Brown, F. 1991. Classification and Nomenclature of Viruses: Fifth Report of the International Committee on Taxonomy of Viruses. Archives of Virology, Supplementum 2. Springer, New York.

Fujiwara, T., Giesman-Cookmeyer, D., Ding, B., Lommel, S. A., and Lucas, W. J. 1993. Cell-to-cell trafficking of macromolecules through plasmodesmata potentiated by the red clover necrotic mosaic virus movement protein. Plant Cell 5:1783-1794.

Gallie, D. R., Le, H., Caldwell, C., and Browning, K. S. 1998. Analysis of translation elongation factors from wheat during development and following heat shock. Biochem Biophys Res Commun. 245:295-300.

Gallie, D. R., Le, H., Caldwell, C., Tanguay, R. L., Hoang, N. X., and Browning, K. S. 1997. The phosphorylation state of translation initiation factors is regulated developmentally and following heat shock in wheat. J Biol Chem. 272:1046-1053.

Ge Z., Hiruki C., and Roy K. L. 1992. A comparative study of the RNA-2 nucleotide sequences of two sweet clover necrotic mosaic virus strains. J. Gen. Virol. 73:2483-2486.

Ge, Z., Hiruki, C., and Roy, K. L. 1993. Nucleotide sequence of sweet clover necrotic mosaic dianthovirus RNA-1. Virus Res. 28:113-124.

Ghoshroy, S., Lartey, E., Sheng, J., and Citovsky, V. 1997. Transport of proteins and nucleic acids through plasmodesmata. Annu. Rev. Plant Physiol. Plant Mo. Biol. 48:27-50.

Giesman-Cookmeyer, D. and Lommel, S. A. 1993. Alanine scanning mutagnesis of a plant virus movement protein identifies three functional domains. Plant Cell 5:973-982.

Giesman-Cookmeyer, D., Kim, K. H. and Lommel, S. A. 1995. Dianthoviruses. In "pathogenesis and host specificity in plant diseases" (Singh, R. P., Singh, U. P., and Kohmoto, K. Eds.), pp157-176. Elsevier, Oxford.

Gilbertson, R. L., and Lucas, W. J. 1996. How do viruses traffic on the vascular highway? Trends Plant Sci. 1:260-267.

Gmyl, A. P., Belousov, E. V., Maslova, S. V., Khitrina, E. V., Chetverin, A. B., and Agol, V. I. 1999. Nonreplictive RNA recombination in poliovirus. J. Virol. 73:8958-8965.

Greene, A. E. and Allison, R. F. 1994. Recombination between viral RNA and transgenic plant transcripts. Science 263:1423-1425.

Guan, H. and Simon, A. E. 2000. Polymerization of nontemplate bases before transcription initiation at th 3' ends of termplates by an RNA-dependent RNA polymerase: an activity involved in 3' end repair of viral RNAs. Proc. Natl. Acad. Sci. USA. 97:12451-12456.

Guan, H., Carpenter, C. D., and Simon, A. E. 2000. Analysis of cis-acting sequences involved in plus-strand synthesis of a trunip crinkle virus-associated satellite RNA identifies a new carmovirus replication element. Virology 268:345-354.

Guan, H., Carpenter, C. D., and Simon, A. E. 2000. Requirement of a 5'-proximal linear sequence on minus strands for plus-strand synthesis of a satellite RNA associated turnip crinkle virus. Virology 268:355-363.

Guan, H., Song, C., and Simon, A. E. 1997. RNA promoters located on (-)-strands of a subviral RNA associated with turnip crinkle virus. RNA 3:1401-1412.

Guilley, H., Carrington, J. C., Balazs, E., Jonard, G., Richards, K., and Morris, T. J. 1985. Nucleotide sequence and genome organization of carnation mottle virus RNA. Nucleic Acids Res. 13:6663-6677.

Haasnoot, P. C., Brederode, F. T., Olsthoorn, R. C., and Bol, J. F. 2000. A conserved hairpin structure in Alfamovirus and Bromovirus subgenomic promoters is required for efficient RNA synthesis in vitro. RNA 6:708-716.

Haasnoot, P. C., Olsthoorn, R. C., and Bol, J. F. 2002. The Brome mosaic virus subgenomic promoter hairpin is structurally similar to the iron-responsive element and functionally equivalent to the minus-strand core promoter stem-loop C. RNA 8:110-122.

Hansen, J. L., Long, A. M., and Schultz, S. C. 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure 5:1109-1122.

Havelda, Z. and Burgyan, J. 1995. 3' Terminal putative stem-loop structure required for the accumulation of cymbidium ringspot viral RNA. Virology 214:269-272.

Heringa, J., and Argos, P. 1994. "The evolutionary biology of viruses" (S. S. Morse, ed.), pp. 87-103. Raven Press, New York.

Hiruki, C. 1987. The dianthoviruses: A distinct group of isometric plant viruses with a bipartite genome. Adv. Virus Res. 33:257-300.

Hobson, S. D., Rosenblum, E. S., Richards, O. C., Richmond, K., Kirkegaard, K., and Schultz, S. C. 2001. Oligomeric structures of poliovirus polymerase are important for function. EMBO J. 20:1153-1163.

Hong, Z., Cameron, C. E., Walker, M. P., Castro, C., Yao, N., Lau, J. Y., and Zhong, W. 2001. A novel mechanism to ensure terminal initiation by hepatitis C virus NS5B polymerase. Virology 285: 6-11.

Hope, D. A., Diamond, S. E., and Kirkegaard, K. 1997. Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB. J. Virol. 71:9490-9498.

Ishikawa, M., Diez, J., Restrepo-Hartwig, M., and Ahlquist, P. 1997. Yeast mutations in multiple complementation groups inhibit brome mosaic virus RNA replication and transcription and perturb regulated expression of the viral polymerase-like gene. Proc. Natl. Acad. Sci. USA. 94:13810-13815.

Jacks, T., and Varmus, H. E. 1985. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science. 230:1237-1242.

Jacoba-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hiza, A., Hughes, S. H., and Arnold, E. 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proc. Natl. Acad. Sci. USA. 90:6320-6324.

Janda, M. and Ahlquist, P. 1998. Brome mosaic virus RNA replication protein 1a dramatically increases *in vivo* stability but not translation of viral genomic RNA-3. Proc. Natl. Acad. Sci. USA. 95:2227-2232.

Kamer, G. and Argos, P. 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucleic Acids Res. 12:7269-7282.

Kao, C. C. and Ahlquist, P. 1992. Identification of the domains required for direct interaction of the helicase-like and polymerase-like RNA replication proteins of brome mosaic virus. J. Virol. 66:7293-7302.

Kao, C. C. and Sun, J. H. 1996. Initiation of minus-strand RNA synthesis by the brome mosaic virus RNA-dependent RNA polymrease: use of oligoribonucleotide primers. J. Virol. 70:6826-6830.

Kao, C. C., Singh, P., and Ecker, D. J. 2001. De novo initiation of viral RNA-dependent RNA synthesis. Virology 287:251-260.

Kendall, T. L. and Lommel S. A. 1992. Nucleotide sequence of carnation ringspot dianthovirus RNA-2. J Gen Virol. 73:2479-2482.

Khatchikian, D., M. Orlich, and R. Rott. 1989. Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. Nature 340:156-157.

Kim, C. H. and Kao, C. C. 2000. RNA motifs that determine specificity between a viral replicase and its promoter. Nat. Struct. Biol. 7:415-423.

Kim, K. H. and Hemenway, C. L. 1999. Long-distance RNA-RNA interactions and conserved sequence elements affect potato virus X plus-strand RNA accumulation. RNA 5:636-645.

Kim, K. H. and Lommel, S. A. 1994. Identification and analysis of the site of -1 ribosomal frameshifting in red clover necrotic mosaic virus. Virology 200:574-582.

Kim, M. J. and Kao, C. C. 2001. Factors regulating template switch *in vitro* by viral RNA-dependent RNA polymerase: Implications for RNA-RNA recombination. Proc. Natl. Acad. Sci. USA 98:4972-4977.

Kirkegaard, K. and Baltimore, D. 1986. The mechanism of RNA recombination in Poliovirus. Cell 47:433-443.

Klovins, J. and van Duin, J. 1999. A long-range pseudoknot in Qβ RNA is essential for replication. J. Mol. Biol. 294:875-884.

Klovins, J., Berzins, V., and van Duin, J. 1998. A long-range interaction in Qbeta RNA that bridges the thousand nucleotides between the M-site and the 3' end is required for replication. RNA 4:948-957.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. 1992. Crystal structure at 3.5 A° resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 256:1783-1790.

Koonin, E. V. 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J. Gen. Virol. 72:2197-2206.

Koonin, E. V. and Dolja, V. V. 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit. Rev. Biochem. Mol. Biol. 28:375-430.

Koonin, E. V., Mushegian, A. R., Ryabov, E. V., and Dolja, V. V. 1991. Diverse groups of plant RNA and DNA viruses share related movement proteins that may possess chaperone-like activity. J. Gen. Virol. 72:2895-2903

Kroner, P. A., Young, B. M., Ahlquist, P. 1990. Analysis of the role of brome mosaic virus 1a protein domains in RNA replication, using linker insertion mutagenesis. J Virol. 64:6110-6120.

Lai, M. M. C. 1992. RNA recombination in animal and plant viruses. Microbiological Reviews 56:61-79.

Ledinko, N. 1963. Virology 180:107-119.

Lee, W. M., Ishikawa, W., and Ahlquist, P. 2001. Mutation of host delta9 fatty acid desaturase inhitits brome mosaic virus RNA replication between template recognition and RNA synthsis. J. Virol. 75:2097-2106.

Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F., and Weber, P. C. 1999. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat Struct Biol. 6:937-943.

Lommel, S. A., Weston-Fina, M., Xiong, Z., and Lomonossoff, G. P. 1988. The nucleotide sequence and gene organization of red clover necrotic mosaic virus RNA-2. Nucleic Acids Res. 16:8587-8602.

Mayo, M. A. and Pringle, C. R. 1998. Virus taxonomy--1997. J Gen Virol. 79:649-657.

Mayo, M. A., and C. A. Jolly. 1991. The 5'-terminal sequence of potato leafroll virus RNA: Evidence of recombination between virus and host RNA. J. Gen. Virol. 72:2591-2595.

Meyers, G., N. Tautz, E. J. Dubovi, and H. J. Thiel. 1991. Viral cytopathogenicity correlated with integration of ubiquitin coding sequences. Virology 180:602-616.

Miller, E. D., Plante, C. A., Kim, K. H., Brown, J. W., and Hemenway, C. 1998. Stem-loop structure in the 5' region of potato virus X genome required for plus-strand RNA accumulation. J. Mol. Biol. 284:591-608.

Miller, W. A., Waterhouse, P. M., and Gerlach, W. L. 1988. Sequence and organization of barley yellow dwarf virus genomic RNA. Nucleic Acids Res. 16:6097-6111.

Munishkin A. V., Boronin L. A., and Chetverin A. B. 1988. An *in vivo* recombinant RNA capable of autocatalytic synthesis by Qβ replicase. Nature 333:473-475.

Murphy, F. A., C. M. Fauquet, M. A. Mayo, A. W. Jarvis, S. A. Ghabrial, M. D. Summers, G. P. Martelli,and D. H. L. Bishop. 1995. The Classification and Nomenclature of Viruses: Sixth Report of the International Committee on Taxonomy of Viruses. Archives of Virology, Supplementum 3. Springer Verlag, New York.

Nagy, P. D. and Bujarski, J. J. 1993. Targeting the site of RNA-RNA recombination in brome mosaic virus with antisense sequences. Proc. Natl. Acad. Sci. U.S.A. 90:6390-6394.

Nagy, P. D. and Bujarski, J. J. 1995. Efficient system of homologous RNA recombination in brome mosaic virus: Sequence and structure requirements and accuracy of crossovers. J. Virol. 69:131-140.

Nagy, P. D. and Bujarski, J. J. 1996. Homologous RNA recombination in brome mosaic virus AU-rich sequences decrease the accuracy of crossovers. J. Virol. 70:415-426.

Nagy, P. D. and Bujarski, J. J. 1997. Engineering of homologous recombination hot-spots with AU-rich sequences in brome mosaic virus. J. Virol. 71:1294-1306.

Nagy, P. D. and Bujarski, J. J. 1998. Silencing homologous RNA recombination hot spots with GC-rich sequences in brome mosaic virus. J. Virol. 72:1122-1130.

Nagy, P. D. and Simon, A. E. 1997. New insights into the mechanisms of RNA recombination. Virology 235:1-9.

Nagy, P. D. and Simon, A. E. 1998a. *In vitro* characterization of late steps of RNA recombination in turnip crinkle virus I. Role of the motif1-hairpin structure. Virology 249:379-392.

Nagy, P. D. and Simon, A. E. 1998b. *In vitro* characterization of late steps of RNA recombination in turnip crinkle virus II. The role of the priming stem and flanking sequences. Virology 249:393-405.

Nagy, P. D., Dzianott, A., Ahlquist, P., and Bujarski, J. J. 1995. Mutations in the helicase-like domain of protein 1a alter the sites of RNA-RNA recombination in brome mosaic virus. J. Virol. 69:2547-2556.

Nagy, P. D., Ogiela, C., and Bujarski, J. J. 1998a. Mapping sequences active in homologous RNA recombination in brome mosaic virus: Prediction of recombination hot-spots. Virology 254:92-104.

Nagy, P. D., Pogany, J., and Simon, A. E. 1999. RNA elements required for RNA recombination function as replication enhancers *in vitro* and *in vivo* in a plus-strand RNA virus. EMBO 18:5653-5665.

Nagy, P. D., Zhang, C., and Simon, A. E. 1998b. Dissecting RNA recombination *in vitro*: Role of RNA sequences and the viral replicase. EMBO 17:2392-2403.

Nutter, R. C., Scheets, K., Panganiban, L. C., and Lommel, S. A. 1989. The complete nucleotide sequence of the maize chlorotic mottle virus genome. Nucleic Acids Res. 17:3163-3177.

O'Reilly, E. K. and Kao., C. C. 1998. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. Virology 252:287-303.

O'Reilly, E. K., Paul, J. D., and Kao, C. C. 1997. Analysis of the interaction of viral RNA replication proteins by using the yeast two-hybrid assay. J. Virol. 71:7526-7532.

O'Reilly, E. K., Tang, N., Ahlqist, P., and Kao, C. C. 1995. Biochemical and genetic analysis of the interaction between the helicase-like and polymerase-like proteins of the brome mosaic virus. Virology 214:59-71.

O'Reilly, E. K., Wang, Z., French, R., and Kao, C. C. 1998. Interactions between the structural domains of the RNA replication proteins of plant-infecting RNA viruses. J. Virol. 72:7160-7169.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. 1985. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. Nature 313: 762-766.

Osman, T. A. M. and Buck, K. W. 1987. Replication of red clover necrotic mosaic virus RNA in cowpea protoplasts: RNA-1 replicates independently of RNA-2. J. Gen. Virol. 68:289-296.

Osman, T. A. M. and Buck, K. W. 1990. Double-stranded RNAs isolated from plant tissue infected with red clover necrotic modaic virus correspond to genomic and subgenomic single-stranded RNAs. J. Gen. Virol. 71:945-948.

Osman, T. A., Hayes, R. J., and Buck, K. W. 1992. Cooperative binding of the red clover necrotic mosaic virus movement protein to single-stranded nucleic acids. J Gen Virol. 73:223-227.

Paje-Manalo, L. and Lommel, S. A. 1993. Idependent replication of red clover necrotic mosaic virus RNA-1 in electroporated host and non-host Nicotiana sp. protoplasts. Phytopathol. 79:457-461.

Pata, J. D., Schultz, S. C., Kirkegaard, K. 1995. Functional oligomerization of poliovirus RNA-dependent RNA polymerase. RNA. 1:466-477.

Poch, O., Sauvaget, I., Delarue, M., and Tordo, N. 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J. 8:3867-3874.

Pogany, J., Romero, J., Huang, Q., Sgro, J. Y., Shang, H., and Bujarski, J. J. 1995. De novo generation of defective interfering-like RNAs in broad bean mottle bromovirus. Virology 212:574-586.

Pogue, G. P. and Hall, T. C. 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.

Pogue, G. P., Marsh, L. E., and Hall, T. C. 1990. Point mutations in the ICR2 motif of brome mosaic virus RNAs debilitate (+)-strand replication. Virology 178:152-160.

Pogue, G. P., Marsh, L. E., Connell, J. JP., and Hall, T. C. 1992. Requirement for ICR-like sequences in the replication of brome mosaic virus genomic RNA. Virology 188:742-753.

Prufer, D., Tacke, E., Schmitz, J., Kull, B., Kaufmann, A., and Rohde, W. 1992. Ribosomal frameshifting in plants: a novel signal directs the -1 frameshift in the synthesis of the putative viral replicase of potato leafroll luteovirus. EMBO J. 11:1111-1117.

Quadt, R., Ishikawa, M., Janda, M., and Ahlquist, P. 1995. Formation of brome mosaic virus RNA-dependent RNA polymerase in yeast requires coexpression of virla proteins and viral RNA. Proc. Natl. Acad. Sci. USA. 92:4892-4896.

Quadt, R., Kao, C. C., Browning, K. S., Hershberger, R. P., and Ahlquist P. 1993. Characterization of a host protein associated with brome mosaic virus RNA-dependent RNA polymerase. Proc. Natl. Acad. Sci. USA. 90:1498-1502.

Restrepo-Hartwig, M. A. and Ahlquist, P. 1996. Brome mosaic virus helicase- and polymerase-like proteins colocalize on the endoplasmic reticulum at sites of viral RNA synthesis. J. Virol. 70:8906-8916.

Restrepo-Hartwig, M. and Ahlquist, P. 1999. Brome mosaic virus RNA replication proteins 1a and 2a colocaloze and 1a independently localizes on the yeast endoplasmic reticulum. J. Virol. 73:10303-10309.

Richards, O. C., and Ehrenfeld, E. 1997. One of two NTP binding sites in poliovirus RNA polymerase required for RNA replication. J Biol Chem. 272:23261-23264.

Richards, O. C., Hanson, J. L., Schultz, S., and Ehrenfeld, E. 1995. Identification of nucleotide binding sites in the poliovirus RNA polymerase. Biochemistry. 34:6288-6295.

Roux, L., Simon, A. E., and Holland, J. J. 1991. Efeects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo*. Adv. Virus Res. 40:181-211.

Ryabov E.V., Generozov E.V., Kendall T.L., Lommel S.A., and Zavriev S.K. 1994. Nucleotide sequence of carnation ringspot dianthovirus RNA-1. J. Gen. Virol. 75:243-247.

Siegel, R. W., Adkins, S., and Kao, C. C. 1997. Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. Proc Natl Acad Sci U S A. 94:11238-11243.

Siegel, R. W., Bellon, L., Beigelman, L., and Kao, C. C. 1998. Moieties in an RNA promoter specifically recognized by a viral RNA-dependent RNA polymerase. Proc Natl Acad Sci U S A. 95:11613-11618.

Simon, A. E. and Bujarski, J. J. 1994. RNA-RNA recombination and evolution in virus infected plants. Annu. Rev. Phytopathology 32: 337-362.

Sit, T. L., Vaewhongs, A. A., and Lommel, S. A. 1998. RNA-mediated trans-activation of transcription from a viral RNA. Science 281:829-832.

Sivakumaran, K. and Kao, C. C. 1999. Initiation of genomic plus-strand RNA synthsis from DNA and RNA templates by a viral RNA-dependent RNA polymerase. J. Virol. 73:6415-6423.

Song, C. and Simon, A. E. 1995. Requirement of a 3'-terminal stem-loop in *in vitro* transcription by an RNA-dependent RNA polymerase. J. Mol. Biol. 254:6-14.

Sousa, R., Chung, U. J., Rose, J. P., and Wang, B. 1993. Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A° resolution. Nature 364:593-599.

Stawicki, S. S. and Kao, C. C. 1999. Spatial perturbations within an RNA promoter specifically recognozed by a viral RNA-dependent RNA polymerase (RdRp) reveal that RdRp can adjust its pormoter binding sites. J. Virol. 73:198-204.

Stupina, V. and Simon, A. E. 1997. Analysis *in vivo* of turnip crinkle virus satellite RNA C variants with mutations in the 3'-terminal minus-strand promoter. Virology 238:470-477.

Sullivan, M. L. and Ahlquist, P. 1999. A brome mosaic virus intergenic RNA-3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA *in vivo*. J. Virol. 73:2622-2632.

Sun, J. H. and Kao, C. C. 1997. Characterization of RNA products associated with or aborted by a viral RNA-dependent RNA polymerase. Virology 236:348-353.

Sun, J. H., Adkins, S., Faurote, G., and Kao, C. C. 1996. Initiation of (-)-strand RNA synthesis catalyzed by the BMV RNA-dependent RNA polymerase: synthesis of oligonucleotides. Virology 226:1-12.

Turner, R. L. and Buck, K. W. 1999. Mutational analysis of cis-acting sequences in the 3'- and 5'-untranslated regions of RNA-2 of red clover necrotic mosaic virus. Virology 252:115-124.

Vaewhongs, A. A. and Lommel, S. A. 1995. Virion formation is required for the long-distance movement of red clover necrotic mosaic virus in movement protein transgenic plants. Virology 212:607-613.

Van Dyke, T. A., Flanegan, J. B. 1980. Identification of poliovirus polypeptide P63 as a soluble RNA-dependent RNA polymerase. J Virol. 35:732-740.

Varmus, H. 1988. Regulation of HIV and HTLV gene expression. Genes Dev. 2:1055-1062.

Varrelmann M, Palkovics L, Maiss E. 2000. Transgenic or plant expression vector-mediated recombination of Plum Pox Virus. J Virol. 74:7462-7469.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1:945-951.

Wang, H., Wang, Y., Giesman-Cookmeyer, D., Lommel, S. A., and Lucas, W. J. 1998. Mutations in viral movement protein alter systemic infection and identify an intercellular barrier to entry into the phloem long-distance transport system. Virology 245:75-89.

Wang, J. and Simon, A. E. 1997. Analysis of the two subgenomic RNA promoters for turnip crinkle virus *in vivo* and *in vitro*. Virology 232:174-186.

Wang, J., Carpenter, C. D., and Simon, A. E. 1999. Minimal sequence and structural requirements of a subgenomic RNA promoter for turnip crinkle virus. Virology 253:327-336.

Weng, Z. and Xiong, Z. 1995. A method for accurate determination of terminal sequences of viral genomic RNA. Genome Research 5:202-207.

Weng, Z. and Xiong, Z. 1997. Genome organization and gene expression of saguaro cactus carmovirus. J. Gen. Virol. 78:525-534.

White, K. A., and T. J. Morris. 1995. RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. RNA 1:1029-40.

Wolf, S., Deom, C. M. Beachy, R. N., and Lucas, W. J. 1989. Movement protein of tobacco mosaic virus modifies plasmodesmatal size exclusion limit. Science 246:377-379.

Xiang, W., Cuconati, A., Hope, D., Kirkegaard, K., and Wimmer, E. 1998. Complete protein linkage map of poliovirus P3 proteins: interaction of polymerase 3Dpol with VPg and with genetic variants of 3AB. J. Virol. 72:6732-6741.

Xiong, Y., and Eickbush, T. H. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 9:3353-3362.

Xiong, Z. 1988. Nucleotide sequence and organization of the red clover necrotic mosaic virus genome. Ph.D dissertation. Kansas State University.

Xiong, Z. and Lommel, S. A. 1989. The complete nucleotide sequence and genome organization of red clover necrotic mosaic virus RNA-1. Virology 171:543-554.

Xiong, Z. and Lommel, S. A. 1991. Red clover necrotic mosaic virus infectious transcripts synthesized *in vitro*. Virology 182:388-392.

Xiong, Z., Kim, K. H., Giesman-Cookmeyer, D., and Lommel, S. A. 1993. The roles of the red clover necrotic mosaic virus capsid and cell-to-cell movement proteins in systemic infection. Virology 192:27-32.

Xiong, Z., Kim, K. H., Kendall, T. L., and Lommel, S. A. 1993. Synthesis of the putative red clover necrotic mosaic virus RNA polymerase by ribosomal frameshifting *in vitro*. Virology 193:213-221.

Yoshinari, S., Nagy, P. D., Simon, A. E., and Dreher, T. W. 2000. CCA initiation boxes without unique promoter elements support *in vitro* transcription by three viral RNA-dependent RNA polymerases. RNA 6:698-707.

Zaccomer, B., Haenni, A. L., and Macaya, G. 1995. The remarkable variety of plant RNA virus genomes. J. Gen. Virol. 76:231-247.

Zavriev, S. K., Hickey, C. M., and Lommel, S. A. 1996. Mapping of the red clover necrotic mosaic virus subgenomic RNA. Virology 216:407-410.

Zhang, C. X., Cascone, P. J., and Simon, A. E. 1991. Recombination between satellite and genomic RNAs of turnip crinkle virus. Virology 184:791-794.

CHAPTER 2.

The N-terminal 14 Amino Acids of RNA Polymerase of *Red Clover Necrotic Mosaic Virus* Are Responsible for Differential Levels of Replication of Two Infectious Clones

ABSTRACT

Red clover necrotic virus (RCNMV) contains bipartite, positive-sense, single-stranded RNA genome. RNA-1 (3.9 kb) encodes three functional proteins, p27, p88 RNA polymerase, and p37 coat protein. RNA-2 (1.45 kb) encodes a single p35 movement protein. RNA-1 can replicate in protoplasts in the absence of RNA-2. By using reverse transcription and polymerase chain reaction (RT-PCR), a series of full-length cDNA clones of RNA-1 and RNA-2 were generated. Among these new clones, the transcripts derived from pRC1IG69 RNA-1 clone and pRC2IG54 RNA-2 clone were as infectious as wild type viral RNAs. RNA-1 transcripts derived from the highly infectious clone pRC1IG69 replicated and accumulated 100-fold higher than those from a poorly infectious clone pRC1IG in *Nicotiana clevelandii* protoplasts. A series of chimerical clones between pRC1IG69 and pRC1IG were subsequently constructed. Analysis of these RNA-1 chimerical clones allowed the localization of the mutation responsible for low infectivity and low RNA replication to the 5'420 nt region of pRC1IG. Direct sequencing of this 420 nt region revealed a single U deletion in the p27 and p88 coding regions of pRC1IG (145 UUUUU 149 changed to 145 UUUU 148) as well as several

other nucleotide substitutions. Addition of an uridylate residue to pRC1IG at position 145-148 by site-directed mutagenesis restored both the RNA replication and infection *in planta*. This single U deletion caused translation frameshifting starting from the N-terminal 9th amino acid (aa) of both p27 and p88 and subsequently premature termination at UAA stop codon at the 14th aa position, resulting in the production of a small 13 aa oligopeptide. However, this U deletion prompted ribosome reinitiation from a downstream AUG start codon located at 15th aa position, which led to the production of p27 and p88 lacking the N-terminal 14 aa, and consequently poor infectivity. Taken together, these data suggest that the N-terminal 14 aa of p88 RdRp and/or p27 is required for efficient RNA replication in RCNMV.

INTRODUCTION

Most plant infecting viruses and many animal viruses are RNA viruses (Murphy *et al.*, 1995). As the current techniques are suitable only for DNA manipulation, highly infectious full-length cDNA (complementary DNA) clones are needed to study the gene functions and the mechanisms involved in virus life cycle. Since the first publication of infectious cDNA clone of Qβ in 1978 (Taniguchi *et al.*, 1978), the infectious full-length cDNA clones from numerous and very diverse RNA viruses have been generated (for reviews see Boyer and Haenni, 1994; Lai, 2000). The basic strategy to construct the infectious full-length cDNA clones is to insert the full-length copy of the double strand cDNA synthesized from a viral genomic RNA into a vector under the control of an RNA polymerase promoter. The infectious transcripts are transcribed either *in vivo* (Mori *et al.*,

1991) or *in vitro* (Ahlquist *et al.*, 1987). Several common problems will be experienced. First, the efficiency of the first strand cDNA synthesis is very low due to secondary structure in RNA template, and the size of the most cloned cDNA is less than 2 kb. Second, the extra non-viral nucleotides present at both termini of viral sequences will dramatically reduce or abolish the viral infectivity. The initial way to overcome small cDNA clones is to screen the overlapping fragments and join them into full-length. The non-viral vector sequences are removed by site-specific mutagenesis. Although these solutions work, they are laborious, and only one or few full-length clones (which may be non- or poorly infectious) will be made each time (Xiong and Lommel, 1991). With the power of polymerase chain reaction (PCR) for amplification of sequences of interest and the availability of the reverse transcriptase and Taq polymerase with increased fidelity, most of these difficulties have been overcome. The low yield of first strand cDNA was overcome by the strong amplification ability of PCR. The nonviral vector sequences were solved by simply linking the polymerase (T7/T3/SP6) promoter to the 5' end of viral sequence in PCR upstream primer and by designing a restrict enzyme site at the viral 3' end in PCR downstream primer. Moreover, multiple full-length clones with various levels of infectivity can be generated simultaneously. Since its first successful application in cloning the infectious full-length cDNA clone of cucumber mosaic virus, RT-PCR has been used in the construction of many of the infectious cDNA clones in last ten years (Hayes and Buck, 1990).

In this study, a highly infectious cDNA clone was generated with reverse-transcription and polymerase chain reaction (RT-PCR) approach. Comparative

characterization between the highly infectious cDNA clone and a previously reported poorly infectious clone (Xiong and Lommel, 1991) identified a region of the pre-fusion portion of RdRp required for efficient replication of *Red clover necrotic mosaic virus* (RCNMV) RNA genome. RCNMV belongs to *Dianthovirus* genus, *Tombusviridae* family (Giesman-Coolmeyer *et al.*, 1995; Mayo and Pringle, 1998). It contains two single-stranded, positive-sense genomic RNAs with sizes of 3.9- and 1.45-kb (Xiong and Lommel, 1989; Lommel *et al.*, 1988). RNA-1 can replicate in the absence of RNA-2 in the infected cell (Paje-Manalo and Lommel, 1993; Osman and Buck, 1987). RNA-1 is composed of three ORFs encoding for proteins p27, p57, and p37 from the 5' to 3'. A p88 fusion protein of p27 and p57 is produced by ribosomal frameshifting (Kim and Lommel, 1994; Xiong *et al.*, 1993b). Both p27 and p88 are required for viral RNA replication (Vaewhongs and Lommel, 1995), and are associated with membrane (Bates *et al.*, 1995). While p88 is a RNA-dependent RNA polymerase (RdRp), the function of p27 is unknown.

Expression of RdRp as a fusion protein of the 5' first two ORFs is a feature conserved in *Tombusviridae*. The N-terminal pre-fusion portion is produced from the first ORF and the C-terminal post-fusion portion is directed from the second ORF (Buck, 1996). The post-fusion portion contains the conserved motifs including the GDD motif identified in all RdRps (Koonin, 1991; Koonin and Dolja, 1993). It also shares a similar order of secondary structure elements as identified in the RdRps of several viruses including poliovirus, hepatitis C virus (HCV), and bacteriophage Φ6 (Bressanelli *et al.*, 1999; Butcher *et al.*, 2001; Hansen *et al.*, 1997; Hong *et al.*, 2001; Lesburg *et al.*, 1999;

O'Reilly and Kao, 1998). In contrast to the post-fusion part, the pre-fusion part of RdRp is unique to tombusviruses, shows no amino acid sequence homology to any known functional proteins in databases. This pre-fusion part is also produced as an independent small protein from the first ORF. Mutagenesis revealed that both the small protein and the fusion RdRp were required for viral RNA replication (Vaewhongs and Lommel, 1995). How this N-terminal element is involved in RNA replication is currently not clear. RCNMV is used in present study to understand the function of the small pre-fusion protein.

We report here that a series of full-length cDNA clones were generated by using RT-PCR strategy and that the transcripts derived from pRC1IG69 of RNA-1 and pRC2IG54 of RNA-2 were as infectious as viral RNAs. The transcripts derived from pRC1IG69 can replicate and accumulate in protoplasts to a 100-fold higher than a previously generated clone pRC1IG (Xiong and Lommel, 1991). By studying RNA-1 chimerical clones of highly infectious clone pRC1IG69 and poorly infectious clone pRC1IG, we present evidence to show that the N-terminal 14 aa of p88 RdRp and/or p27 proteins is required for efficient RCNMV RNA replication.

MATERIALS AND METHODS

Cloning and sequencing of the 5'- and 3'-terminal sequences

A combination of ligation-anchored PCR and anchored cDNA cloning techniques developed previously (Weng and Xiong, 1995) was used to clone the termini of RCNMV genomic RNA-1 and -2. Briefly, RNA-1 and -2 were reverse transcribed into cDNA with

reverse transcriptase (Invitrogen, Carlsbad, CA) using primers RC1c830 and RC2c640 (Table 2.1) that were complementary to nucleotides 834-851 of RNA-1 and 643-659 of RNA-2, respectively. RNA was removed by the addition of NaOH and boiling. After neutralization with HCl, cDNA was precipitated with ethanol. Two complementary primers F500 and R500 (Weng and Xiong, 1995) were used for ligation-anchored PCR. The anchor primer F500 was phosphorylated with polynucleotide kinase and ATP at the 5' end and blocked with TdT and ddATP at the 3' end. The 5'-phosphorylated and 3'-blocked anchor F500 was ligated to the first strand cDNA using T4 RNA ligase (Invitrogen, Carlsbad, CA). The viral 5' termini were amplified by a 40-cycle PCR using primers RC1c300 and R500 for RNA-1 and primers RC2c320 and R500 for RNA-2 (Table 2.1). The PCR products were blunt-ended with the Klenow fragment, phosphorylated with kinase and cloned into pBluscript SK (+) at *Sma*I site. The 5'-terminal clones were then sequenced using T7 or T3 primers based on insertion orientation.

Anchored cDNA cloning was used to clone the 3' termini. The 5'-phosphorylated and 3'-blocked anchor F500 was ligated directly to RCNMV RNA-1&-2 with T4 RNA ligase. The anchor-ligated viral RNA was then reverse transcribed into cDNA using primer R500 followed by second strand cDNA synthesis using a combination of *Escherichia coli* RNase H, *E. coli* DNA polymerase I, *E. coli* DNA ligase and T4 DNA polymerase. Two *Xba*I sites, one within primer R500 and another 417 nt upstream of the 3' end of RNA-2, were used to clone the DNA fragment representing the RNA-2 3' terminus into pBluscript SK (+). The 3'-terminal clones (4 clones) were then sequenced using T7 and T3 primers.

Synthesis of full-length cDNA clones by RT-PCR

The strategy to generate the full-length cDNA clones of RCNMV RNA-1 and –2 was illustrated in Figure 2.1. Two 3' primers, RC13'END and RC23'END (Table 2.1), were used to synthesize the first strand cDNA of RNA-1 and RNA-2, respectively. Primers RC13'END and RC23'END contain 20 nt complementary to the 3'-termini of RNA-1 and RNA-2, respectively, and 8 nt that creates a *SmaI* site for the generation of viral 3' end and protects the viral 3' end from degradation. A 20 µl reverse transcription contains 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM each of dATP, dCTP, dGTP and dUTP, 1 µg RC13'END/RC23'END primer, 3 µg heat-denatured RCNMV RNA, and 500 U reverse transcriptase (Invitrogen, Carlsbad, CA). After one-hour incubation at 42°C, the reaction was diluted to 1:10 with sterile distilled water and used directly as template for PCR.

The full-length cDNAs of RCNMV RNAs were amplified by PCR using two pairs of terminal primers, RC1T75'END and RC13'END for RNA-1 and RC2T75'END and RC23'END for RNA-2 (Table 2.1). Primers RC1T75'END and RC2T75'END contain 18 nt identical to the 5' end of RNA-1 and RNA-2, respectively, T7 promoter sequence upstream viral sequence, and an extraneous non-viral G residue added between 5' viral sequence and the T7 promoter sequence to facilitate the *in vitro* transcription by T7 RNA polymerase. The PCR reaction was carried out in 25 µl of 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% Triton X-100, and 0.2 mM dNTPs containing 1.25 units of a mixture of *Taq* DNA polymerase (Promega, Madison, WI):Pfu DNA polymerase

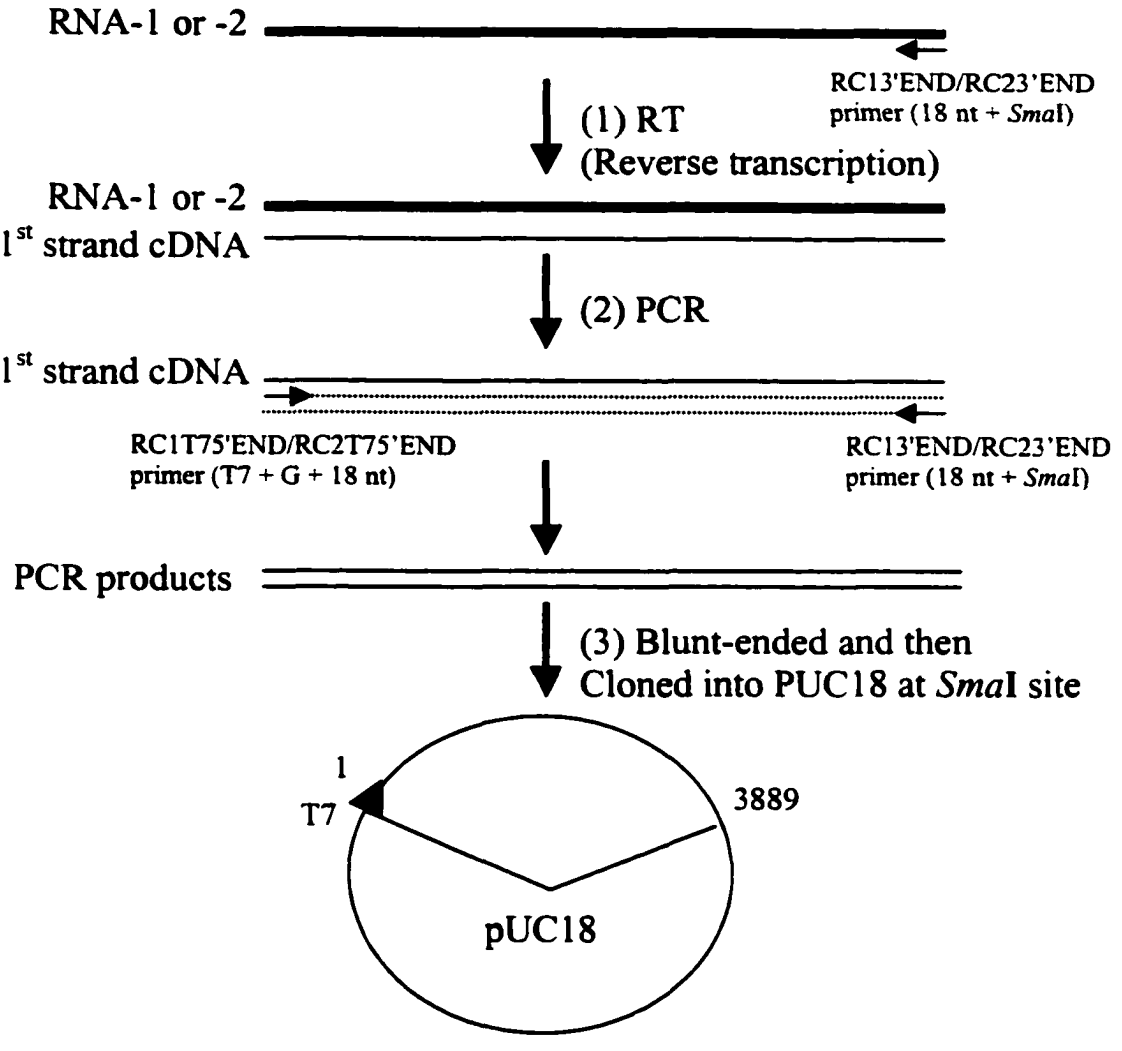


Figure 2.1 RT-PCR strategy for the construction of full-length cDNA clones of RCNMV RNA. (1) RCNMV RNA was reverse transcribed into complementary DNA (cDNA) by a specific primer RC13'END or RC23'END which contains 18 nucleotides complementary to the 3' end of RNA-1 or RNA-2, and a *Sma*I site to facilitate the generation of transcripts with wild type viral 3' end. (2) The first strand cDNA was then amplified by 30 cycles of PCR using two primers, primer RC13'END or RC23'END and primer RC1T75'END or RC2T75'END which has 18 nucleotide of 5' end of RNA-1 or RNA-2, T7 promoter sequence leading the viral sequence and a G between T7 and viral sequences for increasing the transcription efficiency by T7 RNA polymerase.

Table 2.1 Sequences of primers used in this study

Primers	Primer Sequences
RC1T75'END*	5'-GAATTGTAATACGACTCACTATAg**ACAAACGTTTTACCGGTT**-3'
RC2T75'END*	5'-GAATTGTAATACGACTCACTATAg**ACAAACCTCGCTCTATAA**-3'
RC13'END**	5'-GAGCT*CCCGGGGTACCTAGCCGTTATAC*-3'
RC23'END**	5'-GAGCT*CCCGGGGTGCCTAGCCGTTATAC*-3'
RC1c830	5'-GAAAGCTGAGTGGGCCGC-3'
RC1c300	5'-AAGGCATCCACCGCCCAC-3'
RC2c640	5'-TGCACGTAGGCTTCCACTGCC-3'
RC2c320	5'-CTACGGTGACGTGTCCTG-3'
RC1vPOL***	5'-CGTACCAGCCATGGGTTTTATAAATCTTTCGCTTTTTGATGTGG-3'
RC1cPOL	5'-CCACATCAAAAAGCGAAAGATTTATAAAACCCATGGCTGGTACG-3'

*: Lower case g indicates the non-viral nucleotide added between the T7 promoter sequence (underlined) and the viral 5'-terminal sequence (bold) to facilitate the initiation of transcription by T7 RNA polymerase.
**: Twenty nt complementary to the 3'-ends of RNA-1 and RNA-2 are bold. *Sma*I restriction site is underlined.
***: Bold ATG indicates the translation start codon of p27 and p88. The underlined five T indicates the region where one T was added. The underlined C is the substitution for T to generate a *Nco*I restriction site.

(Stratagene, La Jolla, CA) [20:1, (vol/vol)], 0.4 μM each of the two primers, and 15 ng of first strand cDNA. The mixture was overlaid with 25 μl of light mineral oil and incubated at 94°C for 3 minutes in a Temp-tronic thermocycler (Thermolyne, IL) prior to the 30 PCR cycle. Each PCR cycle consisted of denaturation at 94°C for 45 seconds, annealing at 58°C for 60 seconds, and polymerization at 72°C for 3 minutes. The full-length PCR products (~4 kb and 1.5 kb) were blunt-ended with Klenow fragment, phosphorylated with T4 kinase, and ligated into pUC18 at the *Sma*I site.

In vitro transcription

Full-length cDNA clones of RNA-1 and RNA-2 were linearized with *Sma*I followed by *in vitro* transcription with bacteriophage T7 RNA polymerase (Invitrogen, Carlsbad, CA) as described previously (Xiong and Lommel, 1991; Weng and Xiong, 1997). A 10-μl reaction contains 1 μg template DNA, 40 mM Tris-HCl pH 8.0, 8 mM $MgCl_2$, 2 mM spermidine-(HCl)3, 25 mM NaCl, 5 mM dithiothreitol (DTT), 1 mM each of ATP, CTP, GTP and UTP, 8 U of T7 RNA polymerase (Invitrogen, Carlsbad, CA), and 1 U of Ribonuclease inhibitor (Invitrogen, Carlsbad, CA). The amount of transcripts synthesized *in vitro* was determined by agarose gel electrophoresis prior to plant and protoplast assays.

Plant inoculation and infectivity analysis

RNA-1 and RNA-2 transcripts from each of 10-μl *in vitro* transcription (approximately 2 μg RNA) were mixed and diluted at 1:5 ratio with GKP buffer (50 mM glycine, 30 mM K_2HPO_4 pH 9.2, 1% bentonite, 1% celite), and then inoculated to a total

of six primary leaves on three cowpea plants as previously described (Xiong and Lommel, 1991). The inoculated plants were kept at 23°C.The lesion number on the inoculated leaves was counted five days after inoculation.

Protoplast transfection

A protocol described previously (Rao *et al.* 1994; Weng and Xiong, 1998) was modified and used. The young *N. clevelandii* leaves were sliced and digested with cellulase "Onozuka" R-10 (Yakult Honsha Co., Ltd.) and macerozyme R-10 (Yakult Honsha Co., Ltd.). Protoplasts were isolated, purified, and divided into aliquots of 1.25 x 10^6 into each polypropylene tube. The DNA template in the *in vitro* transcription reaction was removed by lithium chloride precipitation. Approximately 10 μg transcripts derived from either pRC1IG or pRC1IG69 were co-transfected with 10 μg transcripts from pRC2IG54 into the purified protoplasts using the polyethylene glycol (PEG: Mr 1450, Sigma) method (Rao *et al.* 1994; Weng and Xiong, 1998). The transfected protoplasts were incubated at room temperature under fluorescent light for 24 hours.

Northern blot analysis

Total RNA was extracted from protoplasts at 0 hour and 24 hours post transfection as described (Rao *et al.* 1994; Weng and Xiong, 1998). Total RNAs (5 μg) was electrophoresed in 1.5 % agarose gels after incubation at 70°C for 10 minutes, and transferred to nylon membranes (Amersham, Arlington Heights, ILs). The blots were

then hybridized with ^{32}P-labeled (-) sense RNA-1- or RNA-2- probes as described (Xiong *et al.*, 1993a).

In vitro translation

In vitro transcripts were translated in a wheat germ *in vitro* translation system (Boehringer Mannheim Biochemica). A 25 µl reaction contains 25 µM each of 19 amino acids except methionine, 1 mM ATP, 20 µM GTP, 8 mM creatine phosphate, 30 µg/ml spermine, 2 mM dithiothreitol, 14 mM Hepes pH7.6, 40 µg/ml creatine kinase, 100 mM potassium acetate, 1 mM magnesium acetate, 1 µg *in vitro* transcripts, 0.8 µM [^{35}S]-methionine, and 7.5 µl of wheat germ extract. After incubation at 30 °C for 1 hour, ^{35}S -labeled translation products were resolved by 12.5% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970), fixed and visualized by fluorography (Xiong *et al.*, 1993a; Weng and Xiong, 1997).

Construction of RNA-1 chimerical clones

Two pairs of restriction enzymes, (*Xho*I & *Sma*I and *Eco*RI & *Sma*I), were used to make the chimerical clones between pRC1IG and pRC1IG69 (Figure 2.3). The *Xho*I site located in the middle and *Sma*I at the 3' terminus of RNA-1 were used to generate RC169XSIG and RC1IGXS69. An *Eco*RI restriction site 420 nt downstream of the 5' terminus and *Sma*I were used to generate two other clones, RC169ESIG and RC1IGES69 (Figure 2.3).

PCR-based, site-directed mutagenesis

QuikChange™ site-directed mutagenesis protocol (Stratagene, La Jolla, CA) was used to generate pRC1IGRS (Figure 2.3) which has a T inserted to the 145-TTTT-148 region of pRC1IG to restore the wild type 145-TTTTT-149 sequence. Two complimentary primers, RC1vPOL and RC1Cpol (Table 2.1), were designed for PCR mutagenesis. Primer RCvPOL contains 44 nt identical to 113-155 nt region of pRC1IG with two exceptions of a single T insertion within 145-148 nt and a single nucleotide substitution of C for T at nucleotide 121. This C for T substitution generated a *Nco*I site upstream AUG start codon of p27 and p88 to facilitate the screen of restoration clones. Primer RC1cPOL is completely complimentary to primer RC1vPOL. A typical 50-µl PCR reaction contains 10 mM KCl, 6 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (PH8.0), 2 mM $MgCl_2$, 0.1% Triton X-100, 10 µg /ml bovine serum albumin (BSA), 40 ng PRC1IG, 0.25 µM each of two primers, 0.2 mM each of dATP, dCTP, dGTP AND dTTP, and 2.5 U Pfu DNA polymerase (Stratagene, La Jolla, CA). The mixture was overlaid with 25 µl of light mineral oil and incubated at 95°C for 30 seconds in a Temp-tronic thermocycler (Thermolyne, IL) prior to the 18 PCR cycle. Each PCR cycle consisted of denaturation at 95°C for 30 seconds, annealing at 55°C for 1 minute, and polymerization at 68 °C for 14 minutes. PCR products were incubated with 10 U *Dpn*I at 37 °C for 1 hour to digest the parental (nonmutated) supercoiled ds DNA. The *Dpn*I-digested PCR products were transformed into *E. coli* DH5α competent cells. Mutant clones were initially screened by *Nco*I digestion and confirmed by direct sequencing.

RESULTS

5'- and 3'-terminal sequences

As both 5' and 3' termini contain the promoters for viral RNA replication, the accuracy of these terminal sequences is the pre-requirement for the generation of highly infectious cDNA clones. Therefore, we decided to re-examine the terminal sequences of both RCNMV RNA-1 and -2. By using a combination of ligation-anchored PCR and anchored cDNA cloning techniques (Weng and Xiong, 1995), the 5'- and 3'-termini were cloned into pBluscript SK (+). A total of 8 clones of RNA-1 5' end, 7 clones of RNA-2 5' end, and 4 clones of RNA-2 3' end were sequenced. Sequences of the 5' ends of both RNA-1 and RNA-2 and the 3' end of RNA-2 determined from these clones were identical to the terminal sequences previously determined (Xiong and Lommel, 1989, 1991). Because the three terminal sequences published earlier were proved to be accurate and the fact that the 3' ends of RNA-1 and RNA-2 are highly conserved, we decided not to repeat the cloning and sequencing of the 3' end of RNA-1, but used the sequence determined previously directly for the construction of RNA-1 full-length cDNA clones.

Infectious full-length cDNA clones of RCNMV RNA-1 and RNA-2

A large number of nearly full-length cDNA clones of RCNMV RNA-1 and RNA-2 were generated by RT-PCR and cloned into pBluscript SK (+) at the *Sma*I site. PCR primers were designed in such a way that transcripts identical to viral RNA would be produced directly by T7 RNA polymerase from *Sma*I linearized clones. The 5' PCR primers contained a fully functional promoter for T7 RNA polymerase immediately

upstream of the RCNMV 5' termini. The 3' PCR primers contained a *Sma*I restriction site immediately downstream of the viral 3' termini. In order to produce infectious transcripts, a clone has to retain both the complete T7 RNA polymerase promoter and the intact 5' and 3' terminal sequences of viral RNA genome. Thus initial screens of the nearly full-length clones focused on the retention of the *Sma*I restriction site immediately downstream of RCNMV RNA 3' termini and on the efficient transcription by T7 RNA polymerase. Among 72 RNA-2 clones screened, 42 clones retained the *Sma*I site at the 3' end. Eleven of these clones were efficiently transcribed by T7 RNA polymerase. Out of 27 near full-length clones of RNA-1, 17 maintained the *Sma*I site at the 3' end and 5 clones can be efficiently transcribed *in vitro* by the T7 RNA polymerase.

The infectivity of the full-length RNA-1 and -2 clones were assayed by co-inoculation of the transcripts derived from known infectious cDNA clones. Transcripts derived from five RNA-2 clones were co-inoculated with the transcripts derived from RNA-1 clone pRC1IG, a previously reported infectious clone (Xiong and Lommel, 1991), to 4 cowpea leaves. Clone pRC2IG47 that resulted in the largest number of lesions 5 days post inoculation (dpi) (data not shown) was then used to determine the relative infectivity of the newly synthesized RNA-1 clones.

Infectivity ranging from 1% to 87% was observed when RNA-1 transcripts were co-inoculated with RNA-2 transcripts derived from pRC2IG47 (Table 2.2A). Two clones pRC1IG69 and pRC1IG291 were nearly as infectious as the wild type viral RNA. Clone pRC1IG69 was then used to screen all 11 RNA-2 full-length clones (Table 2.2B). Five RNA-2 clones (pRC2IG54, pRC2IG59, pRC2IG35, pRC2IG52, and pRC2IG47)

produced transcripts that are nearly as infectious as the wild type viral RNA. Clones pRC2IG6, pRC2IG55, and the original infectious RNA-2 clones were intermediate in infectivity. The remaining clones were non-infectious. These results revealed that infectious cDNA clones with infectivity as high as wild type viral RNA-1 and RNA-2 can be obtained by the techniques described in this paper. The original infectious RNA-1 clone pRC1IG (Xiong and Lommel, 1991) was approximately 100-times less infectious than wt RNA-1 and some of the new RNA-1 clones including clone PRC1IG69 which was selected in further experiments.

RNA replication of pRC1IG69 and pRC1IG

Transcript derived from clone pRC1IG69 was about 100-fold more infectious than that from pRC1IG when co-inoculated with RNA-2 in cowpea leaves (Table 2.2B). To test whether the lower infectivity of pRC1IG is due to poor RNA replication or due to inability of the virus to move from cell-to-cell, RNA-1 replication was assayed in *N. clevelandii* protoplasts. Transcripts derived from pRC1IG and pRC1IG69 were transfected with pRC2IG54 transcripts to protoplasts. Total RNAs isolated from transfected protoplasts 24 hours post inoculation (hpi) were hybridized with (-)-sense, ^{32}P-labeled RNA-1 probe. RNA-1 accumulated to a level of about 97-fold higher in protoplasts inoculated with pRC1IG69 transcripts than in protoplasts inoculated with pRC1IG transcripts (Figure 2.2). Consistent with the RNA-1 replication, RNA-2 accumulated to a similar level of 105-fold higher in protoplast when co-inoculated with pRC1IG69 transcripts than that when co-inoculated with pRC1IG transcripts (Figure 2.2).

Table 2.2 Infectivity of full-length cDNA clones of RCNMV RNA-1 and RNA-2.

A.

Clone	Average lesion number	Percentage (%)
pRC1IG69	349	86.80
pRC1IG102	247	64.0
pRC1IG132	70	17.40
pRC1IG155	24	6.00
pRC1IG291	311	77.40
pRC1IG	4	1.00
RCNMV RNA	402	100

B.

Clone	Lesion number	Percent (%)
pRC2IG6	129	36.50
pRC2IG10	0	0
pRC2IG16	0	0
pRC2IG34	0	0
pRC2IG35	291	82.40
pRC2IG47	250	70.80
pRC2IG52	274	77.60
pRC2IG54	304	86.10
pRC2IG55	157	44.50
pRC2IG59	302	85.60
pRC2IG63	0	0
pRC2IG	67	19.00
RCNMV RNA	353	100

These data implied that the low infectivity of the pRC1IG clone is due to a low level of viral RNA replication.

Construction and infectivity analysis of chimerical clones of pRC1IG69 and pRC1IG

To characterize the molecular determinant(s) responsible for the low level of viral RNA replication in pRC1IG, chimerical clones between the highly infectious pRC1IG69 and poorly infectious pRC1IG were generated (Figure 2.3). Two chimerical clones, pRC169XSIG and pRC1IGXS69, contained one half of the RNA-1 genome from each of the two parental clones. pRC169XSIG consisting of 5' half from pRC1IG69 and 3' half from pRC1IG was as infectious as pRC1IG69. In contrast, the reciprocal clone pRC1IGXS69 composed of 5' half from pRC1IG and 3' half from pRC1IG69 was poorly infectious, similar to pRC1IG clone. This result indicated that the 5' half of RNA-1 genome of pRC1IG contained the molecular determinants responsible for the poor viral RNA replication.

To further delimit the region of the viral genome controlling the poor viral replication level, two additional chimerical clones, pRC169ESIG and pRC1IGES69, were created by exchanging a 420 nt *Eco*RI fragment at the 5' end of pRC1IG69 and pRC1IG (Figure 2.3). pRC169ESIG obtained the 5' 420 nt from pRC1IG69 and the remainder from pRC1IG. The reciprocal pRC1IGES69 clone contained 5' 420 nt from pRC1IG and the rest from pRC1IG69. While the infectivity of pRC169ESIG was similarly high as that of pRC1IG69, pRC1IGES69 was poorly infectious as pRC1IG, suggesting that the mutation(s) responsible for the differential replication was located within the 5' 420 nt of pRC1IG.

Sequence comparison of the 5' 420 nucleotides between pRC1IG69 and pRC1IG

The 5' 420 nt of both pRC1IG69 and pRC1IG were sequenced directly in order to identify specific changes in nucleotide sequences that were responsible for the differential level of viral RNA replication. The nucleotide and the predicted amino acid sequence differences within this region between pRC1IG69 and pRC1IG were summarized in Table 2.3. Five nucleotide substitutions and one nucleotide deletion were found. Three of the five nucleotide substitutions occurred in the 5' untranslated region. The remaining two substitutions were located in the p27 and p88 coding region The U to C transitions at nucleotides 318 and 370 resulted in amino acid substitutions from Tyr to His and Leu to Pro, respectively. The single nucleotide deletion took place between nucleotide 145 and nucleotide 149 of the p27 and p88 coding region, reducing five uracil residues in pRC1IG69 to four uracil residues in pRC1IG. This deletion resulted in a translation frameshifting and premature termination of the p27/p88 protein (Figure 2.4). Translation initiation at the original AUG codon of p27/p88 produces a small 13 amino acid peptide instead of full-length p27 and p88.

Characterization of pRC1IG progenies

It had been observed that pRC1IG progenies recovered its infectivity and became as infectious as wild type RNA (Xiong and Lommel, 1991). To identify which of the six mutations in pRC1IG were responsible for its low infectivity, the progenies of pRC1IG were characterized. The 5' 830 nt fragment of progeny RNA-1 was amplified by RT-PCR

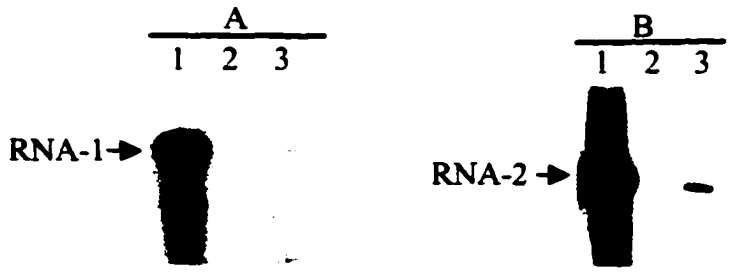


Figure 2.2 Replication and accumulation of RNA-1 and RNA-2 in *N. clevelandii* protoplasts. Protoplasts isolated from *N. clevelandii* plants were coinoculated with either the transcripts derived from pRC1IG69 and pRC2IG54 (lane 1) or the transcripts from pRC1IG and pRC2IG54 (lane 2) or the transcripts from pRC2IG54 only (lane 3). Twenty-four hours after transfection, total RNA was extracted from the protoplasts, electrophoresed in 1.5% agarose gel, bloted to a N+-hybond membranes, and hybridized with RNA probe complementary either to RCNMV RNA-1 (panel A) or to RCNMV RNA-2 (panel B). RNA-1 (panel A) as well as RNA-2 (panel B) replicated and accumulated to a level approximately 100 times higher in protoplasts transfected with transcripts from pRC1IG69 and pRC2IG54 than in protoplasts inoculated with transcripts form pRC1IG and pRC2IG54.

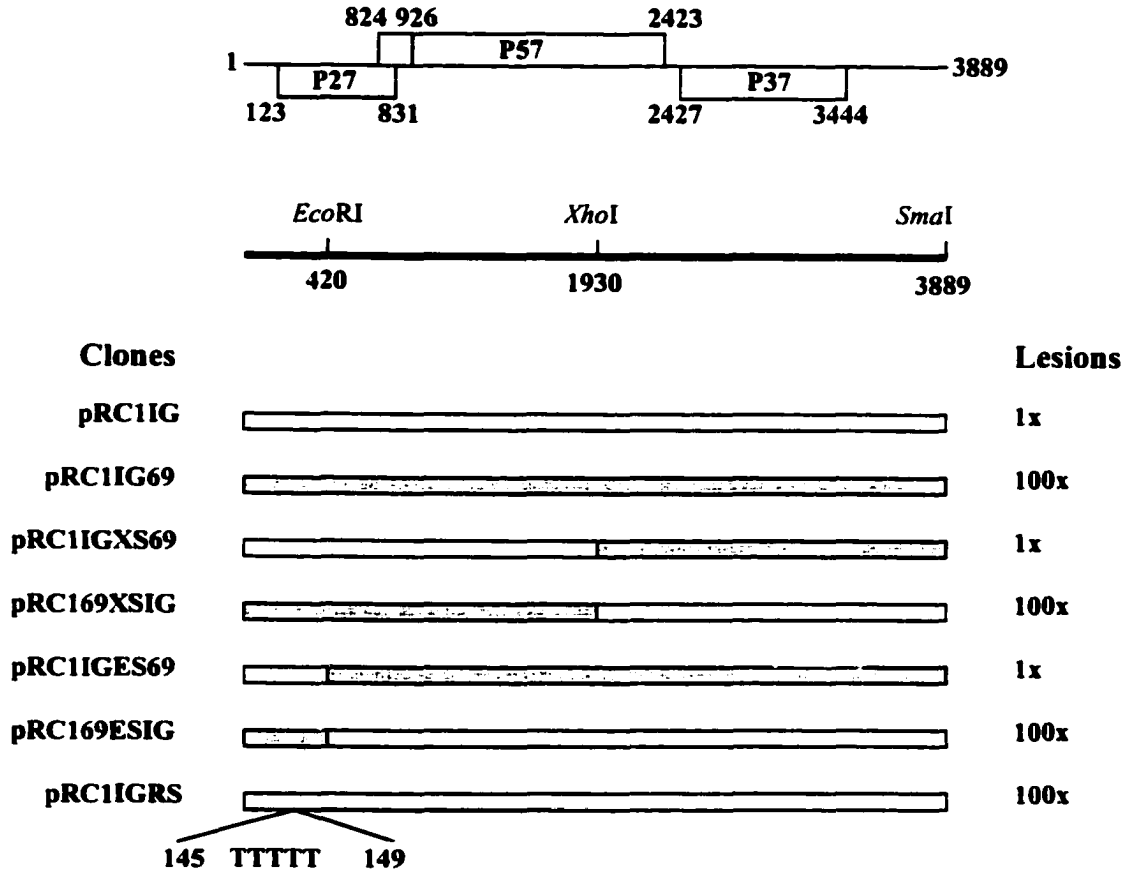


Figure 2.3 Constructs and infectivity of RNA-1 chimerical clones between pRC1IG and pRC1IG69 and restoration clone pRC1IGRS. The infectivity of transcript derived from RNA-1 clone pRC1IG69 is about 100x higher than that of pRC1IG. Chimerical clones pRC1IGXS69 and pRC1IGES69 which contain the 5' 1930 and 420 nt, respectively, from pRC1IG is as poorly infectious as the parental clone pRC1IG. Chimerical clones pRC169XSIG and pRC169ESIG which have the 5' 1930 and 420 nt, respectively, from pRC1IG69 is as highly infectious as the parental clone pRC1IG69. The restoration clone of pRC1IG, pRC1IGRS in which one T was added to the 145-TTTT region, has infectivity similar to pRC1IG69.

Table 2.3 Comparison of 5' 420 nucleotides and the corresponding amino acid sequences between pRC1IG69 and pRC1IG

Nucleotide position	pRC1IG69	pRC1IG	Amino Acid
47	A	U	-
53	U	C	-
78	G	A	-
145	UUUUU	UUUU	Frameshifting
318	U	C	Tyr to His
370	U	C	Leu to Pro

using one pair of primers, RC1C830 and RC1T75'END, and cloned into pBluscript SK (+). Direct sequencing of the 5' 420 nt from several independent clones revealed a single U insertion which reverted the 4 U beginning at nucleotide 145 back to 5 U (data not shown). This result indicated that the low infectivity and poor RNA replication of original pRC1IG was likely caused by one U deletion between nucleotides 145 and 149.

Insertion of a single uracil to pRC1IG restored both infectivity and RNA replication

To unequivocally demonstrate that the single uracil deletion around nucleotide 145 of pRC1IG was responsible for its low infectivity and poor RNA replication, one U was inserted to pRC1IG at nucleotide 145 by PCR-based, site-directed mutagenesis. In addition to the uracil insertion, a *Nco*I site (CCATGG) was engineered as a selection marker by changing U to C at nt 121 immediately upstream of p27/p88 AUG start codon. The resulting clone pRC1IGRS was screened initially by *Nco*I digestion. Desirable mutations were then confirmed by direct sequencing.

Infectivity and RNA replication of pRC1IGRS were assayed on cowpea leaves and *N. clevelandii* protoplasts, respectively. When co-inoculated with RNA-2, transcripts derived from pRC1IGRS produced a similar amount of local lesions on cowpea leaves as those from pRC1IG69 (Figure 2.3). Consistent with infectivity data, the RNA-1 accumulation level in the protoplasts transfected with pRC1IGRS is similar to that of pRC1IG69 (data not shown). These results clearly showed that the single U deletion around nucleotide 145 was responsible for the low infectivity and poor RNA replication of pRC1IG.

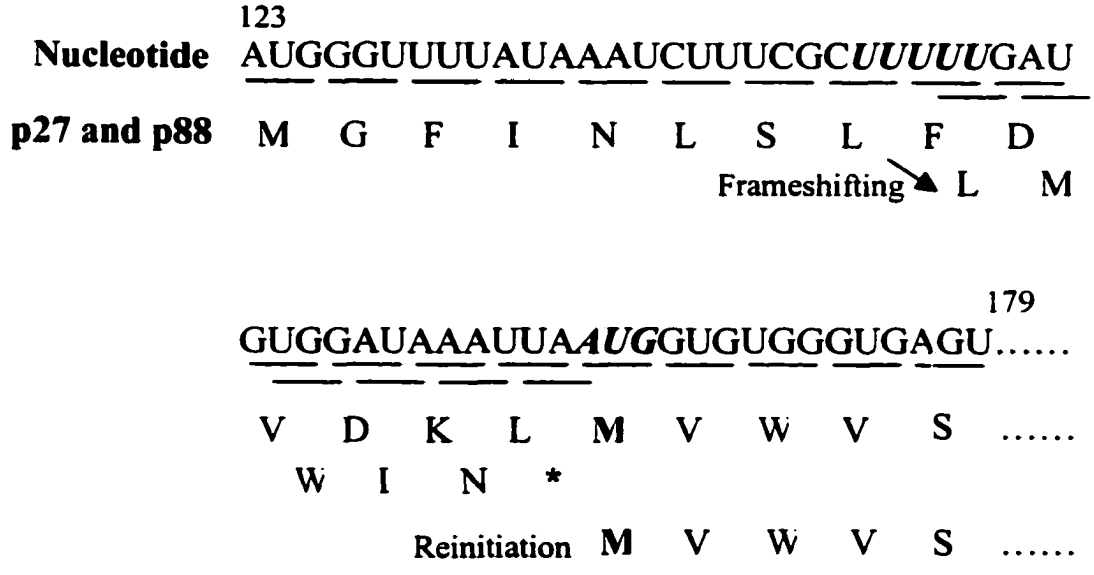


Figure 2.4 Sequence of the N-terminal of p27 and p88 coding region where frameshifting occurs due to a single U deletion. The synthesis of wild type p27 and fusion protein p88 started from the first AUG codon at nucleotide 123. One U deletion at the 145-***UUUUU*** (italic) region caused translation frameshifting, which resulted in the aa sequence changing from the 9th aa and premature termination at the 14th UAA stop codon. The reinitiation of translation from the second ***AUG*** codon (italic) at the 15th aa resulted in the production of p27 and p88 lacking the N-terminal 14 aa.

In vitro translation products of pRC1IG69, pRC1IG, and viral RNA

The one U deletion in pRC1IG should cause translation frameshifting and premature termination which would result in the production of a small (13 aa) peptide instead of full-length p27 and p88. However, closer examination of the nucleotide sequence around the region where a single U was deleted revealed a second inframe AUG methionine codon 15 aa downstream from the original AUG initiation codon (Figure 2.4). This second AUG codon was in a good context for translation initiation and could potentially serve as the AUG initiation codon for the translation of a truncated version of p27 and p88 lacking the N-terminal 14 aa. To test this hypothesis, an equal amount of transcripts derived from pRC1IG69, pRC1IG, and pRC1IGRS were translated in a wheat germ *in vitro* translation system. The ^{35}S -labeled translation products were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970; Xiong *et al.*, 1993a; Weng and Xiong, 1997). As shown in Figure 2.5, pRC1IG transcripts derive the synthesis of p27 and p88 RdRp proteins with sizes similar to wild type proteins synthesized from pRC1IG69 transcripts. The amount of p27 produced from pRC1IG is about 33% of the p27 expressed from pRC1IG69. This result suggested that a significant amount of p27 and p88 lacking the N-terminal 14 aa were produced from the second AUG start codon in pRC1IG. Taken together with the observation that the level of RNA-1 and -2 replication catalyzed by the replicase encoded in pRC1IG transcript was dramatically reduced to about 100 times lower (Figure 2.2), these results indicated that the N-terminal 14 aa of p27 and/or p88 is required for efficient RNA replication.

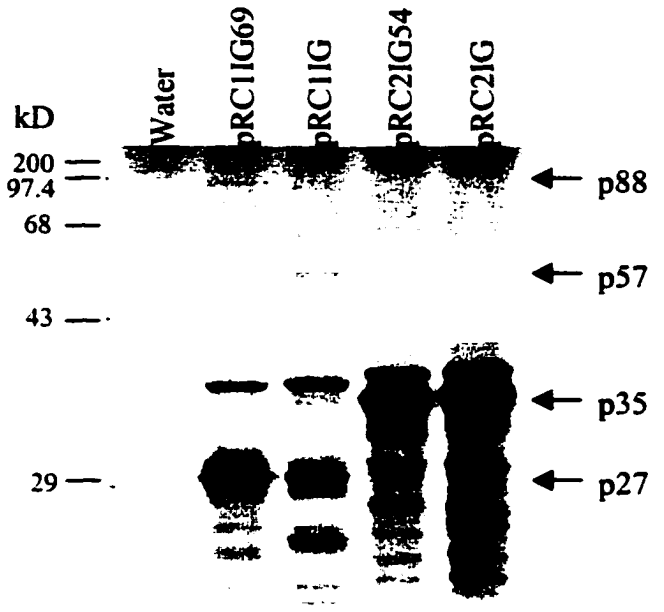


Figure 2.5 *In vitro* translation products of pRC1IG69, pRC1IG, and pRC1IGRS. An equal amount of transcripts derived from RNA-1 clones pRC1IG69 (lane 2), pRC1IG (lane 3), and pRC1IGRS were translated in a wheat germ *in vitro* translation system. The 35S-methionine labeled protein products were electrophoresed in a 12.5% SDS polyacrylamide gel. A similar amount of p27 was synthesized from both pRC1IG69 transcripts and pRC1IG transcripts.

DISCUSSION

Infectious cDNA clones (pRC1IG and pRC2IG) of RCNMV were previously described (Xiong and Lommel, 1991). However, these clones were poorly infectious and capping of the *in vitro* transcripts was required to generate sufficient infection for genetic studies (Xiong *et al.*, 1993a). In this study, we have generated highly infectious cDNA clones of RCNMV RNA-1 (pRC1IG69) and RNA-2 (pRC2IG54) by direct amplification of the entire viral RNA genome with RT-PCR. The infectivity of the transcripts derived from these clones is similar to RCNMV RNAs and is nearly 100-fold more infectious than earlier version of RCNMV infectious cDNA clones. Genetic dissection of the highly infectious pRC1IG69 and low infectious pRC1IG revealed that the N-terminal 14 aa of p88 RdRp and /or p27 protein is required for efficient RCNMV RNA replication.

For any RNA viruses such as RCNMV, highly infectious cDNA clones are essential for genetic dissection and direct manipulation of the viral genome. By using an RT-PCR strategy in which the T7 promoter sequence is directly incorporated into the 5' PCR primer upstream of the viral sequence, a large number of transcribable cDNA clones were rapidly generated without additional sequence manipulation. This strategy makes it possible to screen a large number of full-length clones and consequently identify infectious cDNA clones with infectivity comparable to that of the wild type viral RNA. In this study, twelve RCNMV RNA-2 clones and five RCNMV RNA-1 clones with various levels of infectivity were attained. Several RNA-1 and -2 clones including pRC1IG69 and pRC2IG54 were as infectious as RCNMV RNAs. In contrast, pRC1IG of

RNA-1 and pRC2IG of RNA-2 reported earlier were synthesized by ligation of the overlapping cDNA fragments, followed by insertion of the full-length sequence into a plasmid vector immediately downstream T7 RNA polymerase promoter. The whole process is laborious, including the identification of proper restriction enzyme site within the overlapping region and removal of the non-viral vector sequence between the T7 RNA polymerase promoter and the 5' end of viral sequences by site directed mutagenesis (Xiong and Lommel, 1991). In addition, only one clone was constructed due to the complexity of the process. Therefore, RT-PCR is an easy and fast way to generate highly infectious cDNA clones of RNA viruses, especially those with small RNA genomes.

In order to obtain a highly infectious cDNA clone, a large number of full-length clones must be screened and assayed. The RdRp enzyme that catalyzes the rapid viral RNA genome replication lacks the proof reading function, resulting in a population with sequence variations in individual viral genomes. Some of these variations may abolish their infectivity. The cDNA representing these genomes will be noninfectious. In addition, errors may also be introduced during the cloning procedure including both the reverse transcription and the subsequent cDNA amplification by PCR.

An example of this type of error is a single U deletion within the 5'-420 nt region of poorly infectious clone pRC1IG. The mutation responsible for the low infectivity of pRC1IG is localized to the 5'-420 nt by reciprocal exchanges of the poorly infectious clone pRC1IG and the highly infectious clone pRC1IG69. Sequencing analysis of the 5'-420 nt revealed a total of six nucleotide differences between the two clones (Table 2.3). Among these differences, a single U deletion in the p27 and p88 coding region of

pRC1IG was the ultimate culprit responsible for the low infectivity. This conclusion is supported by two lines of evidence. The reversion of infectivity to the wild type level of the progeny from pRC1IG transcripts is accompanied by one U insertion at nucleotide 145; direct insertion of a single uracil to the same region by site-directed mutagenesis restored the infectivity of pRC1IG to a level similar to that of pRC1IG69. Furthermore, we demonstrated that the low infectivity of pRC1IG is due to a low level of RNA replication. The pRC1IG transcript replicated and accumulated to a level about 100-fold lower than that of pRC1IG69 in protoplasts 24 hours after co-transfected with RNA-2 (Figure 2.2). The difference in viral RNA replication between these two clones is comparable to the difference in their infectivity.

The single U deletion around nucleotide 145 of pRC1IG caused translation frameshifting and premature termination, resulting in the production of a 13-aa peptide instead of p27 and p88. However, the ribosome terminated at 14th UAG stop codon reinitiates translation at the downstream second AUG start codon (at the 15th position) that is in good translation initiation context, and produces the p27 and p88 lacking the N-terminal 14 aa. In eukaryotes, generally only the first AUG from the capped 5' end will be recognized by the ribosome as the translation initiation codon (Kozak, 1999). However, there are some exceptions in which the downstream AUG can also be used as start codon. One of these exceptions is the translation re-initiation in which ribosome restarts the translation from the AUG downstream from the stop codon of the first ORF. Several situations favor the occurrence of re-initiation, including the good context of the AUG (purines at the -3, the third nucleotide position upstream of the AUG, and at +4, the

first nucleotide downstream the AUG) and small size of the first ORF (smaller than 14 aa) (Joshi *et al*, 1997; Kozak, 2001). In this study, the good context around the second AUG for translation reinitiation in pRC1IG is suggested by the G residue (a purine) at the +4 position (Figure 2.4). In addition, the small 13 aa peptide produced from the original AUG in pRC1IG transcript also supports ribosome re-initiation from the downstream second AUG. The production of this truncated p27 and p88 is confirmed by *in vitro* wheat germ translation analysis. A significant amount (33%) of p27 and p88 proteins with size similar to wild type forms were synthesized from pRC1IG transcript compared to those from pRC1IG69 transcript (Figure 2.5). The mobility difference between the wild type p27 and the p27 with 14 aa truncation is not clear in Figure 2.5 due to the low resolution of a small 12.5% polyacrylamide gel. The p88 and p27 lacking the N-terminal 14 aa are apparently capable of catalyzing a low level (1%) of RNA replication *in vivo* (Figure 2.2). While the smaller amount of p27 and p88 produced from pRC1IG contributes about 3-fold reduction of viral RNA replication, the remaining 33-fold reduction is probably caused by the low replication capability of the truncated forms of p27 and p88. This low level of RNA synthesis by the replicase consisting of the truncated forms of p88 and p27 allows the addition of one U to the 145-UUUU region, resulting in the production of wild type forms of p88 and p27 proteins. The production of wild type p88 and p27 restores the wild type level of RNA replication, which in turn results in the restoration of wild type level of infectivity in the progeny. Taken together, these data suggested that the N-terminal 14 aa of p88 and/or p27 is required for efficient RNA replication in RCNMV.

Production of RdRp as a fusion protein is a characteristic feature of *Tombusviridae* family. Tombusvirus RdRp contains a N-terminal pre-fusion part which is encoded in the 5'-proximal ORF with unknown function, and a C-terminal post-fusion part with all RdRp conserved domains and all structure elements as identified in the crystallographic structures of several RdRp (Bressanelli *et al.*, 1999; Butcher *et al.*, 2001; Hansen *et al.*, 1997; Hong *et al.*, 2001; Lesburg *et al.*, 1999; O'Reilly and Kao, 1998). The protein encoded by the 5' proximal ORF of tombusviruses appears as both an independent small protein and as the N-terminus of the RdRp fusion protein. Both of the small protein and the fusion protein are required for viral replication. The fact that the small protein is coded by the very 5' ORF in the viral genome suggests its vital importance. Yet, the functions of this protein and its role in viral RNA replication are virtually unknown. Results from this study clearly indicated that the N-terminal 14 aa region is required for efficient RNA replication.

This result raises the question how the N-terminal 14 aa would affect the function of p27 and p88 and what is the function of p27 involved in viral RNA replication. The amino acid sequence of p27 (236 aa) was analyzed by using different amino acid analysis programs. While no functional domain nor motif was identified, three transmembrane helices and one glycosylation site were predicted. TMpred is a program to predict possible transmembrane domains based on amino acid sequence data of known transmembrane proteins. All three transmembrane helices predicted by TMpred are located at the N-terminus of p27, including one outside-to-inside helices starting from 16 to 38 with score of 278 and two inside-to-outside helices begining from 1 to 18 and from

44 to 67 with scores of 84 and 760 (scores above 500 are considered significant), respectively. The predicted tranmembrane domains in p27 fit well with the observation that purified RCNMV replicase complex containing p27 and p88 are membrane associated (Bates *et al.*, 1995). Deletion of the N-terminal 14 aa may affect the membrane association property of both p27 and p88, leading to dramatic reduction of their efficiency to catalyze viral RNA replication. The membrane association of replicase complex and viral RNA genome replication is a conserved feature in RNA viruses. Brome mosaic virus (BMV) RNA replication occurs on the perinuclear region of the endoplasmic reticulum (ER), both in its natural plant host and in the yeast *Saccharomyces cerevisiae* (Restrepo-Hartwig and Ahlquist, 1999). Viral RNA genome encodes two replication proteins: 1a consisting of a N-terminal capping domain and a C-terminal helicase domain, and 2a containing a central RdRp domain. While protein 1a is able to localize to ER in the absence of other components of replication complex, protein 2a is co-localized to ER via interaction with 1a, and the template RNA is recruited to replication complex at ER by 1a binding to a cis-acting element in the template RNA (Chen and Ahlquist, 2000; den Boon *et al.*, 2001). These data clearly showed that besides its enzymatic functions, 1a is a key component to recruit other components for the assembly of functional replication complex associated with ER membrane. In addition to the transmembrane domains, the Asparagine (Asn or N) at the 5^{th} aa of p27 N-terminus was the predicated glycosylation site. Glycosylation is one of the most frequent and important post-translational modifications. In this process oligosaccharides are covalently attached to Asn in the proteins forming glycoproteins. The glycosylation Asn site must be

in the context of N-X-S/T where X could be either P or D. As carbohydrates are negatively charged and can interact with amino acid side chains, they are important for protein folding. Removing of the glycosylation site in p27 and p88 by deletion of the N-terminal 14 aa may change the folding of p27 and p88, which may in turn alter their RdRp enzymatic activity, leading to low level of RCNMV RNA replication.

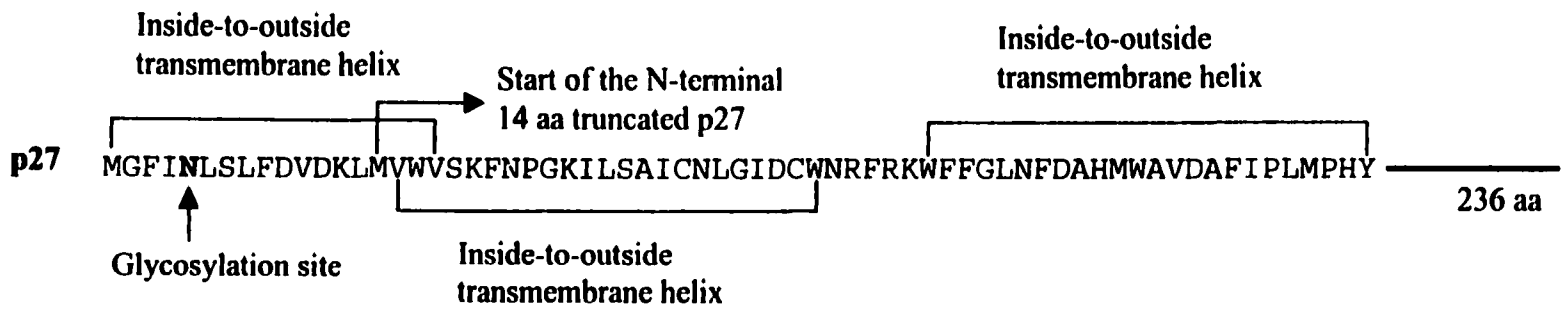


Figure 2.6 Predicted transmembrane helices and glycosylation site at the N-terminus of RCNMV p27 protein. There are two inside-to-outside helices from amino acids of 1 to 18 and from 44 to 67, and one outside-to-inside helix from amino acids of 16 to 38. The 5th N is the predicted glycosylation site. The N-terminal 14 aa truncated p27 eliminates the glycosylation site and one transmembrane helix.

REFERENCES

Ahlquist, P., French, R., and Bujarski, J. J. 1987. Molecular studies of brome mosaic virus using infectious transcripts from cloned cDNA. Adv. Virus Res. 32:215-242.

Bates, H. J., Farjah, M., Osman, T. A., and Buck, K. W. 1995. Isolation and characterization of an RNA-dependent RNA polymerase from Nicotiana clevelandii plants infected with red clover necrotic mosaic virus. J. Gen. Virol. 76:1483-1491.

Boyer, J. C. and Haenni, A. L. 1994. Infectious transcripts and cDNA clones of RNA viruses. Virology 198:415-426.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R., and Rey, F. A. 1999. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Proc Natl Acad Sci U S A. 96:13034-13039.

Buck, K. W. 1996. Comparison of the replication of positive-stranded RNA virus of plants and animals. Adv. Virus Res. 47:159-251.

Butcher, S., Grimes, J., Makeyev, E., Bamford, D., and Stuart, D. 2001. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* 410, 235-240.

Chen, J. and Ahlquist, P. 2000. Brome mosaic virus polymerase-like protein 2a is directed to the endoplasmic reticulum by helicase-like viral protein 1a. J Virol 74:4310-4318.

den Boon, J. A., Chen, J., and Ahlquist, P. 2001. Identification of sequences in Brome mosaic virus replicase protein 1a that mediate association with endoplasmic reticulum membranes. J Virol 75:12370-12381.

Giesman-Cookmeyer, D., Kim, K. H. and Lommel, S. A. 1995. Dianthoviruses. In "pathogenesis and host specificity in plant diseases" (Singh, R. P., Singh, U. P., and Kohmoto, K. Eds.), pp157-176. Elsevier, Oxford.

Hansen, J. L., Long, A. M., and Schultz, S. C. 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure 5:1109-1122.

Hayes, R. J. and Buck, K. W. 1990. Infectious cucumber mosaic virus RNA transcribed *in vitro* from clones obtained from cDNA amplified using the polymerase chain reaction. J. Gen. Virol. 71:2503-2508.

Hong, Z., Cameron, C. E., Walker, M. P., Castro, C., Yao, N., Lau, J. Y., and Zhong, W. 2001. A novel mechanism to ensure terminal initiation by hepatitis C virus NS5B polymerase. *Virology* 285, 6-11.

Joshi, C. P., Zhou, H., Huang, X., and Chiang, V. L. 1997. Context sequences of translation initiation codon in plants. Plant Mol Biol 35:993-1001.

Kim, K. H. and Lommel, S. A. 1994. Identification and analysis of the site of -1 ribosomal frameshifting in red clover necrotic mosaic virus. Virology 200:574-582.

Koonin, E. V. 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J. Gen. Virol. 72:2197-2206.

Koonin, E. V. and Dolja, V. V. 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit. Rev. Biochem. Mol. Biol. 28:375-430.

Kozak, M. 1999. Initiation of translation in prokaryotes and eukaryotes. Gene 234:187-208.

Kozak, M. 2001. Constraints on reinitiation of translation in mammals. Nucleic Acids Research 29:5226-5232.

Laemmli, U. K. 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 227:680-685.

Lai, M. M. C. 2000. The making of infectious viral RNA: no size limit in sight. PNAS 97: 5025-5027.

Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F., and Weber, P. C. 1999. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat Struct Biol. 6:937-943.

Lommel, S. A., Weston-Fina, M., Xiong, Z., and Lomonossoff, G. P. 1988. The nucleotide sequence and gene organization of red clover necrotic mosaic virus RNA-2. Nucleic Acids Res. 16:8587-8602.

Mayo, M.A. and Pringle, C. R. 1998. Virus taxonomy--1997. J Gen Virol. 79:649-657.

Mori, M., Mise, K., Kobayashi, K., Okuno, T., and Furosawa, I. 1991. Infectivity of plasmids containing brome mosaic virus cDNA linked to the cauliflower mosaic virus 35S RNA promoter. J. Gen. Virol. 72:243-246.

Murphy, F. A., C. M. Fauquet, M. A. Mayo, A. W. Jarvis, S. A. Ghabrial, M. D. Summers, G. P. Martelli,and D. H. L. Bishop. 1995. The Classification and Nomenclature of Viruses: Sixth Report of the International Committee on Taxonomy of Viruses. Archives of Virology, Supplementum 3. Springer Verlag, New York.

O'Reilly, E. K. and Kao., C. C. 1998. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. Virology 252:287-303.

Osman, T. A. M. and Buck, K. W. 1987. Replication of red clover necrotic mosaic virus RNA in cowpea protoplasts: RNA-1 replicates independently of RNA-2. J. Gen. Virol. 68:289-296.

Paje-Manalo, L. and Lommel, S. A. 1993. Independent replication of red clover necrotic mosaic virus RNA-1 in electroporated host and non-host Nicotiana sp. protoplasts. Phytopathol. 79:457-461.

Rao A. L. N., Duggal, R., Lahser, F. C., and Hall, T. C. 1994. Analysis of RNA replication in plant viruses. Methods in Molecular Genetics. 4:216-236. Academic Press. Inc.

Restrepo-Hartwig, M. and Ahlquist, P. 1999. Brome mosaic virus RNA replication proteins 1a and 2a colocalize and 1a independently localizes on the yeast endoplasmic reticulum. J Virol 73:10303-10309.

Taniguchi, T., Palmieri, M., and Weissmann, C. 1978. Qβ DNA-containing hybrid plasmids giving rise to Qβ phage formation in the bacterial host. Nature 274:2293-2298.

Vaewhongs, A. A. and Lommel, S. A. 1995. Virion formation is required for the long-distance movement of red clover necrotic mosaic virus in movement protein transgenic plants. Virology 212:607-613.

Weng, Z. and Xiong, Z. 1995. A method for accurate determination of terminal sequences of viral genomic RNA. Genome Research 5:202-207.

Weng, Z and Xiong, Z. 1997. Genome organization and gene expression of saguaro cactus carmovirus. J. General Virol. 78:525-534.

Weng, Z and Xiong, Z. 1998. The role of the 3' terminal stem-loop structure in RNA replication of red clover necrotic mosaic virus. Phytopathology 88:S96. Publication no. P-1998-0697-AMA.

Xiong, Z. and Lommel, S. A. 1989. The complete nucleotide sequence and genome organization of red clover necrotic mosaic virus RNA-1. Virology 171:543-554.

Xiong, Z. and Lommel, S. A. 1991. Red clover necrotic mosaic virus infectious transcripts synthesized *in vitro*. Virology 182:388-392.

Xiong, Z., Kim, K. H., Giesman-Cookmeyer, D., and Lommel, S. A. 1993a. The roles of the red clover necrotic mosaic virus capsid and cell-to-cell movement proteins in systemic infection. Virology 192:27-32.

Xiong, Z., Kim, K. H., Kendall, T. L., and Lommel, S. A. 1993b. Synthesis of the putative red clover necrotic mosaic virus RNA polymerase by ribosomal frameshifting *in vitro*. Virology 193:213-221.

CHAPTER 3.

Identification of Three Nucleotides in the 3'-terminal Stem-loop Required for RNA Replication of *Red Clover Necrotic Mosaic Virus*

ABSTRACT

Red clover necrotic mosaic dianthovirus (RCNMV) contains two positive-sense, single-stranded genomic RNAs. RNA-1 (3.9 kb) encodes p27, p88 fusion protein, and p37 coat protein. p27 and p88 putative RNA-dependent RNA polymerase (RdRp) were the only two viral encoded proteins required for viral RNA replication. RNA-2 (1.45 kb) encodes for a single p35 movement protein and is dispensable for replication. The 3' 29 nucleotides of RNA-1 and RNA-2 are nearly identical and are predicted to form a stable stem-loop structure (SLS). The SLS presumably acts as the promoter for (-) RNA replication and interacts with the viral replicase. Mutations were introduced into the stem and loop regions of RCNMV RNA-2 3' SLS to study their roles in viral RNA replication. Infectivity and replication of these stem and loop mutants were assayed on cowpea leaves and in *Nicotiana clevelandii* protoplasts, respectively. While mutants that disrupted the base pairing in the stem abolished both the infectivity and RNA-2 replication, compensatory mutations that reconstituted the base pairing restored both the infectivity and RNA-2 replication, indicating that the secondary structure of the 3' SLS is essential for RNA replication. However, the primary sequence in the stem region is not important. Multiple nucleotide substitutions within the 5-nucleotide loop region eliminated the

template activity of RNA-1, suggesting that a specific nucleotide sequence in the loop is required for interaction with the viral replicase and for RNA replication. A systematic, single nucleotide substitution analysis of the loop region identified three discontinuous nucleotides (1U, 2A, and 4A) within the 5 nucleotide loop UAUAA that are critical for the promoter activity of the RNA-2. The discontinuous pattern of the critical nucleotides suggests that the 3' SLS of RCNMV is involved in RNA-protein interactions. Substitutions of the nucleotides at these critical positions with other nucleotides allowed the prediction of the base-specific functional groups involved in template-replicase interaction.

INTRODUCTION

The replication of positive-strand RNA virus is initiated by synthesis of minus-strand RNA from the 3'-terminus of the positive strand. The 3' ends of positive strand genomic RNAs are generally highly structured and contain sequences highly conserved within same virus (for reviews see Buck, 1996 and Duggal *et al.*, 1994). These structure and sequence elements interact with viral replicases and serve as promoter elements for the initiation of (-)-strand RNA synthesis. RNA virus replicase is composed of virus-encoded subunits including RNA-dependent RNA polymerase (RdRp) and host-encoded subunits. Either viral subunits or host subunits recognize the viral promoter element and initiate viral RNA replication. The 3' terminal structures serving as the viral promoter elements can be as complicated as the tRNA-like structure (TLS) of *Brome mosaic virus* (BMV) (Dreher and Hall, 1988) or as simple as a single stem loop structure (SLS) of

Turnip crinkle virus (TCV) (Song and Simon, 1995) and *Red clover necrotic mosaic virus* (RCNMV) (Turner and Buck, 1999).

The 3' TLS of BMV consists of several stem-loops and a 4-nt single strand tail (ACCA3'). It is essential and sufficient to direct (-)-strand RNA synthesis (Dreher and Hall, 1988; Felden *et al.*, 1994). Even though several stem loops of the complicated BMV TLS contributed to the initiation efficiency of (-)-strand RNA synthesis *in vivo* (Dreher and Hall, 1988), only one of them, the stem loop C, was required for interaction with BMV replicase (Chapman and Kao, 1999). An artificial minimum promoter capable of directing complementary RNA synthesis *in vitro* could be constructed by joining a single-strand tail (ACCA3') to the stem loop C (Chapman and Kao, 1999). The two C residues of the 3' terminal tail (ACCA) were required for initiation of RNA synthesis but were not involved in the replicase binding. Both the structure and two loop sequences are essential for interaction with BMV replicase. While all 4 nucleotides in the bulge loop contributed to recognition by the replicase complex, only nucleotide A1 in the hairpin loop was most critical for template-replicase interaction.

The fact that a minimum BMV promoter consisted of a simple stem-loop fits well with the structures of promoters of many other RNA viruses including members of the *Tombusviridae* family. These viruses have a simple SLS at the 3' end instead of the elaborated TLS (Buck, 1996; Mayo and Pringle, 1998). The roles of these simple SLS as (-) RNA promoters have been demonstrated in a number of viruses such as *turnip crinkle virus* (TCV) and *Cymbidium ringspot virus* (CyRSV). The promoter for the (-) RNA synthesis of TCV satellite RNA (sat-RNA) C is made of a single, stable SLS consisting

of the 3'-terminal 29 nt (Song and Simon, 1995). The simple 3' SLS was sufficient to direct the minus strand synthesis by purified RdRp *in vitro*. Mutational analysis indicated that only the secondary structure, but not the specific nucleotide sequences in either the stem or the loop, was required to form an active promoter for (-) RNA synthesis (Song and Simon, 1995; Stupina and Simon, 1997; Carpenter and Simon, 1998). Similar conclusions were made with the (-) RNA promoter of CyRSV defective-interfering (DI) RNA (Havelda and Burgyan, 1995). The 3' 77 nt of CyRSV DI RNA was predicted to form a structure composed of three stem-loop structures and two non-basepaired region. Maintenance of the stem base-pairing, but not the stem sequence, was required for RNA replication. These studies suggested that specific nucleotide sequences are not required in the (-) RNA promoters of *Tombusviridae*. It is not clear how viral replicases interact with the viral promoters for (-) RNA synthesis under these circumstances.

A more recent study of RCNMV, another member of *Tombusviridae*, showed that both the secondary structure and the specific loop sequence of the 3' SLS were required for the (-) RNA promoter activity (Turner and Buck, 1999). Substitution of three nucleotides in the loop abolished viral RNA replication. RCNMV is a well characterized member of *Dianthovirus* genus, *Tombusviridae* family. Its genome consists of two positive-sense, single-stranded genomic RNAs. RNA-1 (3.9 kb) encoded three polypeptides of p27, p88 RdRp, and p37 capsid protein from 5' to 3' (Xiong and Lommel, 1989; Xiong *et al.*, 1993a). RNA-2 (1.45 kb) encoded a single p35 movement protein (Lommel *et al.*, 1988). Although all four proteins were necessary for wild type (wt) infection (Xiong *et al.*, 1993b; Vaewhongs and Lommel, 1995; Zavriev *et al.*, 1996; Sit *et*

al., 1998), only p27 and p88 RNA polymerase were required for RCNMV RNA replication (Kim and Lommel, 1994; Bates, *et al.*, 1995). RNA-1 was able to replicate in cowpea and tobacco protoplasts in the absence of RNA-2 (Osman and Buck, 1987; Paje-Manolo and Lommel, 1989). The bipartite nature of RCNMV genome made it a very attractive model for the study of *cis*-acting elements required for RNA replication. Since mutations introduced into RNA-2 do not affect the production of the viral RdRp by viral RNA-1, changes in the accumulation of RNA-2 directly reflect the effect of the mutations in the *cis*-acting elements on viral RNA replication.

There is little sequence homology between the two viral genomic RNAs with the exception of 5 nucleotides at the 5' end and 29 nucleotides at the 3' end. The 3' 29 nucleotides of RNA-1 and RNA-2 are nearly identical and were predicted to form a stable SLS (Lommel *et al.*, 1988; Xiong and Lommel, 1989). Similar 3' SLS with identical loop sequences (UAUAA) was predicted in two other *Dianthoviruses*, *Sweet clover necrotic mosaic virus* (SCNMV) (Ge *et al.*, 1992) and *Carnation ringspot virus* (CRSV) (Kendall and Lommel, 1992), implying that the 3' SLS may play important roles in the RNA replication of *Dianthoviruses*.

In this study, we constructed a series of mutations in the 3'-SLS of RNA-2 including a stem disruption mutant, a stem compensatory mutant, and single nucleotide loop mutants and examined their effects on RNA-2 replication in *N. clevelandii* protoplast and viral infectivity on cowpea plants. The results confirmed and extended a previous report that both of the stem-loop structure and the loop sequence were important for RNA replication (Turner and Buck, 1999). In addition, by systematically substituting each of

the 5 nucleotides in the loop region with other three nucleotides, we identified three essential nucleotides required for RNA-2 replication. These nucleotides could potentially serve as contact sites between the template-replicase interactions. Based on effects of different nucleotide substitutions on viral RNA replication, functional groups in these nucleotides that are potentially involved in template-replicase recognition and interaction are predicted.

MATERIALS AND METHODS

Construction of RNA-2 mutant clones

A highly infectious full-length cDNA clone of RCNMV RNA-2, pRC2IG54 (Chapter 2), was used as template for mutagenesis. A series of RNA-2 stem and loop mutant clones were generated by PCR-based mutagenesis using paired oligonucleotide primers (Table 3.1). Primer vRC2T75'END (Table 3.1) that contains 18 nucleotides identical to the 5' end of RNA-2 was used as the upstream primer for all mutagenesis. The downstream mutagenic primers contain nucleotides complementary to the 3' terminus of RNA-2 with the exception of designated nucleotide substitutions. In addition, seven nucleotides (GATCCCC) consisting of four random nucleotides GATC and nucleotides CCC that together with the 3' viral sequence created a *Sma*I restriction site CCCGGG were added to the 5' end of all the mutagenic primers (Table 3.1). These primers included the stem disruption primer (cRC2SD), the stem restoration primer (cRC2SR), two multiple loop sequence substitution primers (cRC2LEcoRI and cRC2LKpnI), and fifteen

Table 3.1 Primers used to construct RCNMV RNA-2 3'-SLS mutants.

Primers	Orientation	Primer sequences (5' to 3')
vRC2T75'END*	Viral	GAATTGTAATACGACTCACTATAG**ACAAACCTCGCTCTATAA**
wt RNA-23'END**	Complementary	GGGGTGCCTAGCCGTTATACGACTAGGCA
cRC2SD***	Complementary	GATC*CCCGGG*GTGCagctgCGTTATACG
cRC2SR	Complementary	GATC*CCCGGG*GTGCagctgCGTTATACGcagctGCATTATA
cRC2LEcor I	Complementary	GATC*CCCGGG*GTGCCTAGCCGaatTcCGACTAGGCA
cRC2LKpn I	Complementary	GATC*CCCGGG*GTGCCTAGCCGgTAccCGACTAGGCA
cRC2L1A	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTATtCGACTAGGCA
cRC2L1C	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTATgCGACTAGGCA
cRC2L1G	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTATcCGACTAGGCA
cRC2L2C	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTAgACGACTAGGCA
cRC2L2G	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTAcACGACTAGGCA
cRC2L2U	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTAaACGACTAGGCA
cRC2L3A	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTtTACGACTAGGCA
cRC2L3C	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTgTACGACTAGGCA
cRC2L3G	Complementary	GATC*CCCGGG*GTGCCTAGCCGTTcTACGACTAGGCA
cRC2L4C	Complementary	GATC*CCCGGG*GTGCCTAGCCGTgATACGACTAGGCA
cRC2L4G	Complementary	GATC*CCCGGG*GTGCCTAGCCGTcATACGACTAGGCA
cRC2L4U	Complementary	GATC*CCCGGG*GTGCCTAGCCGTaATACGACTAGGCA
cRC2L5C	Complementary	GATC*CCCGGG*GTGCCTAGCCGgTATACGACTAGGCA
cRC2L5G	Complementary	GATC*CCCGGG*GTGCCTAGCCGcTATACGACTAGGCA
cRC2L5U	Complementary	GATC*CCCGGG*GTGCCTAGCCGaTATACGACTAGGCA

* Bold letters indicate the 5' 18 nt of RNA-1. ** Underlined letters indicate the base-paired stem region, five nucleotide (TTATA) between the base-paired region is the 5-nt loop region. *** Lower case letters indicate the altered nucleotide(s); italicized letters represent *SmaI* restriction site.

single loop substitution primers. cRC2LEcoRI and cRC2LKpnI changed the wt loop sequence UAUAA to GAAUU and GGUAC and at the same time generated *Eco*RI and *Kpn*I restriction sites in the cDNA clones, respectively, to facilitate mutant screening. The single loop substitution primers were named according to the position of the target nucleotide in the loop and the resulting substitution. For example, primer cRC2L1A substituted the first nucleotide in the loop, U, to an A residue. For the PCR reaction, 0.2 μm each of the upstream primer and the down stream primer and 0.01 μg of template DNA (pRC2IG54) were added to a standard PCR reaction solution. After 30 cycles of denaturation at 94°C for 45 sec., annealing at 54°C for 60 sec., and extension at 72°C for 90 sec., PCR amplified DNA fragments were purified and then digested with restriction enzymes *Sma*I and *Nco*I. Restriction site *Sma*I was engineered immediately downstream from the 3' end of the viral sequence during mutagenesis and restriction site *Nco*I is endogenous and located at nucleotide 1230. The resulting 219 bp *Sma*I-*Nco*I fragment with the designated mutations in the 3'-SLS was used to replace the corresponding fragment in pRC2IG54 clone to generate mutant RNA-2 full-length infectious clones. Mutations in all mutant clones were confirmed by restrict digestion if applicable and direct sequencing prior to any experiments.

In vitro transcription

The infectious cDNA clones, pRC1IG69 of RNA-1 and pRC2IG54 of RNA-2, and the RNA-2 stem and loop mutant clones were linearized with *Sma*I followed by *in vitro* transcription with bacteriophage T7 RNA polymerase (Invitrogen, Carlsbad, CA) as

described previously (Xiong and Lommel, 1991). Each 10-µl reaction contains 1 µg template DNA, 40 mM Tris-HCl pH 8.0, 8 mM MgCl2, 2 mM spermidine-(HCl)3, 25 mM NaCl, 5 mM dithiothreitol (DTT), 1 mM each of ATP, CTP, GTP and UTP, 8 U of T7 RNA polymerase, and 1 U of Ribonuclease inhibitor (Invitrogen, Carlsbad, CA).

Plant inoculations and infectivity analysis

Equal amount of RNA-1 and RNA-2 transcripts (approximately 5-10 µg RNA per10-µl reaction) were mixed and diluted at 1:5 ratio with GKP buffer (50 mM glycine, 30 mM K2HPO4 pH 9.2, 1% bentonite, 1% celite), and then inoculated onto a total of six primary leaves on three cowpea plants as previously described (Xiong and Lommel, 1991). The inoculated plants were kept at 25°C. Infectivity of each clone was determined by the number of lesions on the inoculated leaves counted 4 days after inoculation.

Protoplasts isolation and transfection

Protoplasts were isolated from leaves of systemic host plant *N. clevelandii* using a protocol modified from Rao *et al.* (1994). Briefly, young leaves of *N. clevelandii* were surface-sterilized, sliced into 1-mm strips, and incubated in enzyme solution containing 1.25% cellulase "Onozuka" R-10 (Yakult Honsha Co., Ltd.), 0.15% macerozyme R-10 (Yakult Honsha Co., Ltd.), 0.1% bovine serum albumin (BSA, fraction V, Sigma, St. Louis, Mo), and 0.55 M mannitol (pH 5.9) for sixteen hours at room temperature in dark. After filtration and sucrose centrifugation, the protoplasts were collected from the

interface between mannitol and sucrose, washed, and divided into aliquots of 1.25 X 10^6 cells in 15 ml polypropylene tubes.

Prior to protoplast transfection, DNA template was removed from the *in vitro* transcripts by lithium chloride precipitation. After phenol-chloroform extraction and ethanol precipitation, transcripts from a 100-µl reaction were dissolved in solution containing 0.1 M NaCl, 10 mM Tris-HCl (pH 7.5), and 1 mM EDTA. Lithium chloride was added to a final concentration of 2.7 M and the mixture was incubated on ice for sixteen hours. After centrifugation, the RNA pellet was resuspended in sterile distilled water at a final concentration of 1µg RNA/µl.

Ten µg each of purified, DNA-free RNA-1 and RNA-2 transcripts were co-transfected into the purified protoplasts using the polyethylene glycol (PEG: Mr 1450, Sigma, St. Louis, Mo) method as previously described (Rao *et al.* 1994). The transfected protoplasts were incubated at room temperature under a fluorescent lamp.

Total RNA extraction and Northern blot analysis

Total RNA was extracted from protoplasts 24 hours post transfection (Rao *et al.* 1994). The protoplasts were pellet and extracted with 250 µl each of RNA extraction buffer (100 mM glycine, 10 mM EDTA, 100 Mm NaCl pH 9.5, 2% SDS, 2.5 mg/ml bentonite) and phenol-chloroform. Following ethanol precipitation, RNA pellet was dissolved in 20 µl of sterile distilled water. After heat denaturation at 70 °C for 10 min., 5 µl of total RNAs were fractionated by electrophoresis in 1.5% agarose gels, transferred to the nylon membrane, and cross-linked by UV. The blots were then hybridized with a ^{32}P-

labeled, (-)-sense riboprobe specific for either RCNMV RNA-1 or RNA-2. Labeling of the (-) RCNMV transcripts with α-^{32}P-UTP was carried out as previously described (Xiong *et al.*, 1993b).

Cloning and sequencing analysis of the progeny viral RNA from mutant clones

To determine the sequences of progeny RNA-2 from stem restoration mutant, progeny viruses were purified from infected *Nicotiana clevelandii* plants 7 days after inoculation, followed by viral RNA extraction. RNA-2 was reverse transcribed into cDNA by reverse transcriptase (Invitrogen, Carlsbad, CA) using cRC23'END primer (GAGCTCCCGGGGTGCCTAGCCGTTATAC) containing 20 nucleotide (underlined) complementary to the 3' end of wild type RNA-2. The cDNA was then amplified by 30 cycle's PCR using *Taq* DNA polymerase (Promega, Madison, WI) and primers cRC23'END and cRC2T75'END. The 1.45 kb PCR products were cloned into pBluscripts SK(+) at *Ecor*V site and sequenced to determine the viral 3'-terminal sequence using T7 or T3 primers based on the insertion orientation.

To determine the sequences of progeny RNA-2 from the loop substitution mutants, total RNAs were extracted from inoculated cowpea leaves 4 days after inoculation using TRIZOL™ solution (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. The viral RNA-2 in the total RNA was then reverse transcribed into cDNA using primer cRC2S3END (5'TGCG*CTGCAG*GGGTGCCTAGCC3') that contains 13 nucleotides (underlined) complementary to the 3' terminus of wild type RNA-2, a *Pst*I site (*italic*), and 4 protecting nucleotides. A 1.5 kb DNA fragment corresponding to the entire

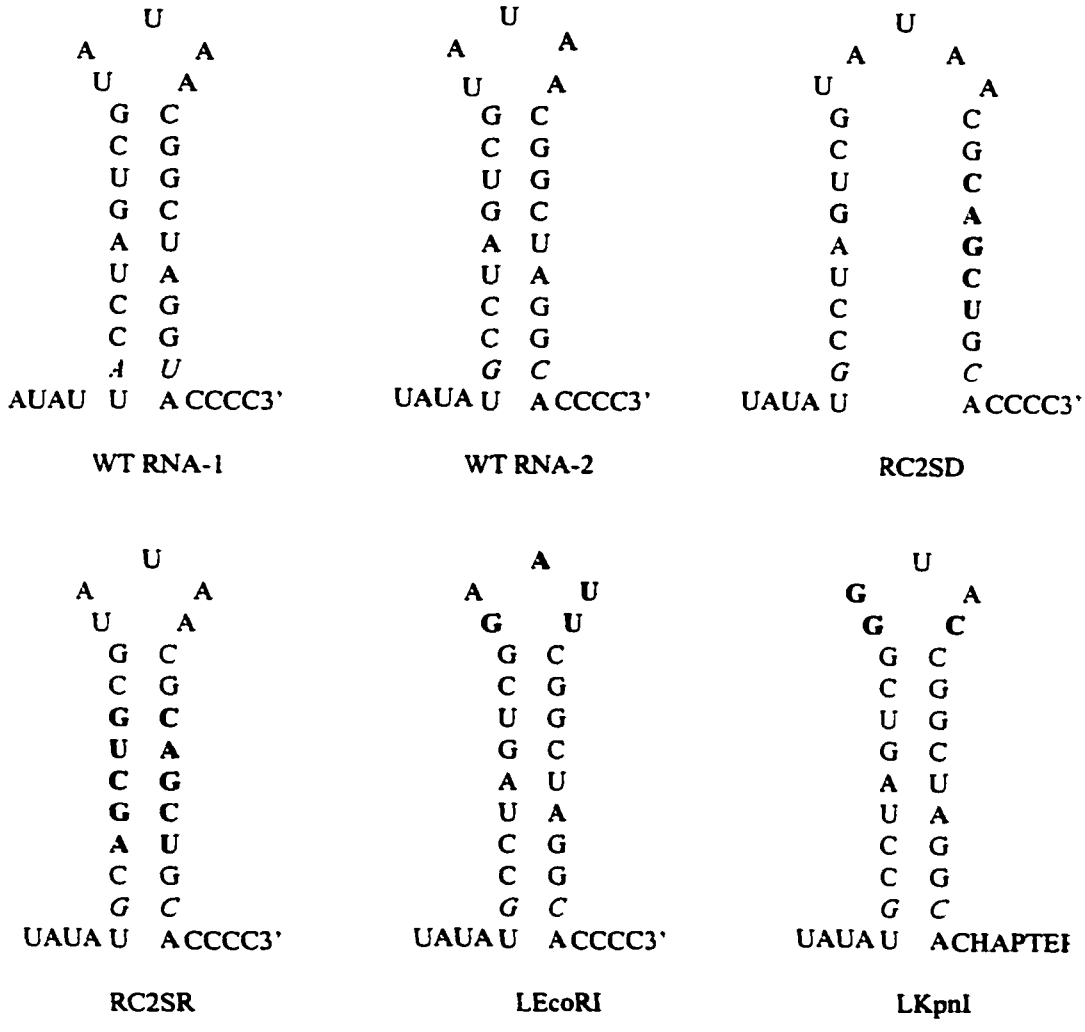
RCNMV RNA-2 was amplified by 30-cycle PCR using primers cRC2S3END and vRC2T75'END. A smaller DNA fragment containing the 3' end of RCNMV RNA-2 was then isolated by restriction digestions at a *Pst*I site in the primer cRC2S3END and at a *Xba*I site 419 nt from the 3' end of RNA-2. The *Pst*I-*Xba*I fragment containing the 3' end was then cloned into pBluscripts SK(+) vector linearized with the same enzymes and sequenced using T7 primer.

RESULTS

3'-terminal SLS is required for the replication of RCNMV RNA-2 and plant infection

RCNMV genome is split into two single-stranded, (+)-sense RNAs. Proteins necessary for replication, p27 and p88 RdRp were encoded in RNA-1. RNA-2 that encodes a single movement protein is dispensable for viral RNA replication. The 3'-terminal 29 nucleotides of RCNMV RNA-1 and RNA-2 were nearly identical and were predicted to form identical SLS except one basepair at the bottom of the stem (A-U in RNA-1 and G-C in RNA-2) as shown in Figure 3.1. To isolate the effect of the 3' SLS on viral RNA replication, mutations were introduced only to RCNMV RNA-2 while RNA-1 encoding for the essential replication proteins were not disturbed.

To determine if the 3' SLS is required for viral RNA replication and infectivity, an initial mutant (RC2SD) with the 3' SLS completely abolished was made. Five nucleotides GCUAG in the 3' side of the stem were replaced with nucleotides CAGCU, resulting in the disruption of stem base-pairing of the 5 nucleotide and consequently the



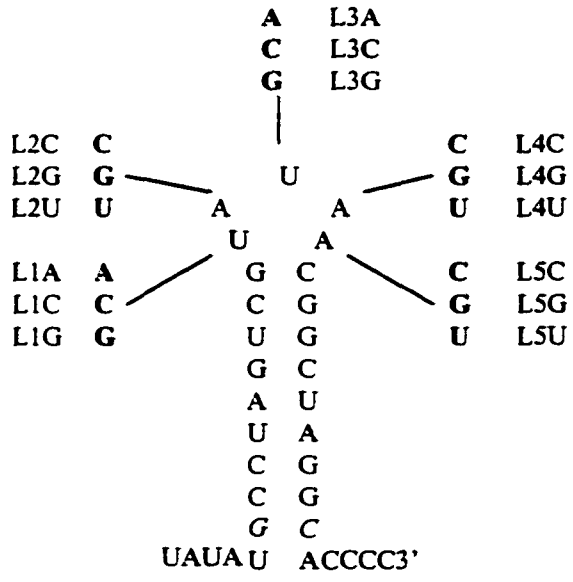


Figure 3.1 Sequence and predicted stem-loop structure of 3' terminal nucleotide of wild type and mutant RNA-2 of RCNMV. The single basepair difference between RNA-1 and RNA-2 is located near the bottom of the stem and *italicized*. Substituted nucleotides in mutants are shown in **bold**.

3' SLS. Infectivity and replication of RNA-2 mutant RC2SD were examined on intact plants and in protoplasts, respectively. Cowpea primary leaves co-inoculated with RC2SD transcripts and wt RNA-1 transcripts did not develop any local lesion five days after inoculation while similar cowpea leaves inoculated with wt RCNMV RNA-1 and RNA-2 transcripts developed several hundred lesions (Table 3.2). No lesions were observed on the mutant-inoculated leaves over the next several weeks, indicating no reversion of the mutant to a replication-competent form.

Either a failure in viral RNA replication or an inability of the mutant to move from cell-to-cell could contribute to the lack of infectivity of the mutant. Replication of the mutant was thus examined in *N. clevelandii* protoplasts by co-transfection of RC2SD transcripts with wt RNA-1 transcripts. Total RNA was extracted from protoplasts twenty-four hours after transfection. Accumulation of RNA-2 in the transfected protoplasts was then determined by Northern hybridization with ^{32}P-labeled probes specific for (+) RNA-2. In protoplasts transfected with wt RNA-1 and RNA-2 transcripts, RNA-2 accumulated to a high level. In contrast, no RNA-2 was detected in the protoplasts inoculated with wt RNA-1 and RC2SD transcripts (Figure 3.2). The undetectable level of mutant RNA-2 can be caused either directly by the inability of RNA-2 to replicate or indirectly by inability of RNA-1 to replicate. RNA-1 encoded two proteins required for viral RNA replication. If RNA-1 failed to replicate and to provide essential proteins for replication, RNA-2 would not be replicated. To rule out this possibility, accumulation of RNA-1 in the transfected protoplasts was assayed by hybridization with ^{32}P-labeled probes specific for (+) RNA-1. Although RNA-2 was not detectable in protoplasts co-transfected with

Table 3.2 Infectivity of RNA-2 stem and loop mutants determined by local lesion numbers produced on inoculated cowpea leaves.

Clones	Average lesion number (%)	Leaves #	Groups
WT RNA-2*	100	36	I
RC2SD	0	6	III
RC2SR	105.5	6	I
RC2LEcoRI	0.1	8	III
RC2LKpnI	0.1	8	III
RC2L1A	0.4	8	III
RC2L1C	38.3	10	II
RC2L1G	0	8	III
RC2L2C	87.1	8	I
RC2L2G	0.4	14	III
RC2L2U	1.6	12	III
RC2L3A	89.1	14	I
RC2L3C	93.7	8	I
RC2L3G	108.6	8	I
RC2L4C	4.1	12	III
RC2L4G	0.9	12	III
RC2L4U	3.8	12	III
RC2L5C	98	8	I
RC2L5G	86.7	8	I
RC2L5U	108.7	8	I

* The average number of local lesions produced by wt RNA-1 and RNA-2 transcripts ranged from 196 to 643, depending on environmental conditions. Results were compiled from eight independent inoculations.

RC2SD and wt RNA-1, RNA-1 accumulated to the wild type level (Figure 3.2; Figure 3.3). These results indicated that the 3'-SLS was required for RNA replication.

The primary sequence in the stem of 3'-SLS is not required for RNA replication and plant infection

Analysis of mutant RC2SD clearly illustrated that the 3' SLS was essential for viral RNA replication. However, complete disruption of the stem in RC2SD would not distinguish possibilities that either specific nucleotides in the stem, RNA secondary structure, or a combination of both was required for viral RNA replication. A stem restoration mutant, RC2SR, was engineered to address this question. RC2SR was made by compensatory mutations of 5 pairs of nucleotides in both sides of the stem. The resulting mutant was able to maintain the 3' SLS structure yet the primary sequences of the stem have been altered (Figure 3.1). The infectivity of RC2SR transcripts was tested on the cowpea leaves. The number of local lesion produced by RC2SR on cowpea leaves was similar to that by wt RNA-2 when co-inoculated with wt RNA-1 (Table 3.2). Coinoculation of RC2SR and wt RNA-1 transcripts resulted in wt systemic infection in both *N. clevelandii* and *N. benthamiana*, indicating that mutations in RC2SR did not alter the viral biological properties (data not shown). Consistent with the *in planta* infectivity, the mutant RNA-2 replicated to a level close to that of the wt RNA-2 in the protoplasts co-transfected with wt RNA-1 (Figure 3.2). To examine the possibility that the lesions on cowpea leaves were resulted from the reversion of the mutant RC2SR RNA-2 back to wild type sequence, RNA-2 progeny of RC2SR as well as wild type RNA-2 was

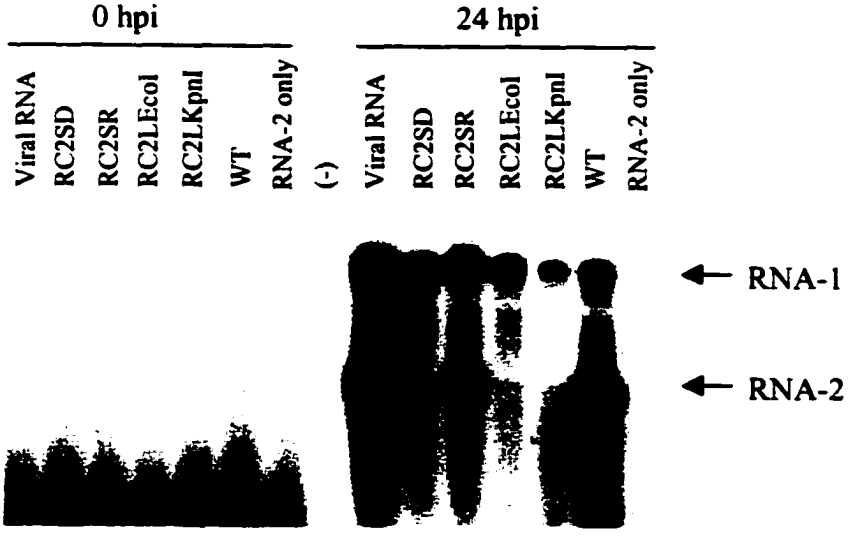


Figure 3.2 Replication and accumulation of stem mutants and multiple substitution loop mutants in *N. clevelandii* protoplasts. Total RNA was extracted from protoplasts at 0 and 24 hours after transfection with wild type RNA-1 and mutant RNA-2 (labeled on the top of each lane), electrophoresed on 1.5% agarose gel, transferred to a nylon membrane, and hybridized with (-)-sense, RNA-2 specific riboprobe.

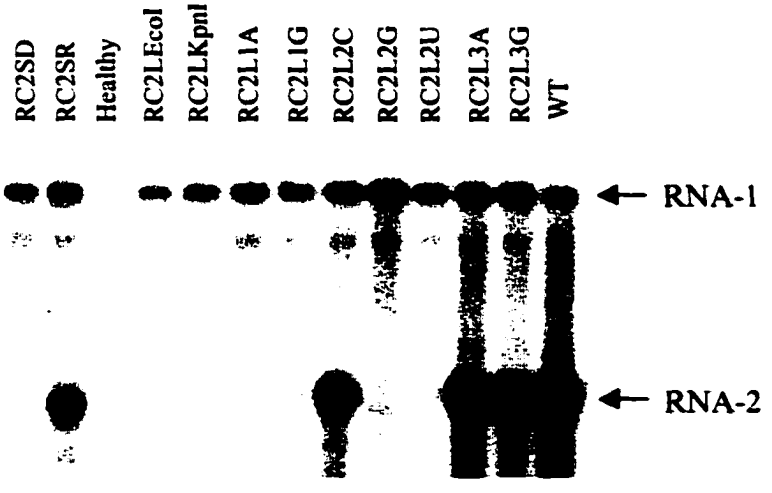


Figure 3.3 Replication and accumulation of RNA-1 and RNA-2 in *N. clevelandii* protoplasts transfected with RNA-1 and mutant RNA-2 transcripts. Total RNA was extracted from protoplasts 24 hours after transfection with wild type RNA-1 and mutant RNA-2 (labeled on the top of each lane), electrophoresed on 1.5% agarose formamide gel, transferred to a nylon membrane, and hybridized with (-)-sense, RNA-2 specific ribprobe.

amplified by RT-PCR, cloned and sequenced. The sequence data showed that the mutations in RC2SR were maintained in its progeny (data not shown). These results suggested that the altered primary sequence of stem was not important for either viral RNA replication or plant infection as along as the 3' SLS was maintained.

Loop sequence is important for RNA replication and plant infection

The five nucleotide sequence (UAUAA) in the predicted 3' SLS loop region was conserved among RNA-1 and RNA-2 of three dianthoviruses: RCNMV, SCNMV and CRSV, indicating that this loop sequence may play an important role in replication. To test this hypothesis, two independent loop mutants with multiple nucleotide substitutions, RC2LEcoRI and RC2LKpnI, were constructed (Figure 3.1). RC2LEcoRI changed four loop nucleotides (**GAAUU**) (altered nucleotides were bold). RC2LKpnI substituted three loop nucleotides (**GGUAC**). When transcripts from either mutant were coinoculated with wt RNA-1 transcripts to cowpea leaves, no local lesions were observed five days after inoculation (Table 3.2). The inability of the mutants RC2LEcoRI or RC2LKpnI RNA-2 to infect cowpea leaves was caused by a failure of the mutant transcripts to replicate as indicated by lack of detectable level of RNA-2 in protoplasts cotransfected with wt RNA-1 transcripts (Figure 3.2; Figure 3.3). These data clearly suggested that the five-nucleotide loop sequence contained *cis*-elements essential for viral RNA replication.

Three residues in the 5-nt loop region is required for RNA replication and plant infection

To further study the role of five individual nucleotide residues in the loop region in viral RNA replication, fifteen single-nucleotide substitution mutants were constructed (Figure 3.1). Each nucleotide of the five-nucleotide UAUAA loop was replaced with the other three nucleotides, resulting in three individual mutants at a single nucleotide position. For example, the nucleotide U at the first position of the loop was replaced with A, C, and G in three independent mutants: RC2L1A, RC2L1C, and RC2L1G.

The *in planta* infectivity on cowpea leaves for each of the single nucleotide substitution mutants was determined by co-inoculation of wild type RNA-1 and mutant RNA-2 transcripts. The average number of local lesions induced by each mutant was recorded and normalized to that induced by wt RNA-1 and RNA-2 as a measure of infectivity (Table 3.2). All fifteen mutants were cataloged into three groups on the basis of lesion numbers. Group I consisted of RC2L2C, all three mutants at loop position 3 (RC2L3A, RC2L3C, RC2L3G), and all three mutants at loop position 5 (RC2L5C, RC2L5G and RC2L5U). The relative lesion numbers produced by this group varied from 91% to 110% and were not significantly different from those produced by wt RNA-1 and RNA-2 transcripts. In addition, these RNA-2 mutants produced wild type systemic infection in two systemic host plants, *N. clevelandii* and *N. benthamiana* (data not shown). Sequence analysis of the progeny RNA confirmed that all mutations were maintained and that no compensatory mutation was detected (Table 3.3). Group II contains only RC2L1C, which produced a lower number of local lesions (37.95% of lesions produced by the wild type viral transcripts). As with the progenies of group I

Table 3.3 Mutation maintaining or reversion in the progeny of mutant RNA-2

Clones	Original mutation	Second site mutation
RC2L1C-1	Maintained	No
RC2L1C-2	Maintained	No
RC2L2C-6	Maintained	No
RC2L2C-7	Maintained	No
RC2L2U-9	Mutated to C	No
RC2L3A-11	Maintained	No
RC2L3C-16	Maintained	No
RC2L3G-22	Maintained	No
RC2L4C-1	Revert to wt	No
RC2L4C-26	Revert to wt	No
RC2L4C-27	Revert to wt	No
RC2L4G-4	Maintained	U to C at L1 and A to G at L2
RC2L4U-1	Maintained	U to A at L1
RC2L4U-31	Maintained	U to A at L1
RC2L4U-32	Revert to wt	No
RC2L5C-37	Maintained	No
RC2L5G-42	Maintained	No
RC2L5U-46	Maintained	No

mutants, sequencing of the progeny viral RNA indicated that the original mutation was maintained and no other sequence changes were observed (Table 3.3). The remaining mutants: RC2L1A, RC2L1G, RC2L2G, RC2L2U, RC2L4C, RC2L4G, and RC2L4U made up group III. These mutants produced less than 5% of the lesions produced by the wt viral RNA on cowpea leaves. Furthermore, progeny from these mutants contained either reversions or second-site mutations in addition to the engineered mutation (Table 3.3). Three independent clones of the RC2L4C progeny contained a reversion of the originally mutated C residue back to the wild type A residue. An interesting reversion occurred in the progeny of RC2L2U. Instead of reverting to the wt A residue, the engineered U mutation changed to a C residue, resulting in a genotype identical to mutant RC2L2C that behaved like the wild type RNA-2 in both infectivity and replication assay. Among three independent clones of the RC2L4U progeny, two reverted to wild type sequence and one maintained the mutated U residue at L4 but contained a second site mutation of U to A at L1. One clone of the RC2L4G progeny contained the original mutation as well as two second-site mutations of U to C at L1 and A to G at L2. Progeny viruses containing reversions or second site mutations exhibited a high level of viral RNA replication and infectivity similar to those of wt RNA-1 and RNA-2 transcripts.

The effects of the single nucleotide substitution loop mutants on RNA replication were examined in protoplasts by Northern hybridization. Total RNA was extracted from protoplasts transfected with mutant RNA-2 transcripts and wt RNA-1 transcripts and hybridized with probes specific for (+) RNA-2. As shown in Figure 3.4, all L3 mutants (RC2L3A, RC2L3C and RC2L3G) and L5 mutants (RC2L5C, RC2L5G and RC2L5U)

supported the wild type level of RNA-2 replication, in consistent with the infectivity data. These results indicated that the nucleotide identity at positions of L3 and L5 were not important for RNA-2 replication and could be substituted by any nucleotides. In contrast, RNA-2 was not detected in protoplasts transfected with all L4 mutants (RC2L4C, RC2L4G and RC2L4U), suggesting that L4 was critical and the specific nucleotide A was required for RNA-2 replication (Figure 3.4). At L1 position, substitution of A (RC2L1A) or G (RC2L1G) residues in place of the wt U residue also eliminated RNA-2 replication, whereas C for U substitution (RC2L1C) only reduced RNA-2 replication to approximately 30% of the wild type level (Figure 3.4, 3.3). This result corresponded well with the 37.9% infectivity displayed by the mutant on cowpea leaves. Among L2 mutants, substitution of G (RC2L2G) or U (RC2L2U) for the wt A residue abolished RNA-2 replication while substitution of C residue (RC2L2C) allowed RNA-2 to replicate to a level similar to wild type (Figure 3.4, 3.3).

Data obtained from the replication assay in protoplast transfection experiments fit very well with the infectivity results determined by local lesion numbers. These results indicated that in the 5-nt loop region, the identities of nucleotides at L3 and L5 were not important; however, specific nucleotides were required at L1, L2, and L4 positions. It was apparent that the A residue at L4 was absolutely required for the 3' SLS promoter activity while partial and wt promoter activities could be restored by substation of C for U at L1 and C for A residue at L2, respectively. Any other substations at these two loop positions were lethal.

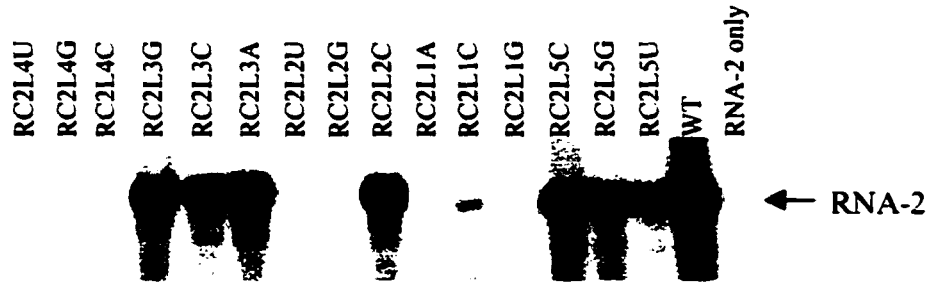


Figure 3.4 Replication and accumulation of RNA-2 mutant with single loop nucleotide-substitution in the 3' SLS in *N. clevelandii* protoplast. Total RNA extracted from protoplasts 24 hours after transfection with wild type RNA-1 and mutant RNA-2 were hybridized with (-)-sense, RNA-2 specific riboprobe.

DISCUSSION

We have constructed a series of mutants in the 3'-terminal SLS and provided experimental evidence to show that both the secondary structure and the loop sequence are required as *cis*-promoter elements for RCNMV RNA-2 replication. In addition, exhaustive substitutions of each nucleotide in the 5-nt loop region with the other nucleotides led to the identification of three critical residues required for RNA replication and potentially for template-replicase interaction.

Requirement of the 3'-SLS for viral RNA replication

The involvement of the predicted 3'-terminal SLS as *cis*-element in the (-) RNA replication of RCNMV was indicated by the inability of stem disruption mutant (RC2SD) to replicate in *N. clevelandii* protoplasts and to produce any lessons on cowpea leaves (Figures 3.1, 3.2, 3.3; Table 3.2). RC2SD had 5 nucleotides in the 3' side of the stem changed to non-basepairing nucleotides, resulting in complete disruption of the 3' SLS. In contrast, the stem restoration mutant (RC2SR) that contained the stem base-pairing consisting of different primary nucleotide sequences in the stem region was able to replicate in protoplasts and produced local lesions on cowpea leaves at a level comparable to wt RNA (Figures 3.1, 3.2, 3.3; Table 3.2). Taken together, analysis of these two mutants suggested that the 3' SLS was required for RCNMV (-) strand synthesis and that the primary sequence in the stem region was not. These results were consistent with those observed recently in RCNMV and were comparable to those in TCV and CyRSV. Turner and Buck (1999) analyzed similar stem-disruption and

restoration mutants of RCNMV RNA-2 and obtained similar results. The major difference was that every single nucleotide in the 10-basepair was altered in their mutants. The substituted nucleotides AUUAGCUUAC in their study is different from GCUAG used in this study, yet stem-restoration mutants containing either sequence were fully infectious, further demonstrating that the sequence specificity of the stem is not required for the (-) RNA promoter. The 3'-terminal 29 nt of the TCV associated sat-RNA C, which contains the complete promoter for directing the minus strand synthesis by purified RdRp *in vitro*, forms a stable stem-loop structure (Song and Simon, 1995). Mutation analysis indicated that the secondary structure, but not the primary stem sequence, was essential for RNA replication *in vitro* and *in vivo* (Song and Simon, 1995; Stupina and Simon, 1997; Carpenter and Simon, 1998). Similarly, the 3' 77 nt of CyRSV DI RNA was predicted to form a structure composed of three stem-loop structures and two non-basepaired region. Maintenance of the stem base-pairing, but not the stem sequence, was required for CyRSV RNA replication (Havelda and Burgyan, 1995). Taken together, ours and previously published studies suggest that a simple, highly stable SLS at the 3' terminus is required for the promoter activity in *Tombusviridae*. Interaction between a simple SLS and viral replicase during replication and transcription could be a universal feature of the (+) RNA viruses. CP subgenomic promoter of BMV, a member of *Bromoviridae*, can be folded into a simple SLS with a 3 nt loop sequence AUA (Haasnoot *et al.*, 2002; 2000). The best supporting evidence for this hypothesis is that the minimum (-) RNA promoter of BMV consists of a simple SLS (Chapman and Kao, 1999), despite the fact that the 3' end of BMV RNA can be folded into a complicated

TLS consisting of several stem-loops and a pseudoknot. Deletion and insertion studies indicated that there was a preferred spacing between the initiation site and the stem loop. This observation led to the speculation that the role of the 3' stem loop may be to bind to the replicase and to ensure a free 3' end at the correct distance for the initiation of RNA synthesis (Chapman and Kao, 1999).

Loop sequence

Two multiple loop substitution mutants, RC2LEcoRI and RC2LKpnI that had 4 and 3 loop nucleotides altered respectively, were not able to replicate in protoplasts and to produce any symptom on cowpea leaves (Figure 3.1, 3.2, 3.3; Table 3.2), indicating that specific loop sequence was required for RNA replication. These data was consistent with the observation Turner and Buck (1999) made in RCNMV. Furthermore, single nucleotide substitutions demonstrated that a single nucleotide change of the loop sequence could completely abolish the viral promoter activity. Requirement of specific loop sequence for the promoter activity is apparently not universal in the *Tombusviridae*. In both TCV and CyRSV, two members of *Tombusviridae*, only base-paring of the SLS was essential. Nucleotide substitutions, deletions or insertions of additional nucleotides in the loop of the 3' SLS of TCV sat-RNA C (Song and Simon, 1995) and CyRSV DI RNA (Havelda and Burgyan, 1995) apparently had no effect on viral RNA replication, suggesting that the loop sequence was not required for replication.

Requirement of specific loop sequence for the promoter-replicase interaction in RCNMV is more similar to that in an unrelated BMV. The minimum promoter of the

BMV (-) genomic RNA promoter consisted of a single stem with a bulge loop and a hairpin (Chapman and Kao, 1999). The nucleotide A1 in the hairpin loop (AUA) was critical for template-replicase interaction while all 4 nucleotides in the bulge loop (UAGA) stabilized this interaction.

It is possible that the stem of the 3'-SLS in RCNMV restrain the loop sequence to a certain conformation that is accessible to replicase for contact and binding, similar to the function of stem C in the 3' tRNA-like structure of BMV (Chapman and Kao, 1999). In contrast, the 3'-SLS structure in TCV sat-RNA C and CyRSV DI RNA may simply function to ensure a free 3' end in the proper position for initiation of RNA synthesis like the pseudoknot A in the 3' TLS in BMV (Chapman and Kao, 1999). The difference in the requirement of the loop sequence in the RNA replication of RCNMV, TCV, and CyRSV is somewhat surprising. All three viruses are members of *Tombusviridae* and share a high degree of homology in the virus-encoded RdRp (Koonine, 1991). The difference perhaps can be explained by the fact that subviral RNA (sat-RNA or DI RNA) was used in the TCV and CyRSV studies.

Three critical nucleotides in loop

Our results clearly showed that not all the nucleotides in the loop of the 3' SLS participated equally in the promoter activity. Alteration of the first, second, and fourth nucleotides in the UAUAA loop to other nucleotides abolished or drastically reduced both RNA replication and viral infectivity, indicating that these three loop positions were critical for viral RNA replication and possibly interacting directly with viral replicase. No

specific nucleotide was required at L3 and L5 positions while varying degrees of nucleotide specificity existed for L1, L2, and L4 (Figure 3.4 and Table 3.2). Nucleotide U at L1 is the preferred nucleotide but substitution of nucleotide C in place of U partially restored the promoter activity of the 3' SLS. This nucleotide substitution was maintained faithfully in the progeny of the mutant, indicating that cytosine can functionally replace uracil, albeit at a lower efficiency. Substitutions with adenosine and guanosine at L1 resulted in non-detectable level of viral RNA replication. The adenosine residue at L2 could be substituted with cytosine without noticeable loss of promoter activity of 3' SLS, but substitutions with guanosine and uracil resulted in no detectable level of viral RNA replication in transfected protoplasts. Infectivity assay showed that mutant RC2L2U produced a few lesions, but sequence analysis of the progeny viral RNA revealed that the introduced U residue mutated into a C residue. The resulting progeny behaved exactly like RC2L2C in both infectivity and replication assay (data not shown). An adenosine residue is absolutely required at L4. Substitution of L4A with any other nucleotides eliminated both viral RNA replication and viral infectivity. Some of the mutations such as L2U, L4C, L4G, and L4U did not completely abolish *in vivo* viral RNA replication although accumulation of RNA-2 was not detectable in protoplasts assays. This is supported by a few local lesions (1 to 4% of the wild type level) induced by these mutants on cowpea leaves. However, the progenies of these mutants either reverted back to wild type sequence or incorporated second site mutations. These data indicated that these substitutions may allow a very low level of viral RNA replication, which provides

an opportunity for the mutant viral RNA to mutate to a more efficient promoter, consequently producing a limited number of local lesions.

The apparent wild type behavior of RC2L2C and RC2L3G in both infectivity and replication conflicted with an earlier observation by Turner and Buck (1999). In their limited single nucleotide substitution study, both substitutions of A to C at L2 and U to G at L3 eliminated the replication and the infectivity of the mutants. It is not clear why completely opposite results were obtained in two different studies. One possibility is that an unintended mutation was introduced at a site distant from the targeted 3' SLS in their mutants, resulting in the lethal mutants. This unintended mutation could escape inspection unless the entire mutated viral genome was sequenced. Indeed, we observed one stem restoration mutant that failed to replicate and to infect cowpea plants. However, sequencing analysis revealed a single nucleotide substitution of A for G at 1262 in addition to the intended nucleotide alteration (data not shown). This additional change was likely introduced by PCR mutagenesis and was likely the cause of the lethal phenotype. In the construction of our RC2L2C and RC2L3G mutants, the entire *Nco*I-*Sma*I fragment containing the engineered mutation was first verified by sequencing analysis and then replaced the same fragment in the wt RNA-2 clone pRC2IG54. Furthermore, transcripts from several independent clones produced the wt infection when coinoculated with the RNA-1 transcripts.

Predicted functional group recognized by RdRp

The discontinuous pattern of critical nucleotides required for the 3' SLS promoter activities suggested that RCNMV replicase might make direct contact with the three key nucleotides in the loop region. Differential effects of single nucleotide substitution at these positions on viral RNA replication allowed a prediction of the functional groups that might be recognized by RCNMV replicase. Both L1A and L1G substitutions abolished RNA replication, whereas substitution with C (which contains a spatially equivalent C2 keto group as in U) still supported replication at 30% wild type level, indicating that the C2 keto group of U at L1 at least contributed to replicase recognition. Similarly, while L2G and L2U changes eliminated replication, substitution with C (which contains spatially equivalent C6 amine and N1 imine as in A) did not affect RNA replication, suggesting that the C6 amine group and the N1 imine group of A at L2 were possibly the functional groups for replicase binding. Changing L4A to every other nucleotide dramatically decreased RNA replication, indicating that multiple functional groups of A at L4 were involved in RNA-replicase recognition. The base-specific functional groups at discontinuous nucleotide locations required for replicase recognition predicted from this study is similar to the replicase-RNA interaction in the BMV subgenomic core promoter (Siegel *et al.*, 1997, 1998; Adkins and Kao, 1998). The specific functional moieties of several nucleotides interacting with BMV replicase was initially predicted by nucleotide substitutions and were later confirmed experimentally using synthetic template RNA containing nucleotide analogues. In the CP subgenomic promoter of BMV, nucleotides at positions -17, -14, -13, and -11 relative to the

subgenomic initiation site were required for RNA synthesis and direct interaction with viral replicase (Siegel *et al.*, 1997, 1998). Nucleotide substitutions at these critical positions allowed them to predict that C6 keto and/or N1 amine at -17 G, multiple groups at -14 A, C4 amine at -13 C, and both C6 keto and N7 imine at -11G were the base-specific functional groups interacting with BMV RdRp. These predicted functional groups were experimentally shown to be required for interaction with BMV replicase by using chemically synthesized RNAs containing base analogs that were modified in these functional groups (Siegel *et al.*, 1998).

Two other members of the *Dianthovirus* genus, SCNMV and CRSV, have a similar 3'-terminal SLS with the 5-nt loop sequence identical to RCNMV 3'-SLS (Ge *et al.*, 1992; Kendall and Lommel, 1992). The sequence-specific recognition mode identified in RCNMV may be common in the *Dianthovirus* genus.

Potential interactions between nucleotides at three key loop positions

Seven single-nucleotide substitution loop mutants at three key positions produced less than 5% lesions of the wt RNA-2 (Table 3.2). However, four of these mutants (L2U, L4C, L4G, and L4U) produced a sufficient amount of local lesions that permitted the analysis of mutant progenies. All of the progenies analyzed harbored mutations generated in *vivo*. In addition to reversion to wt sequence, second site mutations were observed near the original mutation sites in the progenies of two mutants (L4G and L4U) (Table 3.3). In the progeny of mutant L4G, the first two nucleotides of the mutated loop, UAUGA, were replaced with nucleotide C and G, respectively, resulting in a new loop sequence

CGUGA. In two independent progenies of mutant L4U, a single second-site mutation of A for U at L1 generated another new loop sequence, AAUUA. Progeny viruses containing these second site mutations were as infectious as the wt RCNMV. Apparently both CGUGA and AAUUA were functionally equivalent to the wt UAUAA, indicating that they can be recognized and replicated by RCNMV replicase as efficiently as wt UAUAA. These second site mutations thus should be considered as compensatory mutations as they restored the promoter activity of the 3' SLS. Interestingly, all second-site mutations occurred at three key positions, indicating that the nucleotides at these key loop positions potentially interact with each other. Further studies are required to understand this base-base interaction and to determine the effect of this interaction on the template-replicase recognition.

REFERENCES

Adkins, S. and Kao, C. C. 1998. Subgenomic RNA promoters dictate the mode of recognition by bromoviral RNA-dependent RNA polymerases. Virology. 252:1-8.

Bates, H. J., Farjah, M., Osman, T. A., and Buck, K. W. 1995. Isolation and characterization of an RNA-dependent RNA polymerase from Nicotiana clevelandii plants infected with red clover necrotic mosaic virus. J. Gen. Virol. 76:1483-1491.

Buck, K. W. 1996. Comparison of the replication of positive-stranded RNA virus of plants and animals. Adv. Virus Res. 47:159-251.

Carpenter, C. D. and Simon, A. E. 1998. Analysis of sequences and predicted structures required for viral satellite RNA accumulation by *in vivo* genetic selection. Nucleic Acids Res. 26:2426-2432.

Chapman, M. R. and Kao, C. C. 1999. A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. J Mol Biol. 286:709-720.

Dreher, T. W., and Hall, T. C. 1988. Mutational analysis of the sequence and structural requirements in brome mosaic virus RNA for minus strand promoter activity. J. Mol. Biol. 201:31-40.

Duggal, R., Lahser, F., C., and Hall, T. C. 1994. cis-Acting sequences in th replication of plant viruses with plus-sense RNA genomes. Annu. Rev. Phytopathol. 32:287-309.

Felden, B., Florentz, C., Giege, R., and Westhof, E. 1994. Solution structure of the 3'-end of brome mosaic virus genomic RNAs. Conformational mimicry with canonical tRNAs. J Mol Biol. 235:508-531.

Ge Z., Hiruki C., and Roy K. L. 1992. A comparative study of the RNA-2 nucleotide sequences of two sweet clover necrotic mosaic virus strains. J. Gen. Virol. 73:2483-2486.

Haasnoot, P. C., Brederode, F. T., Olsthoorn, R. C., and Bol, J. F. 2000. A conserved hairpin structure in Alfamovirus and Bromovirus subgenomic promoters is required for efficient RNA synthesis in vitro. RNA 6:708-716.

Haasnoot, P. C., Olsthoorn, R. C., and Bol, J. F. 2002. The Brome mosaic virus subgenomic promoter hairpin is structurally similar to the iron-responsive element and functionally equivalent to the minus-strand core promoter stem-loop C. RNA 8:110-122.

Havelda, Z. and Burgyan, J. 1995. 3' Terminal putative stem-loop structure required for the accumulation of cymbidium ringspot viral RNA. Virology 214:269-272.

Kendall, T. L. and Lommel S. A. 1992. Nucleotide sequence of carnation ringspot dianthovirus RNA-2. J Gen Virol. 73:2479-2482.

Kim, K. H. and Lommel, S. A. 1994. Identification and analysis of the site of -1 ribosomal frameshifting in red clover necrotic mosaic virus. Virology 200:574-582.

Koonin, E. V. 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J. Gen. Virol. 72:2197-2206.

Lommel, S. A., Weston-Fina, M., Xiong, Z., and Lomonossoff, G. P. 1988. The nucleotide sequence and gene organization of red clover necrotic mosaic virus RNA-2. Nucleic Acids Res. 16:8587-8602.

Mayo, M. A. and Pringle, C. R. 1998. Virus taxonomy--1997. J Gen Virol. 79:649-657.

Osman, T. A. M. and Buck, K. W. 1987. Replication of red clover necrotic mosaic virus RNA in cowpea protoplasts: RNA-1 replicates independently of RNA-2. J. Gen. Virol. 68:289-296.

Paje-Manalo, L. and Lommel, S. A. 1993. Idependent replication of red clover necrotic mosaic virus RNA-1 in electroporated host and non-host Nicotiana sp. protoplasts. Phytopathol. 79:457-461.

Rao A. L. N., Duggal, R., Lahser, F. C., and Hall, T. C. 1994. Analysis of RNA replication in plant viruses. Methods in Molecular Genetics. 4:216-236. Academic Press. Inc.

Siegel, R. W., Adkins, S., and Kao, C. C. 1997. Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. Proc Natl Acad Sci U S A. 94:11238-11243.

Siegel, R. W., Bellon, L., Beigelman, L., and Kao, C. C. 1998. Moieties in an RNA promoter specifically recognized by a viral RNA-dependent RNA polymerase. Proc Natl Acad Sci U S A. 95:11613-11618.

Sit, T. L., Vaewhongs, A. A., and Lommel, S. A. 1998. RNA-mediated trans-activation of transcription from a viral RNA. Science 281:829-832.

Song, C. and Simon, A. E. 1995. Requirement of a 3'-terminal stem-loop in *in vitro* transcription by an RNA-dependent RNA polymerase. J. Mol. Biol. 254:6-14.

Stupina, V. and Simon, A. E. 1997. Analysis *in vivo* of turnip crinkle virus satellite RNA C variants with mutations in the 3'-terminal minus-strand promoter. Virology 238:470-477.

Turner, R. L. and Buck, K. W. 1999. Mutational analysis of cis-acting sequences in the 3'- and 5'-untranslated regions of RNA-2 of red clover necrotic mosaic virus. Virology 252:115-124.

Vaewhongs, A. A. and Lommel, S. A. 1995. Virion formation is required for the long-distance movement of red clover necrotic mosaic virus in movement protein transgenic plants. Virology 212:607-613.

Xiong, Z. and Lommel, S. A. 1989. The complete nucleotide sequence and genome organization of red clover necrotic mosaic virus RNA-1. Virology 171:543-554.

Xiong, Z. and Lommel, S. A. 1991. Red clover necrotic mosaic virus infectious transcripts synthesized *in vitro*. Virology 182:388-392.

Xiong, Z., Kim, K. H., Giesman-Cookmeyer, D., and Lommel, S. A. 1993b. The roles of the red clover necrotic mosaic virus capsid and cell-to-cell movement proteins in systemic infection. Virology 192:27-32.

Xiong, Z., Kim, K. H., Kendall, T. L., and Lommel, S. A. 1993a. Synthesis of the putative red clover necrotic mosaic virus RNA polymerase by ribosomal frameshifting *in vitro*. Virology 193:213-221.

Zavriev, S. K., Hickey, C. M., and Lommel, S. A. 1996. Mapping of the red clover necrotic mosaic virus subgenomic RNA. Virology 216:407-410.

CHAPTER 4.

Transgenic Recombination of *Red Clover Necrotic Mosaic Virus*

ABSTRACT

Systemic infection by *Red clover necrotic mosaic virus* (RCNMV) consists of two distinct events: cell-to-cell spread through plasmodesmata and long distance transport through vasculature. The cell-to-cell movement is mediated by the movement protein (MP) encoded on RNA-2 while the long distance transport is facilitated by capsid protein encoded on RNA-1. Although RNA-1 has all genes essential for viral RNA replication and can replicate in protoplasts, systemic infection of intact plants by RCNMV requires RNA-2. The requirement of RNA-2 encoded MP for systemic infection was exploited to characterize recombination between transgenically expressed viral RNA sequences and infecting viral RNA. Wild type RCNMV RNA-2 or truncated versions lacking either the first 55 nucleotides in the 5'- untranslated region (UTR) (Δ5), the last 219 nucleotides in the 3' UTR (Δ3), or both (Δ53) were transformed into *Nicotiana benthamiana* plants, under the control of an enhanced 35S promoter. The resulting MP transgenic plants were neither resistant to wild-type RCNMV infection nor capable of complementing RCNMV RNA-1 infection. However, incidence of RCNMV systemic infection was observed 7 days or longer in transgenic plants expressing the Δ5 transcripts inoculated with RCNMV RNA-1. Virions recovered from the systemically infected leaves were infectious on non-transformed *N. benthamiana* and *N. clevelandii*, producing wild type infections. In

contrast, no systemic infection developed in the RNA-1 inoculated transgenic plants expressing either the Δ3 or Δ53 transcripts. Analysis of virion RNA recovered from RNA-1 inoculated Δ5 transgenic plants indicated that RCNMV RNA-1 has recombined with and captured RNA-2 transgene mRNA by providing an essential 5'-terminal replication element. Several distinct species of recombinant RNA-2 molecules were cloned and sequenced. The 5' termini of all recombinant RNAs contained from 8 to 15 nucleotides identical to the 5' terminus of RNA-1. However, sequences at the crossover junctions differ from clone to clone. The recombinant RNAs can be divided into three groups according to recombinant sequences: Group I contains no extraneous nucleotides between the RNA-1 5' terminus and RNA-2 sequences starting from various positions within the 5' UTR of RNA-2; group II contains additional 3 to 29 nucleotides of unknown origins between the RNA-1 and -2 sequences; group III has the 5'-terminus of RNA-1, sequences of unknown origins, and the sequences from the transgene transcripts composed of nonviral vector sequences followed by RNA-2 sequences. These data indicated that recombination between RCNMV RNA-1 and RNA-2 transgenic transcripts resulted in recombinant RNA-2 molecules that could be replicated by RCNMV RNA-1 and provided MP function essential for systemic infection. A viral polymerase-mediated template-switching model was proposed to explain the transgenic recombination in RCNMV.

INTRODUCTION

Recombination in RNA viruses occurs naturally and has been a driving force in RNA virus evolution (see reviews by Alejska *et al.*, 2001; Domingo *et al.*, 1996; Garcia-Arenal *et al.*, 2001; Nagy and Simon 1997). Under experimental conditions, recombination among RNA viruses has been well studied in different genera including *Bromovirus* (see reviews by Figlerowicz and Bujarski, 1998), *Alfamovirus* (Van Der Kuyl, *et al.*, 1991), *Cucumovirus* (Aaziz and Tepfer, 1999a), *Nepovirus* and *Tobravirus* (Le Gall *et al.*, 1995), *Carmovirus* (see reviews by Simon and Bujarski, 1994; Nagy and Simon, 1997), and *Tombusvirus* (White and Morris, 1995). Mechanistically, RNA recombination can be divided into two categories: homologous recombination that occurs between two related RNA molecules and non-homologous recombination that occurs between two unrelated RNA molecules.

Recombination takes place not only between viral RNA molecules, but also between viral RNA and cellular mRNA, leading to acquisition of cellular genes by RNA viruses. RNA viruses are capable of recombining with mRNA, ribosomal RNA, or tRNA as indicated by presence of an ubiquitin-coding sequence in bovine diarrhea virus genome (Meyers *et al.*, 1991); 28S ribosomal RNA sequence in the genome of a human influenza virus (Khatchikian *et al.*, 1989); and tRNA sequences in defective interfering RNA of Sindbis virus (Monroe and Schlesinger, 1983) and in satellite RNAs of Qβ bacteriophage (Munishkin *et al.*, 1988). For plant RNA viruses, a sequence homologous to an exon of tobacco chloroplast DNA was found at the 5' terminus of potato leafroll virus (Mayo and Jolly, 1991).

RNA recombination can also occur between viral RNA and viral transgene mRNA expressed in virus-resistant transgenic plants (For review see Aaziz and Tepfer, 1999b; Rubio *et al.*, 1999). Transgenic plants expressing segments of viral genome have been demonstrated to confer resistance to the cognate as well as closely related viruses (For reviews see Baulcombe, 1994; Beachy, 1997; Kaniewski and Thomas, 1999; Miller and Hemenway, 1998; Ravelonandro *et al.*, 2000; Rovere *et al.*, 2002). In the absence of virus control strategies, transgenic resistance appears to be a viable tool for viral disease management. However, the transgenic plants resist only the cognate or closely related viruses, and unrelated or less related viral RNAs can still replicate in the transgenic plants and can potentially recombine with the viral transgene transcripts. This scenario has been confirmed in several studies. Recombination between capsid protein (CP) deletion mutants and CP transgene transcripts were reported in *Cowpea chlorotic mottle virus* (CCMV) (Greene and Allison, 1994), *Tomato bushy stunt virus* (TBSV) (Borja, *et al.* 1999), *Tobacco mosaic virus* (TMV) (Adair and Kearney, 2000), and *Plum pox virus* (PPV) (Varrelmann *et al.*, 2000). Recombination between a viral transgene and its cognate virus has also been documented in *Cauliflower mosaic virus* (CaMV). Although CaMV is a pararetrovirus, the transgenic recombination was reported to involve the transgene transcript and the RNA replication intermediate of the pararetrovirus (Schoelz and Wintermantel, 1993; Wintermantel and Schoelz, 1996).

The transgenic recombinations reported to date were all homologous recombinations with a crossover site within homologous regions of the transgene transcripts and the cognate mutant viral RNA. However, in natural field conditions, viruses that can infect

the transgenic plants are most likely non-homologous to the transgene transcripts as the resistance provided by the transgene would prevent infections by the homologous viruses. Therefore, field recombination between viral RNAs and transgenes is very likely nonhomologous.

Another common theme emerged from these studies is the high frequency of transgenic recombination. Presence of strong selection pressures favoring recombinant viruses is a contributing factor. In addition, the presence of the 3'-untranslated region (UTR) included in the transgene mRNA in some of the studies has also been hypothesized as another important contributing factor for the high frequency of transgenic recombination (Greene and Allison, 1994). However, no experiment has been carried out to test this possibility. In this paper, we use another RNA virus, *Red clover necrotic mosaic virus* (RCNMV) as model system to investigate whether nonhomologous RNA recombination can occur between a viral transgene transcript and a viral RNA, and whether viral 3' UTR is an important factor for transgenic RNA recombination.

RCNMV is the best-characterized member of the *Dianthovirus* genus (Murphy *et al.*, 1995). The RCNMV genome consists of two single-stranded, positive-sense RNAs. Both RNAs possess 5' terminal m^7GpppA caps and lack 3'-terminal polyA tail (Lommel *et al.*, 1988; Xiong and Lommel, 1989). RNA-1 contains p27 and p88 RNA-dependent RNA polymerase (RdRp) genes essential for virus replication and the CP gene required for encapsidation and long-distance movement (Xiong and Lommel, 1989; Xiong *et al.*, 1993a; 1993b; Kim and Lommel, 1994; Bates *et al.*, 1995, Vaewhongs and Lommel, 1995; Sit *et al.*, 1998). RNA-2 encodes a single p35 MP gene required for RCNMV cell-

to-cell movement (Lommel *et al.*, 1988; Fujiwara *et al.*, 1993; Xiong *et al.*, 1993a). RNA-1 can replicate in protoplasts in the absence of RNA-2 (Osman and Buck, 1987; Paje-Manolo and Lommel, 1993), but systemic infection requires the cell-to-cell movement function provided by the RNA-2 encoded MP and long distance movement function provided by the RNA-1 encoded CP. Sequence analysis indicates that RNA-1 and -2 are essentially nonhomologous except the 29 nucleotides (nt) of the 3' terminus and the 5 nt of the 5'-terminus. The 3' terminal 29 nt is predicted to form a stable stem loop structure (SLS) that is required for the synthesis of (-)-strand RNA (Chapter 3; Turner and Buck, 1999). The 5'-terminal 6 nt (ACAAAC) is required for the synthesis of (+)-strand RNA from the (-)-strand template (Turner and Buck, 1999).

In this study, we generated transgenic plants expressing different versions of RCNMV RNA-2 transcripts that lack either the first 55 nucleotides of the 5' UTR (Δ5), the last 219 nucleotides of the 3' UTR (Δ3), or both termini (Δ53). The transgenic lines were incapable of complementing local and systemic infection by RNA-1. However, systemic infection by RNA-1 was observed in some of the Δ5 transgenic plants after a delay of several days. Analysis of the progeny virus recovered from the systemically infected Δ5 transgenic plants revealed that systemic infection was the result of non-homologous recombination between the infecting RNA-1 and the mRNA transcribed from the chromosomally integrated viral sequence. In contrast, no systemic infection was observed in transgenic plants expressing either Δ3 or Δ53 RNA-2 transcripts after inoculated with RCNMV RNA-1. This report illustrates that non-homologous recombination can occur between viral transgene and infecting RNA viruses and that the

3'-UTR is important for transgenic recombination. This is the first time that transgenically expressed viral MP gene is demonstrated to recombine with infecting viral RNA.

MATERIALS AND METHODS

Construction of RNA-2 clones, pU19RC2IG54 and pBRC2IG54

The orientation of pRC2IG54, an infectious cDNA clone of RCNMV RNA-2 (Chapter 2), was reversed to facilitate the subcloning of RNA-2 transgene into Ti-plasmid pVKGK in the correct orientation (Figure 4.1). After digestion of pRC2IG54 with *Kpn*I & *Bam*HI in the multiple cloning site of pUC18, the 1.5 kb *Kpn*I-*Bam*HI fragment containing full-length RNA-2 was blunt-ended and cloned into pUC19 at *Sma*I site to construct pU19RC2IG54 with RNA-2 3' near to the vector *Sst*I site. pBRC2IG54 clone was then constructed by transferring a 1.5 kb *Hind*III-*Sst*I fragment containing full-length RNA-2 from pUC19 to pBS (+) which contains a M13 intergenic sequence required for mutagenesis using the *ung-dut* system.

Site-directed mutagenesis

The *ung-dut* system (Kunkel *et al.*, 1987) was used to introduce molecular markers into RNA-2 sequences. pBRC2IG54 was transformed into competent cells of *Escherichia coli* strain CJ236. The phagemids containing single-stranded UTP+ DNA was synthesized in the presence of M13KO7 helper phage and precipitated by PEG (800) and NaCl. The single-stranded UTP+ DNA was extracted with phenol:chloroform and

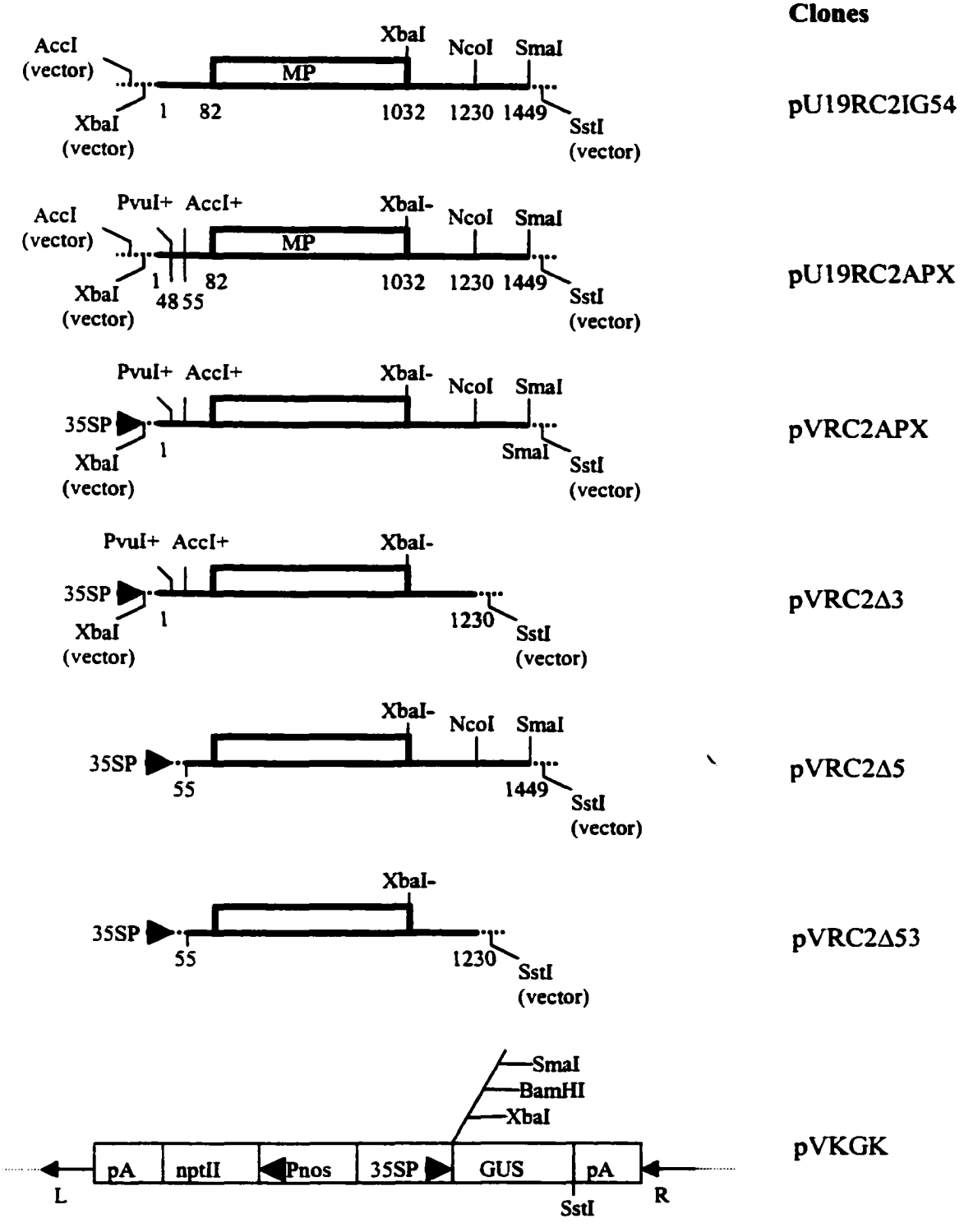

Clones
AccI
(vector)
XbaI
(vector)
MP
XbaI
NcoI
SmaI
1
82
1032
1230
1449
SstI
(vector)
pU19RC2IG54
PvuI+
AccI+
XbaI-
48
55
pU19RC2APX
35SP
pVRC2APX
pVRC2Δ3
pVRC2Δ5
pVRC2Δ53
BamHI
pA
nptII
Pnos
GUS
L
R
pVKGK

Figure 4.1 Schematic representation of RCNMV RNA-2 cDNA constructs and plant transformation vectors. RCNMV RNA-2 sequence is presented as solid line; vector sequences are presented as dashed lines. Restriction sites used in vector construction are labeled above and below the schematic. Clone pU19RC2IG54 contains the full-length RNA-2 sequence. Clone pU19RC2APX contains three molecular markers introduced in the RNA-2 sequence by mutagenesis, *Pvu*I and *Acc*I sites at the 5'-UTR and the disruption of the *Xba*I site at the stop codon of MP ORF. Ti plasmid pVKGK contains kanamycin resistant gene and GUS gene whose expression is controlled by 35S promoter. Clone pVRC2APX which contains the full-length RNA-2 is generated by replacing the GUS gene in the Ti plasmid pVKGK with the *Xba*I-*Sst*I fragment from clone pU19RC2APX; the expression of viral sequence is controlled by 35S promoter. Clone pVRC2Δ3 contains RNA-2 sequence lacking the 3'-terminal *Nco*I-*Sma*I fragment (219 nt). Clone pVRC2Δ5 contains RNA-2 lacking the 5'-terminal *Acc*I fragment (including 55 nt viral sequence). Clone pVRC2□53 contains RNA-2 lacking both the 5'-terminal 55 nt and the 3' 219 nt.

precipitated with ethanol. Mutant DNA was synthesized with T4 DNA polymerase (Invitrogen, Carlsbad, CA) and T4 DNA ligase (Invitrogen, Carlsbad, CA) using two 5'-phosphorylated oligonucleotides, RC2AP and RC2X, as primers. Primer RC2AP (5' TTGTTGGTATACTTACGATCGTGTTTC 3') is complementary to nucleotides 37-65 in the 5' UTR of RNA-2. Three nucleotide substitutions (underlined) created *Pvu*I (CGATCG) and *Acc*I (GTATAC) restriction sites as molecular markers. RC2X primer (5' CGGCTCGTTTAGAGTCTTTC 3') is complementary to nucleotides 1023-1042 around the termination codon of the MP ORF. One nucleotide substitution (underlined nucleotide) disrupted the *Xba*I (TCTAGA) but maintain the stop codon. After transformed to competent *E. coli* DH5α, the synthesized mutant DNA was selected and amplified. The mutant clone, pBRC2APX, was first screened by digestion with *Acc*I, *Pvu*I, or *Xba*I, and then confirmed by direct sequencing.

Infectivity and stability of the marked RNA-2

To prevent the second conflicting T7 primer sequence in vector pBS (+), the 1.5 kb *Hind*III-*Sst*I fragment from pBRC2APX was cloned back to pUC19 to generate pU19RC2APX (Figure 4.1). RNA-1 and RNA-2 transcripts were synthesized from *Sma*I-linearized cDNA clones of pRC1IG69 and pU19RC2APX with T7 RNA polymerase, respectively, and co-inoculated onto cowpea, *N. benthamiana* and *N. clevelandii* as described previously (Xiong and Lommel, 1991). Local lesions on cowpea leaves were counted 5 days post inoculation. Virus was then purified from infected *N. clevelandii* leaves, and the viral RNA was subsequently extracted from the purified virions. RT-PCR

was employed to amplify the full-length RNA-2 using two primers, RC2-T7-5' containing T7 promoter followed by the first 18 nt of 5' end of RNA-2 (chapter 2) and RC23'END containing 20 nucleotide (underlined) complementary to the 3' end of RNA-2 (chapter 2). The full-length PCR products were digested with *Acc*I and *Xba*I to determine whether these markers were maintained in the progeny RNA-2.

Construction of plant transformation vectors

The infectious, molecularly marked RNA-2 clone (pU19RC2APX) was used to produce the 5'-truncated clone (pU19RC2Δ5), 3'-truncated clone (pU19RC2Δ3), and a clone with truncations at both termini (pU19RC2Δ53) (Figure 4.1). pU19RC2Δ5 was constructed by *Acc*I digestion of pU19RC2APX followed by self-ligation to delete the first 55 nt in the 5' UTR. To produce pU19RC2Δ3 with the last 219 nt in the 3' UTR removed, pU19RC2APX was digested with *Nco*I (at nt 1230) and *Sma*I (at the 3' end), blunt-ended by Klenow fragment (Invitrogen, Carlsbad, CA), and then self-ligated. pU19RC2Δ3 was further digested with *Acc*I, followed by self-ligation to remove the first 55 nt at the 5' UTR, generating pU19RC2Δ53.

Intact or terminally modified RNA-2 cDNA inserts were transferred from pUC19 into and replaced the GUS gene in the Ti-plasmid pVKGK to construct a series of plant transformation vectors. pU19RC2APX and pU19RC2Δ3 were digested first with *Xba*I and *Sst*I. The *Xba*I-*Sst*I fragments (1.5 kb and 1.3 kb respectively) containing RNA-2 sequences were then ligated into pVKGK at the same sites to generate pVRC2APX and pVRC2Δ3, respectively. pU19RC2Δ5 and pU19RC2Δ53 were linearized with *Acc*I and

blunt-ended, followed by second digestion with *Sst*I, the resulting *Acc*I (blunt-ended)-*Sst*I fragment was then transferred into pVKGK at *Sma*I and *Sst*I to construct pVRC2Δ5 and pVRC2Δ53, respectively.

Plant transformation

N. benthamiana leaf disks were inoculated with *Agrobacterium tumefaciens* carrying either pVRC2APX, pVRC2Δ5, pVRC2Δ3, or pVRC2Δ53. Kanamycin-resistant calli were regenerated into plants as described by Horsch *et al.* (1985). The regenerated and kanamycin-resistant transgenic plants were transplanted in soil and self fertilized.

PCR and Southern analysis of transgenes

Total DNA was extracted from transgenic and non-transformed *N. benthamiana* plants as described by Murray and Thompson (1980). For the PCR detection of the RNA-2 transgenes, primers RC2V170 (5' GGAAATGCAGGTCTGGAGTTT 3') corresponding to RNA-2 nucleotides 161-178 and RC2C650 (5' TGCACGTAGGCTTCCACTGCC 3') complementary to nucleotides 643-659 were used to amplify a 498 nt fragment from total plant DNA. A 30 cycle PCR reaction (denaturation at 94°C for 45 sec, annealing at 54°C for 60 sec, and extension at 72°C for 30 sec) was carried out as described previously (Weng and Xiong, 1995). The transgenes were also examined by southern hybridization (Sambrook *et al.*, 1989). Total DNA extracted from transformed and non-transformed *N. benthamiana* plants were digested overnight with *Hind*III and *Eco*RI restriction enzymes at 37 °C. The restricted DNA fragments were resolved in 0.8% agarose gel and blotted

into a Hybond-N+ membrane (Amersham, Arlington Heights, IL). DNA fragments containing the RNA-2 transgene sequence were detected by hybridization with a ^{32}P-labeled DNA probe specific for RNA-2. DNA probe was synthesized by Exo(-) Klenow enzyme and random primer as described by manufacturer using the Prime-It II Random Primer Kit (Stratagene, La Jolla, CA).

Northern blot analysis

Transgene mRNA levels in different transgenic lines were examined by Northern hybridization with transgene specific, ^{32}P-labeled probe. Total RNA was extracted from transgenic and non-transformed *N. benthamiana* leaves using TRIzol RNA extraction solution (Invitrogen, Carlsbad, CA). The mRNA was then concentrated from total RNA using PolyATract mRNA Isolation System IV (Promega, Madison, WI). Northern hybridization was carried out essentially as described previously (Xiong *et al.*, 1993a). Five microgram of total RNAs or mRNA were electrophoresed in 1.5 % agarose gels after heat denaturation and transferred to Hybond-N+ membranes (Amersham, Arlington Heights, IL). The blots were then hybridized with (-) sense, ^{32}P-labeled RNA-2 riboprobe that was prepared from a RNA-2 full-length cDNA clone, pRC1IG (Xiong *et al.*, 1991), by *in vitro* transcription using T3 RNA polymerase.

Sequencing analysis of the transgene 5' terminus

To determine whether any additional mutation was introduced into transgenes during transformation and subsequent plant propagation, the 5' terminus of transgenes was

amplified by PCR using 35SP primer (5' TATATAAGGAAGTTCATTTCATTTGGA GAGGA 3') corresponding to the 35S promoter sequence upstream of the RNA-2 transgene and primer RC2C650 complementary to RNA-2 643-659 nt. The PCR products were clone into pBluscript SK(+) and sequenced.

Expression of movement protein (MP)

The T7 primer sequence upstream of viral sequence was deleted in pU19RC2Δ5 and pU19RC2Δ53. In order to synthesize the 5'-truncated RNA-2 transcripts, both pU19RC2Δ5 and pU19RC2Δ53 were digested with *Hind*III and *Sst*I. The 1.5 kb *Hind*III-*Sst*I fragment was subsequently cloned into pBS (+) at the same restriction sites to generate pBRC2Δ5 and pBRC2Δ53. The T3 promoter in pBS (+) was used to direct the transcription of the 5'-truncated RNA-2 transcripts with T3 RNA polymerase.

To test whether the 35-kDa MP can be expressed from all transgene constructs, RNA-2 transcripts with different terminal truncations were transcribed *in vitro* from *Sma*I-linearized pU19RC2APX and *Eco*RI -lineared pU19RC2Δ3 with T7 RNA polymerase, and from *Sma*I-linearized pBRC2Δ5 and *Eco*RI -lineared pBRC2Δ53 with T3 RNA polymerase. One microgram of transcripts was then translated in a wheat germ *in vitro* translation system (Boehringer Mannheim Biochemica, Satellite Beach, FL) as instructed by the manufacturer. After incubation at 30°C for 1 hour, ^{35}S -labeled translation products were electrophoresed on a 12.5% SDS-polyacrylamide gel (Laemmli, 1970), fixed and fluorographed. (Xiong *et al.*, 1993a; Weng and Xiong, 1997).

Plant inoculation, virus and viral RNA purification

RNA-2 transformed and non-transformed *N. benthaniana* plants were inoculated with RNA-1 transcripts derived from pRC1IG69 or co-inoculated with RNA-1 and RNA-2 transcripts derived from PRC1IG69 and PRC2IG54 respectively as described previously (Xiong and Lommel, 1991). The inoculated plants were maintained in greenhouse with temperature between 20 to 25°C. Virus purification from the infected tissues and the subsequently viral RNA extraction from the purified virions were carried out as described by Xiong and Lommel (1989).

Direct RNA sequencing and transcriptional runoff mapping

Direct RNA sequencing and transcriptional runoff mapping with Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) were performed to initially determine the 5' terminal sequence of recombinant RNA-2 molecules using primer RC2C120 (5'-CTGCTACTCCATCATTTGTC-3') complementary to nucleotides 119-138 of RNA-2 as described previously (Xiong and Lommel, 1989; Weng and Xiong, 1997). ^{32}P-labeled sequencing products were electrophoresed on a 6% denaturing polyacrylamide gel and autoradiographed.

Cloning and sequencing of the 5' termini of recombinant RNA

Ligation-anchored PCR was carried out to clone the 5' termini of the recombinant RNA-2 moleculars as previously described (Weng and Xiong, 1995) with the following exceptions. Recombinant RNA-2 was reverse-transcribed into first strand cDNA using

primer RC23'END (5'-GAGCTCCCGGGGTGCCTAGCCGTTATAC-3') which is complementary to RNA-2 3'terminal 20 nucleotides. The anchor primer F500 used was previously described (Weng and Xiong, 1995). A 40-cycle touchdown PCR program was used to amplify the 5'-terminus of recombinant RNA-2 from the anchor-ligated cDNA by two primers, R500 (Weng and Xiong, 1995) complementary to F500 and RC2c320 (5'-CTACGGTGACGTGTCCTG-3') complementary to RNA-2 nucleotides 321-339, with *Taq* DNA polymerase (Invitrogen, Carlsbad, CA). The A-tailed PCR products were ligated into *Ecor*V-linearized and T-tailed pBluscript SK (+) with T4 DNA ligase (Invitrogen, Carlsbad, CA) and directly sequenced.

Computer sequence analysis

FASTA and pileup-alignment computer programs in the GCG software (Genetics Computer Group, Inc., University of Wisconsin) were used to analyze the 5' terminal sequences of recombinant RNA-2 molecules.

RESULTS

Silent markers introduced in RCNMV RNA-2 did no affect its infectivity and were stably maintained

To differentiate recombinant RNA molecules generated from transgenic recombination from contaminating wild type (wt) RCNMV RNA-2, it was necessary to introduce silent molecular markers into the transgenes. Site-directed mutagenesis was used to generate restriction sites *Pvu*I and *Acc*I at the 5' UTR and to disrupt restriction

site *Xba*I at the stop codon of the MP ORF in RNA-2 infectious cDNA clone pRC2IG54 (Figure 4.1). To determine whether the introduced markers affected RNA-2 infectivity and whether the markers were stably maintained, RNA-2 transcripts derived from the resulting clone pU19RC2APX which contained all the molecular markers were co-inoculated with wild type RNA-1 transcripts onto cowpea leaves. The number of local lesions produced by pU19RC2APX transcripts was similar to that produced by the wt RNA-2 transcripts, indicating that pU19RC2APX is as infectious as the wt clone pRC2IG54 (data not shown). The stability of the molecular markers was then examined by determining the presence of these markers in progeny RNA-2. The full-length RNA-2 was amplified from the total RNA extracted from the infected tissue by RT-PCR. The PCR products were subsequently digested with *Acc*I and *Xba*I separately. While wt RNA-2 progeny carried restriction site *Xba*I but not *Acc*I, pU19RC2APX progeny contained restriction site *Acc*I but not *Xba*I (Figure 4.2). These results showed that the silent molecular markers were stable and maintained in the progenies. pU19RC2APX was suitable for construction of terminally truncated RNA-2 clones.

Construction of transformation constructs

Three terminally truncated RNA-2 clones were subsequently constructed. Clone pU19RC2Δ5 lacking the first 55 nt of 5'UTR of RNA-2 was generated by deleting the *Acc*I fragment from pU19RC2APX. Clone pU19RC2Δ3 with the deletion of the last 219 nt of RNA-1 3' UTR was created by removing the *Nco*I-*Sma*I fragment from

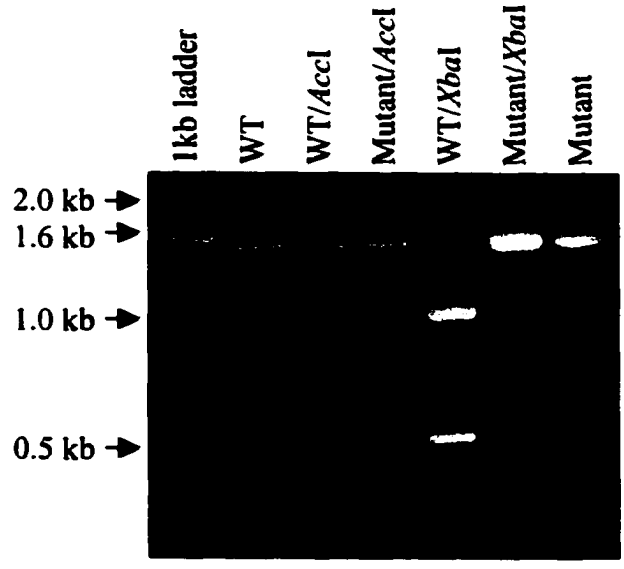


Figure 4.2 Stability of the molecular markers in RNA-2. Total RNA was extracted from *N. clevelandii* inoculated with wild type RCNMV (WT) and mutant RCNMV containing *Acc*I + (at 55 nt.) and *Xba*I – (at 1030 nt.) molecular markers in RNA-2. Full-length cDNA of RNA-2 (1.5 kb) was amplified from the total RNA by RT-PCR, digested with either *Acc*I or *Xba*I restriction enzyme, and electrophoresed in agarose gel. In WT, a single band with full-length size (1.5 kb) and two smaller bands of 1 and 0.5 kb appear after digestion with *Acc*I and *Xba*I, respectively, confirming the absence of *Acc*I and the presence of one *Xba*I in wild type RNA-2 sequence. In the mutant progeny, *Acc*I digestion resulted in two bands, one band is slightly smaller than the full-length size, another band is very small, 55 nt, and has run out of the gel; *Xba*I digestion resulted in one band, indicating that these markers were stable and maintained in the progeny.

pU19RC2APX. Clone pU19RC2Δ53 lacking both termini was made by removing the *Acc*I fragment from pU19RC2Δ3 (Figure 4.1). Four transformation constructs, pVRC2APX, pVRC2Δ5, pVRC2Δ3, and pVRC2Δ53, were then generated by replacing the GUS gene in the Ti-plasmid pVKGK with each of the full-length or terminally truncated RNA-2 fragments from pUC19 clones pU19RC2APX, pU19RC2Δ5, pU19RC2Δ3, and pU19RC2Δ53, respectively. As shown in Figure 4.1, pVRC2APX contained full-length RNA-2 sequence with three molecular marks, a *Pvu*I and an *Acc*I sites at the 5' UTR and an *Xba*I disruption at the end of the MP ORF. pVRC2Δ5 had the 5'-untranslated 55 nucleotides (upstream of the *Acc*I) removed, but retained all the 3' end sequence and the 3' marker. pVRC2Δ3 contained the complete 5' end and the 5' markers but had the 3'-untranslated 215 nucleotides deleted; pVRC2Δ53 had deletion at both termini and contained the 3' *Xba*I disruption marker.

Characterization of transgenic Nicotiana benthamiana lines

Using *Agrobacterium tumefaciens* mediated gene transformation method (Horsch *et al.*, 1985), we generated numerous lines of transgenic *Nicotiana benthamiana* plants expressing vector sequence only (NB), full-length RNA-2 (FL), 5'-truncated RNA-2 (Δ5), 3'-truncated RNA-2 (Δ3), and both termini-truncated RNA-2 (Δ53). RNA-2 transgene in the transformed plants was initially screened by PCR using a pair of RNA-2 specific primers RC2V170 and RC2C650. All 9 lines of FL, 12 of 13 lines of Δ5, 16 of 18 lines of Δ3, and 14 of 16 lines of Δ53 harbored the transgene as indicated by the

presence of a 498 bp DNA fragment in the PCR products (Figure 4.3). In contrast, no RNA-2 sequence was detected in the vector-transformed NB lines.

The presence of the transgene in transformed plants was confirmed and the copy number of transgene was estimated by Southern blot using a RNA-2-specific probe (Figure 4.4). Total DNA isolated from transformed and non-transformed plants was digested with restriction enzymes *Hind* III and *Eco*RI separately. Both restriction enzymes cut once within the vector DNA sequence transferred from the transformation vectors to the plant genome, but not in the viral sequence. Theoretically, each hybridizing DNA fragment would represent a single copy of the inserted transgene. The hybridization results revealed that the transgenic lines contained from one copy (Figure 4.4, FL-7A and Δ3-14A) to a maximum of 4 copies (Figure 4.4, Δ5-12A) of the transgene. There was no apparent correlation between the transgene copy number and the transformation construct. The copy number determined from *Hind*III digestion was consistent with that from *Eco*RI digestion, indicating the accuracy of this result.

The transgene mRNA levels were then assayed by Northern hybridization. mRNA or total RNA from plants transformed with RNA-2 sequences, as well as those from vector transformed plants and non-transformed plants was blotted and hybridized with the (-)-sense, ^{32}P-labeled RNA-2 riboprobe. The RNA-2 specific probe hybridized to an RNA species of approximately 1.2-1.5 kb, the expected sizes of mRNA transcribed from the transgene, in both the total RNA and the polyadenylated RNA preparations from the transgene-transformed *N. benthamiana* plants (Figure 4.5). In contract, total or polyadenylated RNA preparations from vector-transformed and non-transformed plants

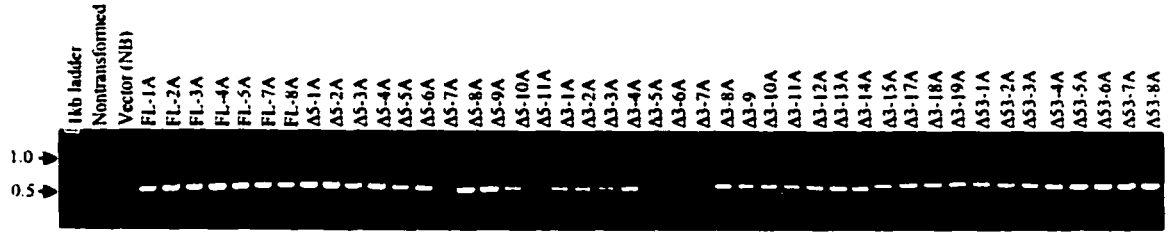


Figure 4.3 MP transgene detection by PCR. Total DNA was extracted from nontransformed, vector transformed, or a series of MP transformed *N. benthamiana* plants. A fragment of MP transgene (0.5 kb) was amplified from the total DNA by PCR using two MP specific primers around positions of 160 nt. and 660 nt. The resulting PCR product was electrophoresed in agarose gel. While no band was observed in nontransformed and vector transformed plants, a 0.5 kb band was amplified from all except 4 of the MP transformed lines, indicating that MP transgene is present in most of the transgenic lines.

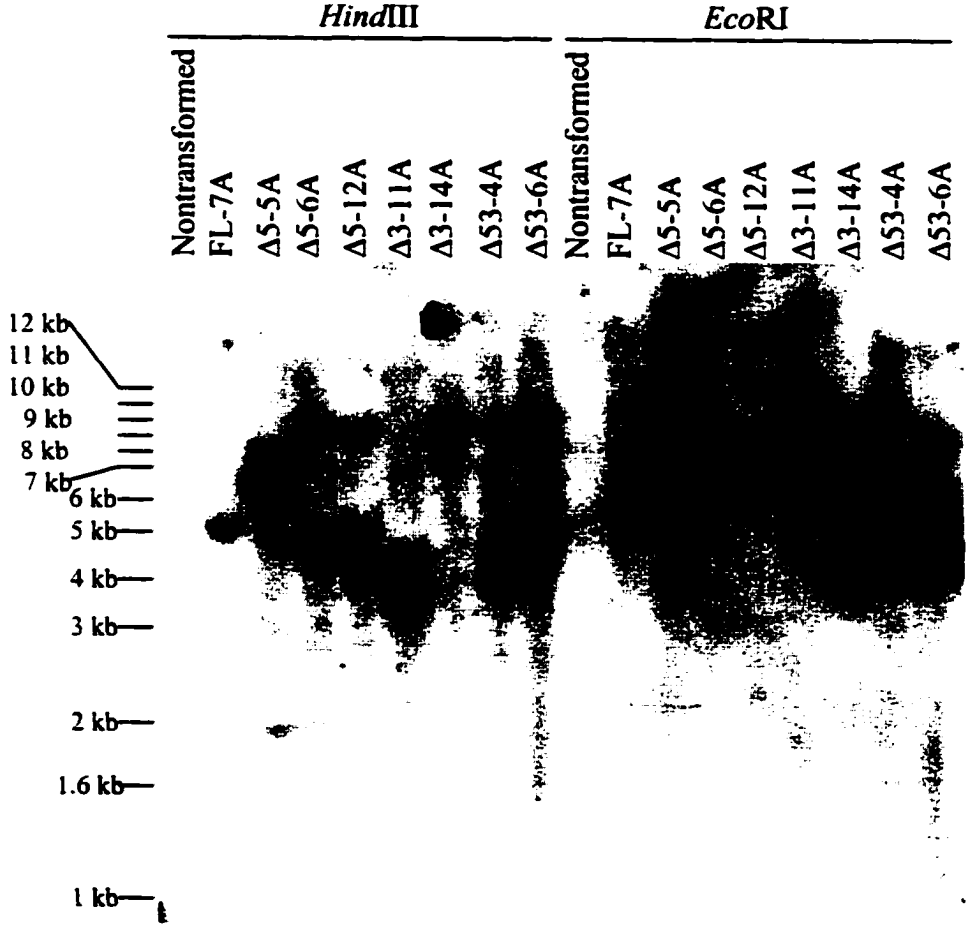


Figure 4.4 Southern hybridization of MP transgene. Total DNA was extracted from nontransformed and transgenic plants expressing RCNMV RNA-2 transcripts of full-length (FL) or lacking either the 5'-terminal 55 nt. (Δ5), the 3'-terminal 215 nt. (Δ3), or both termini (Δ53). The purified total DNA was then digested with either *Hind*III or *Eco*RI restriction enzyme. Both enzymes cut once in the vector sequence but not in the RCNMV RNA-2 sequence. The digested DNA was electrophoresed, blotted, and hybridized with a 32P-labeled probe specific for (+) RNA-2. While no hybridization was observed in total DNA isolated from nontransformed plant, 1-4 hybridizing bands were detected in total DNA purified from different MP transgenic lines.

did not hybridize with the same probe. The transgene mRNA levels varied dramatically among different lines (Figure 4.5). In nine FL lines tested, two lines (FL-7A and FL-9A) had a high level of the transgene mRNA, 2 lines (FL-5A and FL-8A) had a low to medium level, and the other five lines (FL-1A, FL-2A, FL-3A, FL-4A, FL-10A) had a very low or non-detectable level. Three of 12 Δ5 lines, Δ5-5A, Δ5-6A and Δ5-12A, had a high level of transgene mRNA, 4 lines had a low to medium level, and the other 5 lines had the mRNA level too low to be detected. Three Δ3 lines, Δ3-9A, Δ3-11A, and Δ3-14A, had a high level of transgene mRNA, three lines had a low to medium level, and 9 lines had a very low or non-detectable level. In plants transformed with pVRC2Δ53, 2 lines (Δ53-4A and Δ53-6A) had high transgene mRNA level, 6 lines had low-medium level, and 7 lines had very low or non-detectable level. The Northern blot results from the representative lines expressing transgene mRNA at high, low to medium, and very low to undetectable levels were shown in Figure 4.5. Transgenic lines with high levels of transgene expression were selected for further study.

The expression of MP in transgenic plants that accumulated a high level of transgene mRNA was assayed by Western blot with MP antibody. Two MP antibodies, one against a synthetic oligopeptide representing the carboxyl-terminal 17 amino acids of the MP (Xiong *et al.*, 1993a) and one against a recombinant MP expressed in *E. coli* (Lommel, personal communication), were used. Both antibodies were previously reported to immuno-precipitate *in vitro* translated MP. However, none of them was capable of detecting MP in total protein isolated from either transgenic plants or RCNMV infected non-transformed plant due to high background (data not shown). To verify whether the

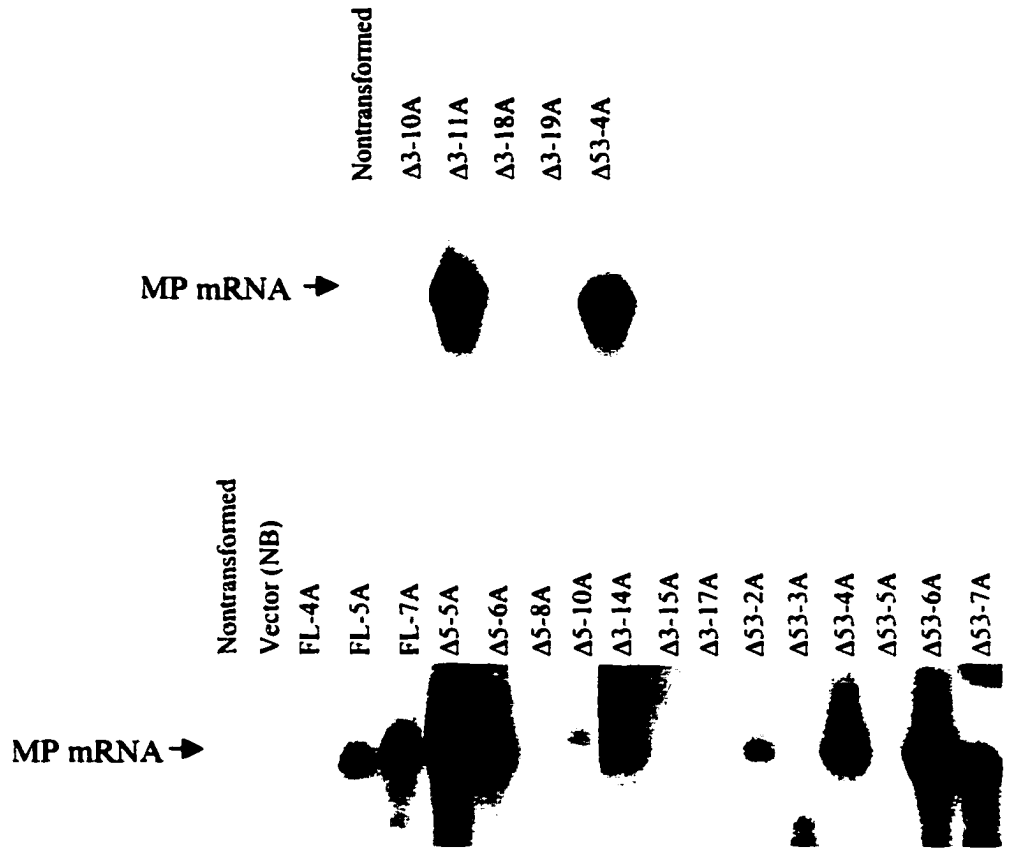


Figure 4.5 MP transgene mRNA levels detected by Northern hybridization. Either mRNA (top panel) or total RNA (bottom panel) isolated from nontransformed, vector transformed and MP transgenic plants (FL, Δ5, Δ3, and Δ53) were hybridized with ^{32}P-labeled probe specific for (+) RCNMV RNA-2. No hybridization was detected in total RNA purified from nontransformed or vector transformed plants. Varying levels of transgene mRNA were observed in different transgenic lines.

transgene constructs were functional for MP expression, *in vitro* transcripts from clones pU19RC2APX, pBRC2□5, pU19RC2□3 and pBRC2□53, which contains full-length, 5'-truncated, 3'-truncated, and both termini-truncated RNA-2 sequences, respectively, were translated in a wheat germ *in vitro* translation system. The translation products were analyzed by SDS-polyacrylamide gel electrophoresis. As shown in Figure 4.6, similar amount of the 35-kDa protein was expressed from all 4 constructs as well as from wild type RNA-2. These data indicated that transgenes delivered by four transformation constructs contained all necessary elements for efficient translation of MP. To further demonstrate that undetectable MP expression did not result from any mutations introduced into transgenes during plant transformation, the transgenes from the transgenic plants were amplified by PCR using primer 35Sprimer corresponding to the 35S promoter sequence upstream of the transgene and a RNA-2 primer RC2C650 complementary RNA-2 in MP coding region. PCR products were cloned and sequenced. While transgenes from the transgenic plants FL and Δ3 maintained complete RNA-2 5' sequence, those from the plants Δ5 and Δ53 lacked the first 55 nt of 5' UTR as designed (data not shown). No additional mutation was observed. Combining with the observation that MP is not detected from RCNMV infected plants, these results suggest that failure to detect MP in RNA-2 transgenic plants is very likely due to the inability of the antibodies to bind MP specifically and strongly in the Western blot.

RNA-2 transgenic plants were not resistant to RCNMV infection

RCNMV infected non-transgenic *N. benthamiana* plants initially develop necrotic

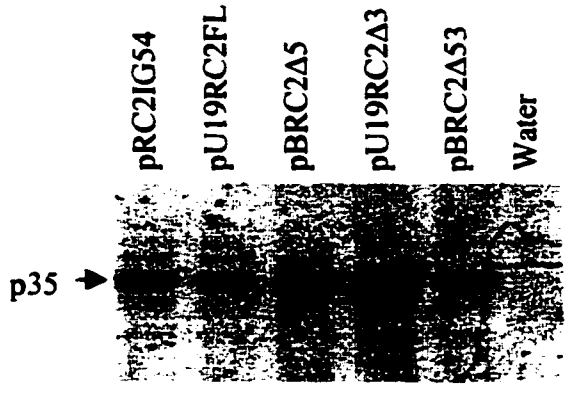


Figure 4.6 Wheat germ *in vitro* translation of RNA-2 transcripts. Equal amount of RNA-2 transcripts from either original full-length clone (pRC2IG54), full-length clone containing molecular markers (pU19RC2FL), 5'-terminal deletion clone (pBRC2Δ5), 3'-terminal deletion clone (pU19RC2Δ3), or both termini deletion clone (pBRC2Δ53) was added to a wheat germ *in vitro* translation system. The 35S-methionin labeled translation product was electrophoresed in a 12.5% SDS polyacrylamide gel and fluorographed. Similar amounts of p35 MP were produced from all transcripts tested.

lesions on the inoculated leave three days post inoculation (dpi) and then spread to upper un-inoculated leaves to produce systemically mosaic symptom 5 dpi. When the transgenic lines were inoculated with RCNMV virions or RNA transcripts, all of the lines developed symptoms identical to those on non-transgenic plants. The timing of symptom development on both the transgenic and non-transgneic *N. benthamiana* was also indistinguishable (data not shown). These results indicated that the transgenic lines expressing either full-length RNA-2 transcripts, 5' terminally truncated RNA-2 transcripts, 3' terminally truncated RNA-2 transcripts, or both termini-truncated RNA-2 transcripts were not resistant to RCNMV infection.

Replication capturing of full-length RNA-2 transgene by RCNMV RNA-1

Non-transformed plants inoculated with RCNMV RNA-1 transcripts do not show any symptoms because cell-to-cell movement and the subsequent systemic infection require RNA-2 encoded MP. However, the FL transgenic plants expressing the full-length RNA-2 transcripts showed wt systemic symptom 5 days after inoculation with RNA-1 transcripts. The timing of symptom development was similar to that on non-transformed *N. benthamiana* plants co-inoculated with RNA-1 and RNA-2 (data not shown). This data suggest that mRNA of full-length RNA-2 transgene could be captured by RNA-1. The captured RNA-2 could then provide the cell-to-cell movement function essential for local and systemic viral spread. The transgene mRNA in the FL transgenic lines contained 48 vector nucleotides upstream of the RNA-2 sequence and 28 vector nucleotides followed by poly(A) tail downstream the RNA-2 sequence. In order to capture this full-length

transgene mRNA, RdRp encoded by RNA-1 must be able to access and bind the viral (-) strand promoter and initiate (-) RNA synthesis at the viral 3' end located internally on the transgene mRNA. The elongating RdRp then must terminate (-) RNA synthesis internally at the viral 5' end to stop copying the extraneous nucleotides at the 5' end of the transgene mRNA. Alternatively, RdRp could keep moving to the 5' end of mRNA and copy the vector sequence into the 3' end of (-) RNA. The nonviral, vector sequence could then be removed by internal initiation of the subsequent progeny (+) RNA synthesis.

Two factors, transgene mRNA level and RNA-1 inoculum concentration, were examined for their effects on the frequency of capturing RNA-2 transgene mRNA by RNA-1. Three different concentration of viral RNA-1 inoculum (16.7 ng/μl, 3.3 ng/μl, and 0.5 ng/μl) were inoculated onto FL transgenic plants expressing high, medium, and low level of full-length RNA-2 transcripts. Two plants each from high expression lines FL-7A and FL-9A, medium expression lines FL-5A and FL-8A, and low expression lines FL-1A and FL-2A were inoculated. Systemically infected transgenic plants were counted as successful captures of the transgene mRNA by viral RNA-1. As shown in Table 4.1, capture frequency increased from the low expression line to the high expression line. Similarly, the capture frequency increased when the RNA-1 inoculum increased. In the high expression lines, capturing of MP-transgene by RCNMV RNA-1 occurred regardless the level of RNA inoculum. In contrast, MP transgene capturing did not occur in the low and medium expression lines when they were inoculated with 10 μl of 0.5

Table 4.1 Effect of the level of transgene mRNA and the concentration of RNA-1 inoculum on the capture of the transgene by the infecting RNA-1.

Level of transgene mRNA	Concentration of RNA-1		
	16.7 ng/ul	3.3 ng/ul	0.5 ng/ul
Low	2/4	1/4	0/5
Medium	3/4	2/4	0/4
High	4/4	4/4	2/4

ng/μl RNA-1 transcripts as none of the inoculated plants developed systemic infection. These data indicated that both transgene expression level and inoculum level influenced and contributed equally to the capture of transgene by RNA-1.

Systemic infection of Δ5 transgenic lines inoculated with RNA-1

Non-transformed *N. benthamiana* plants inoculated with wt RCNMV develop systemic infections within 5 dpi, but none of the Δ5 transgenic plants inoculated with RNA-1 transcripts developed any visible symptoms within 5 dpi, suggesting that the mRNA of the 5' truncated RNA-2 transgene was unable to complement RNA-1 infection. However, 15 of 19 (78.9%) RNA-1 inoculated, high expression Δ5 transgenic lines (Δ5-5A, Δ5-6A and Δ5-12A) developed systemic symptom 1-2 weeks after inoculation, while RNA-1 inoculated non-transformed *N. benthamiana* did not produce any symptom 2 months after inoculation. Two types of systemic symptoms, milder symptoms and symptoms identical to wt RCNMV infection, were observed. Virus recovered from infected transgenic plants was infectious when back inoculated to non-transformed *N. benthamiana* and *N. clevelandii*, indicating that the virus contained a replication-competent RNA-2 molecule. These results demonstrated that the systemic infections did not result from the functional complementation of MP expressed from 5'-truncated RNA-2 transgene. The systemic infection was the result of capturing and replication of the 5' truncated MP transgene as a part of the viral genome. This conclusion was confirmed by the detection of RNA-2 like molecules in virions purified from the systemically infected transgenic plants. When virion RNA was extracted and fractionated by agarose gel

electrophoresis, two RNA species, RNA-1 and another RNA species with size similar to RNA-2, were observed (data not shown). As the MP transgene lacked the 5' promoter necessary for viral RNA replication, a reasonable explanation is that the RCNMV RNA-1 has restored or reconstituted the 5' promoter on RNA-2, possibly by recombination.

To ensure that the RNA-2 like recombinant molecule detected in the virion came from the Δ5 transgene and not from wt RNA-2 contamination, the presence of the silent molecular markers originally introduced to RNA-2 transgene were determined. wt RCNMV RNA-2 contains an *Xba*I site at nucleotide 1032. This restriction site was removed by site-directed mutagenesis during the construction of transgenes. To check for the presence of the *Xba*I site, the 3' 1.3 kb fragment of the RNA-2 was amplified by RT-PCR using a pair of primers: RC23END complementary to 3' end 20 nucleotides of RNA-2 and RC2V170 corresponding to RNA-2 nucleotides 161-178. The PCR fragment was then digested with *Xba*I and resolved by agarose gel electrophoresis. Only one band was observed in PCR products amplified from the recombinant RNA-2 while two bands of expected sizes (0.5 kb and 0.8 kb) were detected from wt RCNMV (data not shown). These results showed that the recombinant RNA-2 molecules originated from the transgene transcripts, not from contamination.

In contrast, none of 116 transgenic Δ3 plants of high expression lines (Δ3-9A, Δ3-11A, and Δ3-14A) expressing 3'-terminally truncated RNA-2 transcripts and none of 108 transgenic Δ53 plants of high expression lines (Δ53-4A and Δ53-6A) expressing both terminally truncated RNA-2 transcripts developed any symptom 2 months after inoculation with RNA-1. Absence of RCNMV RNA-1 infection was confirmed by

Northern dot blot analysis of total RNA isolated from either inoculated or systemic leaves of transgenic plants with (-)-sense, ^{32}P-labeled RNA-1 probe. RCNMV RNA-1 was not detected in any of the inoculated plants (data not shown). These data further confirm that the RNA-2 MP transgene was unable to complement RNA-1 infection, as previously demonstrated in transgenic Δ5. They also suggested that capturing of the Δ3 and Δ53 MP mRNA by RCNMV RNA-1 did not occur or occurred at a frequency below the detection limit of this experiment.

Characterization of recombinant RNA-2 molecules

Out of 15 Δ5 transgenic plants that developed systemic infection after RNA-1 inoculation, five (Δ5-6A-1, Δ5-12A-1, Δ5-12A-2, Δ5-12A-4, Δ5-12A-6) were further analyzed for the origin of recombinant RNA-2. Recombinant virus was purified from RNA-1 inoculated transgenic plants. The 5'-termini of the encapsidated recombinant RNA-2 molecules were then determined by runoff transcriptional mapping and direct RNA sequencing with reverse transcriptase using primer RC2C120 complementary to nucleotides 119-138 of RNA-2. The sequencing products and transcriptional runoff products were resolved in denatured polyacrylamide gel. Although the resolved RNA sequences were not clear and the last few nucleotides were blocked by a heavy band resulting from the termination of sequencing products at the ultimate nucleotide of the RNA 5' terminus (data not shown), different RNA species and their sizes were clearly defined by reverse transcriptional runoff (Figure 4.7). In general, the recombinant RNA-2 molecules were smaller than the wt RNA-2. The recombinant RNA-2s recovered from

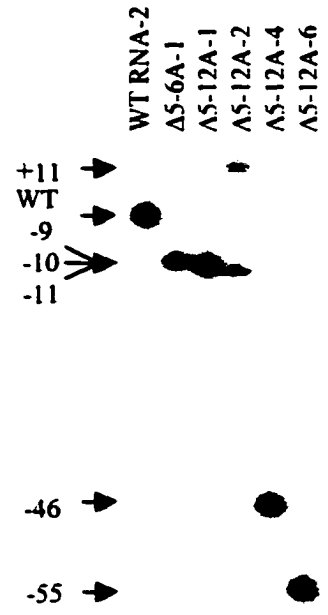


Figure 4.7 Mapping of the 5' termini of recombinant RNA-2 molecules. Viral RNA was extracted from five systemically infected Δ5 plants (Δ5-6A-1, Δ5-12A-1, Δ5-12A-2, Δ5-12A-4, Δ5-12A-2). The purified viral RNA was then reverse transcribed into cDNA using Superscript II reverse transcriptase using primer RC2C120 complementary to nucleotides 119-138 of RNA-2 (Xiong and Lommel, 1989; Weng and Xiong, 1997). ^{32}P-labeled cDNA was electrophoresed on a 6% denaturing polyacrylamide gel and autoradiographed.

plants Δ5-6A-1 and Δ5-12A-1 were 9 nt. and 10 nt. shorter than wt RNA-2, respectively. In comparison, the sizes of recombinant RNA-2 molecules from plants Δ5-12A-4 and Δ5-12A-6 were 46 and 55 nt. shorter than RNA-2, respectively. Plant Δ5-12A-2 contained two species: a smaller RNA 11 nt. shorter and a larger RNA 11 nt. longer than RNA-2, indicating that more than one recombination event could occur in a single transgenic plant.

To determine the nucleotide sequences at the 5' termini and at the cross-over junctions, recombinant RNA-2 molecules were cloned by a combination of reverse transcription and ligation-anchored PCR (Weng and Xiong, 1995). Thirteen clones of recombinant RNA-2 molecules from five RNA-1 inoculated Δ5 transgenic plants were sequenced. The alignment of the 5' terminal sequences of recombinant RNA molecules with those of wt RNA-1, wt RNA-2, and the transgene mRNA is presented in Figure 4.8. A total of 7 unique recombinant RNA-2 species were identified, one species each from transgenic plants Δ5-12A-4 (RΔ512A4, two independent clones sequenced), Δ5-12A-6 (RΔ512A6, three independent clones sequenced), Δ5-12A-1 (RΔ512A1, three independent clone sequenced), and Δ5-6A-1 (RΔ56A1, two independent clones sequenced), and three species from transgenic plant Δ5-12A-2 (RΔ512A2S, RΔ512A2M, and RΔ512A2L). While all of these recombinant molecules contain RNA-2 sequence starting at nucleotides upstream of MP AUG initiation codon and at least first 8 nucleotides from the 5' end of RNA-1, the sequences around the crossover sites are different from each other. Both RΔ512A4 and RΔ512A6 consisted of the 5' terminus of RNA-1 joint with the viral RNA-2 sequence derived from the transgene, but they are

RNA species	Sequences
WT RNA-1	ACAAACGTTTTACCGGTTTGTAAATAGGAG.....
RΔ512A4	ACAAACGCTTTAC_____CAACAATACAAAGAGGTTTGAGATG
RΔ512A6	ACAAACGTTTTAC_____AAAGAGGTTTGAGATG
RΔ512A2S	ACAAACGT***CTAATAAGAGTG***_____ACAAANAGGTTTGAGATG
RΔ512A2M	ACAAACGTCTTAC***TGT***_____gggggactctagaanatcgatccccgGTATACCAACAATACAAAGAGGTTTGANATG
RΔ512A2L	ACAAACGT***CTAATAAGAGTTACAAACGAAAGCATTGT***acggggactctagaggatcgatccccgGTATACCAACAATACAAAGAGGTTTGAGATG
RΔ512A1	ACAAACGT***CTAGCGAG***_____cgggggactctanangatcgatccccgGTATACCAACAATACAAAGAGGTTTGAGATG
RΔ56A1	ACAAACGT***CGCACACAAG***_____gggggactctagaanatcgatccccgGTATACCAACAATACAAAGAGGTTTGANATG
Transgene mRNA	acacgggggactctagaggatcgatccccgGTATACCAACAATACAAAGAGGTTTGAGATG
WT RNA-2	ACAAACCTCGCTCTATAAACAGATTAATTGATTGTAAGCAGACAAGATCGTAAGACTACCAACAATACAAAGAGGTTTGAGATG

Figure 4.8 Alignment of the 5' sequences of recombinant RNA-2 with WT RNA-1, RNA-2, and transgene mRNA. Low case letter represents transcribed vector sequence. Italic letter indicates the sequence of unknown origin. Underlined letter indicates the silent marker introduced into the transgene.

slightly different from each other. Three nucleotides at the crossover junction, UAC, are present in both the RNA-1 sequence and the RNA-2 sequence. Depending on the precise crossover nucleotide, the 5'-termini of these two molecules contained either 10 or 13 nucleotides from RNA-1 and RNA-2 transgene sequence starting at nucleotide 57/61 and nucleotide 66/69, respectively. Four species of recombinant molecules, RΔ512A1, RΔ56A1, RΔ512A2L, and RΔ512A2M, contained non-viral sequences of the vector origin and of unknown origin between RNA-1 and RNA-2 sequences. The vector sequence is transcribed from the full-length transgene into mRNA, therefore, it is considered as part of the donor RNA. The vector sequences in the recovered recombinant molecules cover almost the entire 5' end of transgene mRNA except the first 2-4 nt (Figure 4.8). The unknown sequences range from 3 nt to 29 nt and have no homology to any sequences in database. The first 8 nt (ACAAACGT) of the 5'-terminal sequence of RNA-1 is conserved in all four recombinant species. The recombinant molecule, RΔ512A2S, recovered from plant Δ5-12A-2 contains only some unknown nucleotides but no vector sequence between RNA-1 and RNA-2 sequences. However, closer examination of the sequence around the crossover site revealed that RΔ512A2S is similar to RΔ512A2L isolated from the same plant, except a deletion of 57 nt fragment including the partial unknown sequence, entire vector sequence, and partial RNA-2 sequence. Therefore, RΔ512A2S may have been evolved from RΔ512A2L *via* further RNA recombination and deletion. In addition to the silent molecular marker (disruption of *Xba*I site) at the MP stop codon, the presence of another marker *Acc*I as well as the vector

sequence in some recovered molecules further confirmed that these RNA-2-like molecules are products of heterologous, intermolecular RNA recombination.

Environmental condition affecting RCNMV transgenic recombination

The effect of temperature on transgenic recombination was assessed by comparing the incidence of systemic infection of Δ5 plants inoculated with RNA-1 in different temperature regimes: 30°C day/20°C night and 20°C day/20°C night. As shown in Table 4.2, while a high frequency of transgenic recombination was observed at 20°C/20°C condition, no recombination was observed at 30°C/20°C condition, suggesting that transgenic recombination could be permitted at a low temperature but inhibited at a high temperature.

DISCUSSION

In this study, we have generated a series of transgenic *N. benthamiana* plants expressing different levels of RCNMV RNA-2 that lacked either the 5'-terminus, 3'-terminus, or both termini. None of these transgenic plants was resistant to RCNMV infection, or capable of complementing RNA-1 infection. Five days after inoculated with RNA-1, the time required for systemic infection of RCNMV in non-transformed *N. benthamiana* plants, no symptom was observed in these transgenic plants. However, after two days or a longer delay, systemic infection was developed in the RNA-1 inoculated transgenic plants that expressed the 5'-terminally truncated RNA-2 transcripts. Recombinant RNA-2 molecules between the MP transgene mRNA and the infecting

Table 4.2 Effect of the temperature on the frequency of transgenic recombination.

Temperature	20°C day/20°C night	30°C day/20°C night
Recombination frequency	16/19	0/20

RNA-1 were recovered from these systemically infected Δ5 transgenic plants. In contrast, no systemic infection was observed in transgenic plants expressed the 3'-terminally or both-termini truncated RNA-2 transcripts after inoculated with RNA-1 up to 8 weeks.

MP gene expressed in transgenic plants is not capable of complementing RCNMV RNA-1 infection

The MP transgene was highly transcribed in the transgenic lines chosen for recombination study in this paper. However, neither MP nor MP function (mediating viral cell-to-cell movement) was detected by Western blot of the total protein purified from MP transgenic plants with MP antiserum and complementation of RNA-1 infection, respectively. The same MP antiserum used for detection of MP expressed from uninoculated transgenic plants was not capable of detecting the MP in the nontransformed plant systemically infected with wt RCNMV (data not shown), indicating that these antisera were not able to bind MP strongly and specifically in Western blot, therefore, not suitable for MP detection. Although MP was not detected, all the available evidence suggests that MP should be expressed in the transgenic plants. The *in vitro* transcripts from all four MP transformation constructs were translated as efficiently as wt RNA-2 in an *in vitro* wheat germ translation system, suggesting that they were functional templates for the plant translation system. The ability of the transgene mRNA to express fully functional MP was further demonstrated by the recombinant RNA-2 captured by RCNMV RNA-1. The recombinant RNA-2 was fully functional and capable of mediating the cell-to-cell movement of the recombinant virus, indicating there was no debilitating

mutation in the transgenes. Furthermore, sequencing of the 5' end of the transgenes taken from the transgenic plants did not show a single nucleotide change. Nevertheless, the MP-expressing transgenic plants failed to functionally complement infection by RCNMV RNA-1. This might be explained by the differential localizations between viral RNA replication and MP production from transgene mRNA which could physically prevent the MP to interact with RNA-1 and complement its infection, or/and the improper timing of MP expression. Alternatively, the amount of MP translated from transgenic plant could be much lower than that expressed from replicating viral RNA-2 and beyond the detection limit in this study, as the viral RNA-2 could replicate and accumulate to a concentration much higher than the MP transgene mRNA in the transgenic plants (data not shown).

Systemic infection of Δ5 transgenic plants inoculated with RNA-1 resulted from transgenic recombination

The systemic infection observed in the RCNMV RNA-1 inoculated Δ5 transgenic plants in this study was clearly a result of transgenic recombination between the MP transgene mRNA and the infecting RNA-1. The virus recovered from the systemic infected Δ5 transgenic plant contained the RNA-2 like molecule and was able to systemically infect non-transformed plants, indicating that the systemic infection in RNA-1 inoculated Δ5 transgenic plants resulted from the RNA-2 like molecule that was captured, replicated, and encapsidated by RNA-1. The presence of RNA-1 sequence of 5' terminus and the silent markers (absence of *Xba*I at 1032 nt and presence of *Acc*I at 55nt)

specific to the MP transgene, in addition to RNA-2 sequence, in the recovered RNA-2 like molecules unambiguously demonstrated that these RNA-2 molecules were the products of transgenic recombination between the infecting RNA-1 and RNA-2 transgene mRNA. In construct pVKΔ5 that was used to generate Δ5 transgenic plants, the transcription initiation site directed by 35S promoter is 30 nt upstream from the RNA-2 sequence. These 30 nt were inherited from the vector and were transcribed as the 5' terminus of the transgene mRNA. The presence of this vector sequence in the recovered RNA-2 molecules further confirmed their recombinant nature.

Transgenic recombination under selection pressure

A high frequency of systemic infection (78.9%) resulting from transgenic recombination was observed in Δ5 transgenic plants inoculated with RCNMV RNA-1. The observed high recombination frequency was not entirely unexpected, considering the extremely high selection pressure for recombinant RNA-2 molecules in these plants.

RCNMV contains two genomic RNAs. RNA-1 encodes all proteins required for viral RNA replication (Xiong and Lommel, 1989). RNA-2 encodes MP protein required for cell-to cell- movement (Lommel *et al.*, 1988). Without RNA-2 or MP, RNA-1 can replicate by itself in single cell but is restricted in the original infected cell (Paje-Manalo and Lommel, 1993; Osman and Buck, 1987). The first 6 nt (ACAAAC) of 5' end and the 3'-terminal stem-loop structure composed of the last 29 nt of 3' end are conserved in both RNA-1 and RNA-2, and are required for progeny (+)-strand RNA synthesis and for complementary (-)-strand RNA synthesis, respectively (chapter 3; Turner and Buck,

1999). As MP transgene mRNA expressed in all of Δ5, Δ3, and Δ53 transgenic plants lacked one or both of these terminal elements required for RNA replication, they could not be replicated by RCNMV RdRp. Moreover, MP transgenic plants could not functionally complement RNA-1 infection. Therefore, a high selection pressure for acquisition of these terminal sequences in order for virus to move cell-to-cell and systemically was exerted in the RNA-1 inoculated MP transgenic plants.

Transgenic recombination in other viruses was detected under a similarly strong selection pressure (systemic or cell-to-cell movement), including CCMV (Greene and Allison, 1994), TBSV (Borja, *et al.*, 1999), TMV (Adair and Kearney, 2000), PPV (Varrelmann *et al.*, 2000), and CaMV (Schoelz and Wintermantel, 1993; Wintermantel and Schoelz, 1996). In CCMV, CP is required for systemic infection. The transgenic plants expressing the 3' two-thirds of the CCMV CP gene and the 3' UTR of RNA-3 developed systemic infection after inoculated with a CCMV deletion mutant lacking the 3' one-third of the CP gene. This systemic infection resulted from aberrant homologous recombinations between the transgenic mRNA and the infecting viral RNA mutant at the central one-third common region (Greene and Allison, 1994). In TBSV, lethal systemic symptom requiring functional CP was observed in transgenic plants expressing TBSV CP without 3'-UTR after inoculated with mutant TBSV having the 46 nucleotide within the 5' of CP gene deleted (Borja, *et al.*, 1999). The systemic infection is caused by double homologous recombination between the transgenic mRNA and the infecting viral RNA (Borja, *et al.*, 1999). In TMV, CP is also required for systemic infection. After inoculated with mutant TMV having a 168 nt deletion in the central region of CP gene, the CP

transgenic plants showed efficient systemic infection through homologous recombination (Adair and Kearney, 2000). In the CaMV studies, gene IV, which determines systemic infection in solanaceous hosts, was expressed in transgenic *N. bigelovii* plants. In one case, CaMV isolates that were unable to systemically infect *N. bigelovii* acquired the transgene restoring systemic infection function in the process (Schoelz and Wintermantel, 1993). In another case, a less aggressive CaMV isolate recombined with the transgene in *N. bigelovii* and became more aggressive (Wintermantel and Schoelz, 1996). Although CaMV is a DNA virus, the transgenic recombination in CaMV was reported to involve the transgene transcript and the RNA replication intermediate of the pararetrovirus (Schoelz and Wintermantel, 1993; Wintermantel and Schoelz, 1996).

In these examples, an infecting mutant virus has recombined with transgenically expressed, cognate viral sequences that are required for the viral long distance movement. Our data showed that a transgene mRNA for a viral MP, which provides a cell-to-cell movement function, can also be a potential template for viral RNA recombination, without apparent sequence homology.

In the field condition, the infecting viruses unrelated or less closely related to the transgene can systemically infect the virus-resistant transgenic plant; there would be little or no selection pressure for the possible recombinant RNA molecules between the infecting virus and the transgene mRNA. On the other hand, a fully functional virus that can systemically infect a transgenic plant would be infect most of the cells in the plant and dramatically increases potential contact between viral RNA and the transgene mRNA. In our current system, infection by RCNMV RNA-1 would be restricted to the

initially inoculated cells in the absence of recombination. In order to assess the natural transgenic recombination under the field condition, further research on recombination between a viral transgene and a fully functional virus in the absence of selection pressure will be required.

RCNMV transgenic recombination is non-homologous

RNA-1 and -2 of RCNMV are non-homologous except the 5'-terminal 6 nucleotides (ACAAAC) and the 3'-terminal SLS composed of 29 nt (Lommel *et al.*, 1988; Xiong and Lommel, 1989). Although the MP transgene in plants Δ5, Δ3, and Δ53 were derived from the same virus as the infecting RCNMV RNA-1, they should be considered as non-homologous RNA molecules for RNA recombination as there was no sequence homology around the recombination crossover junction. Therefore, the seven species of recombinant RNA molecules recovered from Δ5 transgenic plants inoculated with RNA-1 were the products of non-homologous recombination. With the exception of two recombinant molecules that were generated by a crossover event at a UACC or UAC overlap, there was no sequence overlap at the crossover junction between the donor and receptor RNA molecules. This is the first report to show that transgenic recombination could be mediated by non-homologous RNA recombination in addition to homologous recombination observed in other viruses (Adair and Kearney, 2000; Borja *et al.*, 1999; Greene and Allison, 1994; Varrelmann *et al.*, 2000).

RCNMV transgenic recombination is via a viral polymerase-mediated template-switching mechanism

High frequency of systemic infection was observed in Δ5 transgenic plants inoculated with RNA-1 *via* non-homologous RNA recombination between RNA-1 and MP transgene mRNA lacking the first 55 nt of 5' end of RNA-2. All seven species of recombinant RNA-2 molecules isolated from the systemically infected plants contained RNA-1 5'-terminus sequence at the 5' end and RNA-2 sequence including the entire MP ORF and the 3' UTR at the 3' end (Figure 4.8). In contrast to the high accidence of systemic infection produced in Δ5 transgenic plants inoculated with RNA-1, no systemic infection was developed in Δ3 transgenic plants under the same condition. The difference between plants Δ5 and Δ3 was that the Δ5 transgene mRNA contained the entire 3' end of RNA-2 while the Δ3 transgene mRNA lacked the last 219 nt of the 3' UTR, which contains the replication element essential for the initiation of (-) RNA synthesis (Chapter 3). These data suggested the involvement of RCNMV RdRp in the observed transgenic recombination. Based on the studies of RNA virus recombination (Kirkegaard and Baltimore, 1986; Jarvis and Kirkegaard, 1992; Nagy and Simon, 1997; Figlerowicz and Bujarski, 1998; Chetverin, 1999), a viral polymerase-mediated template-switching model was proposed to explain the formation of recombinant RNA-2 molecules observed in this study (Figure 4.9). As the 5'-truncated transgene mRNA expressed in the Δ5 transgenic plants contains the complete RNA-2 3' sequence including the 3' replication element, the RCNMV polymerase can recognize and bind to this terminal element and initiate the synthesis of (-)- strand RNA-2. When RNA synthesis approaches the 5' terminus of the

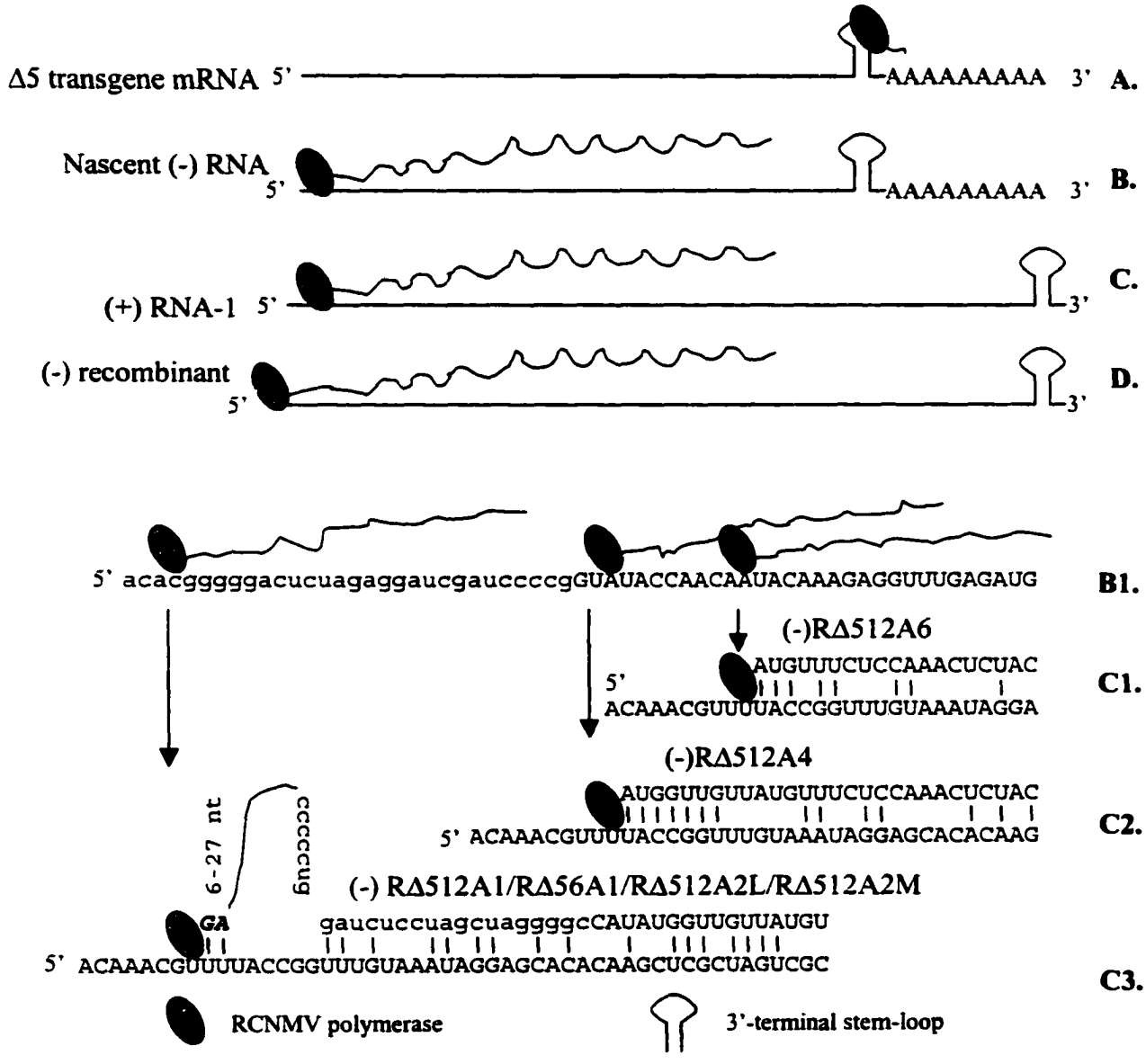


Figure 4.9 Polymerase-mediated template-switching model of transgenic recombination. Δ5 transgene mRNA contains poly(A) tail at 3' end (A) and 30 nt vector sequence (B1, lower case) upstream of RNA-2 sequence (B1, Upper case) at 5' end. RCNMV polymerase recognizes and initiates the nascent (-) RNA synthesis from the 3'SLS located upstream of the poly(A) tail of Δ5 transgene mRNA (A). When approaching to the 5' end of trangene mRNA, polymerase dissociated from the template at different nt (B1) and switched to (+) RNA-1 (receptor) downstream the 5' end (C, C1, C2, C3) with the assistance of partial base pairing between the nascent (-) RNA and (+) RNA-1. RNA synthesis was resumed and the 5' replication element was acquired in the (-) recombinant RNA from RNA-1 (D). No extraneous nt was inserted where 3 or more base pairing formed between the 5' end of nascent (-) RNA and acceptor (+) RNA-1 (C1, C2). However, in the absence of base pairing between the 5' end of nascent (-) RNA and the receptor (+) RNA-1, extraneous nucleotides (italicized) would be randomly added by the TNTase activity of polymerase (C3). Until two randomly added nucleotides (5'GA in this study) formed base pairing with RNA-1, RNA-1 could be used as template by RCNMV polymerase (C3).

MP mRNA, the polymerase complex along with the attached nascent minus RNA dissociates from its original template (transgenic mRNA/donor RNA) and switches template to the 5' end of the RCNMV RNA-1 positive stand (acceptor RNA) where (-) RNA synthesis is resumed and completed. Different species of recombinant RNA-2 were then formed dependent on crossover junctions on the donor and receptor RNAs. In the case of molecules RΔ512A4 and RΔ512A6 which did not contain any non-viral sequences but contained 4 nt (TACC) or 3 nt (TAC) overlaps at the crossover sites in both parental RNAs, polymerase dissociated from donor RNA at the position of 21-25 nt and position of 13-16 nt, respectively, upstream of the AUG start codon of MP, and re-associated with and resumed RNA synthesis at receptor RNA (RNA-1) at nucleotides between 10 to 14 and between 10 to 13 downstream of the 5' end of RNA-1, respectively. The common 4 or 3 nt at the crossover sites may help the template switching event by providing weak but potentially critical local hybridization between the nascent (-) RNA and the receptor (+) RNA-1. In molecules RΔ56A1, RΔ512A1, RΔ512A2L, and RΔ512A2S, non-viral sequences including vector sequence and sequences of unknown origin present and locate between RNA-1 and RNA-2 sequences. The unknown origin sequences are different from each other and show no homology with any sequence in the Genbank Database. These sequences may be randomly incorporated into recombinant RNA due to the terminal nucleotidyl transferase activity of polymerase (TNTase). This activity has been reported in the RdRps of hepatitis C virus (HCV) (Ranjith-Kumar *et al.*, 2001). The recombinant HCV RdRp purified from *Escherichia coli* was capable of adding nontemplated nucleotides (1 to 5 nt) to the 3' end of viral RNA. Mutation of the

highly conserved GDD catalytic motif in HCV RdRp resulted in the abolishment of both viral RNA synthesis and TNTase activity, indicating that TNTase activity is an inherent function of HCV RdRp. Similar activity has also been reported during the replication of a plant virus, *Turnip crinkle virus* (TCV), a member of the *Tombusviridae* of which RCNMV is also a member (Guan and Simon, 2000). Transcripts with short deletions (4 or 5 nt) of the conserved 3' end of satellite (sat) C RNA associated with TCV were repaired by the addition of wt or non-wt 3' nucleotides *in vivo*. RdRp purified from TCV infected plants was able to polymerize nontemplate nucleotides to the 3' end of the templates before using them to synthesize complementary RNA. A similar TNTase activity of RCNMV RdRp could be speculated to add non-template nucleotides during the synthesis of the (-) MP transgene mRNA, resulting in the incorporation of these non-template nucleotides to the recombinant RNA molecules. As all the recombinant RNA molecules with non-template nucleotides contained the nearly complete transgene mRNA sequence at the 5' end, we can further speculate that the RCNMV TNTase activity could be stimulated due to no base pairing between the 3' end of the nascent (-) RNA and the receptor RNA (RNA-1). The switching over to the receptor RNA could then take place when some of the randomly incorporated nucleotides can form weak basepairing with RCNMV RNA-1. Our data did not rule out the possibility that the 5' 8-13 nt of the recombinant RNA-2 identical to RNA-1 sequences are generated randomly by the TNTase activity of RdRp in the same fashion as the generation of those unknown sequences, but this seems very unlikely.

The importance of proposed weak basepairing between the nascent (-) RNA and the acceptor RNA could also explain the effect of temperature on the frequency of transgenic recombination. A high frequency of transgenic recombination was observed when the Δ5 transgenic plants inoculated with RNA-1 were kept at low temperature (20°C day/20°C night), whereas no transgenic recombination was detected when the inoculated plants were kept at high temperature (30°C day/20°C night). Partial local hybridization between the nascent RNA and the acceptor RNA should be more stable at a lower temperature than at a higher temperature.

Although all the recovered recombinant RNA molecules contained the 5' terminal sequences from RCNMV RNA-1, crossover events during the transgenic recombination could occur all over the parental RNAs of transgene mRNA and the infecting RCNMV RNA-1. However, our assay would only detect those recombinant molecules that can replicate and provide the cell-to-cell movement function. These molecules must possess the entire MP ORF and has acquired the 5' terminal replication element. There were also indications that the recombinant molecules we detected could have rapidly evolved from unidentified intermediates. Sequence comparison around the crossover sites in molecules RΔ512A2L and RΔ512A2S revealed that they were almost identical except an internal 57-nt fragment was deleted in RΔ512A2S. Thus, RΔ512A2S could have been evolved from RΔ512A2L by an internal deletion; and RΔ512A2L might be an intermediate for the production of the final recombinant molecule RΔ512A2S. Our results showed that different recombinant RNA molecules, RΔ512A1, RΔ512A2L, RΔ512A4, and RΔ512A6, were produced in different plants of the same transgenic line Δ5-12A, indicating that

different recombination events can occur with the same set of parental RNA molecules. Furthermore, different recombinant molecules, RΔ512A2L and RΔ512A2M, were recovered from the same transgenic plant Δ5-12A-2, showing that more than one recombination event could occur in the same plant.

RCNMV transgenic recombination can lead to the production of virus causing different symptoms

The incorporation of non-viral sequences in the recombinant RNA-2 changed the viral symptom. Five of the seven recombinant RNA-2 molecules that contained non-viral sequences (vector or unknown origin) caused milder mosaic symptom, whereas two other molecules without any non-viral sequences produced wild type symptom (data not shown). It indicates that transgenic recombination can result in altered viral symptoms. Although the transgenic recombination observed in this study produced disease with milder symptom, it is possible that virus with high virulence may be produced *via* transgenic recombination.

Factors Affecting the Transgenic Recombination

The occurrence of transgenic recombination may be affected by many factors, such as selection pressure, the presence of viral terminal sequences in transgenes, the homology between the transgenic mRNA and the infecting viral RNA, the level of the transgenic mRNA and the infecting viral RNA, and the environmental conditions.

The high rate of recombination in Δ5 transgenic plant was detected under strong selection pressure (cell-to-cell movement). A similar strong selection pressure (systemic or cell-to-cell movement) is the key factor for the detection of transgenic recombination in CP transgenic plant in other systems, including CCMV (Greene and Allison, 1994), TBSV (Borja, *et al.*, 1999), TMV (Adair and Kearney, 2000), and PPV (Varrelmann *et al.*, 2000). In contrast, when GFP (with AUG start codon disrupted) transgenic plants were inoculated with TMV fused with mutant GFP (containing wild type AUG start codon), no selective pressure was present for detecting the TMV with wild type GFP. In this case, no recombinant TMV fused with wild type GFP was detected (Adair and Kearney, 2000).

Under the same conditions, a high frequency of recombination was detected in Δ5 transgenic plants expressing the 5'-terminally truncated RNA-2 transcripts, whereas no recombination was found in Δ3 transgenic plants expressing the 3'-terminally truncated RNA-2. This result indicates that the 3'-terminal sequence that is required for the initiation of minus-strand RNA synthesis is another important factor for transgenic recombination. This result supports the hypothesis that the 3' UTR of viral RNA was important for transgenic recombination Greene and Allison (1994) proposed from the transgenic recombination study in CCMV.

Another factor governing frequencies of transgenic recombination is the accessibility of transgene mRNA to virus replication machinery for recombination. Compartmentalization of viral RNA replication and preferential localization of mRNA to endoplasmic reticulum membranes could potentially create physical separation between

the two. To measures the accessibility of transgene transcripts to viral RNA polymerase, transgene mRNA containing the full-length RNA-2 sequence was expressed in the FL transgenic plants. The viral RNA in the transgene mRNA was flanked by 30 non-viral nucleotides at the 5' terminus and 28 non-viral nucleotides in addition to poly (A) tail at the 3' end. To capture the transgenic RNA-2, RCNMV RdRp needs to directly contact with the transgene mRNA, initiates (-) RNA synthesis internally at the correct nucleotide, and terminates at the 5' terminus of RNA-2 to remove the extraneous vector sequence. The process of replication capturing of a transgene by viral RNA does not require a RNA recombination event. The frequency of replication capturing is a simple measure of the accessibility of transgene mRNA to viral RdRp, the ability of RdRp to internally initiated (-) viral RNA synthesis and to internally terminate (-) RNA synthesis at the correct nucleotide. Our results clearly showed that replication capturing of transgene mRNA, as measured by the frequency of systemic infection (successful capture of transgene mRNA) of the FL transgenic lines inoculated with RNA-1, was positively correlated with the level of viral RNA inoculum and with the amount of the transgene mRNA in the plants (Table 4.1). Both viral RNA inoculum and the amount of the transgene mRNA could be the limiting factor in this interaction. Since physical contact between mRNA and viral RdRp is also a prerequisite for transgenic recombination, the amount of viral RNA and transgene mRNA could also influence the recombination frequency as have been demonstrated in this study. In practice, choosing of transgene silenced virus-resistant lines could dramatically reduce the potential transgenic recombination, which in turn would decrease the risk of using virus-resistant, transgenic plants.

REFERENCES

Aaziz R, Tepfer M. 1999a. Recombination between genomic RNAs of two cucumoviruses under conditions of minimal selection pressure. Virology 263:282-289.

Aaziz R, Tepfer M. 1999b. Recombination in RNA viruses and in virus-resistant transgenic plants. J Gen Virol. 80: 1339-1346.

Adair, T.L., and Kearney, C. M. 2000. Recombination between a 3-kilobase tobacco mosaic virus transgene and a homologous viral construct in the restoration of viral and nonviral genes. Arch Virol145:1867-83.

Alejska, M., Kurzyniska-Kokorniak, A., Broda, M., Kierzek, R., and Figlerowicz, M. 2001. How RNA viruses exchange their genetic material. Acta Biochim Pol. 48:391-407.

Bates, H. J., Farjah, M., Osman, T. A., and Buck, K. W. 1995. Isolation and characterization of an RNA-dependent RNA polymerase from Nicotiana clevelandii plants infected with red clover necrotic mosaic virus. J. Gen. Virol. 76:1483-1491.

Baulcombe D. 1994. Replicase-mediated resistance: a novel type of virus resistance in transgenic plants? Trends Microbiol. 2: 60-63.

Beachy, R. N. 1997. Mechanisms and applications of pathogen-derived resistance in transgenic plants. Curr Opin Biotechnol. 8:215-220.

Borja, M., Rubio, T., Scholthof, H. B., and Jackson, A. O. 1999. Restoration of wild-type virus by double recombination of tombusvirus mutants with a host transgene. Mol. Plant Microbe. Interact. 12:153-162.

Chetverin, A. B. 1999. The puzzle of RNA recombination. FEBS Letters 460:1-5.

Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S. F., Quer, J., Novella, I. S., and Holland, J. J. 1996. Basic concepts in RNA virus evolution. FASEB J. 10:859-864.

Figlerowicz, M. and Bujarski, J. J. 1998. RNA recombination in brome mosaic virus, a model plus strand RNA virus. Acta Biochimica Polonica 45:847-868.

Fujiwara, T., Giesman-Cookmeyer, D., Ding, B., Lommel, S. A., and Lucas, W. J. 1993. Cell-to-cell trafficking of macromolecules through plasmodesmata potentiated by the red clover necrotic mosaic virus movement protein. Plant Cell 5:1783-1794.

Garcia-Arenal F, Fraile A, Malpica JM. 2001. Variability and genetic structure of plant virus populations. Annu. Rev. Phytopathol. 39:157-86.

Greene, A. E. and Allison, R. F. 1994. Recombination between viral RNA and transgenic plant transcripts. Science 263:1423-1425.

Guan H, Simon AE. 2000. Polymerization of nontemplate bases before transcription initiation at the 3' ends of templates by an RNA-dependent RNA polymerase: an activity involved in 3' end repair of viral RNAs. Proc Natl Acad Sci U S A. 97:12451-12456.

Horsch, R. B., J. E. Fry, N. L. Hoffman, S. G. Rogers, and R. T. Fraley. 1985. A simple and general method for transferring genes into plants. Science 277:1229-1231.

Jarvis, T. C., and K. Kirkegaard. 1992. Poliovirus RNA recombination - Mechanistic studies in the absence of selection. EMBO J. 8:3135-3145.

Kaniewski, W. K., and Thomas, P. E. 1999. Field testing for virus resistance and agronomic performance in transgenic plants. Mol Biotechnol. 12:101-115.

Khatchikian, D., M. Orlich, and R. Rott. 1989. Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. Nature 340:156-157.

Kim, K. H. and Lommel, S. A. 1994. Identification and analysis of the site of -1 ribosomal frameshifting in red clover necrotic mosaic virus. Virology 200:574-582.

Kirkegaard, K., and D. Baltimore. 1986. The mechanism of RNA recombination in Poliovirus. Cell 47:433-443.

Kunkel, T. A., Roberts, J. D., and Zakour, R. A. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. 154:367-82.

Laemmli, U. K. 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 227:680-685.

Le Gall, O., Candresse, T., and Dunez, J. 1995. Transfer of the 3' non-translated region of grapevine chrome mosaic virus RNA-1 by recombination to tomato black ring virus RNA-2 in pseudorecombinant isolates. J Gen Virol. 76:1285-1289.

Lommel, S. A., M. Weston-Fina, Z. Xiong, and G. P. Lomonossoff. 1988. The nucleotide sequence and gene organization of red clover necrotic mosaic virus RNA-2. Nucl. Acids Res. 16:8587- 8602.

Mayo, M. A., and C. A. Jolly. 1991. The 5'-terminal sequence of potato leafroll virus RNA: Evidence of recombination between virus and host RNA. J. Gen. Virol. 72:2591-2595.

Meyers, G., N. Tautz, E. J. Dubovi, and H. J. Thiel. 1991. Viral cytopathogenicity correlated with integration of ubiquitin coding sequences. Virology 180:602-616.

Miller ED, Hemenway C. 1998. History of coat protein-mediated protection. Methods Mol Biol. 81: 25-38.

Monroe, S. and Schlesinger, S. 1983. RNAs from two independently isolated defective interfering particles of Sindbis virus contain a cellular TRNA sequence at their 5' ends. Proc. Natl. Acad. Sci. U.S.A. 80:3279-3283.

Munishkin A. V., Boronin L. A., and Chetverin A. B. 1988. An in vivo recombinant RNA capable of autocatalytic synthesis by Qβ replicase. Nature 333:473-475.

Murphy, F. A., C. M. Fauquet, M. A. Mayo, A. W. Jarvis, S. A. Ghabrial, M. D. Summers, G. P. Martelli,and D. H. L. Bishop. 1995. The Classification and Nomenclature of Viruses: Sixth Report of the International Committee on Taxonomy of Viruses. Arch. Virol., Supplementum 3. Springer Verlag, New York.

Murray, M. G., and W. F. Thompson. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8:4321-4325.

Nagy, P. D. and Simon, A. E. 1997. New insights into the mechanisms of RNA recombination. Virology 235:1-9.

Osman, T. A. M. and Buck, K. W. 1987. Replication of red clover necrotic mosaic virus RNA in cowpea protoplasts: RNA-1 replicates independently of RNA-2. J. Gen. Virol. 68:289-296.

Paje-Manalo, L. and Lommel, S. A. 1993. Idependent replication of red clover necrotic mosaic virus RNA-1 in electroporated host and non-host Nicotiana sp. protoplasts. Phytopathol. 79:457-461.

Ranjith-Kumar CT, Gajewski J, Gutshall L, Maley D, Sarisky RT, Kao CC. 2001. Terminal nucleotidyl transferase activity of recombinant Flaviviridae RNA-dependent RNA polymerases: implication for viral RNA synthesis. J Virol. 75:8615-8623.

Ravelonandro M, Scorza R, Callahan A, Levy L, Jacquet C, Monsion M, Damsteegt V. 2000. The use of transgenic fruit trees as a resistance strategy for virus epidemics: the plum pox (sharka) model. Virus Res. 71: 63-69.

Rovere, C. V , del Vas, M., and Hopp, H. E. 2002. RNA-mediated virus resistance. Curr Opin Biotechnol. 13:167-172.

Rubio, T., Borja., M., Scholthoff, H. B., and Jackson, A. O. 1999. Recombination with host transgenes and effect on virus evolution: an overview and opinion. Mol. Plant-Microbe Interact. 12:87-92.

Sambrook, J., E. F. Fritisch, and T. Maniatis. 1989. Molecular cloning, A laboratory manual. 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Schoelz, J. E., and W. M. Wintermantel. 1993. Expansion of viral host range through complementation and recombination in transgenic plants. Plant Cell 5:1669-1679.

Simon, A. E. and Bujarski, J. J. 1994. RNA-RNA recombination and evolution in virus infected plants. Annu. Rev. Phytopathology 32: 337-362.

Sit, T. L., Vaewhongs, A. A., and Lommel, S. A. 1998. RNA-mediated trans-activation of transcription from a viral RNA. Science 281:829-832.

Turner, R. L. and Buck, K. W. 1999. Mutational analysis of cis-acting sequences in the 3'- and 5'-untranslated regions of RNA-2 of red clover necrotic mosaic virus. Virology 252:115-124.

Vaewhongs, A. A. and Lommel, S. A. 1995. Virion formation is required for the long-distance movement of red clover necrotic mosaic virus in movement protein transgenic plants. Virology 212:607-613.

van der Kuyl, A. C., Neeleman, L., and Bol J. F. 1991. Complementation and recombination between alfalfa mosaic virus RNA3 mutants in tobacco plants. Virology 183:731-738.

Varrelmann M, Palkovics L, and Maiss E. 2000. Transgenic or plant expression vector-mediated recombination of Plum Pox Virus. J Virol. 74:7462-7469.

Weng, Z. and Xiong, Z. 1995. A method for accurate determination of terminal sequences of viral genomic RNA. Genome Research **5**:202-207.

Weng, Z. and Xiong, Z. 1997. Genome organization and gene expression of saguaro cactus carmovirus. J. Gen. Virol. 78:525-534.

White, K. A., and T. J. Morris. 1995. RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. RNA 1:1029-40.

Wintermantel, W. M., and J. E. Schoelz. 1996. Isolation of recombinant viruses between cauliflower mosaic virus and a viral gene in transgenic plants under conditions of moderate selection pressure. Virology 223:156-164.

Xiong, Z., and S. A. Lommel. 1989. The complete nucleotide sequence and genome organization of red clover necrotic mosaic virus RNA-1. Virology 171:543-554.

Xiong, Z., and S. A. Lommel. 1991. Red clover necrotic mosaic virus infectious transcripts synthesized *in vitro*. Virology 182:388-392.

Xiong, Z., D. Giesman-Cookmeyer, K. H. Kim, and S. A. Lommel. 1993a. The roles of the red clover necrotic mosaic virus capsid and cell-to-cell movement proteins in systemic infection. Virology 192:27-32.

Xiong, Z., K. H. Kim, T. L. Kendall, and S. A. Lommel. 1993b. Synthesis of the putative red clover necrotic mosaic virus RNA polymerase by ribosomal frameshifting. Virology 193:213-221.