

## ABSTRACT

Title of dissertation: RNA ELEMENTS INVOLVED IN TURNIP CRINKLE  
VIRUS REPLICATION AND VIRION ACCUMULATION

Xiaoping Sun, Doctor of Philosophy, 2006

Dissertation directed by: Professor Anne E. Simon  
Department of Cell Biology and Molecular Genetics

*Turnip crinkle virus* (TCV) (family *Tombusviridae*, genus *Carmovirus*) is a positive-strand RNA virus. Satellite RNA C (satC), a TCV-associated non-coding subviral RNA that depends on virus-encoded RNA-dependent RNA polymerase (RdRp) for replication, interferes with virion accumulation in vivo. Previous results suggested that the motif 1-hairpin (M1H) of satC, a replication enhancer on minus strands, forms a plus-strand hairpin that might be involved in virion assembly interference (G. Zhang and A. E. Simon. 2003. *J. Mol. Biol.* 326:35-48). Using in vivo functional SELEX (systematic evolution of ligands by exponential enrichment) and mutational analyses, I determined that M1H functions on plus strands by bringing together flanking cytidylate and adenylate (CA)-rich sequences that are involved in both satC replication and virion assembly interference.

Further investigation of the M1H-flanking CA-rich sequences using a second SELEX revealed three conserved sequence-specific elements. Examination of these three elements in the context of a control RNA suggested that the CCCA and CGGCGG elements are involved in satC replication with a synergistic effect in the presence of both. The CCCA and CAAAA elements are involved in the satC-mediated virion assembly interference.

Hairpin 4 (H4) is a hairpin located within the 3'-untranslated region of TCV genome. RNA structural probing of H4 in full-length TCV suggested that H4 folds into hairpin structures in both TCV plus and minus strands. In vivo mutagenesis determined that H4 is important for efficient accumulation of TCV in protoplasts, with a 98% reduction of genomic RNA levels when H4 was deleted. In vitro RNA transcription and gel mobility shift assays using p88 (TCV RdRp) suggested that H4 functions in TCV replication by enhancing synthesis of both strands through attracting the RdRp to the template.

RNA ELEMENTS INVOLVED IN TURNIP CRINKLE VIRUS REPLICATION AND  
VIRION ACCUMULATION

by

Xiaoping Sun

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2006

Advisory Committee:

Professor Anne E. Simon, Chair  
Associate Professor Douglas A. Julin  
Associate Professor James N. Culver  
Associate Professor Jonathan D. Dinman  
Investigator Kim Y. Green  
Research Chemist Robert A. Owens

©Copyright by

Xiaoping Sun

2006

## ACKNOWLEDGEMENTS

I would like to thank all people who helped me during the completion of this dissertation. First of all, I want to thank my supervisor Dr. Anne Simon for providing me with this great opportunity to work in her lab. In the past five years, Anne helped me know how to be a good scientist. Her enthusiasm and dedication to science and her encouragement inspired me and will continue to influence me in my whole life. Anne also gave me enormous help in improving my spoken and written English as well as presentation skills. I also appreciate that she supported me to attend several national and international virology conferences, which largely widened my view and gave me excellent chances to communicate with other scientists. I am also deeply indebted to my committee members, Dr. Douglas A. Julin, Dr. James N. Culver, Dr. Jonathan D. Dinman, Dr. Kim Y. Green, and Dr. Robert A. Owens, for their valuable support and advices.

I would like to thank Dr. Guohua Zhang for great technical assistance, especially for performing the 10-base SELEX. I would like to thank my colleagues Dr. Sohrab Bodaghi, Dr. Vera Stupina, Dr. Jiuchun Zhang, John C. McCormack, and Rong Guo for their valuable discussion and friendship. I want to thank Catherine P. Gretschel and Nancee Soni for their help in managing lab business. I also want to thank Dr. Jianlong Wang, Dr. Feng Qu, and Dr. Peter Nagy for answering many scientific questions.

I am grateful to my wife Fengli Zhang for her love and encouragement. I would like to thank my son Qinwei Sun, whose interests in science always encourage me. I also want to thank my parents for their care, love, encouragement, and support.

## TABLE OF CONTENTS

LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
LIST OF ABBREVIATIONS.....	ix
CHAPTER I: RNA ELEMENTS INVOLVED IN REPLICATION OF POSITIVE STRAND RNA VIRUSES: AN OVERVIEW.....	1
Introduction.....	1
<i>Cis</i> -acting RNA replication elements.....	3
Core promoters.....	4
Elements involved in genome circulation.....	9
Template recruitment elements.....	10
Elements serving as template for primer formation.....	12
Replication enhancers.....	13
Replication Repressors.....	15
3'-terminal molecular switch.....	16
<i>Trans</i> -acting RNA replication elements.....	18
TCV as a useful system for studying viral RNA elements involved in replication.....	19
CHAPTER II: FITNESS OF SATELLITE RNA C CORRELATES WITH M1H(+) AND FLANKING SEQUENCES THAT ENHANCE REPLICATION AND REPRESS THE ACCUMULATION OF VIRIONS.....	30
Introduction.....	30
Materials and Methods.....	33
Large-scale plasmid DNA isolation.....	33
Small-scale plasmid DNA isolation.....	34
DNA sequencing.....	35
In vitro RNA synthesis using T7 polymerase.....	35
Plant Growth and Inoculations.....	36
Small-scale RNA extraction from turnip leaves.....	36
In vivo functional SELEX using 10-base M1H replacement Sequences.....	37
Secondary structure analysis of SELEX winners.....	41
Fitness comparison of the 3 <sup>rd</sup> round SELEX winners in turnip plants.....	41
Construction of satC variants.....	42
Protoplast preparation and inoculation.....	42
Total RNA extraction from protoplasts.....	44
Northern blot hybridization.....	44
Protein extraction from protoplasts and electrophoresis analysis.....	45
Virion isolation from protoplasts and Western blotting analysis.....	46

Results.....	47
In vivo functional SELEX of satC containing 10 random bases replacing M1H.....	47
Sequence and structural composition of the 2 <sup>nd</sup> round winning sequences.....	52
Effect of 10-base winning sequences on satC replication in protoplasts.....	60
Two most fit 3 <sup>rd</sup> round winners interfere with TCV virion accumulation in protoplasts.....	63
Role of CA-rich sequence flanking M1H in satC replication and virion accumulation.....	64
Discussion.....	69
CHAPTER III:    SHORT INTERNAL SEQUENCES INVOLVED IN REPLICATION OF SATELLITE RNA C AND VIRION ACCUMULATION.....	73
Introduction.....	73
Materials and Methods.....	74
In vivo functional SELEX.....	74
Fitness comparison of SELEX winners in turnip plants.....	75
Construction of satC variants.....	78
Protoplast inoculation and RNA and virion preparation.....	78
Northern blot analysis.....	78
Western blot analysis.....	79
Results.....	79
In vivo functional SELEX of the satC LS and RS regions.....	79
Winning sequences revealed by in vivo SELEX contribute to satC replication and repression of virion accumulation.....	87
CCCA and CGGCGG enhance replication of satC.....	90
CCCA and CAAA affect virion accumulation.....	95
Discussion.....	95
CHAPTER IV:    H4 FUNCTIONS IN BOTH ORIENTATIONS TO ENHANCE VIRAL REPLICATION.....	99
Introduction.....	99
Materials and Methods.....	100
RNA solution structure probing.....	100
Construction of TCV mutants.....	102
Protoplast inoculation.....	108
Northern blot hybridization.....	108
In vitro RNA translation using wheat germ extracts.....	108
Purification of p88 from <i>E. coli</i> .....	109
In vitro RNA transcription using purified recombinant TCV p88.....	111
p88 binding assays.....	112
Results.....	113
Solution structure probing of the H4 region in TCV plus and minus stands.....	113

H4 is important for efficient accumulation of TCV in vivo.....	119
H4 enhances replication of TCV RNAs.....	124
H4(+) and H4(-) have enhancer activity in vitro.....	128
p88 binds to H4(+) and H4(-).....	134
Discussion.....	135
CONCLUSIONS.....	144
REFERENCES.....	149



## LIST OF TABLES

Table	Page
2.1 Oligonucleotides used in Chapter II.....	40
2.2 2 <sup>nd</sup> round SELEX winners.....	50
3.1 Oligonucleotides used in Chapter III.....	77
3.2 Results of in vivo functional SELEX.....	82
3.3 Competition of SELEX winners for fitness in plant.....	86
4.1 Oligonucleotides used in Chapter IV.....	103

## LIST OF FIGURES

Figure	Page
1.1 Core promoters for viral RNA syntheses.....	5
1.2 Genomic and subviral RNAs in the TCV system.....	20
1.3 Sequence and structure of the 3' regions of satC and TCV.....	23
1.4 Minus-strand elements involved in replication of satC.....	27
2.1 Structure folded in the M1H region of satC.....	32
2.2 Flow chart for performing the 10-base SELEX of satC M1H.....	38
2.3 Forty-three M1H replacement sequences recovered from the 1st round selection of the 10-base SELEX.....	49
2.4 Motifs found in 10-base SELEX M1H replacement sequences.....	54
2.5 Computer predicted plus-strand structures encompassing the M1H replacement sequences.....	56
2.6 Accumulation of viral RNAs and virions in protoplasts.....	61
2.7 Effect of the A5C insert in 3rd round SELEX winner II-18 on replication and virion accumulation.....	65
3.1 Flow chart for performing the SELEX of satC LS and RS.....	76
3.2 Sequence and structure of a portion of satC 3' region.....	81
3.3 Replication and virion repression of 2nd round and selected 1st round SELEX winners.....	88
3.4 Effect on replication and virion repression of individual elements identified by in vivo SELEX.....	92
4.1 Structures of TCV H4(+) and CCFV H4(+).....	101
4.2 Flow chart for the construction procedure of M1.....	106
4.3 The pMAL/TCVp88 plasmid.....	110
4.4 Chemical and enzymatic probing of H4(+).....	114

4.5	Chemical and enzymatic probing of H4(-).....	117
4.6	Mutational analysis of H4.....	120
4.7	Effect of H4 on replication of artificial non-coding subviral RNAs.....	126
4.8	Analysis of in vitro translation products of viral RNA templates in wheat germ extracts.....	129
4.9	Effect of H4 on transcription in vitro using recombinant p88.....	130
4.10	Preferential binding of p88 to H4(+) and H4(-).....	136
5.1	Model for the connection between virion repression and satC replication.....	147

## LIST OF ABBREVIATIONS

- 3' CSE: 3' Conserved sequence element
- 3' NCR: 3' Noncoding region
- 3' PE: 3' Proximal element
- AMV: Alfalfa mosaic virus
- BcoV: Bovine coronavirus
- BMV: Brome mosaic virus
- BYDV: Barley yellow dwarf virus
- CCFV: Cardamine chlorotic fleck virus
- CCS: Carmovirus consensus sequence
- CP: Coat protein
- CRE: cis-replication element
- FHV: Flock house virus
- GDD: Glycine/Aspartic acid/Aspartic acid
- H4: Hairpin 4
- H5: Hairpin 5
- IPTG: isopropyl- $\beta$ -D-thiogalactopyranoside
- M1H: Motif1-hairpin
- M3H: Motif3-hairpin
- MB: Growth medium
- MBP: Maltose-binding protein
- MHV: Mouse hepatitis virus

PABP1: Poly(A) binding protein 1

PCBP: Poly(rC) binding protein

PCM: Protoplast culture medium

PIM: Protoplast isolation medium

Pr: Promoter

RCNMV: Red clover necrotic mosaic virus

RdRp: RNA-dependent RNA polymerase

satC: Satellite RNA C

SELEX: Systematic evolution of ligands by exponential enrichment

SIN: Sindbis virus

TBSV: Tomato bushy stunt virus

TCV: Turnip crinkle virus

TLS: tRNA-like structure

TMV: Tomato mosaic virus

UTR: Untranslated region

VPg-pUpU: Uridylylated genome-linked protein

$\Psi$ 1: Pseudoknot

## **CHAPTER I**

### **RNA ELEMENTS INVOLVED IN REPLICATION OF POSITIVE-STRAND**

#### **RNA VIRUSES: AN OVERVIEW**

##### **Introduction**

A majority of serious human, animal, and plant viruses contain positive-strand RNA genomes that encode genetic information required for viral replication, infection, and virion assembly. After invading the host cell, the viral genome is released from the capsid, recruited by ribosomes, and translated to produce proteins including RNA-dependent RNA polymerase (RdRp), the key enzyme required for virus replication. The replicated progeny genomic RNAs are then encapsidated by virus-encoded coat protein (CP) and/or other structural proteins to form virions.

Despite apparent differences in genome organization, virion morphology, and host range, all positive-strand RNA viruses share a similar replication strategy, i.e., the genomic RNA serves as template in transcription of complementary minus-strand RNAs, and the newly synthesized minus strands are then used as templates for synthesis of large quantities of progeny plus strands. The relative levels of the two strands in host cells are often highly asymmetric, with ratios of up to 1000 plus strands for every minus strand produced (Buck, 1996). Viral replication is mediated by a replicase complex comprising the viral RdRp, possibly together with other viral proteins and/or host factors (Lai, 1998).

High-resolution crystal structure analyses of several representative viral RdRps show that RdRp have a similar overall shape resembling a right hand with three subdomains known as “thumb”, “fingers”, and “palm”, similar to DNA-dependent DNA polymerases, RNA-dependent DNA polymerases, and DNA-dependent RNA polymerases (Baker and Bell, 1998; Bruenn, 2003; Ollis et al., 1985; O’Reilly and Kao, 1998; van Dijk et al., 2004). Consistent with these basic structural features, all viral RdRps catalyze RNA syntheses via similar polymerization reactions. Copying of the template begins from the 3’ end of the RNA molecule for genomic RNA synthesis.

Most positive-strand RNA viruses initiate synthesis *de novo* (i.e., without the need of a primer) that requires specific nucleotides positioned at the initiation site (Buck, 1996). Some positive-strand RNA viruses direct primer-dependent synthesis initiation, such as poliovirus (family *Picornaviridae*, genus *Poliovirus*), which needs uridylylated genome-linked protein (VPg-pUpU) as primer for synthesis of the genomic RNA (Xiang et al., 1997). Viral RdRp-directed RNA synthesis initiates from an established initiation complex generally comprising template, RdRp, initiation nucleotide, and a second nucleotide (Kao et al., 2001; van Dijk et al., 2004). After incorporation of several (usually 8-10) nucleotides, the synthesis transits from the initiation step to the elongation step, which results in successful copying of the full-length template, or abortive synthesis by premature termination (Holstege et al., 1997).

Prior to establishment of the initiation complex, viral RdRp must selectively and specifically associate with its cognate RNA and be properly positioned at the transcription initiation site. This RNA-protein recognition is mediated by the interaction between RdRp and specific RNA sequence and/or structural features, known as RNA

replication elements. During the past several decades, a large number of RNA replication elements have been identified for positive-strand RNA viruses using several approaches such as assaying virus accumulation in whole plants or heterologous systems like *Saccharomyces cerevisiae*, transfection of genomic RNAs into protoplasts, and in vitro RNA synthesis using RdRp purified from virus-infected hosts or expressed from *E. coli*. Characterization of these RNA elements is useful for facilitating our understanding of the mechanisms underlying virus replication. In this chapter, I will present the status of our knowledge of these RNA elements with an emphasis on how they regulate viral replication. It is also a goal of this overview to stimulate ideas regarding potential mechanisms involving these elements, which will help in future studies.

### ***Cis-acting* RNA replication elements**

Most RNA replication elements identified to date are present in the RNA template and therefore function in *cis*. Although the RdRp directs *de novo* or primer-dependent initiation of RNA synthesis from the 3' end of the template, RNA replication elements are found at variable positions relative to the initiation sites, such as throughout the 3'-untranslated regions (UTRs), 5' UTRs, open reading frames, and intercistronic regions. These elements participate in regulation of replication via diverse mechanisms, which will be discussed below.

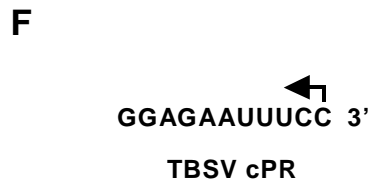
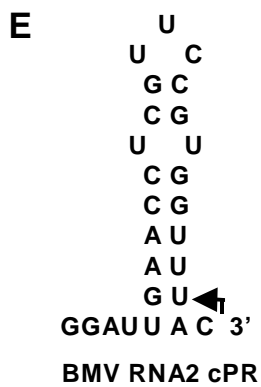
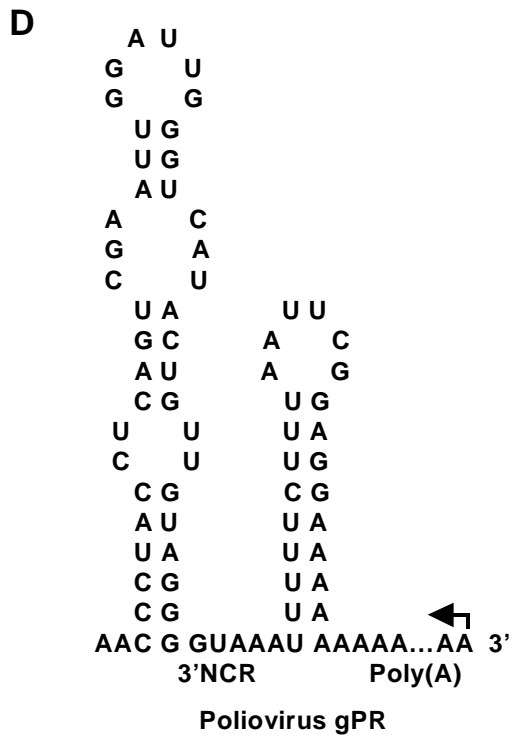
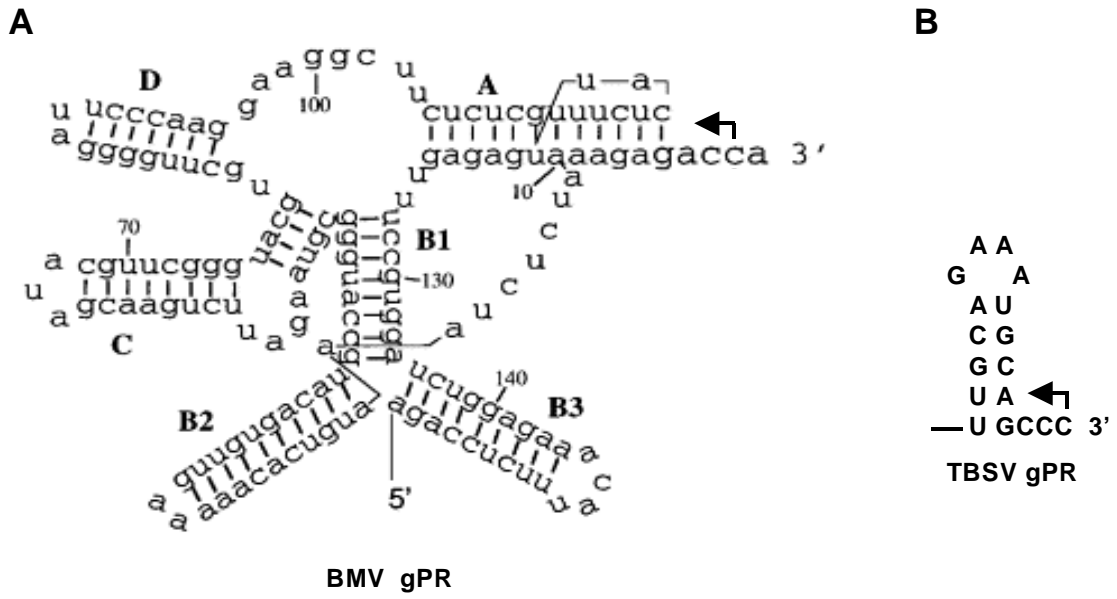


## Core promoters

*Cis*-acting elements required for basal-level replication or transcription are termed core promoters, which, for most RNA viruses, contain the RNA transcription initiation sites (Buck, 1996; Chapman and Kao, 1999; de Graaff and Jaspars, 1994; Dreher, 1999; Duggal et al., 1994; Turner and Buck, 1999). Core promoters are generally located at the 3' ends of the templates and must be recognized by the RdRp to promote synthesis of complementary strands, whereas other elements might be responsible for the initial attraction of RdRp to the RNA molecule. Core promoters usually contain secondary or tertiary structural elements flanked on the 3' side by a short single-stranded sequence (Dreher, 1999).

One of the best-characterized examples of a core promoter is that of *Brome mosaic virus* (BMV) (family *Bromoviridae*, genus *Bromovirus*). The BMV core promoter is an approximately 150-nucleotide (nt) tRNA-like structure (TLS) present at the 3' ends of all three genomic RNAs (Chapman and Kao, 1999; Dreher and Hall, 1988; Dreher et al., 1984; Sullivan and Ahlquist, 1997; Figure 1.1A). Mutational analyses suggested that the TLS is required for *in vivo* minus-strand synthesis (Dreher and Hall, 1988). In the absence of other sequence, the TLS alone is sufficient for directing minus-strand synthesis in an RNA transcription assay using RdRp partially purified from BMV-infected plants (Dreher et al., 1984). Extensive mutational analyses of the TLS indicated that alterations in the loop regions of stem C reduced the ability of the TLS to interact with the RdRp (Chapman and Kao, 1999). Stem C, when directly linked to the 3' terminal initiation sequence (ACCA), can serve as an efficient template to direct RNA

**Figure 1.1** Core promoters for viral RNA syntheses. The arrows indicate RNA synthesis initiation sites. (A-D) Minus-strand initiation promoters (termed gPR) of BMV (Chapman and Kao, 1999, with permission from Elsevier), TBSV (Fabian et al., 2003; Nagy and Pogany, 2000), SIN (Levis et al., 1986; Hardy, 2006), and poliovirus (Sarnow, 1989; Silvestri et al., 2006), respectively. The gPRs are located at the 3' ends of the genomic RNAs. Specific sequence and structural features contained in these promoters are described in the text. (E-F) Plus-strand initiation promoters (termed cPR) of BMV RNA2 (Sivakumaran et al., 1999) and TBSV (Panavas et al., 2002), respectively. The cPRs are located at the 3' ends of minus strands.



synthesis in vitro, suggesting that stem C is an essential component of the TLS for its promoter activity (Chapman and Kao, 1999).

For viruses lacking a terminal TLS, such as *Tomato bushy stunt virus* (TBSV), the type member of the genus *Tombusvirus* (Family *Tombusviridae*), the core promoter for minus-strand synthesis initiation is composed of the 3'-terminal unpaired tail (CCC<sub>OH</sub>) and a hairpin positioned just upstream (Figure 1.1B). Disruption of the stem of the hairpin abolished RNA replication in protoplasts, whereas compensatory mutations resulted in approximately 55% recovery, suggesting that structure of this hairpin is required for efficient viral replication (Fabian et al., 2003). Sequence changes in the terminal loop of the hairpin also dramatically inhibited RNA replication, leading to a reduction of 85% of wt activity (Fabian et al., 2003). This promoter can also direct efficient in vitro minus-strand synthesis using partially purified tombusviral RdRp (Nagy and Pogany, 2000).

In the case of viruses containing a poly(A) tail, the poly(A) tail and nearby sequences are usually required for minus-strand synthesis. For example, the poly(A) tail of *Sindbis virus* (SIN) (family *Togaviridae*, genus *Alphavirus*), along with the upstream 19-nt conserved sequence element (3' CSE), can fully support genome replication in vitro (Levis et al., 1986; Figure 1.1C). Recent in vitro RNA synthesis assays indicated that the poly (A) tail and the adjacent three CSE residues are responsible for localization of the initiation site, with the residue immediately preceding the poly(A) tail representing the predominant initiation site (Hardy, 2006). The poly (A) tail may be added in vivo by cellular cytoplasmic polyadenylation machinery in a template-independent manner (Hardy, 2006). In the case of poliovirus, the poly(A) tail and

adjacent 3' noncoding region (3' NCR) are also required for optimal promoter activity (Sarnow, 1989; Figure 1.1D). The 3' NCR is proposed to participate in selection of the initiation site at or near the end of the poly(A) tail (Brown et al., 2005). The length of the poly(A) tail is also important (at least 20-nt) for achieving wt or nearly wt levels of replication (Silvestri et al., 2006).

Core promoters required for plus-strand synthesis are usually located at the 3' ends of minus strands. Unlike those of plus strands, the 3' ends of minus strands are generally less structured. For example, biochemical, nuclear magnetic resonance, and thermodynamic analyses of BMV RNA2 indicated that the secondary structure of the promoter required for plus-strand synthesis, which is located at the 3' ends of minus strands, can be disrupted without significantly affecting promoter activity (Sivakumaran et al., 1999; Figure 1.1E). In contrast, position-specific nucleotide alterations resulted in reduced replicase recognition and decreased RNA2 accumulation, suggesting that the core promoter for RNA2 plus-strand synthesis is sequence-specific (Sivakumaran et al., 1999). For tombusviruses, the core promoter for plus-strand synthesis is the 3'-terminal 11-nt sequence in minus strands, which directs efficient RNA synthesis *in vitro* (Panavas et al., 2002; Figure 1.1F).

As described above, core promoters are required for viral replication *in vivo*. In the absence of other sequences, core promoters alone are capable of efficiently directing *in vitro* RNA syntheses using purified RdRp. However, some viral RNAs with deletion of core promoters can support a low level of replication *in vivo* (Todd et al., 1997; Wu et al., 2001), suggesting that additional RNA elements are involved in viral replication, which can replace the function of core promoters to some degree.

## **Elements involved in genome circularization**

Although RNA synthesis initiates at the 3' end of the template, RNA elements located at the 5' end can be important for replication of at least some viruses. One example is the 5'-terminal cloverleaf-like structure of poliovirus, which comprises stem A and stem-loops B, C, and D (Andino et al., 1990). Stem-loops B and D are bound by cellular poly(rC) binding protein (PCBP) and the uncleaved viral protease polymerase precursor 3CD<sup>pro</sup>, respectively, to form a ribonucleoprotein complex (Andino et al., 1990, 1993; Gamarnik and Andino, 1997; Parsley et al., 1997; Silvera et al., 1999). Destruction of the cloverleaf-like structure by deleting 4 nt from the stem region of stem-loop D resulted in undetectable RNA synthesis in vitro, indicating that the cloverleaf-like structure is important for poliovirus replication (Barton et al., 2001). In addition, insertion of 4 nt into the loop region of stem-loop D dramatically inhibited minus-strand synthesis, suggesting that the loop region is sequence-specific (Barton et al., 2001).

The involvement of the 5' cloverleaf structure in minus-strand synthesis that initiates at the 3' end suggests a possible interaction between the 5' and 3' ends, resulting in genome circularization. Genome circularization is also supported by finding that poly(A) binding protein 1 (PABP1) interacts with both the 3' poly(A) tail and the 5' ribonucleoprotein complex containing the cloverleaf-like structure, PCBP, and 3CD<sup>pro</sup> (Barton et al., 2001; Herold and Andino, 2001). Therefore, 3CD<sup>pro</sup> is proposed to interact with the 3' poly(A) tail through binding to the 5'-terminal cloverleaf-like structure of the circularized genome. In addition, genome circularization requires a wild-type 5' end, which enhances the fidelity of viral replication by ensuring that templates with defective 5' ends cannot be copied.

In the case of *Viral bovine diarrhoea virus* (family *flaviviridae*), which lacks a poly(A) tail, genome circulation is mediated by the host NFAR protein, which specifically binds to the 5' and 3' UTRs (Isken et al., 2003). In addition, genome circularization mediated by direct RNA-RNA interaction is found for flaviviruses (family *Flavivirida*), including *West Nile virus*, *Dengue virus*, *Japanese encephalitis virus*, *Murray Valley virus*, and *Yellow fever virus*. A putative circulation sequence (10 to 18 nucleotides) located within the capsid protein-coding region near the 5' end is thought to base-pair with its complementary sequence in the 3' end, therefore mediating genome circulation (Corver et al., 2003; Hahn et al., 1987; Khromykh et al., 2001; You et al., 2001). Circulation of *Dengue virus* has been recently confirmed by direct observation of individual viral RNA molecules in a circular conformation by atomic force microscopy (Alvarez et al., 2005).

### **Template recruitment elements**

Specific replication of RNA viruses in host cells is dependent on proper selection and recruitment of the cognate RNA template into membrane-associated compartments, where ribosomes and most cellular RNAs might be excluded (Bolten et al, 1998; Restrepo-Hartwig and Ahlquist, 1996). In vitro and in vivo studies revealed that an approximately 150-nt sequence located within the BMV RNA3 intergenic region is necessary for template selection by the replication protein 1a, which contains a N-terminal domain with methyltransferase activity and a C-terminal domain with helicase activity (Sullivan and Ahlquist, 1999). RNA structural probing indicated that this 150-nt RNA selection signal folds into a stable stem-loop structure with the terminal loop

sequence matching the box B consensus sequence of cellular RNA polymerase III promoters and thus also the conserved T $\psi$ C loop of tRNAs (Baumstark and Ahlquist, 2001; French and Ahlquist, 1987; Marsh and Hall, 1987). Deletion of the box B element dramatically reduced RNA3 replication in both yeast (Sullivan and Ahlquist, 1999) and plant cells (Pogue et al., 1992; Smirnyagina et al., 1994).

The box B sequence is also found in the 5' UTRs of RNA1 and RNA2 of BMV (Chen et al., 2001). In *Saccharomyces cerevisiae* expressing BMV proteins 1a and 2a, which contains a large central domain conserved in RdRps, levels of the 1a-induced recruitment of various RNA2 box B mutants correlated with their abilities to function as templates during replication (Chen et al., 2001). Furthermore, the box B-containing 5' UTR is sufficient to mediate 1a-dependent recruitment of nonviral RNA to endoplasmic reticulum membranes, where BMV RNAs are synthesized (Chen et al., 2001).

Using mobility shift assays, it has been recently determined that TBSV replication protein p33 specifically binds to RII(+)-SL, a stem-loop structure located within the p92-coding region. p92 is the TBSV RdRp and is translated as a ribosomal readthrough product of p33 (Pogany et al., 2005). This RNA-protein interaction is functionally relevant to viral replication in vivo; mutations in RII(+)-SL that disrupt protein binding in vitro also dramatically inhibited RNA replication in plant and *Saccharomyces cerevisiae* cells (Pogany et al., 2005). In vivo analyses of novel RNA-based temperature-sensitive genomic mutants revealed that RII(+)-SL functions mainly at an early step such as template recruitment and/or replicase complex assembly, but is not important at later stages of infection (Monkewich et al., 2005).

Like TBSV p33, the 126-kDa replication protein of *Tomato mosaic virus* (TMV) is



much more abundant in host cells than its ribosomal readthrough product, the 183-kDa protein that contains the polymerase domain. The 3'-terminal TLS of TMV is specifically bound by the 126-kDa protein as assayed by in vitro RNA binding and UV cross-linking (Osman and Buck, 2003; Osman et al., 2000). Therefore, it is possible that the TMV 3'-terminal TLS may also function in template recruitment by interacting with the 126-kDa protein (Osman and Buck, 2003).

### **Elements serving as template for primer formation**

Polioviral replication needs uridylylated genome-linked protein (VPg-pUpU) as primer (Xiang et al., 1997). Uridylylation of VPg for use in minus-strand synthesis is performed by the viral polymerase 3D<sup>pol</sup> using the 3'-terminal poly(A) tail as template (Paul et al., 1998). In addition to the poly(A) tail, an internal stem-loop structure located within the 2C coding region, termed *cis*-replication element (CRE), is also capable of acting as template for formation of VPg-pUpU (Goodfellow et al., 2000; Morasco et al., 2003; Murray et al., 2003; Paul et al. 2000; Rieder et al., 2000). Although both the poly(A) tail and CRE can act as template for VPg uridylylation, there are several pieces of evidence suggesting that CRE has apparent advantages over the poly(A) tail in the process (Paul et al. 2000; Rieder et al., 2000). First, as examined in vitro, the template activity of CRE is much higher than that of the poly(A) tail. Second, using a specific, internally positioned CRE as the template for VPg uridylylation reduces the possibility that cellular mRNAs bearing a poly(A) tail are copied. Third, the presence of another viral protein, 3CD<sup>pro</sup>, largely enhances the template activity of CRE, which might provide additional specificity for viral replication. Finally, when the poly(A) tail serves as

template, uridylylation of VPg primarily uses  $Mn^{2+}$  as cofactor of  $3D^{pol}$ , while the reaction uses  $Mg^{2+}$  as cofactor when CRE is used as template. This feature would also increase the specificity of RNA synthesis, because  $Mn^{2+}$  is generally known to decrease polymerase specificity for their templates (Arnold et al., 1999; Tabor and Richardson, 1989).

As a highly conserved element, CRE is also found in various positions of other picornaviruses such as the polyprotein VP1 coding sequence of *Human rhinovirus 14*, VP2 coding sequence of *Theiler's murine encephalitis virus*, 2A coding sequence of *Human rhinovirus 2*, and the 5' UTR of *Foot-and-mouth disease virus* (Gerber et al., 2001; Goodfellow et al., 2000; Lobert et al., 1999; Mason et al., 2002; McKnight and Lemon et al., 1996; Paul et al., 2000; Yang et al., 2002). Although containing different nucleotide sequences and being positioned in different regions of the viral genomes, these CREs presumably function similar to that of poliovirus.

### **Replication enhancers**

In host cells, the amounts of plus and minus strands of positive-strand RNA viruses are highly asymmetric with ratios up to 1000-fold (Buck, 1996). However, studies of some viruses indicated that the strength distinction between the plus- and minus-strand promoters is not sufficient for achieving this high asymmetric ratio, suggesting that additional elements may be involved (Nagy and Pogany, 2000; Panavas et al., 2002a; 2002b; Sivakumaran et al., 2000). Consistent with this consideration, replication enhancers, which are not required for basal levels of RNA synthesis but effectively enhance RNA synthesis, have been recently identified for a number of positive-strand

RNA viruses (Kim and Makino, 1995; Nagy et al., 1999; Pogany et al., 2003; Ray and White, 1999; 2003).

Using partially purified tombusviral RdRp, one replication enhancer, termed RIII(-), has recently been identified for TBSV. This enhancer, which is internally located on minus strands, stimulates promoter-directed plus-strand RNA synthesis up to 20-fold in vitro (Panavas and Nagy, 2003; 2005). Deletion of RIII(-) reduced RNA accumulation by 10-fold in *Nicotiana benthamiana* protoplasts (Ray and White, 1999; 2003). Computer predictions and RNA structural probing indicated that RIII(-) is composed of two hairpins linked together by a short bridge sequence. Further deletion analyses indicated that these two hairpins have interchangeable roles, with either of them, in the presence of the bridge sequence, being capable of efficiently stimulating RNA synthesis (Panavas and Nagy, 2003). Gel mobility shift assays indicated that the two hairpins are replicase binding sites, suggesting that RIII(-) enhances plus-strand synthesis by facilitating recognition of the template by the replicase (Panavas and Nagy, 2005). Deletion of the bridge sequence resulted in a 30% reduction in plus-strand accumulation in protoplasts. Further studies indicated that the bridge sequence basepairs with sequences near the 3'-terminal promoter for plus-strand synthesis. Therefore, it is possible that the bridge sequence contributes to RNA synthesis by bringing the promoter in close proximity with the replicase-binding site of the enhancer, where the RdRp is recruited to the template (Panavas and Nagy, 2005).

## Replication Repressors

Replication repressors are elements that repress transcription in vitro via RNA-RNA interactions and are required for in vivo replication. Such elements were recently found for tombusviruses. RNA structural probing and phylogenetic analysis suggested that the asymmetric internal loop sequence (5'GGGCU) of a tombusviral 3'-proximal hairpin (SL3) base pairs with the 3' terminus (AGCCC-OH) of the RNA (Pogany et al., 2003). In vitro RNA transcription indicated that changing the GGGCU to GGGGU or alteration of AGCCC to ACCCC, which weakens or disrupts base pairing, increased RNA syntheses by 3- and 9-fold, respectively. The compensatory mutation reduced RNA synthesis to a level near the wt. These results suggest that the RNA-RNA interaction may mediate RNA synthesis repression by altering the 3'-terminal RNA structures that shield the terminal transcription initiation site from recognition of the RdRp (Pogany et al., 2003). Recent in vitro RNA transcription assays using tombusviral RdRp partially purified from yeast revealed that the interaction between SL3 and the 3' terminus plays a role in the replicase assembly at the 3'-terminal initiation site prior to minus-strand synthesis initiation (Panaviene et al., 2005). Closely related carmoviruses have a hairpin that is located in similar position relative to the 3' end and is also proposed to be necessary for correct replicase assembly in vivo (McCormack and Simon, 2004), while the internal loop that pairs with the 3' terminus is symmetric (Zhang et al., 2004a; 2004b). The reason for symmetry in the internal loop of the carmoviral hairpin is not known.

### **3'-terminal molecular switch**

As described above, replication enhancers and putative repressors positioned on opposite strands of TBSV may contribute to regulation of highly asymmetric syntheses of plus and minus strands in host cells (Pogany et al., 2003). However, at least some positive-strand RNA viruses require a 3'-terminal conformational switch to temporally coordinate plus- and minus-strand synthesis as well as translation. One of the best examples is the switch between two mutually exclusive conformations, a stem-loop structure and a TLS, found within the 3' UTR of *Alfalfa mosaic virus* (AMV) and closely related ilarviruses (family *Bromoviridae*) (Olsthoorn et al., 1999). The stem-loop structure contains coat protein (CP)-binding sites (Houser-Scott et al., 1994; Reusken et al., 1996) and the binding of CP to the inoculum RNAs is required for infection initiation (Bol, 1999; Jaspars, 1999). The TLS is recognized by the RdRp and is essential for minus-strand synthesis (Olsthoorn et al., 1999). Several pieces of evidence suggest that under physiological conditions, the 3' end of the genomic RNA is predominantly in the TLS conformation (Olsthoorn et al., 1999). Later in the infection process, the newly synthesized CP shifts the balance toward the stem-loop structure by competing with the RdRp for binding to the 3' UTR, resulting in transition from the TLS to the stem-loop structure. This transition shuts off minus-strand synthesis, thereby contributing to the asymmetric ratio between plus and minus strands (Olsthoorn et al., 1999). This conformational switch model is strongly supported by biochemical and functional assays using transgenic plants and protoplasts that constitutively express viral replication proteins (Olsthoorn et al., 1999), while it is not confirmed by recent experiments using

nontransgenic protoplasts (Petrillo et al., 2005). The inconsistent results may be caused by intracellular environment differences between transgenic and nontransgenic cells.

The coronavirus *Mouse hepatitis virus* (MHV) (family *Coronaviridae*) and its closely related *Bovine coronavirus* (BcoV) share more than 90% sequence similarity in the 5' regions of their 3' UTRs. Within this region, a bulged stem-loop and an adjacent downstream pseudoknot structure, which are mutually exclusive due to partial overlapping of the stem regions, are both essential for viral replication (Hsue and Masters, 1997; Hsue et al., 2000; Williams et al., 1999). Genetic analyses suggested that these two structures are alternate states of a molecular switch regulating a possible transition occurring during viral replication (e.g. formation of one of them may turn off the function of the other by disrupting its structure since they are mutually exclusive) (Goebel et al., 2004).

Computer predictions, RNA structural probing, and phylogenetic analyses suggested that the 3'-terminal 109-nt sequence of *Barley yellow dwarf virus* (BYDV) (family *Luteoviridae*, genus *Luteovirus*) folds into a structure that incorporates the 3' terminus into a coaxially stacked helix. Replication assays in oat protoplasts indicated that disruption of the 3'-terminal structure causes increased production of less-than-full-length minus strands and undetectable production of plus strands, suggesting that free 3' terminus facilitates minus-strand synthesis. Therefore, the embedded 3'-end structure and the alternative structure with a free 3' terminus may be components of a conformational switch regulating minus- and plus-strand synthesis during viral replication. In addition, mutations destabilizing the long-distance base pairing between the 3'-terminal five bases

of Q $\beta$  bacteriophage and an internal sequence allows for access of the RdRp to the 3' end, thereby resulting in increased RNA synthesis (Schuppli et al., 2000).

### ***Trans-acting RNA replication elements***

In addition to the *cis*-acting elements discussed above, some replication elements function in *trans*, such as the element in RNA-2 of *Red clover necrotic mosaic virus* (RCNMV) (family *Tombusviridae*, genus *Dianthovirus*) (Guenther, et al., 2004; Sit et al., 1998). RCNMV contains two single-stranded genomic RNAs, RNA-1 and RNA-2. Binding of a 34-nt element on RNA-2 to the CP subgenomic RNA promoter located on RNA-1, results in a structure that sterically prevents the replicase from completing synthesis of full-length complementary strands of RNA-1 (Sit et al., 1998). The prematurely terminated complementary strand RNAs serve as templates for production of the subgenomic RNA (Sit et al., 1998). Therefore, the 34-nt element of RNA-2 acts as a *trans*-activator of the subgenomic RNA (Sit et al., 1998).

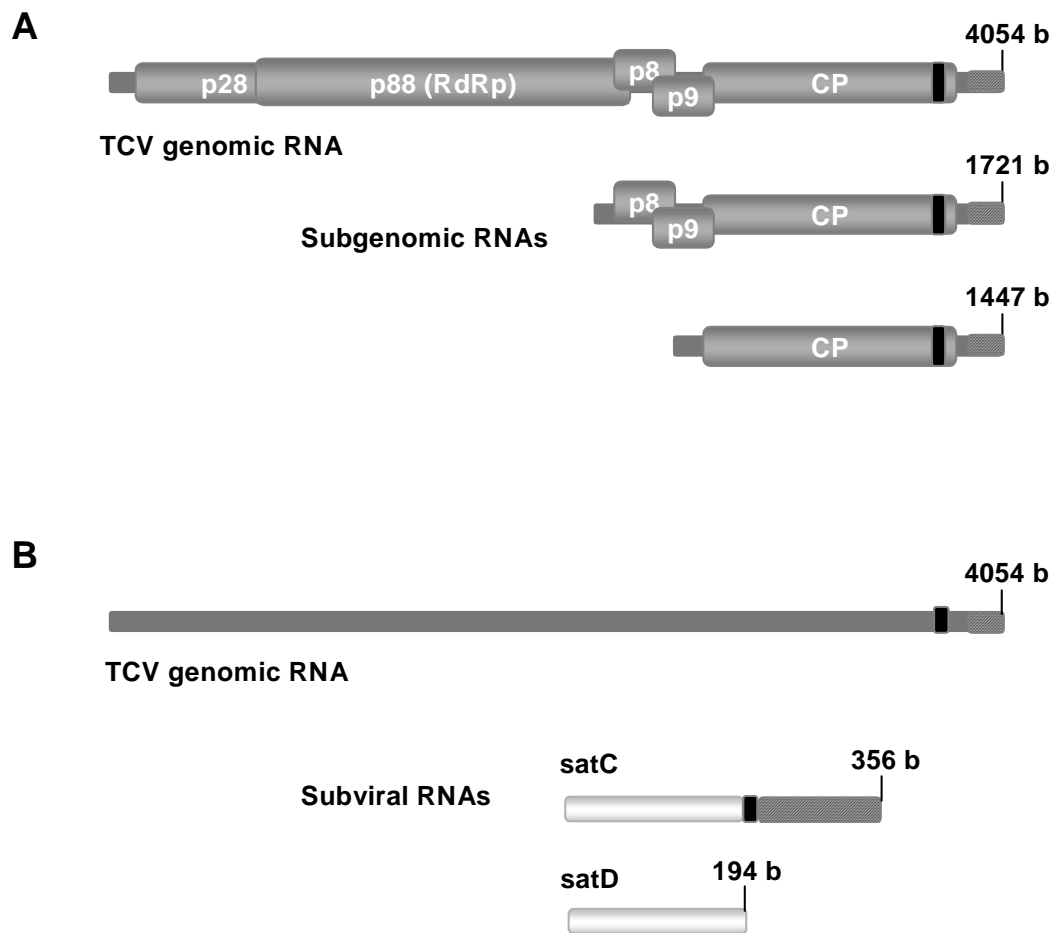
*Flock House virus* (FHV), a member of the family *Nodaviridae*, also has a bipartite genome comprising RNA1 and RNA2, which encode the viral RdRp and capsid protein precursor, respectively. A 3'-coterminal subgenomic RNA (RNA3) is transcribed from RNA1 minus strand (Gallagher et al., 1983; Guarino et al., 1984). It was reported that defective synthesis of RNA3 abolishes RNA2 replication (Eckerle and Ball, 2002). The finding that this effect is not caused by alterations in proteins, and that the *trans*-supplied RNA3 can restore RNA2 replication, suggested that RNA3 transactivates replication of RNA2 (Eckerle and Ball, 2002). Further studies indicated that transactivation of RNA2 replication is related to the 3' end of RNA2 (Albarino et al., 2003) and is dependent on

replication of RNA3 (Eckerle et al., 2003). It is not known how RNA3 replication stimulates an otherwise silent 3'-terminal replication signal of RNA2. RNA3-activation of RNA2 replication would act to coordinate synthesis of the two viral genome segments, since RNA3 is transcribed from RNA1 (Eckerle and Ball, 2002). A final example is that of a subgenomic RNA of BYDV that has recently been identified as a *trans*-acting replication regulator (Shen and Miller, 2004). It functions by inhibiting translation of the RdRp from the genomic RNA, thereby negatively regulating RNA replication (Shen and Miller, 2004).

### **TCV as a useful system for studying viral RNA elements involved in replication**

TCV (family *Tombusviridae*, genus *Carmovirus*) is a 30-nm icosahedral virus with a single-stranded RNA genome of 4054 bases, which encodes five proteins (Hacker, et al., 1992; Figure 1.2A). p28 and p88 (a translational readthrough product of p28 that contains the GDD polymerase active site consensus sequence) are translated from the genomic RNA. These two proteins comprise the viral RdRp, and both are required for replication of TCV and its associated RNAs *in vivo*, while purified p88 expressed in *E. coli* can correctly transcribe TCV-associated templates into complementary strands *in vitro* (Rajendran, et al., 2002). p8 and p9, required for cell-to-cell virus movement, are translated from the 1.7 kb subgenomic RNA and are dispensable for replication (Hacker, et al., 1992; Li, et al., 1998). The CP, which is translated from the 1.45 kb subgenomic RNA, packages TCV-associated RNAs into a 180-subunit, T=3 icosahedral virion.





**Figure 1.2** Genomic and subviral RNAs in the TCV system. (A) Genomic organization of TCV. p28 and p88 are expressed from the genomic RNA. p8 and p9 are expressed from the 1.7 kb subgenomic RNA. CP is translated from the 1.45 kb subgenomic RNA. (B) satC and satD are two satellite RNAs associated with TCV. satC is a chimeric RNA composed of satD and TCV sequences. Similar regions are shaded alike.

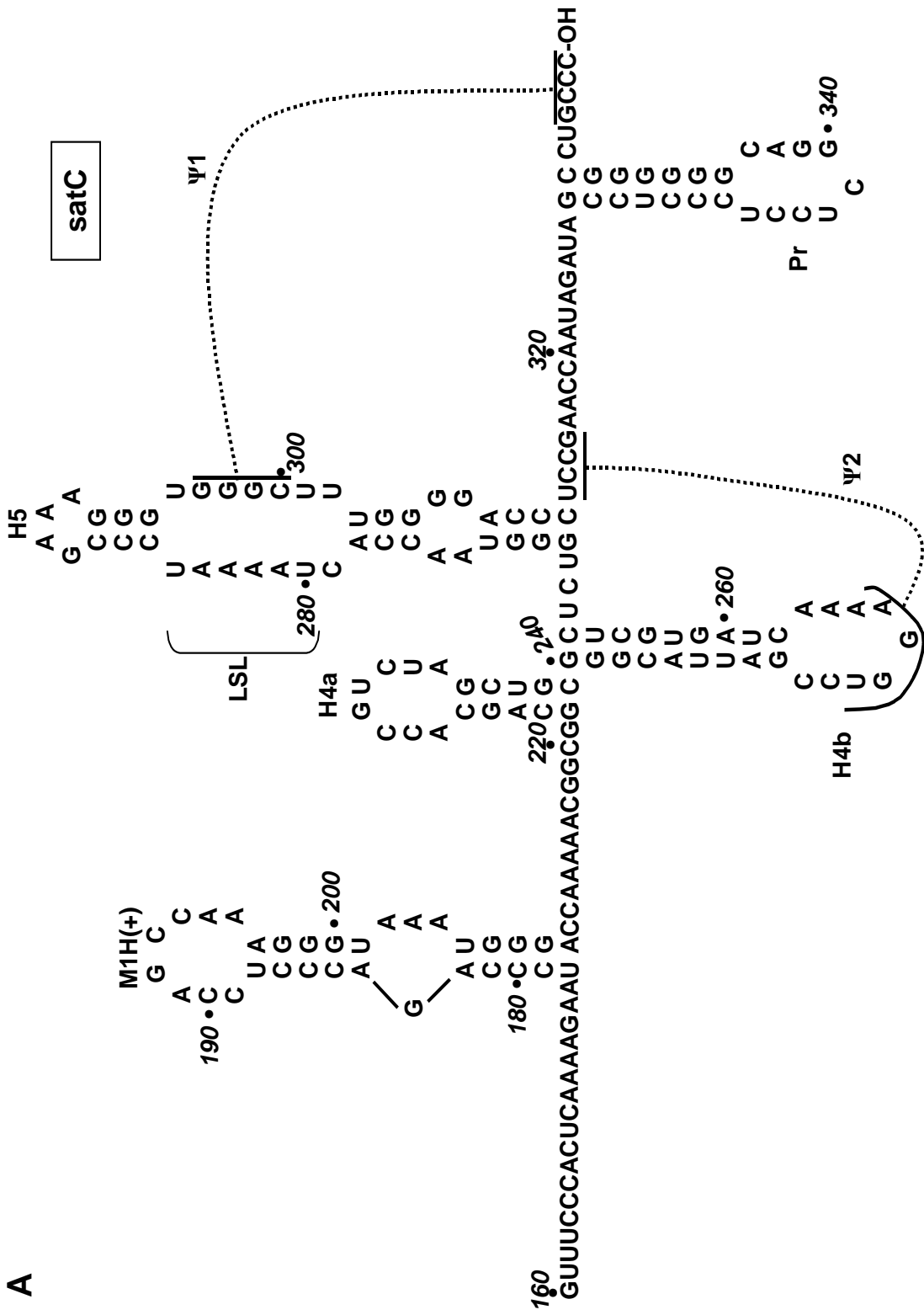
TCV is naturally associated with several subviral RNAs including satD (194 bases) and satC (356 bases), which are not templates for translation and therefore dependent on TCV-encoded proteins for accumulation in host cells (Li et al., 1989; Simon, 1999; Simon and Howell, 1986; Simon et al., 2004). SatD appears to have originated from numerous short non-contiguous stretches of TCV genomic RNA sequence (Carpenter and Simon, 1996). satC, a chimeric molecule containing nearly full-length satD at its 5' end and two discontinuous segments from TCV genomic RNA at its 3' end (Simon and Howell, 1986; Figure 1.2B), was generated by recombination events between satD and TCV. satD has no discernible effect on symptoms associated with TCV on hosts, while satC is virulent, strongly intensifying the symptoms of TCV by changing the mild stunting and chlorotic symptoms of TCV-infected turnip leaves to severely stunted, dark green, and crinkled leaves (Li and Simon, 1990; Simon and Howell, 1986). These findings suggest that satC enhances the ability of the virus to colonize the plant and interact with cellular factors (Li and Simon, 1990; Zhang and Simon, 2003a), while it interferes with TCV replication in protoplasts (Kong et al., 1995) and accumulation in plants (Li and Simon, 1990). Thus, fitness of satC is likely a function of attributes that allow the satRNA to both replicate efficiently and also interact with the helper virus in a way that enhances systemic infection. The correlation between enhanced virulence of TCV and reduced levels of virions in the presence of satC (Zhang and Simon, 2003a) suggests that satC fitness may reflect in part the ability of the satRNA to interfere with virion formation.

Recent reports demonstrating the multifunctional nature of the TCV CP provides a possible explanation for how inhibition of TCV virion formation contributes to the fitness

of satC. Most plant viruses encode repressors of virus induced gene silencing (VIGS), an anti-viral protective system in plants (Baulcombe, 2002; Li and Ding, 2001; Voinnet et al., 1999), which might serve a similar function in other organisms (Plasterek, 2002). The TCV CP was recently shown to be a very strong suppressor of VIGS when assayed independent of the virus (Qu et al., 2003; Thomas et al., 2003). However, when expressed from the virus genome, the CP is a weak suppressor, possibly due to suppression of an early step in gene silencing while the CP is expressed mainly later in infection (Qu et al., 2003). Furthermore, the N-terminus of the CP, required for suppressor activity, is also the RNA binding domain and unavailable for suppressor function when sequestered within the virus capsid (Thomas et al., 2003). By interfering with virion formation, satC could be enhancing the abundance of free CP leading to more efficient suppression of VIGS and thus facilitating the systemic infection of TCV (Zhang and Simon, 2003a).

Since the non-translated small satRNAs must contain *cis*-acting elements recognized by the virus-encoded RdRp, they are excellent templates for examining RNA sequence and structural elements involved in viral replication. Analysis of satC replication in plant protoplasts and transcription of complementary strands *in vitro* using partially purified RdRp preparation from TCV-infected plants or p88 expressed in *E. coli* led to identification of a number of *cis*-acting elements required for efficient accumulation of the satRNA. The elements located at the 3' end of plus strands regulate minus-strand synthesis by participating in an RNA conformational switch, which comprises a pre-active structure stabilized by a pseudoknot ( $\Psi$ 2) and an active structure that includes another pseudoknot ( $\Psi$ 1) and four hairpins (Zhang et al., 2006a; Zhang et al., unpublished data; Figure 1.3A). These four hairpins are (from 3' to 5'; Figure 1.3A):

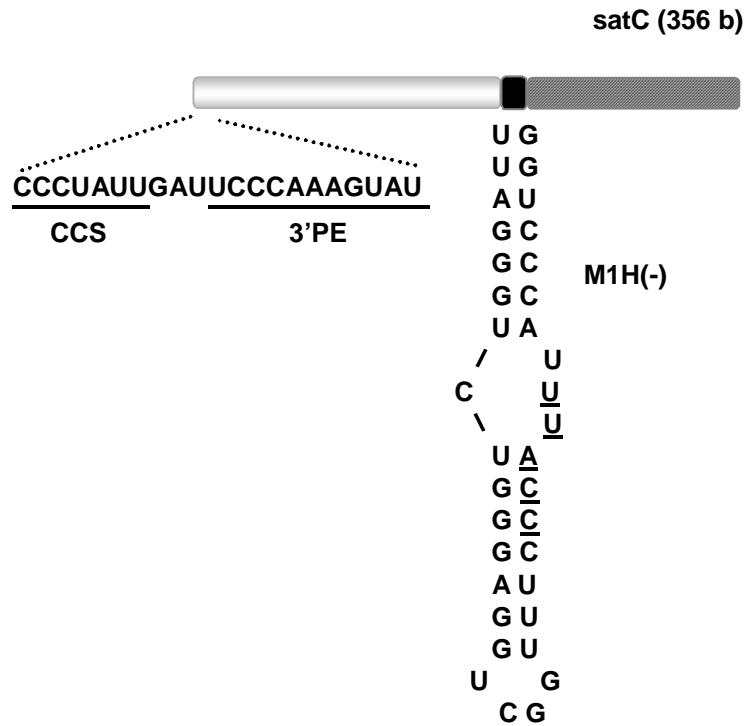
**Figure 1.3** Sequence and structure of the 3' regions of satC and TCV. The structures are phylogenetically conserved and predicted by mFold computer prediction (Zucker, 2003). Names of the hairpins are indicated. Numbering is from the 5' end. (A) The 3' region of satC. Two experimentally confirmed pseudoknots ( $\Psi$ 1 and  $\Psi$ 2) are shown. (B) The 3' region of TCV. The arrow indicates that the second recombination event that produced satC occurred in the lower stem of TCV H4. TCV and satC share related sequences downstream of the recombination site. The majority of TCV H4 and 5' sequence are unrelated to sequence in satC. Base differences between TCV and satC 3' ends are boxed.  $\Psi$ 1 is recently confirmed in TCV (Zhang et al., 2006).





Pr, the core promoter required for synthesis of complementary minus strands (Carpenter and Simon, 1998; Song and Simon, 1995; Stupina and Simon, 1997); H5, a hairpin that contains a large internal symmetrical loop (LSL) that interacts with 3' terminal bases to form  $\Psi$ 1 (Zhang et al., 2004a; 2004b); H4b, a hairpin with terminal loop sequences interacting with the sequence flanking the 3' base of H5 to form  $\Psi$ 2 (Zhang et al., unpublished data); and H4a comprising a single unit with H4b that is important for satC replication in vivo (Zhang et al., 2006a). All of these four hairpins are required for replication of satC in vivo.

Two linear elements termed the 3'CCS and 3'PE were identified on minus strands (Guan et al., 1997; 2000a; 2000b; Figure 1.4). The CCS (Carmovirus Consensus Sequences: C<sub>2-3</sub>A/U<sub>3-7</sub>) is found at the 3' ends of all Carmovirus genomic, subgenomic and satRNA minus-strands identified to date and is required for replication in vivo (Guan et al., 2000b). The 3'PE (Proximal Element; positions 11 to 21), which can function as an independent promoter of complementary RNA synthesis in vitro, also contains a CCS (Guan et al., 1997). A 30-base internal satC hairpin folds into similar structures on both strands, termed M1H(+) (Figure 1.3A) and M1H(-) (Figure 1.4), respectively. M1H(-) was originally identified as a hotspot for recombination between satC and satD and was proposed to help recruitment of the RdRp to the minus-strand satC acceptor template during recombination-mediated template switching (Cascone et al., 1993; Nagy et al., 1998). The RdRp recruitment ability of M1H(-) is supported by having a short motif identical to sequence in the TCV 3'CCS (Zhang and Simon, 2003b; Figure 1.4). Although not required for basal RNA transcription, M1H(-) is necessary for wt levels of satC accumulation in plants and protoplasts and is capable of enhancing transcription



**Figure 1.4** Minus-strand elements involved in replication of satC (shown in their minus-sense, 3'-to-5' orientation). satC is composed of satD sequence at the 5' end (gray box) and two regions of TCV at the 3' end (solid box is derived from the CP ORF and hatched box is from the 3' untranslated region). CCS and 3'PE are described in the text. M1H(-) is a replication enhancer that spans all three satC segments and contains a short motif found in TCV replication elements. The underlined sequence in M1H(-) is identical to the TCV 3'CCS.



from the 3' PE promoter by nearly 10-fold in vitro (Nagy et al., 1999). Therefore, M1H(-) is termed a replication enhancer (Nagy et al., 1999; 2001).

The region of satC including the 3 bases comprising the 3' base of M1H and the sequence downstream of M1H is derived from the 3' end of TCV (positions 3900 to 4054; Figure 1.3B). Computer prediction using mFold indicates that the 3' end of TCV is structurally similar to the analogous satC region despite containing differences (15 bases) at only eight positions (Zhang et al., 2004a; Zucker, 2003; Figure 1.3B). TCV Pr is a 5-fold weaker core promoter than the satC Pr when assayed in vitro (Zhang et al., unpublished data). TCV H5 interacts with the 3' terminal 4 bases to form a structure like  $\Psi$ 1 of satC (Zhang et al., 2006b). While a detailed investigation, especially in the H4a and H4b regions, remains to be performed, similarity with satC in this region suggests that a 3'-end conformational switch is possibly needed for TCV to regulate replication. Since satC is not translated while TCV genome is translated, the viral switch may function by converting the template from a translation-competent form to a replication-competent form prior to minus-strand synthesis initiation.

H4(+) is a TCV hairpin located just beyond the shared regions of TCV and satC, which is similar to M1H(+) in location relative to the 3' end of the RNA (Figure 1.3A). Like M1H(-), H4 on its minus-sense orientation, termed H4(-), is also a recombination hotspot in the absence of another nearby hotspot, M3H (Carpenter et al., 1995). While satC and TCV share related downstream sequences, the majority of H4 and 5' sequence is unrelated to sequence in satC. Introduction of H4(+) into a poorly viable satRNA, which is composed of satD and the 3' end of TCV but lacks M1H, can stimulate RNA accumulation by 6-fold in protoplasts (Nagy et al., 1999). H4(-), together with M3H, can

enhance RNA transcription from the satC 3' PE promoter by nearly 13-fold, while M3H alone results in an 11-fold enhancement of transcription (Nagy et al., 1999).

In this dissertation, I report my results and findings about RNA elements required for efficient virus accumulation. In chapter II, I show that M1H in satC plus strands brings together the flanking CA-rich sequences that are involved in satC replication and virion interference. In chapter III, I identify three individual sequence-specific elements located within the M1H-flanking CA-rich regions that could be specifically assigned roles in satC replication (CGGCGG) or virion repression (CAAAA) or both (CCCA). In Chapter IV, I show that H4 enhances syntheses of both TCV plus and minus strands by attracting the viral RdRp to the template.

## CHAPTER II

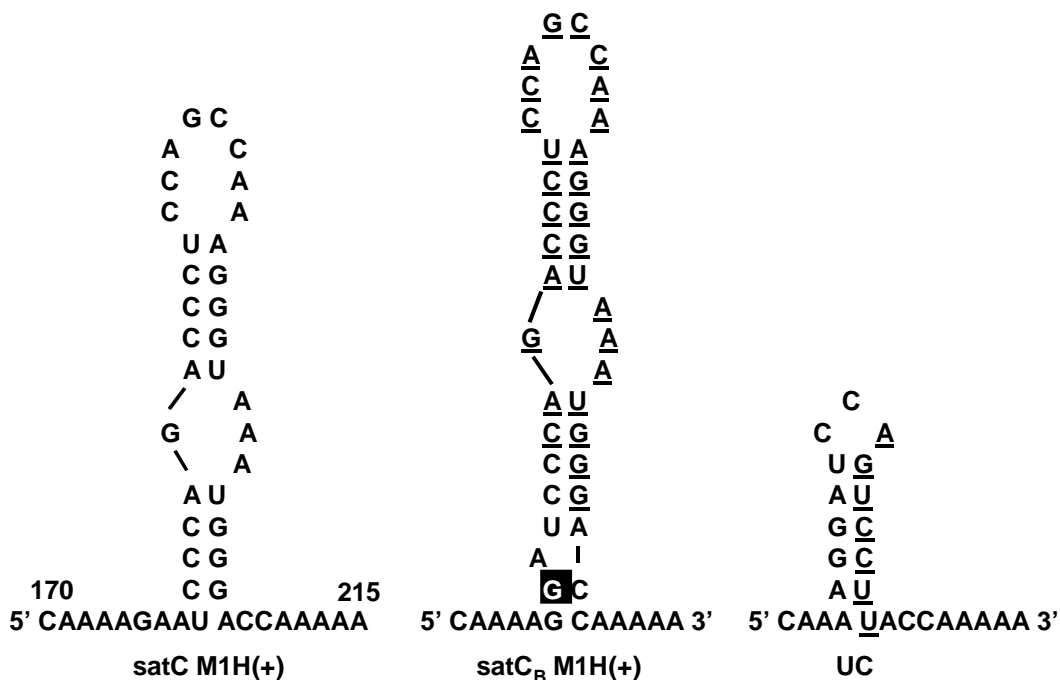
### FITNESS OF SATELLITE RNA C CORRELATES WITH M1H(+) AND FLANKING SEQUENCES THAT ENHANCE REPLICATION AND REPRESS THE ACCUMULATION OF VIRIONS

#### Introduction

Replication of positive-strand RNA viruses is mediated by a replicase complex that in general comprises the viral RdRp, other virally encoded proteins, and possibly host factors (Lai, 1998). The replicase must recognize its cognate RNA through direct or indirect interaction with specific *cis*-acting elements located in the RNA template, such as core promoters that are required for initiation of RNA synthesis (Buck, 1996; de Graaff et al., 1994). In addition to core promoter elements, many RNA virus genomes contain replication enhancers that enhance but are not required for basal-level transcription (Nagy et al., 1999; Ray and White, 1999; 2003). These RNA replication enhancers, which can be found in both plus- and minus-strands of viral RNAs, are located at variable positions in relation to the initiation site for RNA transcription and are thought to facilitate the recruitment of RdRp to the template (Lai, 1998). Replication enhancers have been characterized for a number of RNA viruses such as bacteriophage Q $\beta$  (Barrera et al., 1993), AMV (van Rossum et al., 1997), TBSV (Ray and White, 1999), and BMV (French and Ahlquist, 1987; Quadt et al., 1995; Ranjith-Kumar et al, 2003).

satC is a TCV-associated non-coding subviral RNA depending on virus-encoded proteins for accumulation in host cells (Simon and Howell, 1986; Figure 1.2B). As described in Chapter I, satC intensifies the symptoms of TCV in plants, while it interferes with TCV replication, suggesting that fitness of satC is likely a function of attributes that allow it to both replicate efficiently and interact with TCV in a way that enhances systemic infection. The fitness may reflect in part the ability of satC to interfere with virion formation, since in the presence of satC, enhanced virulence of TCV is related to reduced levels of virions (Zhang and Simon, 2003a).

As described in Chapter I, M1H is an internal hairpin of satC, which folds into similar structures on both strands. M1H(-) is a replication enhancer and a recombination hotspot. Recent results using an *in vivo* functional SELEX (systematic evolution of ligands by exponential enrichment) assay, where 28 random bases replaced 28 bases of M1H, resulted in the identification of winners that replicated to higher levels in protoplasts compared with satC containing non-selected 28-base sequences (Zhang and Simon, 2003b). The replacement sequences enhanced replication of satC in protoplasts to levels that correlated with fitness of the satRNAs to accumulate in plants with one exception. This exception, clone UC (Figure 2.1), contained only a 7-base replacement sequence indicating that a deletion of 21 bases had occurred. Although competition assays determined that UC was the second most fit winner, it did not replicate more effectively in protoplasts than random 28-base replacement sequences (Zhang and Simon, 2003b). The 7-base UC replacement sequence, together with downstream sequence, folded into a hairpin that was 16% more stable on plus-strands than minus-strands and was, like the wt M1H, flanked by CA-rich sequences (Figure 2.1). Surprisingly, the



**Figure 2.1** Structure folded in the M1H region of satC. Left, structure of the wt satC M1H(+) and surrounding sequence. Middle, structure of M1H(+) in satC<sub>B</sub>. The single site alteration required for the construction of satC with randomized sequence is boxed. Residues that were replaced with 10 random bases are underlined. Right, plus-strand structure formed by the M1H(+) replacement sequence in clone UC, generated during the previous SELEX analysis (Zhang and Simon, 2003). The 7-base replacement sequence is underlined.

replacement sequences of the other 2<sup>nd</sup> round winners (and most of the 1<sup>st</sup> round winners), also folded into plus-strand hairpins flanked by CA-rich sequences that were predicted to be more stable than minus-sense hairpins, suggesting that a sequence nonspecific plus-strand hairpin might contribute to satRNA fitness. UC was substantially better at reducing TCV virion accumulation compared with other M1H replacement sequence winners that replicated to higher levels, suggesting that ability to reduce virion levels contributed to the fitness of the satRNA. In addition, it seemed possible that the role of the hairpin was to bring the flanking CA-rich sequences into proximity, contributing to virion reduction.

In this chapter, I tested this hypothesis and found that 10-base M1H replacement sequences also folded into hairpins predicted to be more stable on plus-strands than minus-strands. Several of the most fit satRNAs contained inserts of adenylates and cytidylates at the base of the hairpins whose presence correlated with enhanced replication and reduction of virion levels. These results confirm that a sequence non-specific plus-strand hairpin flanked by CA-rich sequence in the M1H region confers fitness to satC by reducing virion assembly.

## **Materials and Methods**

### **Large-scale plasmid DNA isolation**

Bacteria (*E. coli*) were grown in 250 ml L-broth culture at 37°C for 16 hours with continuous shaking. The cells were collected by centrifugation in a Sorvall GSA rotor at 6000 rpm for 10 minutes and resuspended in 2.5 ml of suspension buffer (25% sucrose,

50 mM Tris-HCl, pH 7.5). To lyse the cell wall, 0.4 ml of freshly prepared 10 mg/ml lysozyme was added. The mixture was swirled and incubated on ice for 10 min, followed by addition of 0.7 ml of 0.5 M EDTA (pH 8.0). The mixture was incubated again on ice for 10 minutes, followed by addition of 5.3 ml of lysis buffer (0.1% Triton X-100, 62.5 mM EDTA, pH 8.0, and 50 mM Tris-HCl, pH 8.0). After incubation at 42°C for 10 minutes, the mixture was centrifuged in a Sorvall SS34 rotor at 17,000 rpm for 20 minutes. The supernatant (yellowish and sticky) was collected into a 15 ml centrifuge tube and 8.8g CsCl and 200 µl of ethidium bromide (10 mg/ml) were added. The mixture was adjusted to a volume of 12 ml with distilled H<sub>2</sub>O and then transferred into a quick-seal tube (Beckman, 16 mm X 76 mm). After centrifugation at 20°C in a VTi 65.1 rotor (Beckman) at 45,000 rpm for 13 hours or at 65,000 rpm for 4.2 hours, the lower DNA band was recovered with a 5 ml syringe and extracted 3 times with NaCl-saturated isopropanol or until no remaining pink color was visible. The solution was diluted with 2 volumes of distilled H<sub>2</sub>O and then mixed with 6 volumes (original volume) of 95% ethanol. The mixture was incubated at -20°C for at least 2 hours and then centrifuged in a Sorvall SS34 rotor at 10,000 rpm for 10 minutes. The DNA pellet was redissolved in 0.4 ml of distilled H<sub>2</sub>O in a 1.5 ml eppendorf tube. The dissolved DNA was precipitated with 2.5 volume of 5 M NaOAc/ethanol (1:25), and washed with 70% ethanol. The pellet was dried and dissolved in an appropriate amount of distilled H<sub>2</sub>O. The concentration of DNA was estimated by measuring the absorbance at 260 nm.

### **Small-scale plasmid DNA isolation**

Bacterial cells (*E. coli*) from 1.5 ml of overnight culture were collected in an eppendorf tube by centrifugation at 13,000 rpm for 12 seconds. The pellet was

resuspended in 140  $\mu$ l of STET buffer (8% sucrose, 5% Triton X-100, 50 mM EDTA, and 50 mM Tris-HCl, pH 8.0) and mixed well. The mixture was then boiled for 1 minute, and centrifuged at 13,000 for 14 minutes at room temperature. The pellet of cell debris was removed with a toothpick. The remained solution was mixed with 140  $\mu$ l of isopropyl alcohol, incubated on ice for 5 minutes, and then centrifuged at 13,000 rpm for 5 min at 4°C. The DNA pellet was rinsed with 70% ethanol, dried, and dissolved in 25  $\mu$ l of distilled H<sub>2</sub>O.

### **DNA sequencing**

Five microliters of plasmid DNA prepared from 1.5 ml culture using the small-scale plasmid DNA isolation method as described above was mixed with 1 pmol of a sequencing primer, 2  $\mu$ l of 2M NaOH, and distilled H<sub>2</sub>O to give a volume of 22  $\mu$ l. The mixture was boiled in a water bath for 2 minutes. The plasmid DNA was precipitated with 2.5 volume of 5 M NaOAc/ethanol (1:25). The dideoxynucleotide sequencing was carried out using DNA sequenase version 2.0 (USB) as recommended by the manufacturer.

### **In vitro RNA synthesis using T7 polymerase**

Plasmid DNA was digested with the appropriate restriction enzyme, extracted with phenol/chloroform, precipitated with 2.5 volume of 5 M NaOAc/ethanol (1:25), and washed with 70% ethanol. Linearized DNA template (8  $\mu$ g) was mixed with 6  $\mu$ l of dithiothreitol, 12  $\mu$ l of 5 mM each of ATP, GTP, CTP, UTP, 12  $\mu$ l of 5X T7 buffer (125 mM NaCl, 40 mM MgCl<sub>2</sub>, 2mM spermidine, 40 mM Tris-HCl, pH8.0), 60 units of



RNAsin (Invitrogen), 80 units of T7 polymerase (Invitrogen), and distilled H<sub>2</sub>O to give a final volume of 60 µl. The mixture was incubated at 37°C for 1 hour (genomic RNA) or 2 hours (satRNA) or until the solution was turbid.

The synthesized RNA transcripts were directly used for inoculation of plants. For inoculation of protoplasts, the RNA transcripts were first extracted with phenol/chloroform, precipitated with 2.5 volume of 5M NH<sub>4</sub>OAc (pH5.3)/isopropanol (1:5), and washed with 70% ethanol. Plus-strand transcripts of TCV were synthesized using T7 polymerase from *Sma*I-linearized pTCV66, which contains a T7 RNA polymerase promoter upstream of TCV full-length plus-strand sequence (Oh et al., 1995). The wt satC transcripts were synthesized using T7 polymerase from *Sma*I-linearized pT7C+, which contains a T7 RNA polymerase promoter upstream of satC full-length plus-strand sequence (Song and Simon, 1994).

### **Plant Growth and Inoculations**

Turnip plants were grown in growth chambers at 20°C as described by Li and Simon (1990). Two true leaves of plant seedlings at the 5-leaf stage (about 2 weeks after seed sowing) were mechanically inoculated with 10 µl (per leaf) of inoculation buffer containing TCV RNA transcripts (0.15 µg/µl) with or without satC transcripts (0.015 µg/µl). The inoculation buffer contains 0.05 M glycine, 0.03 M K<sub>2</sub>HPO<sub>4</sub>, pH 9.2, 1% Bentonite (clean), and 1% Celite.

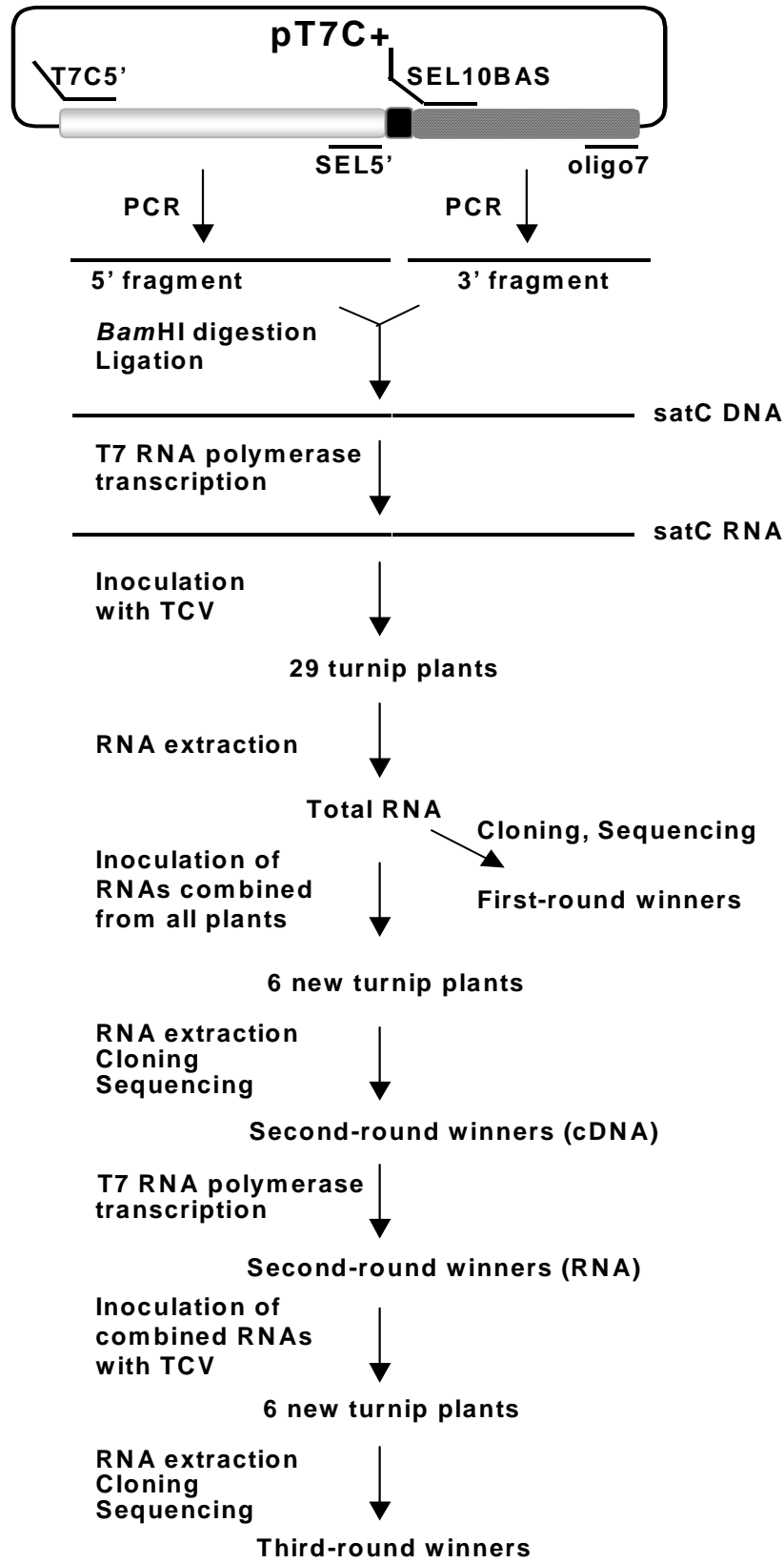
### **Small-scale RNA extraction from turnip leaves**

One gram of uninoculated turnip leaves was ground in liquid nitrogen in a 50 ml beaker with a small pestle. Leaf powder was transferred into a 1.5 ml eppendorf tube and extracted with 0.5 ml of phenol and 0.5 ml of extraction buffer containing 0.2 M Tris-HCl, pH 9.0; 0.4 M LiCl, 25 mM EDTA; and 1% SDS. After centrifugation at 13,000 rpm for 5 minutes at 4°C, the aqueous phase was extracted one more time with phenol/chloroform and precipitated with 2.5 volume of 5 M NaOAc/ethanol (1:25). The pellet was resuspended in 0.3 ml of 2 M LiCl and vortexed well. The solution was centrifuged at 13,000 rpm for 5 minutes at 4°C and the pellet was dissolved in 0.3 ml of distilled H<sub>2</sub>O. The RNA was precipitated with 2.5 volume of 5 M NaOAc/ethanol (1:25), rinsed with 70% ethanol, dried, dissolved in 50 µl H<sub>2</sub>O, and stored at -80°C.

### **In vivo functional SELEX using 10-base M1H replacement sequences**

To produce full-length satC RNAs containing 10 random bases replacing 28 bases of the M1H, two fragments were generated by separate PCR using pT7C+ (see Figure 2.2 for procedure flow chart), which contains a T7 RNA polymerase promoter upstream of satC full-length plus-strand sequence (Song and Simon, 1994). The 5' fragment was produced using primers T7C5' (oligonucleotides used in this Chapter are listed in Table 2.1), which contains a T7 polymerase promoter at the 5' end, and SEL5'. The 3' fragment containing 10 random bases was generated using primers SEL10BAS and oligo7. In these two PCRs, a new *Bam*HI site was generated in satC cDNA by changing a U to C at position 176 for convenient ligation of the purified PCR fragments. satC with this new *Bam*HI site is referred to as satC<sub>B</sub>. Both 5' and 3' PCR fragments were treated with *Bam*HI, purified through 1.2% agarose gels and then ligated to produce full-length

**Figure 2.2** Flow chart for performing the 10-base SELEX of satC M1H. pT7C+ is pUC19 that contains a T7 RNA polymerase promoter upstream of satC full-length plus-strand sequence (Song and Simon, 1994) inserted into the *Sma*I site of the plasmid. The primers used for PCR are given in Table 2.1. See the text for details.



**Table 2.1** Oligonucleotides used in Chapter II

Application /Construct	Name	Position <sup>a</sup>	Sequence <sup>b</sup>	Polarity <sup>c</sup>
SELEX	T7C5'	1-14	5' <i>GTAATACGACTCACTATA</i> GGGATAACT AAGGG	+
	SEL5'	158-175	5'GACT <b>GGATCCT</b> TTTTGAGTGGGAAACAG	-
	SEL10BAS	209-228	5'AAA <b>AGGATCC</b> <u>NNNNNNNNNN</u> NACAAA AACGGCGGCAGCAC-3'	+
	oligo7	338-356	5'GGGCAGGCCCCCGTCCGA	-
satC variants	OligoΔA <sub>5</sub> Cf	209-228	5'GACT <b>GGATCCT</b> TTTTACGGGAACCAAAA ACGGCGGCAGCAC	+
	Oligo2xA <sub>5</sub> Cf	212-228	5'GACT <b>GGATCCT</b> TTTTACGGGAACCA <u>AAAA</u> <u>ACAAAAACAAAA</u> ACGGCGGCAGCAC	+
	Oligo R1/A <sub>5</sub> Cf	212-228	5'GACT <b>GGATCC</b> ATCCGGACCAAC CAAAAACAAAAACGGCGGCAGCAC	+
	Oligo MOV/A <sub>5</sub> Cr	156-175	5'GACT <b>GGATCC</b> GTTTTTTTTTGA GTGGGAAACAGCC	-
Northern blotting	Oligo13	250-269	5'-GTTACCCAAAGAGCACTAGTT	-

<sup>a</sup> Coordinates correspond to those of satC. Oligo13 also corresponds to positions 3950-3970 of TCV genomic RNA.

<sup>b</sup> Bases in italics indicate T7 RNA polymerase promoter sequence. Bold residues denote bases changed to generate a *Bam*HI site. Mutant bases are underlined.

<sup>c</sup> “+” and “-” polarities refer to homology and complementarity with satC plus-strands, respectively.

cDNA. satC transcripts were directly synthesized from the ligated product using T7 RNA polymerase as described above. Five micrograms of satC transcripts and 2 µg of TCV transcripts were co-inoculated into each of 29 turnip seedlings. Total RNA was extracted from uninoculated leaves of each plant at 21-days postinoculation (dpi) as described above, amplified by RT-PCR using primers T7C5' and oligo7, and then cloned into the *SmaI* site of pUC19, and sequenced as described above. For the 2<sup>nd</sup> round SELEX, equal amounts of total RNA extracted from all 1<sup>st</sup> round plants were pooled and approximately 5 µg of the pooled RNA was used to inoculate each of 6 new turnip seedlings. Total RNA was extracted from uninoculated leaves at 21 dpi and satC cloned and sequenced. The 3<sup>rd</sup> round SELEX was accomplished by combining in vitro transcripts of all recovered 2<sup>nd</sup> round clones and then inoculating 5 µg of the pooled transcripts along with 2 µg of TCV transcripts onto 6 new turnip seedlings. satC RNAs accumulating in uninoculated leaves were cloned and sequenced.

### **Secondary structure analysis of SELEX winners**

The 10-base M1H replacement sequences (along with 11 upstream and 8 downstream bases) of all 2<sup>nd</sup> round SELEX winners were analyzed for secondary structures using mFold (Mathews et al., 1999). Each sequence was also subjected to three randomizations using the Shuffle program from Arizona Research Labs and the stability of the folded sequences determined and the values averaged.

### **Fitness comparison of the 3<sup>rd</sup> round SELEX winners in turnip plants**

To compare the fitness of the 3<sup>rd</sup> round SELEX winners for accumulation in plants, equal amounts of transcripts were combined and used to inoculate a single leaf of 3 turnip seedlings along with TCV genomic RNA transcripts. Total RNA was extracted at 21 dpi and satRNAs were cloned and sequenced as described above.

### **Construction of satC variants**

Constructs  $\Delta A_5C$ ,  $2xA_5C$ , and  $R1/A_5C$  were generated by ligation of two PCR fragments. The 5' fragment was obtained by PCR using pT7C+ and primers T7C5' and SEL5', while the 3' fragments were amplified using pT7C+ and primers oligo7 and either Oligo $\Delta A_5C$ f, Oligo $2xA_5C$ f, or Oligo $R1/A_5C$ f, respectively. For MOV/ $A_5C$ , the 5' fragment was obtained by PCR using pT7C+ and primers T7C5' and OligoMOV/ $A_5C$ r and the 3' fragment was identical to that for construction of  $\Delta A_5C$ . 5' and 3' PCR fragments were treated with *Bam*HI, gel purified, ligated together, and inserted into the *Sma*I site of pUC19. All clones were confirmed by sequencing.

### **Protoplast preparation and inoculation**

Protoplasts were prepared from callus cultures of *Arabidopsis thaliana* ecotype Col-0. The calli were generated from sterilized seeds placed on 1% MS (Murashige-Skoog salts, Gibco BRL) agar media (pH 5.8) supplemented with 2 mg/ml kinetin, 2 mg/ml 2,4-D (2,4-dichlorophenoxyacetic acid, Sigma), and 1X vitamins/glycine containing 1  $\mu$ g/ml nicotinic acid, 10  $\mu$ g/ml thiamine HCl, 1  $\mu$ g/ml pyridoxine, 0.1 mg/ml myoinositol, and 4  $\mu$ g/ml glycine. The calli were incubated in a growth chamber

at 20°C under 35  $\mu\text{mol m}^{-2} \text{sec}^{-1}$  lights with a cycle of 16-hour light and 8-hour-dark. Cultures were passaged every 21 days.

To prepare protoplasts, calli in the 4<sup>th</sup> or 5<sup>th</sup> passage were collected and soaked in 40 ml of 0.6 M mannitol at 25°C for 20 minutes with shaking. The calli were recovered by centrifugation (Beckman GPR-type swinging bucket) at 930 rpm for 5 minutes at 4°C, and then suspended in 50 ml of freshly prepared protoplast isolation medium (PIM, pH 5.8) containing 0.5 g of cellulase (11,900 U/g) and 0.1 g of pectinase (3,140 U/g) (Calbiochem, La Jolla, CA). One liter of PIM contains 1 ml of 1000X vitamin stock (1 mg/ml thiamine-HCl, 0.5 mg/ml pyridoxine-HCl, 0.5 mg/ml nicotinic acid, and 0.1 g/ml myo-inositol), 0.5 ml of 2000X hormone stock (0.4 mg/ml 2,4-D, 0.4 mg/ml kinetin, and 50 mM KOH), 4.3 g of MS plant salts, 0.1 M sucrose, 3 mM MES, 0.5 M mannitol, and 5 mM CaCl<sub>2</sub>. The calli/PIM/enzyme mixture was incubated at 26°C in dark for 4 hours with shaking at 75 rpm. The solution, which should become turbid after the incubation, was filtered through a 53- $\mu\text{m}$  nylon mesh (Small Parts, Miami Lakes, FL), followed by centrifugation at 930 rpm for 5 minutes at 4°C. The precipitated protoplasts were washed 3 times with 20 ml of 0.6 M mannitol (pre-cooled on ice). The number of cells was calculated using microscope and a haemocytometer.

Protoplasts ( $5 \times 10^6$  cells in a volume of 100  $\mu\text{l}$ ) were swirled well with 20  $\mu\text{l}$  of TCX genomic RNA transcripts (1 $\mu\text{g}/\mu\text{l}$ ), 2  $\mu\text{l}$  of satRNAs (1 $\mu\text{g}/\mu\text{l}$ ), 8  $\mu\text{l}$  of 1 M CaCl<sub>2</sub>, 400  $\mu\text{l}$  of distilled H<sub>2</sub>O, and 2.17 ml of 50% PEG (prepared in 50 mM Tris-HCl, pH 7.5). The mixture was incubated at 25°C for 30 seconds, followed by addition of cold 0.6 M mannitol/1 mM CaCl<sub>2</sub> and incubation on ice for 20 minutes. The protoplasts were collected by centrifugation at 930 rpm for 5 minutes at 4°C. After washing 3 times with



20 ml of cold 0.6 M mannitol/1 mM CaCl<sub>2</sub>, protoplasts were resuspended in protoplast culture medium (PCM, pH 5.8) and incubated at 25°C for 36 to 40 hours in the dark. One liter of PCM contained 1 ml of 1000X vitamin stock (1 mg/ml thiamine-HCl, 0.5 mg/ml pyridoxine-HCl, 0.5 mg/ml nicotinic acid, and 0.1 g/ml myo-inositol), 0.5 ml of 2000X hormone stock (0.4 mg/ml 2,4-D, 0.4 mg/ml kinetin, and 50 mM KOH), 4.3 g of MS plant salts, 0.1 M sucrose, 3 mM MES, and 0.4 M mannitol.

### **Total RNA extraction from protoplasts**

Protoplasts in 1 ml of PCM (~1.67 X 10<sup>6</sup> cells) were collected in an eppendorf tube at 36 or 40 hours postinoculation (hpi) and resuspended in 200 µl of extraction buffer (50 mM Tris-HCl, pH 7.5, 5 mM EDTA, 100 mM NaCl, and 1% SDS) and 200 µl of phenol/chloroform (1:1), followed by vigorous vortexing. The solution was then centrifuged at 13,000 rpm for 5 minutes at 4°C. The supernatant was precipitated with 2.5 volumes of 5 M NaOAc/ethanol (1:25), and washed with 70% ethanol. Pellets were dried and resuspended in 20 µl of distilled water.

### **Northern blot hybridization**

Four micrograms of total RNA extracted from protoplasts were denatured by heating in 50% formamide and then subjected to electrophoresis through a nondenaturing 1.2% agarose gel. The gel was incubated for 1 hour in 6% formaldehyde with gentle shaking, and then soaked in 10X SSC containing 0.15 M NaCl and 0.015 M sodium citrate for 25 minutes followed by transferring of the RNAs to a NitroPlus membrane (Micron Separations Inc., Westboro, MA). The blot was placed on a UV light box (310

nm, Fotodyne Inc.) for 2 minutes to crosslink the RNA and then dried at 80°C for 5 minutes.

An oligonucleotide complementary to both positions 3950-3970 of TCV genomic RNA and positions 250-269 of satC was labeled with [ $\gamma$ - $^{32}$ P] ATP using T4 polynucleotide kinase and used as a probe for simultaneous detection of TCV genomic RNA and satC. Prehybridization for 1 hour and hybridization for 2 hours at 42°C were performed in a hybridization buffer containing 5X SSPE, 10X Denhardt's reagent (Ausubel et al., 1987), 0.2% SDS, 0.2 mg/ml freshly denatured salmon sperm DNA, and 50% formamide (every 1% of formamide reduces the  $T_m$  by 0.7°C). After hybridization, the blot was washed in a high salt solution containing 6X SSPE and 0.1% SDS for 12 minutes, then washed two times in a low salt solution containing 0.1X SSPE and 0.1% SDS for 20 and 15 minutes, respectively. The blot was covered with a saran wrap and subjected to autoradiography. One liter of 20X SSPE contains 175.3 g NaCl, 27.6 g  $\text{NaH}_2\text{PO}_4$ , 40 ml of 0.5 M EDTA (pH 8.0) and is adjusted to pH 7.4 with 10 M NaOH.

### **Protein extraction from protoplasts and electrophoresis analysis**

Protoplasts in 1 ml of PCM ( $\sim 1.67 \times 10^6$  cells) were collected in an eppendorf tube at 40 hpi and resuspended in 30  $\mu\text{l}$  of protein analysis buffer containing 30 mM Tris-HCl, pH 6.8, 1.5% SDS, 5% glycerol, and 2.5%  $\beta$ -mercaptoethanol, followed by vortexing for 5 minutes. The mixture was centrifuged at 13,000 rpm for 5 minutes to collect the supernatant. Total proteins were separated on a 12% SDS-PAGE gel in a running buffer containing 50 mM Tris base and 38 mM glycine, pH 8.3. The gel was stained using Coomassie brilliant blue.

### **Virion isolation from protoplasts and Western blotting analysis**

Protoplasts in 1 ml of PCM ( $\sim 1.67 \times 10^6$  cells) were collected in an eppendorf tube at 40 hpi and resuspended in 200  $\mu$ l of 0.2 M NaOAc, pH 5.2, followed by addition of 30  $\mu$ l of sterile glass beads (0.1-0.2 mm in diameter). The mixture was vortexed for three 15-second intervals and kept on ice during each interval. The aqueous phase was recovered by centrifugation at 13,000 rpm for 1 minute. The solid phase was reextracted twice with 200  $\mu$ l of 0.2 M NaOAc, pH 5.2. The aqueous phases were combined and incubated on ice for 1 hour, followed by centrifugation at 13,200 rpm for 2 minutes at 4°C. The supernatant was mixed with 1/4 volume of 40% polyethylene glycol (PEG, MW 8000)/1 M NaCl. The mixture was incubated on ice for 12 hours, followed by centrifugation at 13,200 rpm for 1 hour. The pellet (virions) was dissolved in 30  $\mu$ l of 10 mM NaOAc, pH 5.5, stored at 4°C.

For Western blotting analysis, 1  $\mu$ l of virions were separated on a 1% agarose gel prepared in 50 mM Tris base/ 38 mM glycine, pH 8.3. The gel was soaked in 50 mM NaOH for 20 minutes and then in 0.2 M NaOAc pH 5.5, for 20 minutes before blotting to a NitroPlus membrane (Micron Separations Inc., Westborough, MA). Western blotting analysis was performed using the anti-TCV CP polyclonal antibody generated in rabbits injected with gel-purified TCVP. The anti-CP antibody and the second antibody (anti-rabbit IgG horseradish peroxidase; Gibco BRL) were used in 1:5000 and 1:7500 dilutions, respectively. Dilutions were made in phosphate buffered saline (PBS) containing 6% milk (w/v). Chemiluminescent staining was performed with the Western Lighting Chemiluminescence Reagent kit (Perkin Elmer Life Sciences). Membrane

incubated with the substrate was covered with plastic saran wrap and exposed to an X-ray film for 20-30 seconds before developing the film.

## **Results**

### **In vivo functional SELEX of satC containing 10 random bases replacing M1H**

Previous results suggested that M1H participates in two aspects of satC fitness: by serving as a sequence-specific enhancer of satC replication in its minus-sense orientation and possibly by interfering with stable assembly of TCV virions mediated by the hairpin structure in its plus-strand orientation (Zhang and Simon, 2003a). To test this hypothesis, in vivo functional SELEX was conducted by replacing 28 bases of M1H with 10 random bases using a PCR strategy outlined in Materials and Methods. Function-based in vivo SELEX identifies sequences that impart fitness to the molecule regardless of the mechanism of action. This differs from the more traditional in vitro SELEX, which selects for molecules with a specific function such as binding to a particular polypeptide or ability to catalyze a specific reaction (Wilson and Szostak, 1999). The replacement of 28 bases with 10 bases was chosen for two reasons: i) a shorter sequence should lead to the recovery of satC more similar to UC with its 7-base replacement sequence, and thus might contain hairpins that could be compared with UC; and ii) the promoter-like motifs found in the minus-sense 28-base M1H replacement sequences were generally between 6 and 10 bases, suggesting that similar motifs that enhanced replication might be found when using a 10-base replacement sequence.

SatC transcripts containing 10 random bases replacing M1H were co-inoculated onto 29 turnip seedlings along with transcripts of TCV genomic RNA. At 21 dpi, total RNA was isolated from uninoculated leaves. Examination of RNA following gel electrophoresis and ethidium bromide staining revealed that all plants contained detectable satRNA (wt satC normally accumulates to levels similar to 5S ribosomal RNA and thus is readily detectable by this method). Sixty-four satC species were cloned from 19 of the 29 inoculated plants. Sequencing the clones revealed 43 different 1<sup>st</sup> round winning sequences (Figure 2.3).

These and other 1<sup>st</sup> round winners were subjected to further competition by combining equal portions of total RNA extracted from all 29 infected plants and then inoculating six new seedlings. At 21 dpi, total RNA was extracted from uninoculated leaves and satC molecules were cloned. Sequencing 8 to 11 clones from each plant revealed 25 different sequences (the 2<sup>nd</sup> round winners, Table 2.2). The sequences are presented in their minus-sense 3' to 5' orientations since promoter-like motifs are present in sequence of M1H(-) (Figure 1.4) and were also found in the minus-sense 28-base replacement sequences from the previous study (Zhang and Simon, 2003b). Six of the 2<sup>nd</sup> round winning sequences were also found in the 1<sup>st</sup> round (II-6, II-10, II-11, II-15, II-17, II-20; Figure 2.3). Several sequences (II-1a/II-1b and II-3a/II-3b) differed by only one or two bases, respectively, and thus were possibly derived from the same original transcript. Since the II-1a/II-1b and II-3a/II-3b replacement sequences folded into hairpins with different stabilities (see below), they are listed separately in Table 2.2. One of the 2<sup>nd</sup> round winners, II-16, was unusual in having an 18-base replacement sequence instead of the original 10 random bases, indicating that additional modification of replacement

GUUUGCUC	CC	UCACCAUAG
UUAUGCAAC	CC	UUGUCAGGCG
CUAACAGGCC		CCCUCUCUAG
UCGGCUCAUC		CGUUCGCGG
CUUAGGUGUU		UAAAAGCUCU
UUCUAUGCGA		AUAAGUCCU
UCCAUGCGA		CAAUCCAAGG
CUUAGUAGGG		UAUUCAGACC
UCGGAGCCUG		CUUAAGUGUU
GAGUUUCCU		CACGUUCACC
CAUUCCCAAG		UGAGCUAGGU
AAUGCUGAU		GUGAGUAAU
AAUACCAGAU		AUAUAUGUCC
GUGUUUAUGC		CGUAUCUUC
GUGGGUAUGC		CCCUCUCUAG
UUUGCAGGUU		ACCAGGAACA
AUAGUUGUCC		CAAUGCCGCC
CAAGGCUACC		UCAUGUCAAC
CUUACCAGGU		UACAGUAGCG
UCAUGUCAAC		UAGAGCCCC
UAGGCCUGGU		UUAACAAAUG
GUUAAUAGG		

**Figure 2.3** Forty-three M1H replacement sequences recovered from the 1st round selection of the 10-base SELEX. All sequences are shown in their minus-sense, 3' to 5' orientation. The sequences shown in orange were also recovered in the 2nd round selection.

**Table 2.2** 2<sup>nd</sup> round SELEX winners

Name	M1H replacement sequence <sup>a</sup>	Total <sup>d</sup>	$\Delta G$ (kcal/mol)		Randomized <sup>e</sup> $\Delta G$ (kcal/mol)	
			(+)-strand	(-)-strand	(+)-strand	(-)-strand
II-1a	CUUCCCCAAG	5/59	-9.2	-0.7	-5.0	-2.7
II-1b	CUCCCCAAG	1/59	-6.7	-0.7	-5.2	-5.0
II-2	UACUCAAGGC	8/59	-2.1	-1.0	-2.1	-2.3
II-3a	UUAAACCAGG	1/59	-5.7	-3.9	-0.5	-1.6
II-3b	UUAUACUAGG	1/59	-6.3	-6.4	-0.7	-1.6
II-4	CAAUCCAGG	4/59	-5.9	-2.5	-2.8	-3.7
II-5	ACUCUUUCCU	5/59	-7.5	-5.2	-2.5	-2.5
II-6	AUAAGUCCCU	2/59	-6.6	-3.8	-0.7	-2.0
II-7	GGUACCCCUA	2/59	-11.0	-9.5	-2.2	-2.3
II-8	AAUGCUAGAU	1/59	-3.5	-2.7	-0.2	-0.6
II-9	AUAACCCCGG	3/59	-9.2	-4.2	-5.5	-4.4
II-10	CGUUCCGCCG	1/59	-3.7	-2.4	-3.1	-2.5
II-11	AUAGUUGUCC	1/59	-6.1	-8.1	-0.5	-1.5
II-12	CGAAGCCCGC	1/59	-4.3	-4.4	-3.4	-4.1
II-13	UUUGCCUGGC	1/59	-5.0	-5.8	-1.8	-3.8
II-14	CAUUAGCCCA	2/59	-8.1	-4.9	-1.9	-2.3
II-15	CUUACCAGGU	2/59	-5.5	-3.1	-3.8	-4.0
II-16	ACUAGCCCGUAGCCCGU	1/59	-7.8	-6.2	-7.6	-6.6
II-17	AUAUAUGUCC <sup>b</sup>	3/59	-6.1	-8.1	-1.7	-3.1
II-18	AAAAUGCCCU <sup>c</sup>	1/59	-6.2	-4.0	-1.9	-2.5
II-19	UACCGGACCU	1/59	-6.4	-5.4	-4.4	-4.3
II-20	UAUUCAGACC	3/59	-6.1	-6.8	-1.6	-3.3
II-21	UUGAGUCC	7/59	-5.2	-5.5	-1.0	-2.3
II-22	AUAACCUCGG	1/59	-6.0	-3.7	-2.8	-5.0

II-23	CUAUUUGAGC	1/59	-1.8	-3.2	-0.5	-0.4
Mean $\pm$ SD			-6.1 $\pm$ 2.1	-4.5 $\pm$ 2.3	-2.5 $\pm$ 1.9	-3.0 $\pm$ 1.4

<sup>a</sup> Sequences are shown in their minus-sense, 3' to 5' orientation.

<sup>b</sup> Clones containing this sequence also had an insertion of A<sub>10</sub>C (see Figure 2.3C).

<sup>c</sup> Clones containing this sequence also had an insertion of A<sub>5</sub>C (see Figure 2.3C).

<sup>d</sup> Total recovered from six plants

<sup>e</sup> Minus-strand sequences were shuffled and then folded three times with the averages presented as described in Materials and Methods. The complements of the shuffled minus-strand sequences were also folded to give the averaged plus-strand randomized values.



sequences is occurring in planta. The replacement sequence in II-16 (ACUAGCCCGUUAGCCCGU) contains a 10-base sequence (underlined) followed by a repeat of the terminal 8 bases (double underline). In addition, one 2<sup>nd</sup> round winner (II-21) contained only an 8-base replacement sequence.

The 3<sup>rd</sup> round of competition was conducted by combining equal portions of full-length transcripts synthesized for each of the 25 2<sup>nd</sup> round winners and then inoculating the mixture with TCV genomic RNA transcripts onto six seedlings. Of the 46 clones recovered from five of these plants at 21 dpi, 29 were clone II-18 and 10 were either II-3a (7 clones) or II-3b (3 clones). Two additional clones, II-2 and II-17, were isolated from only single plants. All these clones were designated as 3<sup>rd</sup> round winners. To determine if the number of clones recovered in the 3<sup>rd</sup> round reflected the fitness of the individual clones to accumulate in plants, equal amounts of transcripts of pairs of 3<sup>rd</sup> round winners were combined and inoculated with TCV genomic RNA onto three turnip seedlings. At 21 dpi, total RNA was extracted from an uninoculated leaf and approximately equal numbers of clones sequenced from each of the three plants. In direct competition between II-18 and II-3a, 70% of the clones recovered (21/30) were II-18. When plants were inoculated with II-18 and II-2, 92% (24/26) of the recovered clones were II-18. The second most prevalent winner, II-3a out competed II-2 (22/34), while II-2 and II-17 were of similar fitness when assayed together (19/35 clones were II-2). These results indicate that the number of clones accumulating in the 3<sup>rd</sup> round SELEX plants reflect the fitness of the clones under the infection conditions used.

### **Sequence and structural composition of the 2<sup>nd</sup> round winning sequences**

Analysis of the minus-strand M1H replacement sequences in the 2<sup>nd</sup> round 10-base SELEX winners indicated a disproportionate number of cytidylates (33% of the total bases) with a strong preference for multiple consecutive cytidylate residues (5.2-fold greater than expected for random sequences). Cytidylates also comprised 35% of the total number of residues found for the 28-base replacement sequence winners (Zhang and Simon, 2003b). Two of the 10-base SELEX 2<sup>nd</sup> round winners contained motifs found in TCV promoter elements (Figure 2.4, Class I) and two 2<sup>nd</sup> round winners contained sequence similar to a motif found in numerous 28-base SELEX winners (AACCCCU) but not present in any known TCV promoter-like element (Figure 2.4, Class II). Seven winners from the 1<sup>st</sup> round and six winners from the 2<sup>nd</sup> round contained a new motif consisting of an AU-rich sequence followed by C<sub>1-4</sub>, A<sub>1-2</sub>, G<sub>1-2</sub> and C/U (Figure 2.4, Class III). This motif was similar to the 7-base sequence in the 28-base SELEX winner UC (UCAGGAA). Several of the winners also contained base changes or inserted sequence outside of the replacement sequence region. Two of the 3<sup>rd</sup> round winners, II-17 and II-18, had the sequences “UUUUUUUUUG” and “UUUUUG”, respectively (minus-sense orientation), inserted downstream of their replacement sequences while 2<sup>nd</sup> round winner II-9 had an upstream A to U alteration (see Figure 2.5B and C).

To determine if the 10-base M1H replacement sequences in the 2<sup>nd</sup> round winners formed hairpins in conjunction with nearby sequences, mFold structural predictions (Mathews et al., 1999) were generated for all 25 winning sequences along with 11 upstream and 8 downstream bases. Sixty-eight percent of the winning sequences formed local hairpins that were more stable in their plus-sense orientation than in their minus-sense orientation, with the plus-strand structures, on average, 36% more stable than

**Figure 2.4** Motifs found in 10-base SELEX M1H replacement sequences. All sequences are shown in their minus-sense, 3' to 5' orientation. TCV and satC minus-strand promoter elements (3'CCS) are described in the text. Common motifs (classes I through IV) are underlined. Round 1, winners from the 1st round of the SELEX. Asterisk denotes that the clone was also a 3rd round winner.

**Class I. TCV and satC 3'CCS motifs**

TCV 3'CCS CCAUUAGUCGU  
II-14 CAUUAGCCCA  
satC 3'CCS CCCUAUAGAUUCCC  
II-25 CUAUUUGAGC

**Class II. AACCCCU motif**

II-9 AUAAACCCGG  
II-7 GGUAACCCUA

**Class III. C U/A<sub>1-5</sub> C<sub>1-4</sub> A<sub>1-2</sub> G<sub>1-2</sub> C/U**

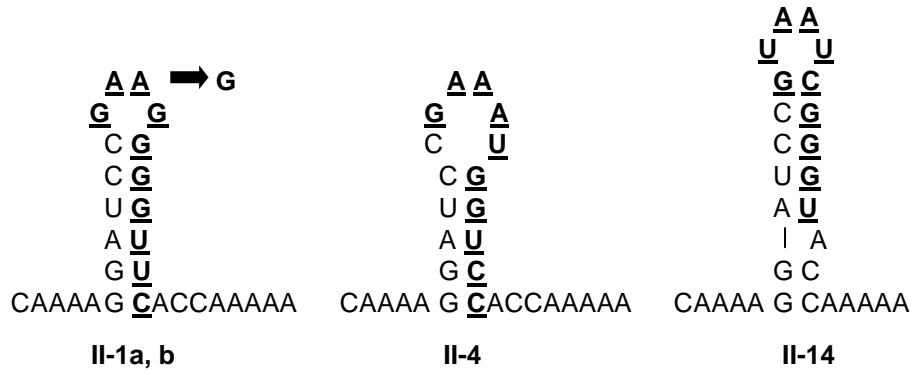
II-2\* UACUCAAGGC  
Round 1 CAAGGCUACC  
Round 1 CAAUCCAAGG  
Round 1 CAUUCCAAG  
II-1a CUUCCCAAG  
II-1b CUCCCCAAG  
  
II-3a\* UUAAACCAGG  
Round 1 ACCAGGAACA  
II-4 CAAAUCCAGG  
II-15 CUUACCAGGU  
Round 1 AAUACCAGAU  
Round 1 UUUGCAGGUU  
Round 1 UUGUCAGGCG

**Class IV. Additional miscellaneous similarities**

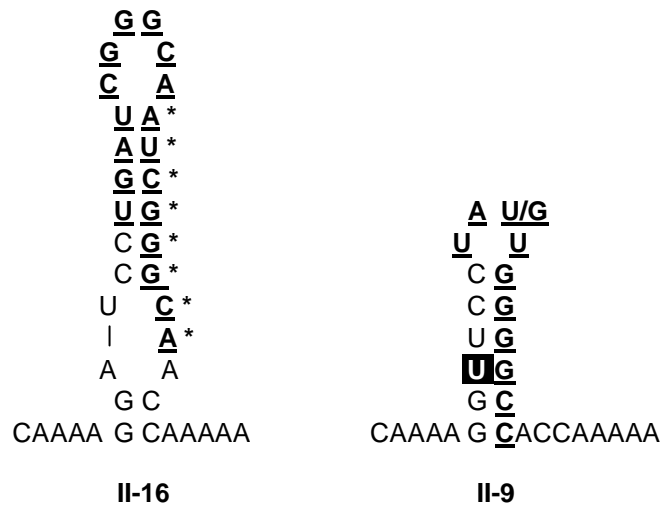
II-12 CGAAGCCCGC  
II-16 ACUAGCCCGUUAGCCCGU  
II-14 CAUUAGCCCA  
Round 1 UAGAGCCCCC  
  
II-6 AUAAGUCCCU  
II-21 UUGAGUCC  
II-11 AUAGUGUCC  
  
Round 1 UAGGCCUGGU  
II-13 UUUGCCUGGC

**Figure 2.5** Computer predicted plus-strand structures encompassing the M1H replacement sequences. Replacement sequences are underlined. (A) Structures formed by the complement of Class III (II-1a, II-1b, II-4) and Class IV (II-14) replacement sequences in 2<sup>nd</sup> round winners (Figure 2.2). II-1a and II-1b differ by either an adenylate or guanylate in the 4-base loop. (B) Two 2<sup>nd</sup> round winners, II-16 and II-9, have additional alterations (a repeat of 8 bases [denoted by asterisks] or an adenylate to uridylate transversion [boxed]) that increase the stability of the plus-strand hairpins. II-9 clones contained either a uridylate or a guanylate in the 4-base loop, as shown. (C) Structures of the five 2<sup>nd</sup> round winners that were recovered in the 3<sup>rd</sup> round. II-3b contains 2 base differences (U to A and G to A) compared with II-3a as shown. For II-18 and II-17, inserted bases outside the replacement sequences that are not found in wt satC are denoted by asterisks.

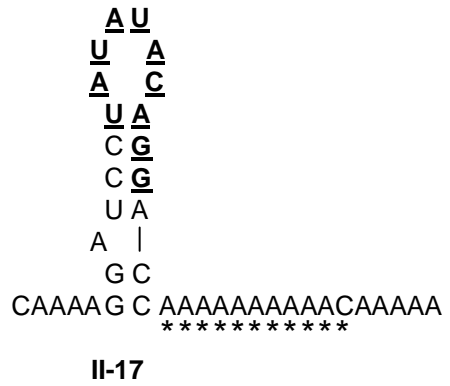
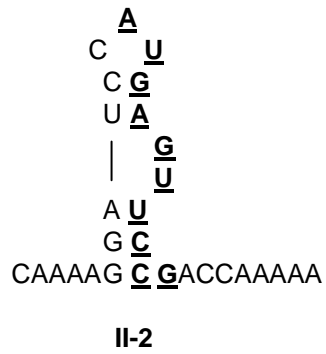
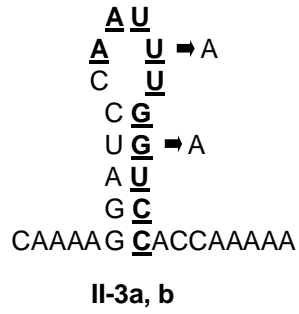
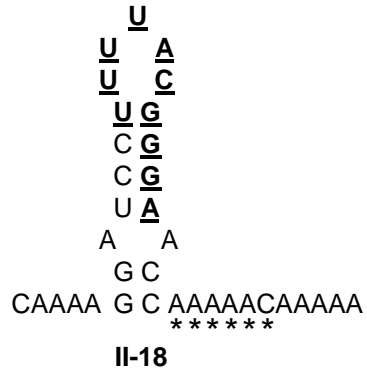
**A**



**B**



C



minus-strand structures (Table 2.2). All hairpins were present in the most stable structures predicted for full-length plus-strands of 3<sup>rd</sup> round winners, while only two of five 3<sup>rd</sup> round winners had hairpins that were present in the most stable full-length satC minus-strand structures.

The plus-strand hairpins predicted to form for nine of the 2<sup>nd</sup> round winners, including the four 3<sup>rd</sup> round winners, are shown in Figure 2.5, A-C. The plus-strand structures of the related clones II-1a and b were not altered by the base differences. For II-3a and II-3b, the sequence differences did not affect the structure of the plus-sense hairpins (U:G or U:A pairing in the stem), but would affect the structure of the minus-strand hairpins. The additional 8-base repeat in II-16, which contained the unusual 18 base replacement sequence, increased the stability of the plus-sense hairpin by 70%. The second site alteration in II-9 increased the stability of the predicted plus-sense hairpin, but not the minus-strand hairpin, by 64%. Most of the Class III motifs found in 1<sup>st</sup> and 2<sup>nd</sup> round winners allowed for the formation of hairpins with AU-rich loops and perfectly base-paired stems (for example, see II-4 in Figure 2.5A and II-3a/b in Figure 2.5C), with the replacement sequence extending down one side to the base of the stem. Class IV motifs formed structures similar to that shown for II-14, where unaltered satC sequence formed the base on both sides of the stem with the replacement sequence forming the loop and upper portion of the stem. All these structures were similar to the wt M1H and the UC hairpin (Figure 2.1), by consisting of hairpins flanked by CA-rich sequences. The possible importance to satC fitness of maintaining CA-rich sequences at the base of a hairpin was also suggested by two of the five 3<sup>rd</sup> round winners (II-17 and II-18),



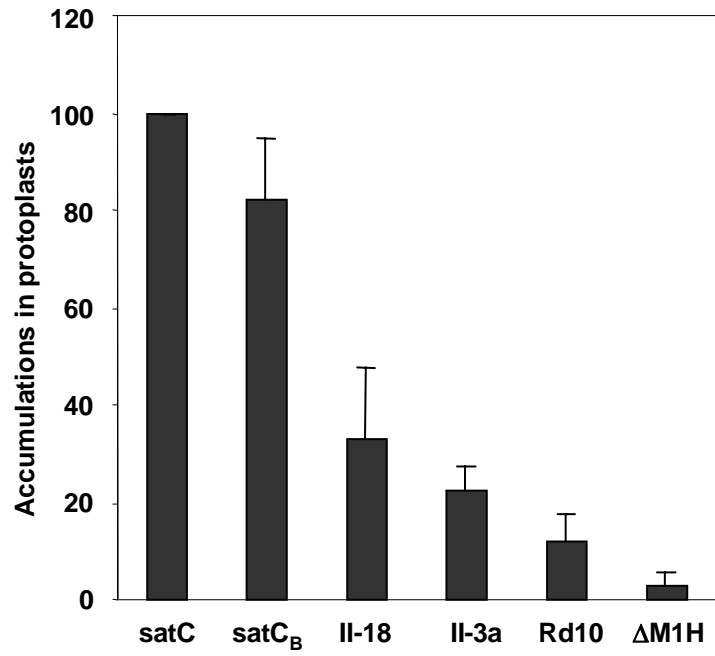
containing additional CA-rich sequence in their plus-strands not found in the parental transcripts (Figure 2.5C).

### **Effect of 10-base winning sequences on satC replication in protoplasts**

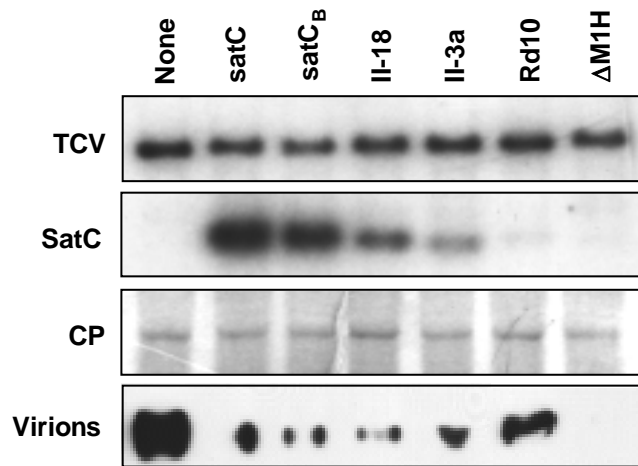
As described in the Introduction, fitness of satC to accumulate in plants depends on several factors including the ability of the satRNA to replicate and its ability to positively affect the systemic infection of the helper virus. Replication of satC can be directly assayed using host *Arabidopsis thaliana* protoplasts, since alterations to M1H, including deletion of the hairpin, have no effect on stability of the RNA (Zhang and Simon, 2003b). To determine how well the selected 10-base sequences contribute to replacing the replication enhancer function of M1H, replication levels in protoplasts was determined for II-18 and II-3a, the two most fit 3<sup>rd</sup> round winners and compared with that of wt satC and satC<sub>B</sub>. SatC<sub>B</sub>, the parental satRNA of the SELEX constructs, contains a single base alteration near the base of M1H resulting in the incorporation of several bases of flanking sequence into the hairpin according to computer models (Figure 2.1). As shown in Figure 2.6A, satC<sub>B</sub> replicates to 82% of wt satC levels at 36 hpi. Deletion of the 28-base M1H ( $\Delta$ M1H) resulted in a 97% decrease in satC levels. SatC with a randomly selected 10-base sequence replacing M1H (Rd10) replicated at 10% of wt satC levels. This increase over levels obtained with  $\Delta$ M1H may be due to size effects, since previous results have shown that satC and other TCV subviral RNAs replicate more poorly when reduced in size (Carpenter et al., 1991; Li and Simon, 1991; Zhang and Simon, 1994). Third round winners II-18 and II-3a replicated to levels 2.7-fold and 1.8-fold higher than Rd10, respectively, with II-18 reaching levels that were 33% more than the less fit II-3a.

**Figure 2.6** Accumulation of viral RNAs and virions in protoplasts. *Arabidopsis* protoplasts were co-inoculated with TCV genomic RNA transcripts and transcripts of various satRNAs. (A) Total RNA was extracted at 36 hpi and subjected to RNA gel blot analysis using an oligonucleotide probe complementary to both TCV and satC. Values shown were calculated from at least three independent experiments. (B) Effect of different M1H replacement sequences on virion accumulation in protoplasts at 40 hpi. CP and virions were subjected to electrophoresis on SDS-PAGE gels and detected by Coomassie brilliant blue staining or chemiluminescence following treatment with anti-TCV CP antibody, respectively. satC, wt satC; satCB, the parental satRNA used to generate satC with randomized sequences; II-18 and II-3a, 3rd round SELEX winners; Rd10, satC with a randomly chosen 10-base sequence replacing the 28 bases of the M1H ; DM1H, satC with a deletion of the 28 bases of the M1H.

**A**



**B**



However, II-18 accumulated 2.5-fold less than parental satC<sub>B</sub>, suggesting that the replacement sequence in II-18 is unable to enhance replication of satC comparable to the natural M1H enhancer. The 18-base size difference between II-18 and satC<sub>B</sub> might also contribute to the difference in levels of the satRNAs. II-18 and II-3a accumulated to similar levels as UC, the 7-base replacement sequence winner from the previous M1H SELEX (Zhang and Simon, 2003b). These results suggest that the 10-base M1H replacement sequences are weakly enhancing the replication of satC in the absence of the natural enhancer. The contribution to replication of the A<sub>5</sub>C insert 3 bases downstream of the replacement sequence in II-18 is presented below.

### **Two most fit 3<sup>rd</sup> round winners interfere with TCV virion accumulation in protoplasts**

Previous results suggested that the 7-base replacement sequence in UC might contribute to the fitness of the satRNA by interfering with TCV virion accumulation (Zhang and Simon, 2003b). To determine if the replacement sequences of the 3<sup>rd</sup> round winners also repress virion accumulation, protoplasts were inoculated with TCV and various satRNA species and levels of viral RNA, total CP and virions determined at 40 hpi. As shown in Figure 2.6B, co-inoculation of TCV with satRNAs did not reproducibly affect the total amount of CP. Wt satC and satC<sub>B</sub> replicated to higher levels than II-18 and II-3a, which replicated better than Rd10 and ΔM1H as previously shown when protoplasts were assayed at 36 hpi (Figure 2.6A). Wt satC, satC<sub>B</sub>, II-18 and II-3a all reduced virion accumulation by similar amounts, although II-18 and II-3a replicated to levels that were 3- to 4.5-fold less than wt satC. II-18 and II-3a were more efficient at

reducing virion levels than Rd10, with a 7-fold greater reduction in virion levels for II-18 compared with Rd10 while satRNA levels varied by less than 3-fold. Surprisingly, the most efficient satRNA at reducing virion levels was  $\Delta$ M1H, with no virions detected despite very low satRNA accumulation. This result suggests that a sequence non-specific hairpin in this plus-strand location may function to bring into proximity the CA-rich sequences that flank the hairpin, which would also result from a deletion of the hairpin sequence.

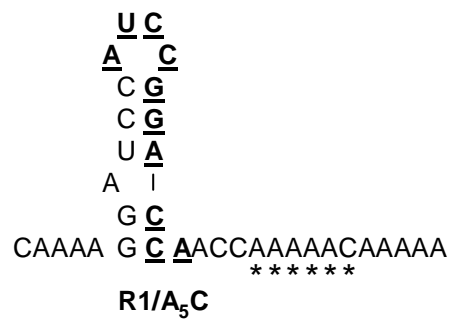
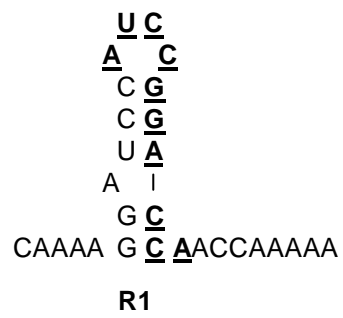
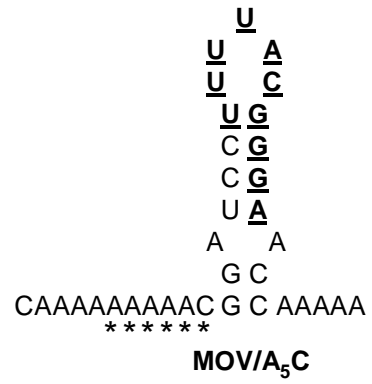
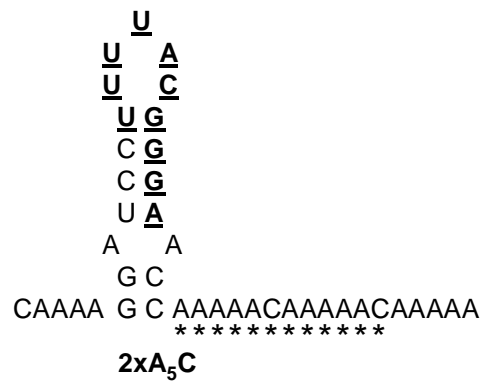
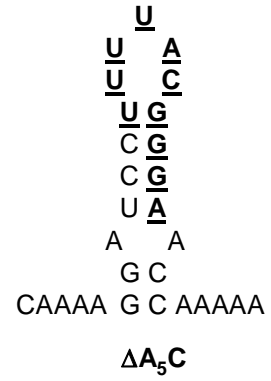
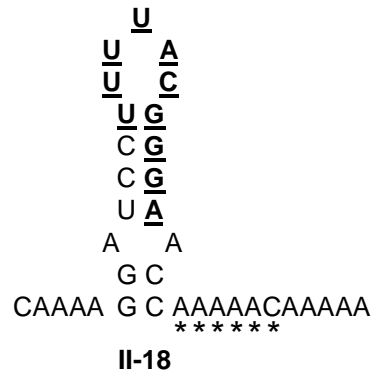
### **Role of CA-rich sequence flanking M1H in satC replication and virion accumulation**

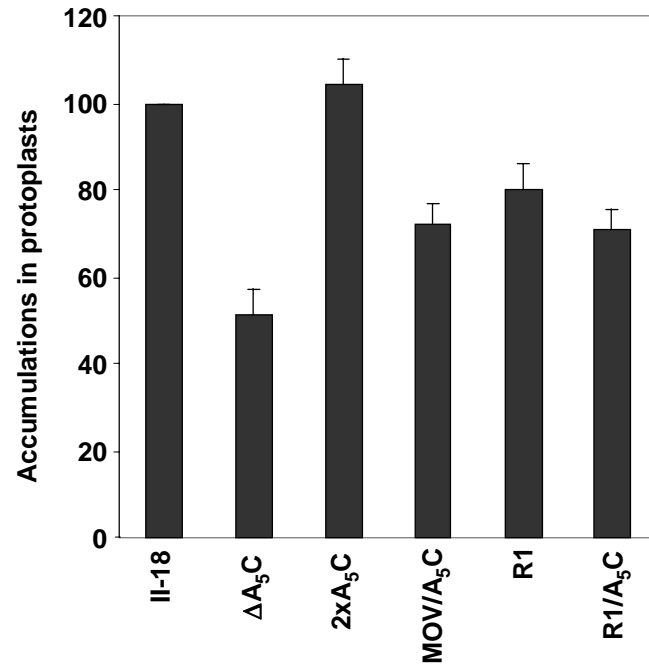
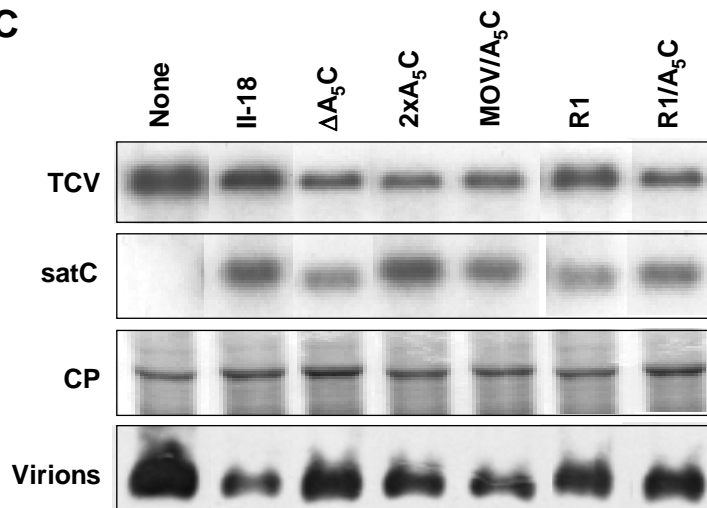
Previous results indicated that deletion of the CA-rich single-stranded sequences flanking M1H reduced satC levels in protoplasts (Nagy et al., 2001). Thus, these sequences may serve a dual function by inhibiting virion levels and enhancing satC replication. Since two of the five 3<sup>rd</sup> round winners had insertions of additional CA-rich sequence just downstream of M1H, we tested if such inserts conferred additional fitness to the satRNA. Mutations were introduced into II-18 to generate a series of derivatives in which the A<sub>5</sub>C insert was deleted, duplicated, or moved upstream of the hairpin (Figure 2.7A). In addition, the 10-base M1H replacement sequence in II-18 was exchanged with the replacement sequence from a randomly selected 1<sup>st</sup> round winner (R1), which was not isolated in further rounds and thus judged less fit than 3<sup>rd</sup> round winning sequences.

Transcripts of the different II-18 derivatives were co-inoculated with TCV genomic RNA into protoplasts and total RNA was extracted at 36 hpi and analyzed by Northern hybridization (Figure 2.7B). Deletion of the A<sub>5</sub>C insert from II-18 reduced RNA accumulation of the resultant satRNA ( $\Delta$ A<sub>5</sub>C) by nearly 50%, suggesting a role for

**Figure 2.7** Effect of the A5C insert in 3<sup>rd</sup> round SELEX winner II-18 on replication and virion accumulation. (A) Predicted plus-strand structures encompassing the M1H replacement sequence and flanking regions. The M1H replacement sequences are underlined and inserted sequences are denoted by asterisks. II-18, most fit 3<sup>rd</sup> round winner;  $\Delta$ A5C, II-18 with a deletion of the A5C insertion; 2xA5C, II-18 with a duplication of the A5C insert; MOV/A5C, II-18 with the A5C insert deleted and moved upstream of the predicted hairpin; R1, randomly selected sequence from a 1<sup>st</sup> round winner; R1/A5C, R1 with an insert of A5C. (B) Accumulation of satRNAs in *Arabidopsis*. Total RNA was extracted at 36 hpi and subjected to RNA gel blot analysis using an oligonucleotide probe complementary to both TCV and satC. Values shown were calculated from at least three independent experiments. (C) Accumulation of viral RNAs, coat protein, and virions in protoplasts inoculated with TCV transcripts alone or together with transcripts of satRNAs. TCV genomic RNA and satRNA were detected by RNA gel blot analysis using an oligonucleotide probe complementary to both TCV and satC. CP and virions were subjected to electrophoresis on SDS-PAGE gels and detected by Coomassie brilliant blue staining or chemiluminescence following treatment with anti-TCV CP antibody, respectively.

**A**



**B****C**



the additional sequence in enhancing replication of II-18. Relocating the A<sub>5</sub>C insert upstream of the hairpin increased replication by 30% over  $\Delta$ A<sub>5</sub>C, suggesting that the CA-rich sequence on either side of the hairpin can contribute to satRNA levels in protoplasts. In contrast, replication of 1<sup>st</sup> round winner R1 was not enhanced by the addition of A<sub>5</sub>C downstream of the hairpin. However, the R1 replacement sequence terminates with "ACCA", which may already be contributing towards enhancement of replication not impacted by additional CA-rich sequence. This possibility was supported by finding that tandem duplication of the A<sub>5</sub>C sequence in II-18 (2xA<sub>5</sub>C) also did not further enhance replication.

To examine the effect of the A<sub>5</sub>C insert on virion accumulation, TCV genomic RNA transcripts along with transcripts of II-18 and its derivatives were inoculated onto protoplasts and total RNA, protein and virions were extracted at 40 hpi. As previously shown in Figure 2.6B, association of TCV with a satRNA did not affect the total amount of CP accumulating in infected protoplasts (Figure 2.7C). Movement of the A<sub>5</sub>C insert upstream of the hairpin did not appreciably affect virion levels compared with II-18. However, deletion of the A<sub>5</sub>C insert from II-18 increased virion accumulation by nearly 2.5-fold, suggesting that an insert of A<sub>5</sub>C upstream or downstream of the hairpin enhanced the satRNA's ability to repress the accumulation of virions. Additional CA-rich sequence (2xA<sub>5</sub>C) did not further contribute to virion repression compared with II-18. The 1<sup>st</sup> round winner R1 was less effective at reducing virion accumulation compared with II-18 and addition of A<sub>5</sub>C did not enhance the ability of R1 to reduce virion levels. However, R1 was almost twice as efficient at reducing virions compared with  $\Delta$ A<sub>5</sub>C (i.e., the II-18 replacement sequence alone). Since R1 also accumulated in protoplasts to 38%

higher levels than  $\Delta A_5C$ , the enhanced fitness of II-18 is primarily due to the  $A_5C$  insert and not its M1H replacement sequence. These results strongly suggest that a sequence non-specific plus-strand hairpin flanked by CA-rich sequences can contribute to satRNA fitness by repressing the accumulation of virions.

## Discussion

The previous correlation between fitness and repression of virions was based on only a single SELEX winner, clone UC with a 7 base M1H replacement sequence (Zhang and Simon, 2003b). I repeated the in vivo functional SELEX using a 10-base replacement sequence. While several winning replacement sequences in the 28-base SELEX were able to enhance satC replication to near satC<sub>B</sub> levels (Zhang and Simon, 2003b), the 10-base replacement sequences in the most fit satRNA of the current study were only weakly effective at enhancing replication. II-18, the most fit 10-base SELEX winner, accumulated to levels that were 2.7-fold greater than a randomly selected sequence (Rd10), but only 40% of parental satC<sub>B</sub> levels (Figure 2.6A). In addition, while 48% of 1<sup>st</sup> and 2<sup>nd</sup> round winners of the 28-base SELEX winners contained one or more minus-sense motifs found in a variety of replication-like elements, only three of 83 1<sup>st</sup> and 2<sup>nd</sup> round winners of the current 10-base SELEX had similar replication-like motifs and none were found in 3<sup>rd</sup> round winners. However, all 25 of the 10-base SELEX 2<sup>nd</sup> round winners contained replacement sequences that were predicted to fold into similar hairpins, flanked by unaltered single-stranded CA-rich sequences. The principal motifs (Class III and IV) found in a number of 10-base SELEX winner differed primarily on

whether the replacement sequences extended midway down, or to the base, of the hairpin (compare IIa/b and II-4 with II-14, Figure 2.5A). This result suggests that a 10-base region is insufficient space to specify both replication-like motifs and the sequence required to form plus-strand hairpins. Since all the 2<sup>nd</sup> round winning sequences formed similar hairpins, this suggests that the major selective pressure on the 10-base M1H replacement sequences was for plus-strand hairpin formation and not replication enhancement. It cannot be ruled out, however, that the high percentage of selected sequences with multiple C residues in the minus-sense orientation in both the 28-base and 10-base SELEX may also weakly enhance replication since recent reports indicated that CCA or CCCA repeats are capable of serving as independent promoters of transcription in vitro using RdRp from several viruses including TCV (Tretheway et al., 2001; Yoshinari and Dreher, 2000; Yoshinari, et al., 2000).

The replacement sequences of two of the 3<sup>rd</sup> round winners, II-17 and II-18, formed hairpins such that a downstream, originally single stranded "ACC" formed the 3' base of the hairpin. Both of these winners contained unusual downstream inserts of A<sub>10</sub>C or A<sub>5</sub>C, respectively, which extended the single-stranded CA-rich sequences at the 3' base of the hairpin. For II-18, A<sub>5</sub>C inserted either upstream or downstream of the hairpin contributed to both the replication of the satRNA as well as the ability of the satRNA to reduce virions. While it is currently unknown if the plus-sense CA-rich sequence or its complementary minus-strand sequence is responsible for enhancement of replication, the strong preference for plus-strand hairpins in the 2<sup>nd</sup> and 3<sup>rd</sup> round winners suggest that the A<sub>5</sub>C insert in its plus-sense orientation enhanced virion reduction and was responsible for the selection of II-18 as the most fit of the 10-base SELEX winners.

$\Delta$ M1H replicated very poorly in protoplasts yet was highly effective at reducing virion accumulation. This result suggested that a hairpin in this location primarily functions to bring together the flanking CA-rich sequences, which would also be achieved by deletion of the region between the CA-rich sequences. The mechanism by which the CA-rich sequences repress virion accumulation is not known. Qu and Morris (1997) determined that a bulged hairpin located near the 3' end of the CP ORF was a specific packaging signal for TCV genomic RNA based on its ability to independently promote the packaging of a heterologous RNA by the TCV CP in protoplasts. Interestingly, this hairpin in the TCV-M isolate used in my study contains an A-rich (five of nine residues) loop. While the sites within the hairpin that specifically interact with the CP are not known, hairpin loops are frequently targets of RNA-binding proteins (Nowakowski and Tinoco, 1997). It is therefore possible that the multiple consecutive adenylates within the CA-rich satC sequences compete for CP binding thus disrupting virion assembly.

A second possibility is that virion assembly requires sequences in addition to the hairpin in the CP ORF. The lack of a similar hairpin in packaged subviral RNAs satC, satD and diG, a defective interfering RNA containing no CP ORF sequence, is indicative that other packaging signal sequences have yet to be identified or the RNA interacts with the genomic RNA for co-packaging. As described in Chapter I, H4(+) is similar to M1H(+) in structure, location relative to the 3' end of the RNA, and flanking CA-rich sequences. Since the second recombination event producing satC occurred near the 3' base of H4(+), satC and TCV share related downstream sequences including the 3' side CA-rich sequence at the base of M1H(+) (see Figure 1.3). However, the majority of TCV

H4(+) and 5' CA-rich sequence is unrelated to sequence in satC. It is possible that the U-rich loop of H4(+) and flanking CA-rich sequences, elements also found in diG, are involved in virion assembly, which might be inhibited by similar CA-rich sequences in satC.

## **CHAPTER III**

### **SHORT INTERNAL SEQUENCES INVOLVED IN REPLICATION OF SATELLITE RNA C AND VIRION ACCUMULATION**

#### **Introduction**

Replication of positive-strand RNA viruses is mediated by the viral RdRp, which together with possibly other viral or host factors (Lai, 1998) forms an initiation complex at the promoter and initiates de novo or primer-directed transcription of the complementary minus strand. The minus-strand intermediate is then used as a template for synthesis of plus strands. The relative levels of the two strands are often highly asymmetric (Buck, 1996), suggesting the existence of a replication machinery switch, which represses minus-strand synthesis thereby constraining the RdRp to synthesize plus strands. As described in Chapter I, some viruses regulate the switch by changing the conformation of 3'-proximal structures to allow or deny access of the RdRp to the promoter or 3'-terminal sequences, which may be mediated by one or more unstable base-pairs occurring between complementary short sequences (Koev et al., 2002; Olsthoorn et al., 1999; Pogany et al., 2003; Zhang et al., 2004). Sequence and structural elements involved in efficient viral RNA synthesis are not limited to the 3'-proximal regions of viral RNAs. RNA elements located in variable positions on both strands that function as transcriptional enhancers or repressors may also be involved in 3' structural conformation changes and/or sequestration of 3'-terminal bases (Pogany et al., 2003; Ranjith-Kumar et al., 2003; Ray and White, 1999; 2003; van Rossum et al., 1997; Zhang et al., 2004).

M1H of satC is a hot spot for recombination between satC and satD in vivo (Cascone et al., 1993; Nagy et al., 1998) and a replication enhancer in its minus-sense orientation in vivo and in vitro (Nagy et al., 1999; 2001). As described in Chapter II, a plus-strand hairpin in this location is important for satC fitness, possibly functioning to bring together flanking sequences. By presumptively juxtaposing these flanking sequences, satC is able to interfere efficiently with TCV virion formation, resulting in accumulation of additional free CP to better suppress RNA silencing (Qu et al., 2003; Zhang and Simon, 2003a). As described in Chapter II, sequences flanking M1H have been implicated in the replication of satC in addition to a role in suppressing virion accumulation (Nagy et al., 2001).

In this Chapter, I demonstrate the importance of specific sequences flanking M1H for satC fitness and establish their role in replication and virion repression by sequence randomization followed by in vivo functional SELEX. Analyses of winning (functional) satRNAs revealed three different conserved elements within the regions. Examination of these three elements in the context of a control RNA suggested that they are involved in satC replication and/or virion repression.

## **Materials and Methods**

### **In vivo functional SELEX**

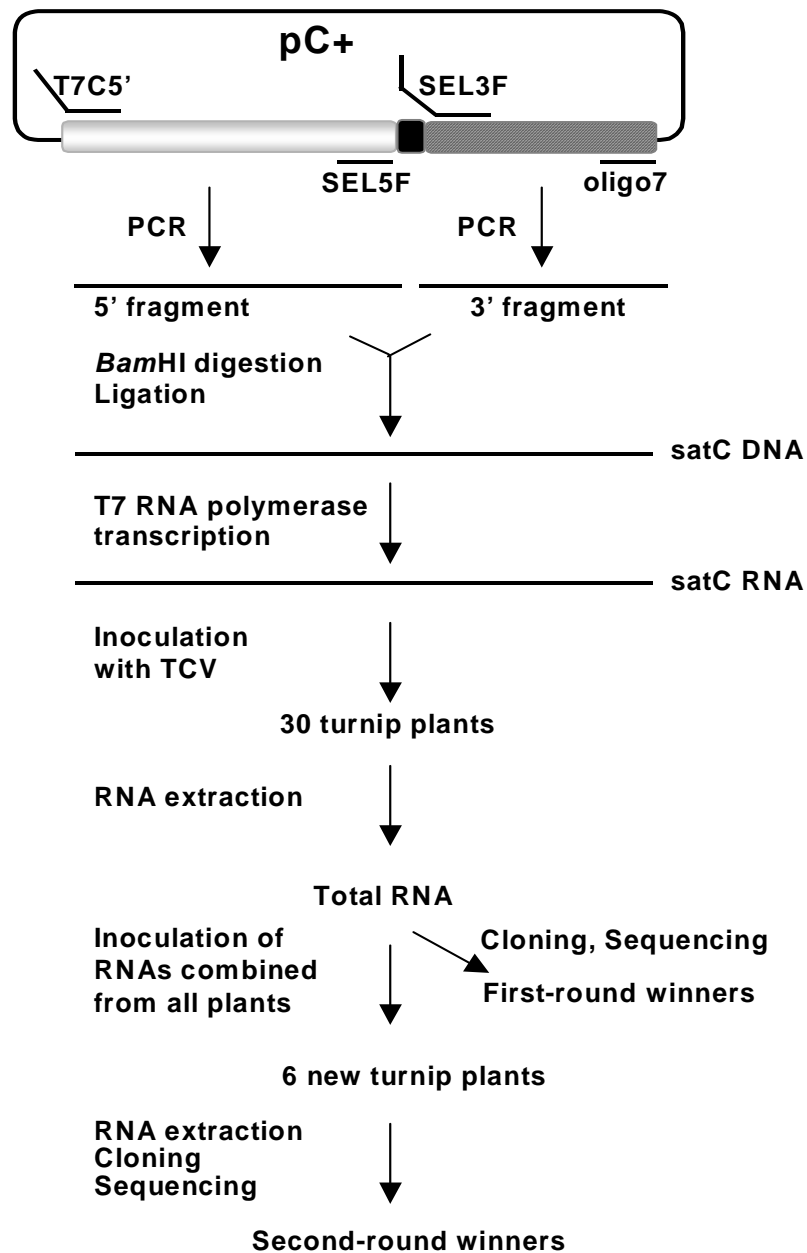
In vivo functional SELEX was performed as described in Chapter II. To produce full-length satC containing random bases replacing the sequence flanking M1H, two

fragments were generated by separate PCRs using pC+ (pUC19 containing full-length satC cDNA) as template (see Figure 3.1 for procedure flow chart). The 5' fragment was produced using primers T7C5' (Table 2.1), which contains a T7 polymerase promoter at its 5' end, and SEL5F (Table 3.1). The 3' fragment was generated using primers SEL3F (Table 3.1) and oligo7 (Table 2.1), which is complementary to satC 3'-terminal sequence. In these two PCRs, a new *Bam*HI site was generated in satC cDNA by changing a U to C at position 176 for convenient ligation of the purified PCR fragments. As described in Chapter II, satC with this new *Bam*HI site is referred to as satC<sub>B</sub>. The 5' and 3' PCR fragments were treated with *Bam*HI, purified through 1.5% agarose gels and then ligated to produce full-length cDNA. satC<sub>B</sub> transcripts containing randomized sequences were directly synthesized from the ligation products using T7 RNA polymerase. For the first-round of selection, 5 µg of these satC<sub>B</sub> transcripts and 2 µg of TCV transcripts were coinoculated onto each of 30 turnip seedlings. Total RNA was extracted from uninoculated leaves at 21 dpi. Full-length satC was amplified by RT-PCR using primers T7C5' and oligo7, and then cloned into the *Sma*I site of pUC19 and sequenced. For the second-round of selection, equal amounts of total RNAs extracted from all 1<sup>st</sup> round plants were pooled. Approximately 5 µg of the pooled RNAs was inoculated onto each of six new turnip seedlings. Total RNA was extracted from uninoculated leaves at 21 dpi and the recovered satC cloned and sequenced.

### **Fitness comparison of SELEX winners in turnip plants**

Fitness competitions between winners were performed as described in Chapter II. Equal amounts of winning RNA transcripts were combined and inoculated onto a single





**Figure 3.1** Flow chart for performing the SELEX of satC LS and RS. pC+ is pUC19 that contains satC full-length plus-strand sequence (Song and Simon, 1994) inserted into the *Sma*I site of the plasmid. The primers used for PCR are given in Table 3.1. See the text for details.

**Table 3.1** Oligonucleotides used in Chapter III

Application /Construct	Name	Position <sup>a</sup>	Sequence <sup>b</sup>	Polarity <sup>c</sup>
SELEX	SEL5F	144-163	5'AGTC <b>GGATCC</b> <u>NNNNNNNNNN</u> NAAACAG CCAGGUUUUCACGC	-
	SEL3F	221-240	5'AGTC <b>GGATCC</b> CAGACCCTCCAGCCAAAG GGTAAATGGG <u>NNNNNNNNNN</u> GGCAGC ACCGTCTAGCTGCG	+
satC variants	OligoLS1r	144-163	5'AGTC <b>GGATCCT</b> CAATGCTGGGAAACAG CCAGGUUUUCACGC	-
	OligoLS2r	144-163	5'AGTC <b>GGATCCT</b> TTTTGGCTCGGAAACAG CCAGGUUUUCACGC	-
	OligoLSf	221-240	5'AGTC <b>GGATCC</b> CAGACCCTCCAGCCAAA GGGTAAATGGGTATGTGAAATGAGGCAG CACCGTCTAGCTGCG	+
	OligoRSr	144-163	5'AGTC <b>GGATCCT</b> CAATGCTCGGAAACAG CCAGGUUUUCACGC	-
	OligoRS1f	221-240	5'AGTC <b>GGATCC</b> CAGACCCTCCAGCCAAAG GGTAAATGGG <u>ACCAAAAAATG</u> AGGCAGCA CCGTCTAGCTGCG	+
	OligoRS2f	221-240	5'AGTC <b>GGATCC</b> CAGACCCTCCAGCCAAAG GGTAAATGGGTATGTGAACGGCGGCAGCA CCGTCTAGCTGCG	+

<sup>a</sup> Coordinates corresponds to those of satC.

<sup>b</sup> Bold residues denote bases changed to generate a *Bam*HI site. Mutant bases are underlined.

<sup>c</sup> “+” and “-” polarities refer to homology and complementarity with satC plus-strands, respectively.

leaf of three turnip seedlings together with TCV genomic RNA transcripts. Total RNA was extracted at 21 dpi, and satC cloned and sequenced.

### **Construction of satC variants**

Constructs LS1, LS2, RS1, RS2, and LS1RS1 were produced by ligation of two PCR fragments amplified using plasmid pC+ as template. The 5' fragments of LS1 and LS2 were amplified using primers T7C5' and OligoLS1r or OligoLS2r, respectively (Table 3.1) while the 3' fragments were produced by PCR using primers oligo7 and OligoLSf (Table 3.1). For RS1 and RS2, the 5' fragments were generated by PCR using primers T7C5' and OligoRSr (Table 3.1); the 3' fragments were obtained using primers oligo7 and OligoRS1f or OligoRS2f (Table 3.1). The 5' fragment for construction of LS1 and the 3' fragment for construction of RS1 were used to produce LS1RS1. 5' and 3' PCR fragments were digested with *Bam*HI, gel-purified, ligated together, and inserted into the *Sma*I site of pUC19. All clones were confirmed by sequencing.

### **Protoplast inoculation and RNA and virion preparation**

Protoplasts ( $5 \times 10^6$ ) prepared from callus cultures of *Arabidopsis thaliana* ecotype Col-0 were inoculated with 20  $\mu$ g of TCV genomic RNA transcripts with or without 2  $\mu$ g of wt satC or satC variant transcripts. Total RNA and virions were extracted from protoplasts at 40 hpi as described in Chapter II.

### **Northern blot analysis**

Total RNA was denatured with formamide and separated by non-denaturing

agarose gel electrophoresis. The blot analysis was performed as described in Chapter II. Plus-strand RNAs were hybridized with [ $\gamma$ - $^{32}$ P] ATP-labeled probe complementary to positions 3950 to 3970 of TCV genomic RNA and positions 250 to 269 of satC. Minus-strand RNAs were detected using a [ $\alpha$ - $^{32}$ P] UTP-labeled probe complementary to the 5'-terminal 151 bases of satC minus-strands and the 5'-terminal 155 bases of TCV minus-strands with nine single-base mismatches.

### **Western blot analysis**

Virions were separated on a 1% agarose gel. Western blotting was performed using polyclonal antibody against TCV CP and detected by chemiluminescence (Pierce Biotech) as described in Chapter II.

## **Results**

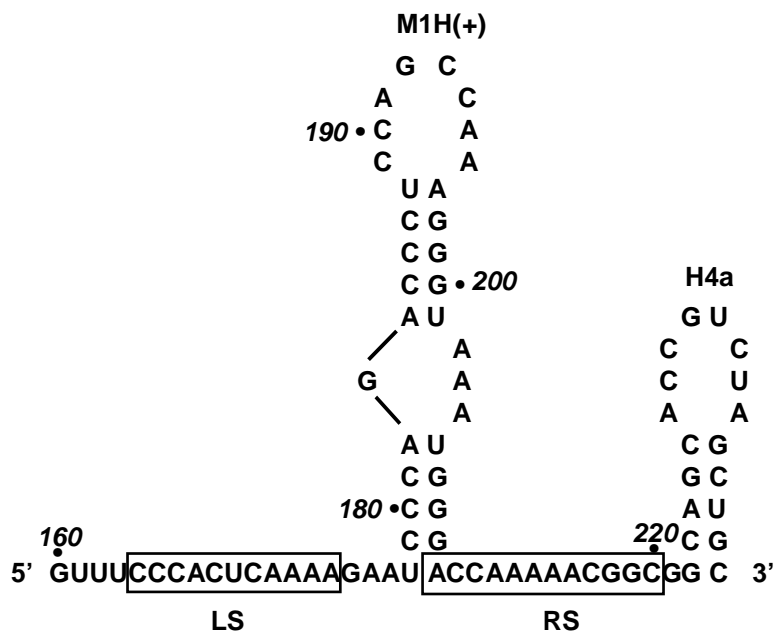
### **In vivo functional SELEX of the satC LS and RS regions**

As described in Chapter II, study on sequence and structural requirements of the satC M1H replication enhancer suggested that flanking sequences contribute to both replication and interference with TCV virion accumulation. Understanding the underlying mechanisms responsible for these properties requires first identifying specific functional sequences within these regions. To identify such elements, in vivo functional SELEX was performed, which allows for selection of sequences important for fitness regardless of function. SELEX was performed by replacing the 11-base region on the 5' side of M1H (labeled LS) and the 12-base region on the M1H 3' side (labeled RS) with random

sequences (Figure 3.2). This construction required an adenylate to guanylate transition at position 176 to create a new restriction site. This parental satRNA, satC<sub>B</sub>, replicates to 82% of wt satC levels in protoplasts and inhibits virion accumulation as efficiently as wt satC as described in Chapter II.

SatC<sub>B</sub> transcripts containing randomized LS and RS sequences were co-inoculated onto 30 turnip seedlings along with TCV genomic RNA transcripts. All inoculated plants displayed satC symptoms (dark green, highly crinkled and stunted leaves) within two weeks and total RNA was extracted from uninoculated leaves at 21 dpi. Agarose gel electrophoresis of the extracted RNA indicated that all plants contained visible satC species as detected by ethidium bromide staining. Full-length satC was recovered by RT-PCR and 37 clones were sequenced from six randomly selected plants. Twenty-five clones were unique and were designated as 1<sup>st</sup> round winners (Table 3.2; all sequences in this Chapter are presented in their plus-sense, 5' to 3' orientation, unless otherwise noted). No second-site alterations beyond the selected regions were observed.

The selected LS sequences were enriched for cytidylates (40%) and adenylates (32%) (Table 3.2). Only 18 guanylates were present in 275 possible positions (7%). While no 1<sup>st</sup> round winners contained wt LS, the LS of S1-13 differed from wt at only three of 11 positions suggesting that most bases in this region of wt satC are sequence specific (Table 3.2). Six of 25 1<sup>st</sup> round selected LS contained the wt sequence CCCA, with clone S1-24 containing two copies of this sequence. Six other winners contained CCUA (repeated twice in S1-18). Four winners had the related sequence CCUG. Based on the LS sequences of the 1<sup>st</sup> round winners, the most common recovered elements can be summarized as CCCA and CCUA/G. Two winners also contained the wt element



**Figure 3.2** Sequence and structure of a portion of satC 3' region. M1H is a hairpin composed of sequence from satD and the two discontinuous regions of TCV (see the text for details). Boxed sequences were subjected to simultaneous in vivo SELEX. Numbering is from the 5' end.

**Table 3.2** Results of in vivo functional SELEX

	Clone	Winning sequences <sup>a</sup>		Plants <sup>b</sup>						Total	
		LS	RS	1	2	3	4	5	6		
Round 1	S1-1	CCUACAAAACC	GUGGACCGUGGCgg	2						2	
	S1-2	CCUACAAAACC	GAGCUCGUCGGCgg	1						1	
	S1-3	UCCCCUGUUAC	CGACCUACGCGGgg	1						1	
	S1-4	UCCUGAUUUCU	ACUACCGGAGGUgg	1						1	
	S1-5	CCCAUAUAAA	GAUUAACCGGUCgg	2						2	
	S1-6	UGCCCUUCCCC	AGGAUCCUGAUCgg	1						1	
	S1-7	UCCUAAAACCC	UUCUUAGCGGUCgg		1					1	
	S1-8	GACCAAAAGAC	AAGGACACGGCAgg		1					1	
	S1-9	UUGCCCAUUUC	AGGAAUAACGGCgg		1					1	
	S1-10	UUCUAAAACCG	GUUUCGCACCGUgg		1					1	
	S1-11	UCUAACACUUC	GAAGUCCCGGCAgg		1					1	
	S1-12	UCCUAAAACCG	UUCUUAGCGGUCgg		2					2	
	S1-13	CCCACUCCGAC	ACCAAAAACGGCgg			1				1	
	S1-14	UCCCAACCUCU	CUUCCGAACGUGgg			1				1	
	S1-15	ACCAUAAUACA	UAGUCCAACGGGgg			1				1	
	S1-16	UCCUCAUACU	UCUAGAAGCGGCgg			1				1	
	S1-17	UUUCCAAAACA	GCCUGGCAGGUgg			1				1	
	S1-18	CCUAAGCCUAU	GACUAGGAGGGUgg				4			4	
	S1-19	UCCUGACCCU	CCUAUCCCUGGCgg				1			1	
	S1-20	ACCAUAUAUC	GACAGCAUCGGCgg				1			1	
	S1-21	UCGAAGAAGAC	CGAGAACAAAGUgg					1		1	
	S1-22	AUCUCAAGCAC	UAGUGUACCGGGgg					3		3	
	S1-23	UCCCCGAACCU	CGUUCGGCGGUgg					1		1	
	S1-24	CCCACCCAGAC	CUGUCUCUGCGGgg						5	5	
	S1-25	CCUGCGAGAAC	GUCUGGUGGUUAUgg						1	1	
				Total	8	7	5	6	5	6	37
Round 2	S2-1	CCCACACAAUA	CAUCGAUUCGGCgg	3	2	2	5	2	6	20	
	S2-2	UUCCCAACAUC	CUCUAAGCGGUgg					2		2	
	S2-3	UCCACUCAAAA	CAUCGAUUCGGCgg		1					1	
	wtC	CCCACUCAAAA	ACCAAAAACGGCgg	3	4	5	2	2	1	17	
				Total	6	7	7	7	6	7	40

<sup>a</sup> Sequences are shown in 5' to 3' orientation. Wt satC (wtC) LS and RS sequences are shown in the bottom row and were recovered in the second round. Bases in lower case were not selected but may be part of the recovered elements (see text). The C-rich LS elements (CCCA, CCUA/G) are in red. The G-rich RS elements U/CGGCGG, UCGGCAGG and GC/UGG are in blue. Additional A-rich elements (CAAAA, ACCAAAA) are in green.

<sup>b</sup> Number of clones found in individual plants at 21 dpi.

CAAAA located in the mid to 3' region of the LS, similar to its location in wt satC. Lastly, two winners contained the wt element ACCAAAA, one in the LS and one in the RS.

Whereas the 1<sup>st</sup> round winning RS sequences contained nucleotide ratios that were more evenly distributed (29% guanylates, 25% cytidylates, 23% adenylates, 23% uridylates), guanylates were mainly located in the 3' portion of the RS (Table 3.2). One of the 1<sup>st</sup> round winners contained the wt RS (S1-13) indicating that the RS is also sequence-specific in wt satC. Eight of the selected RS contained the wt sequence CGGC and three contained the related sequence UGGC. Eight of these eleven related sequences were located at the 3' edge of the RS, suggesting that additional non-selected 3' bases (e.g., guanylates G221 and G222) may form part of the element. This possibility is supported by the location in clone S1-23 of an internal CGGC sequence upstream of two guanylates. Three of the eleven related sequences were preceded by "AGG" (for two winners, this included G221 and G222). A second sequence, "GC/UGG" was found in twelve RS from 1<sup>st</sup> round winners (not including occurrences as part of the major U/CGGCGG recovered sequence). GC/UGG was mainly located near, but not necessarily adjacent to, the RS 3' end. The results of the 1<sup>st</sup> round RS SELEX therefore suggest a requirement for one of three major G-rich elements (U/CGGCGG, U/CGGCAGG and GC/UGG). Interestingly, two 1<sup>st</sup> round winners derived from the same plant (S1-1 and S1-2) had identical LS but very different RS. Since a well-studied satC recombination hot-spot is located just downstream of the LS at the base of the M1H (Cascone et al., 1993), RNA recombination may have occurred to generate these satC species. The recovered RS and LS of two other winners (S1-7 and S1-12) differed by only a single



base. Since these clones were isolated from the same plant, one sequence likely evolved from the other by a single base alteration.

Additional competition between satC species accumulating in 1<sup>st</sup> round plants was carried out by inoculating six seedlings with a mixture of equal amounts of total RNA extracted from all 30 1<sup>st</sup> round plants. At 21 dpi, total RNA was isolated from each 2<sup>nd</sup> round plant and clones generated by RT-PCR. Only four different satC species were recovered in the forty clones sequenced, one of which was identical to wt satC. The majority of the recovered clones from each of the six plants contained a new sequence (S2-1) that differed from wt satC by two bases in the LS and nine of 13 bases in the RS (Table 3.2). S2-1 contained both of the most prevalent elements identified in the first round, CCCA at the 5' terminus of the LS and CGGCGG at the 3' terminus of the RS (includes non-selected G221 and G222). Wt satC, the next most prevalent species, may have been generated by recombination between S1-13 and S2-1, the former of which contains a wt RS and the latter contains a nearly wt LS, followed by sequence evolution. S2-2 and S2-3 were each found in only one of the six plants. S2-2 contained the major CCCA element in the LS and the prevalent RS element, GCGG. S2-3 did not contain either of the major LS elements. The presence of identical RS sequences in S2-3 and S2-1 was likely due to RNA recombination. In addition, the adenylate to guanylate transition at the base of the M1H necessary for construction of the SELEX cDNAs had reverted in S2-3 and in satC with wt LS and RS.

The higher recovery of S2-1 compared to wt satC in 2<sup>nd</sup> round plants did not necessarily reflect enhanced fitness of S2-1, since generation of wt satC would first require multiple events. To determine the relative fitness of 2<sup>nd</sup> round winners and also

selected 1<sup>st</sup> round winners, pair-wise competitions for fitness were conducted. Two 1<sup>st</sup> round winners were selected for inclusion in this experiment for the following reasons: S1-9 contained the prevalent CGGCGG (includes G221 and G222) in its RS and CCCA in the LS yet was not recovered in the second round. The LS of S1-17 did not contain either of the major LS C-rich elements but did contain CAAAA and the prevalent RS element U/CGGCAGG.

Equal amounts of satC transcripts were combined and inoculated together with TCV onto three turnip seedlings. At 21 dpi, 18 to 24 clones were sequenced in total from the three plants used for each pair-wise competition (Table 3.3). Nearly equivalent numbers of clones were recovered in co-inoculations with 1<sup>st</sup> round winners S1-9 and S1-17 (12 and 10, respectively). S2-3 (the least prevalent 2<sup>nd</sup> round winner) was considerably more fit than S1-9 (16 of 18 clones). This suggests that even low level 2<sup>nd</sup> round winners were more fit than some 1<sup>st</sup> round winners. In addition, this result indicates that the simple presence of the wt elements CGGCGG and CCCA in S1-9 were insufficient to out-compete the 2<sup>nd</sup> round winner. S2-2, recovered in 20 of 21 clones, was substantially more fit than S2-3. The most prevalent winner in the 2<sup>nd</sup> round, S2-1, out-competed S2-2 (14 of 20 clones) but was not recovered in direct competition with wt satC. These results indicate the following fitness order: wt satC - S2-1 - S2-2 - S2-3 - S1-9 and S1-17. Since S2-1 and S2-3 contained identical RS sequences, the fitness levels must reflect differences in their LS sequence. The LS of S2-3 differed from wt satC at only one position, strongly suggesting the importance of the wt CCCA for fitness compared with the S2-3 UCCA.

**Table 3.3** Competition of SELEX winners for fitness in planta

Co-inoculation	Number recovered at 21 dpi
wt satC S2-1	24 0
S2-1 S2-2	14 6
S2-2 S2-3	20 1
S2-3 S1-9	16 2
S1-9 S1-17	12 10

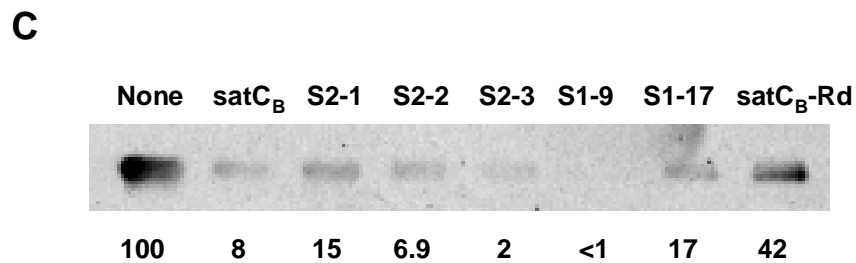
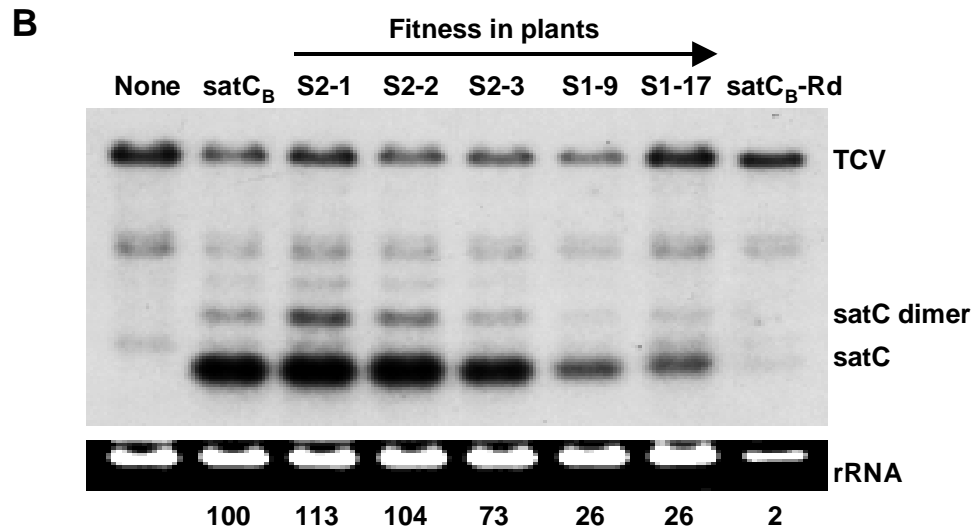
## **Winning sequences revealed by in vivo SELEX contribute to satC replication and repression of virion accumulation**

satC fitness to accumulate depends on replication competence and the ability to restrict TCV virion accumulation, since high levels of free CP are better able to suppress the host's innate RNA silencing defense system. This is advantageous for TCV since it allows the virus to spread more rapidly within a plant (Zhang and Simon, 2003a; 2003b). To determine how well 2<sup>nd</sup> round winners (S2-1, S2-2, S2-3) and selected 1<sup>st</sup> round winners (S1-9 and S1-17) replicate and if they interfere with virion accumulation, protoplasts were inoculated with transcripts prepared from these clones, parental construct satC<sub>B</sub>, and satC<sub>B</sub> with non-selected ("random") sequence in the LS and RS regions (satC<sub>B</sub>-Rd; Figure 3.3A). RNA gel blot analysis of total RNA extracted from protoplasts at 40 hpi showed that satC<sub>B</sub>-Rd accumulated to only 2% of satC<sub>B</sub> levels (Figure 3.3B). First round winners S1-9 and S1-17 accumulated 13-fold more efficiently than satC<sub>B</sub>-Rd, reaching levels that were 26% of satC<sub>B</sub> levels. This demonstrates that even after one round of selection, satC replication was significantly improved. The most-fit second round winner S2-1 replicated to levels similar to wt satC (i.e., ~20% higher than satC<sub>B</sub>) and the second most-fit winner (S2-2) replicated to levels similar to satC<sub>B</sub>. The least fit second round winner accumulated to 73% of satC<sub>B</sub> levels. These results indicate that recovered sequences contribute to fitness of the winners by effectively increasing the ability of the satRNA to replicate. In addition, since there was a substantial difference in fitness between S2-1, S2-2 and wt satC based on direct competition assays in plants (Table 3.3), distinctions between these three satC species likely involve parameters other than simple replication efficiency.

**Figure 3.3** Replication and virion repression of 2<sup>nd</sup> round and selected 1<sup>st</sup> round SELEX winners. (A) Sequence in the LS and RS regions in the parental satRNA (satCB), 2<sup>nd</sup> round winners (S2-1, S2-2 and S2-3), 1<sup>st</sup> round winners (S1-9, S1-17) and satCB containing randomly selected LS and RS sequences (satCB-Rd). Elements identified in the text and in Table 3.2 are boxed. Bases in lower case were outside of the selected sequences but likely comprise part of some identified elements. (B) RNA gel blot of total RNA extracted 40 hours after inoculation of protoplasts with TCV and the satRNAs listed above each lane. The blot was probed with an oligonucleotide specific for both TCV and satC. Positions of TCV and satC are shown. 26S rRNA from the ethidium stained gel is shown and was used as a loading control. Values below each lane are normalized levels of satC. The order of fitness of the satRNAs (indicated by an arrow above the lanes) is from Table 3.3. (C) Virion levels in infected protoplasts. Virions were detected by chemiluminescence following treatment with anti-TCV CP antibody. Values given below each lane represent relative levels of virions.

**A**

	LS	RS
satC <sub>B</sub>	CCCACUC <sup>1</sup> CAAAA <sup>2</sup>	ACCAAAA <sup>1</sup> ACGGC <sup>2</sup> gg
S2-1	CCCACACAAUA	CAUCGAUU <sup>1</sup> CGGC <sup>2</sup> gg
S2-2	UU <sup>1</sup> CCCAACAUC	CUCUAA <sup>1</sup> GCGG <sup>2</sup> UUgg
S2-3	UCCACUC <sup>1</sup> CAAAA <sup>2</sup>	CAUCGAUU <sup>1</sup> CGGC <sup>2</sup> gg
S1-9	UUG <sup>1</sup> CCCA <sup>2</sup> UUUC	AGGAAUAA <sup>1</sup> CGGC <sup>2</sup> gg
S1-17	UUUC <sup>1</sup> CAAAA <sup>2</sup> CA	GCC <sup>1</sup> UGGCAGG <sup>2</sup> UUgg
satC <sub>B</sub> -Rd	CCGAGCAUUGA	UAUGUGAAAUGAgg



To ascertain if the 1<sup>st</sup> and 2<sup>nd</sup> round winners interfered with virion accumulation, virions were extracted from infected protoplasts and analyzed by Western blotting using antibody against TCV CP (Figure 3.3C). Compared with virion levels present during infection with TCV alone, satC<sub>B</sub>-Rd reduced the accumulation of virions by 58% while satC<sub>B</sub> reduced levels by 92%. This suggests that sequences outside the LS/RS region in satC<sub>B</sub>-Rd also contribute to reducing the levels of virions. All SELEX winners examined interfered with virion accumulation more effectively than satC<sub>B</sub>-Rd, ranging from an 85% reduction for the most-fit 2<sup>nd</sup> round winner S2-1, to undetectable levels for S1-9. Although the winners accumulated to different levels, which could affect the extent of virion interference both directly and by differentially influencing the level of TCV genomic RNA accumulation (and hence the level of CP), the most efficient satRNA at reducing virion levels, ΔM1H, had a deletion of the M1H enhancer, and accumulated to only 3% of wt as determined in Chapter II. These findings therefore suggest that, while recovered sequences in the LS and/or RS regions contribute to fitness of satRNAs by inhibition of virion accumulation, other parameters are more important in determining satC fitness.

### **CCCA and CGGCGG enhance replication of satC**

To determine if the LS CCCA and RS CGGCGG elements in wt satC and SELEX winners contribute to replication and/or virion repression, the elements were introduced individually into the LS and RS of satC<sub>B</sub>-Rd by replacing corresponding random bases as follows: CCCA was placed at the 5' edge of the satC<sub>B</sub>-Rd LS next to U163 generating construct LS1 and CGGC was placed at the 3' edge of the RS, upstream of G221 and

G222, generating construct RS1 (Figure 3.4A).

Transcripts of individual constructs were inoculated onto protoplasts with TCV genomic RNA. satC<sub>B</sub>-Rd plus-strands averaged 6% of satC<sub>B</sub> levels in two independent experiments, while minus-strand satC<sub>B</sub>-Rd accumulation was much higher, reaching 27% of satC<sub>B</sub> levels. Plus-strands of LS1 and RS1 accumulated 3.6-fold and 2.8-fold higher than satC<sub>B</sub>-Rd plus-strands, respectively, reaching levels that were 22% and 18% of satC<sub>B</sub> (Figure 3.4B and C). Minus-strands of LS1 and RS1 were only slightly affected by the addition of the elements, accumulating to levels only 1.4- or 1.6- fold greater than satC<sub>B</sub>-Rd. Interestingly, mutations that disrupt H5 structure also have a greater influence on plus-strand levels than minus-strand levels (Zhang and Simon, 2005).

A new construct was generated combining CCCA and CGGCGG (construct LS1RS1) to determine if together these sequences have an additive or synergistic effect on satRNA replication. At 40 hpi of protoplasts, LS1RS1 accumulated to 63% of satC<sub>B</sub> plus-strand levels and 67% of satC<sub>B</sub> minus-strand levels (Figure 3.4B and C). This synergistic effect on plus-strand accumulation suggests participation of the sequences in a related replicative function. However, these results also suggest that additional factors in the region are required to achieve wt levels of replication.

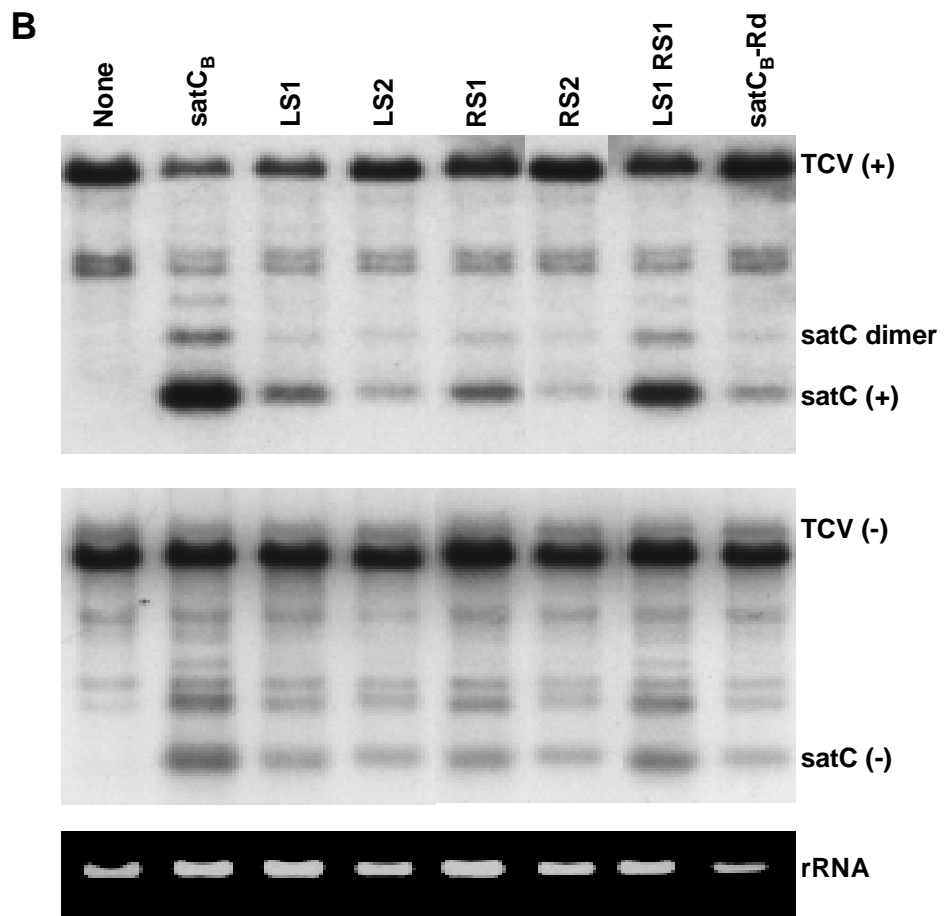
Roles in replication of CAAAA and ACCAAAA elements were also assayed by replacing corresponding random bases in satC<sub>B</sub>-Rd with these sequences in their wt positions (generating constructs LS2 and RS2, respectively; Figure 3.4A). In contrast to LS1 and RS1, plus- and minus-strands of LS2 and RS2 did not accumulate to levels greater than satC<sub>B</sub>-Rd (Figure 3.4B and C). This result was surprising since the results presented in Chapter II show that the best winner of the 10-base SELEX of the M1H had



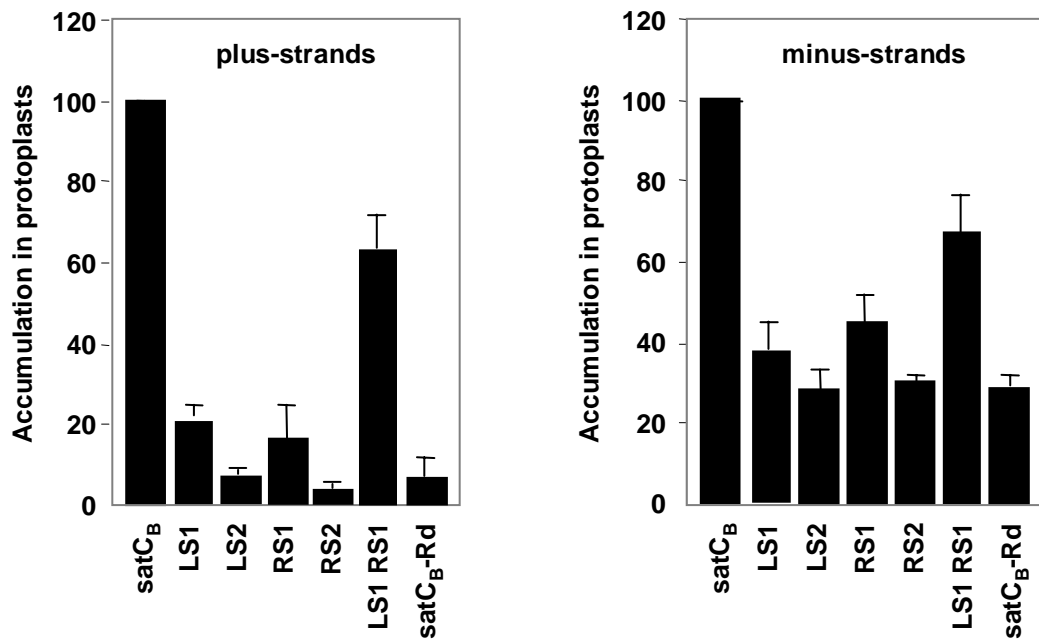
**Figure 3.4** Effect on replication and virion repression of individual elements identified by in vivo SELEX. (A) Elements (boxed) were inserted into the LS or RS regions of satCB-Rd, replacing corresponding randomly selected LS and/or RS sequences. Names of the constructs are given. satCB, parental construct containing wt satC LS and RS sequence. Bases in lower case were outside of the selected sequences but likely comprise part of the identified elements. (B) RNA gel blot of accumulating TCV and satC plus-strands (upper) and minus-strands (middle) at 40 hpi of protoplasts. SatC constructs included in the inoculum are given above each lane. The blot was probed with an oligonucleotide specific for both TCV and satC. Positions of TCV and satC are shown. The strong band migrating slightly faster than TCV minus-strands is found in preparations from uninfected protoplasts and co-migrates with 26S rRNA (data not shown). 26S rRNA from the ethidium stained gel is shown and was used as a loading control. (C) Normalized levels of accumulating satRNAs in replicate experiments. (D) Virion levels in infected protoplasts at 40 hpi. Virions were detected by chemiluminescence following treatment with anti-TCV CP antibody. Values given below each lane represent relative levels of virions.

**A**

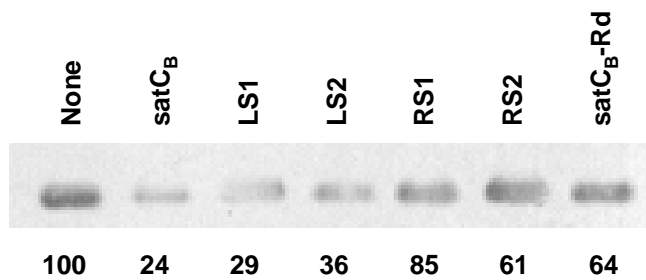
Name	LS	RS
satC <sub>B</sub>	CCCACUCAAAA	ACCAAAAACGGCgg
satC <sub>B</sub> -Rd	CCGAGCAUUGA	UAUGUGAAAUGA <sub>gg</sub>
LS1	CCCAGCAUUGA	UAUGUGAAAUGA <sub>gg</sub>
LS2	CCGAGCCAAAA	UAUGUGAAAUGA <sub>gg</sub>
RS1	CCGAGCAUUGA	UAUGUGAAACGGC <sub>gg</sub>
RS2	CCGAGCAUUGA	ACCAAAA <sub>gg</sub> UAUGA <sub>gg</sub>
LS1 RS1	CCCAGCAUUGA	UAUGUGAAACGGC <sub>gg</sub>



C



D



a duplication of CAAAA in the RS region that substantially contributed to replication and virion repression of the winner (Sun and Simon, 2003). Combined with our current results, this suggests that CAAAA cannot enhance the replication of satC independent of other elements in the LS and/or RS but may enhance replication when combined with other sequences not present in satC<sub>B</sub>-Rd.

### **CCCA and CAAAA affect virion accumulation**

To determine if the most frequently recovered elements affect virion accumulation, virions were isolated from protoplasts infected with TCV alone or accompanied by satC<sub>B</sub>, LS1, LS2, RS1, RS2 and analyzed by Western blotting using antibody specific to TCV CP (Figure 3.4D). In this experiment, satC<sub>B</sub> reduced virion levels to 24% of TCV alone levels while virion levels were 64% in the presence of satC<sub>B</sub>-Rd. RS1 and RS2 did not further reduce virion levels compared with satC<sub>B</sub>-Rd indicating that ACCAAAA and CGGCGG are not independently involved in virion repression. However, LS1 and LS2 reduced virion levels nearly as effectively as satC<sub>B</sub>. These results confirm my finding in Chapter II that duplication of CAAAA in the region enhanced virion repression. In addition, these results indicate that CCCA is multifunctional, enhancing satRNA replication and the satRNA's ability to repress virions.

## **Discussion**

Results described in Chapter II suggest that the CA-rich sequences flanking M1H are involved in satC replication and virion accumulation. In this Chapter, further examination of these sequences by *in vivo* functional SEELX revealed 3 elements

recovered in many of replacement sequences. Protoplast assays indicated that CCCA and CGGCGG are involved in satC replication, with a synergistic effect in the presence of both. Western blot results revealed that the presence of CCCA or CAAAA contributes to the ability of satC to repress virion accumulation.

As described in the results section, 20 of 29 1<sup>st</sup> and 2<sup>nd</sup> round winners contained CCCA or CCUA/G in the 5' region of their selected LS sequence. The nature of these recovered bases suggested that this element might contribute to satC replication through a base-pairing interaction with a sequence such as 5' UGGG, which would account for differences between the recovered sequences by allowing G:U pairings. Examination of the nine winning LS sequences that did not contain CCCA or CCUA/G revealed that seven contained sequence in a similar 5' side location that could also base-pair with 5' UGGG (UCUA[2], UCCA [2], UUCG, CUCA, CCCG).

Full-length satC contains five UGGG sequences: one is at position 106; one is at the base of M1H; and three are within H5. Of the elements within H5, two are in stem regions in the phylogenetically conserved structure and one is in the LSL (Figure 1.3A). The LSL UGGG partially overlaps  $\Psi$ 1 sequence (GGGC; overlapped bases are underlined), the pseudoknot contained in the 3' active conformation of satC as described in Chapter I (Zhang et al., 2004a; 2004b). A connection between CCCA and the structural rearrangement occurring in the H5 region is supported by recent RNA solution structural probing results.

Since TCV H5 also contains one UGGG in its LSL (Figure 1.3B), it is possible that CCCA basepairs with the TCV UGGG, thereby bringing satC and TCV molecules together. The interaction between satC and TCV may facilitate the recruitment of p88

translated from the genomic RNA to satC, resulting in enhanced satC replication. A TCV:satC heterodimer may be occasionally, successfully packaged by CP to form a virion. However, since TCV virion assembly is strictly limited by the size of the packaged molecule (Qu and Morris, 1997), it is more likely that formation of the TCV:satC heterodimer leads to incomplete or unstable virion assembly, resulting in decreased virion accumulation. Such a *trans*-acting role of CCCA can explain how CCCA is involved in both satC replication and virion assembly repression.

How the CAAAA element in the LS region interferes with virion accumulation is not known. It is also not known why ACCAAAA in the RS region, which contains an embedded CAAAA element, had no effect on virion repression. Since satC<sub>B</sub>-Rd was able to repress virion accumulation by 36% in the absence of selected LS and RS sequences, it seems likely that sequences external to these regions influence virion assembly and thus the position of the CAAAA element within the LS may be important.

The SELEX results and protoplast analyses showed that the CGGCGG element is important for satC accumulation in vivo. Recent RNA solution structural probing and in vitro transcription assays indicated that the CGGCGG element (also termed the DR element) may interact directly or indirectly with  $\Psi$ 2, the pseudoknot stabilizing the pre-active structure of satC (Zhang et al., unpublished data). It has been found that a mutation converting the CGGCGG element to CCGCGG changes the structure of  $\Psi$ 2 and mutations that disrupt either  $\Psi$ 2 partner cause similar structural changes in the DR (Zhang et al., unpublished data). Because the combined presence of CCCA and CGGCGG is synergistic (Figure 3.4B and C), the CGGCGG element is also proposed to contribute to the regulation of minus-strand synthesis, possibly by facilitating the

conformational transition from the pre-active to active structure (Zhang et al., 2006a).

The requirement for distal elements to achieve an active 3'-end RNA structure allowing access of the RdRp is similar to a report of factors necessary for initiation of bacteriophage Q $\beta$  minus-strand synthesis (Schuppli et al., 2000). The 3' terminal five bases of Q $\beta$  are involved in a long distance base-pairing that does not permit efficient access to the polymerase. A bacterial protein (Hfq) or a series of mutations including alterations to the 3' end and interacting sequence, are required for efficient replication and are thought to destabilize the secondary structure in the region and allow access of the polymerase to the 3' end. Interestingly, a number of structural and sequence similarities exist between Q $\beta$  and TCV genomic RNAs including three guanylates at the 5' terminus, similar 3' terminal sequence (TCV: CCUGCCC-OH; Q $\beta$ : CCUCCC), and a stable hairpin located just upstream from the 3' sequence that contains stacked pyrimidines on one side and stacked purines (mainly guanylates) on the other side (Schuppli et al., 2000; Zhang et al., 2004a). The presence of alternative structural conformations of some viral 3' ends may be one mechanism needed to expose the 3' end to the RdRp for accurate synthesis of wt levels of minus strands.

## CHAPTER IV

### H4 FUNCTIONS IN BOTH ORIENTATIONS TO ENHANCE VIRAL REPLICATION

#### Introduction

During replication of positive-strand RNA viruses, viral RdRps must recognize their cognate RNA through direct or indirect interaction with specific sequence or structural elements located on the template. The RNA elements can be hairpins, pseudoknots, tRNA-like structures, or cloverleaf-like structures (Dreher, 1999), which may play diverse roles in regulation of RNA synthesis, such as serving as core promoters or enhancers as described in Chapter I. H4(+) is a TCV hairpin located just beyond the shared regions of TCV and satC, which is similar to M1H(+) of satC in location relative to the 3' end of the RNA, although the majority of sequence is unrelated (see Figure 1.3). Introduction of either H4(+) or M1H(+) into a poorly viable satRNA stimulated RNA replication by nearly 6-fold in *Arabidopsis* protoplasts (Nagy et al., 1999). M1H(-) is a stronger enhancer of transcription from a linear minus-sense promoter compared with M1H(+) transcriptional enhancement from the plus-sense satC Pr promoter. M1H(-) and H4(-) are also hot spots for RNA recombination in vivo (Carpenter et al., 1995; Cascone et al., 1993). Altogether, these results led to the suggestion that M1H functions primarily in its minus-sense orientation as a replication enhancer.

While preliminary evidence indicates that M1H and H4 are functionally related, the proximity of H4(+) to the 3' proximal elements H4a/H4b/H5/Pr suggested that H4(+)



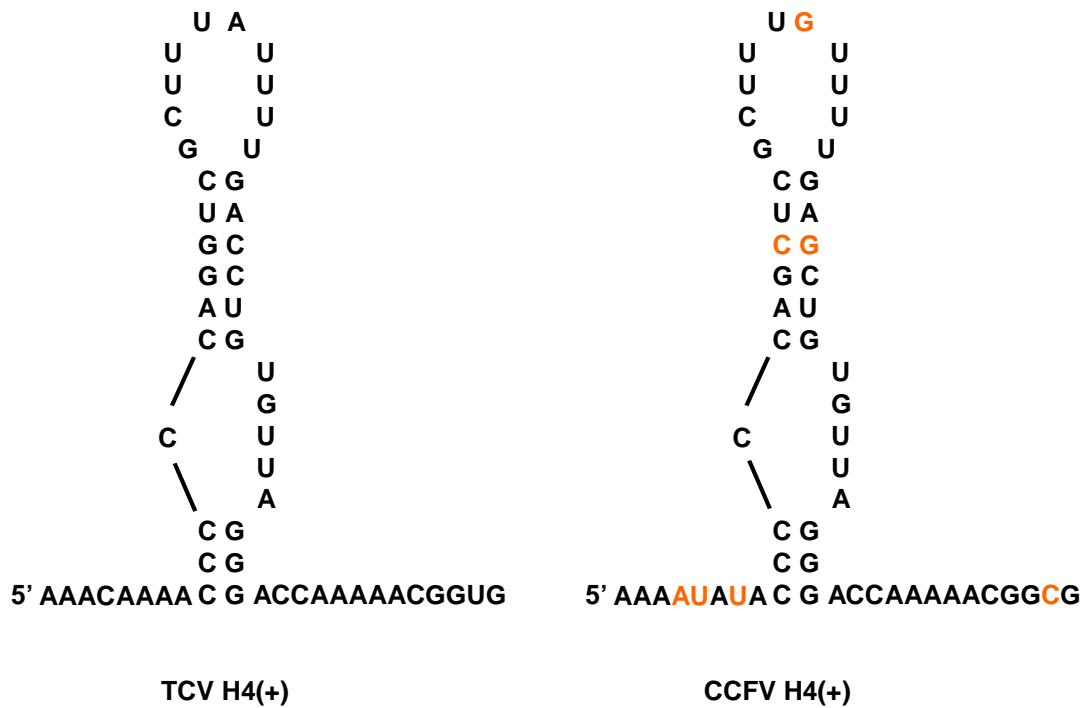
might also be important for minus-strand synthesis. *Cardamine chlorotic fleck virus* (CCFV) (family *Tombusviridae*, genus *Carmovirus*), which shares 41% sequence similarity with TCV in their 3' UTR, has a nearly identical H4 in the same location, except for a GC to CG exchange in the middle stem and a U to G transversion in the terminal loop (Figure 4.1). These observations suggest an important sequence-specific role for H4 in TCV and CCFV accumulation that might reflect functions in both orientations.

In this Chapter, I provide evidence for involvement of H4(+) and H4(-) in replication of TCV. Disruption of its structure or alteration of its internal or terminal loop sequences substantially reduced accumulation of TCV in *Arabidopsis* protoplasts. In vitro RNA transcription and gel mobility shift assays using *E. coli*-expressed p88 indicated that both H4(+) and H4(-) have enhancer activity and both bind RdRp with an affinity that was substantially greater than the binding affinity of the RdRp for the core promoters. This suggests that H4(+) and H4(-) function to enhance plus- and minus-strand synthesis by attracting the RdRp to the TCV genomic RNA template.

## **Materials and Methods**

### **RNA solution structure probing**

Plus-strand transcripts of TCV were synthesized using T7 RNA polymerase from *Sma*I-linearized pTCV66 (Oh et al., 1995). Minus-strand transcripts were synthesized from *Xba*I-linearized pT7TCV(-) containing a T7 RNA polymerase promoter upstream of



**Figure 4.1** Structures of TCV H4(+) and CCFV H4(+). Left, structure of TCV H4(+) and surrounding sequence. Right, Computer-predicted structure of CCFV H4(+) and surrounding sequence. Bases that differ from TCV H4(+) are shown in orange.

the TCV full-length minus-strand sequence (Carpenter et al., 1995). Solution structure probing was performed as previously described (Carpenter et al., 1995; Wang et al., 1999). Eleven microliters of TCV plus- or minus-strand transcripts (1 $\mu$ g/ $\mu$ l) were mixed with 11  $\mu$ l of yeast tRNA (10 $\mu$ g/ $\mu$ l) and 675  $\mu$ l of modification buffer (70 mM HEPES, pH 7.5, 10 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 100 mM KCl). The mixture was heated to 90°C, slowly cooled to 35°C, and incubated at 25°C for 20 minutes. Fifty microliters of samples were added to an equal volume of modification buffer containing either no additional reagents (control) or one of the following: 1% (v/v) dimethylsulfate (DMS; Sigma), 0.05 units of RNase T1 (Ambion), 0.03 units of RNase V1 (Ambion), or 0.04 units of RNase A (Ambion). After treatment with DMS (10 and 20 minutes) or enzymes (5 and 10 minutes), reactions were phenol extracted and ethanol precipitated followed by primer extension using 1 pmol of oligo13 (oligonucleotides used in this Chapter are listed in Table 4.1) for plus-strand TCV or oligo3773(-) for minus-strand TCV, MMLV reverse transcriptase and [ $\alpha$ -<sup>35</sup>S]-radiolabeled dATP. Samples were subjected to electrophoresis on a 6% Long-Ranger sequencing gel (FMC BioProducts), followed by autoradiography.

### **Construction of TCV mutants**

Construct M1 was generated by a three-step method (see Figure 4.2 for procedure flow chart). First, a 5' PCR fragment was obtained using primers Oligo3241F and OligoH4del with pTCV66 as template. Second, a 3' PCR fragment was obtained using primers Oligo3911(-) and Oligo4005R with pTCV66 as template. Third, both the 5' PCR fragment and the 3' PCR fragment were gel purified, ligated together using T4 DNA

**Table 4.1** Oligonucleotides used in Chapter IV

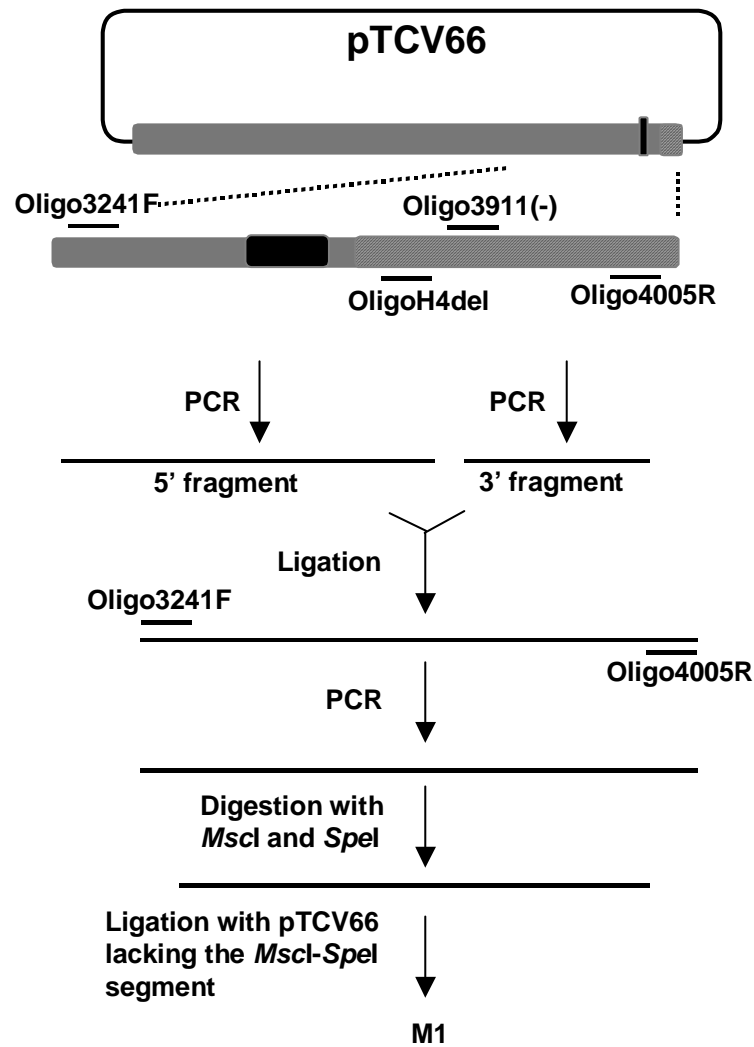
Application /Construct	Name	Position <sup>a</sup>	Sequence <sup>b</sup>	Polarity <sup>c</sup>
Structure probing	Oligo13	3950-3970	5'GTTACCCAAAGAGCACTAGTT	-
	Oligo3773(-)	3773-3789	5'GGTAAATGGCAAGCAC	+
Mutagenesis in TCV	Oligo3241F	3241-3260	5'CGCTTCCCTC TACAACATAG	+
	OligoH4del	3854-3910	5'TTTTTGGTCGTTTTGTTTTCTTTTC	-
	Oligo3911(-)	3911-3924	5'CGGTGGCAGCACTG	+
	Oligo4005R	4005-4025	5'AGGCTATCTTTTAGTTCGGAG	-
	OligoC3891G	3875-3891	5' <u>C</u> TCAAAAATAAAgCgACC	-
	OligoG3876C	3864-3876	5' <u>G</u> CTGGGGGTTTTG	-
	Oligo3873	3854-3873	5'GGGGGTTTTGTTTTCTTTTC	-
	OligoL2UUR	3876-3896	5'CACAGGTCAAAAATAAAGCGAC	-
	OligoPA5mut	3869-3891	5'GTCA <u>TATATA</u> AAGCGACCTGGGGG	-
	Oligo3869	3845-3869	5'GTTTTGTTTTCTTTTCTTAATAT	-
	OligoRA5mut	3892-3910	5' <u>ATATAG</u> GTCCCTAACACAG	-
	Oligo3892(+)	3892-3905	5'GGTCCCTAACACAG-3'	-
	Oligo3892(-)	3892-3905	5'CTGTGTTAGGGACC-3'	+
	Oligo3877(-)	3877-3889	5'TCGCTTTATTTTG	+
	Oligo3877 /3905	3877-3905	5'TCGCTTTATTTTGAG <u>CTGTGTTAGGG</u> ACC	+
	OligoTAat	3874-3911	5' <u>TGGAC</u> GCCTTTATTTTGACCTGTGTTAGGGACCAAAAAC	+
	OligoatTA	3874-3911	5'AGGTCGCCTTTATTTTG <u>ICCA</u> GTGTTAGGGACCAAAAAC	+
	OligoTATA	3874-3911	5' <u>TGGAC</u> GCCTTTATTTTG <u>ICCA</u> GTGTTAGGGACCAAAAAC	+
	OligoGtGt	3874-3911	5' <u>GGGTC</u> GCCTTTATTTTG <u>GCCT</u> GTGTTAGGGACCAAAAAC	+
	OligoCaC	3874-3911	5'AGG <u>CCG</u> CCTTTATTTTGAC <u>CCG</u> GTGTTAGGGACCAAAAAC	+
	OligoL2UUF	3897-3920	5' <u>AAAGGG</u> GACCAAAAACGGTGGCAGC	+
	OligoPRA5mut	3892-3922	5'CTGTGTTAGGGACCT <u>TATAT</u> CGGTGGCAGCAC	+
OligoCC	3870-3919	5'CCCCAGGTCGCCTTTATTTTGACCTGTGTAG <u>CG</u> GACCAAAAACGGTGGCAG	+	

	OligoGG	3870-3919	5' <u>G</u> CCCAGGTCGCTTTATTTTGACCTGTG TTAGGGACCAAAAACGGTGCCAG	+
	OligoGC	3870-3919	5' <u>G</u> CCCAGGTCGCTTTATTTTGACCTGTG TTAG <u>C</u> GACCAAAAACGGTGCCAG	+
Mutagenesis in TT and CT	OligoT70001	1-19	5' <i>GTAATACGACTCACTATAGGTAATCTGC</i> AAATCCCTG	+
	OligoTCV180 R	161-178	5'GCAT <u>GGATC</u> CCTTTTCATGTGACCCAC GT	-
	OligoCxdelH4	3864-3913	5'AGCT <u>GGATC</u> CAAAAACGACCAAAAAC GG	+
	OligoCXgg	3864-3881	5'AGCT <u>GGATC</u> CAAAAACCCCAGGTTCGC T	+
	Oligo8	4035-4054	5'GGGCAGGCCCCCCCCCCGCG	-
	OligoT7C5	1-14	5' <i>GTAATACGACTCACTATAGGGATAACTA</i> AGGG	+
	OligosatC17R	158-175	5'GCAT <u>GGATC</u> CCTTTTGAGTGGGAAACG	-
Constructs for in vitro experiments	OligoT7RdF	3960-4028	5' <i>GTAATACGACTCACTATAGGCACGAGAA</i> GGUGCUAUUUCGCAAGAAUAGUCCUGC CUUCAUCCAUCAGAATAAAAGATAG CCTCCC 3'	+
	OligoT7H4F	3960-4028	5' <i>GTAATACGACTCACTATAGGCAAAAACCC</i> CCAGGTCGCTTTATTTTGACCTGTGTTG GGACCAAAAACGAATAAAAGATAGCC TCCC	+
	OligoT7L2mF	3960-4028	5' <i>GTAATACGACTCACTATAGGCAAAAACCC</i> CCAGGTCGCTT <u>ATATAT</u> AGCCTGTGTTA GGGACCAAAAACGAATAAAAGATAGC CTCCC	+
	OligoT7M5F	3960-4028	5' <i>GTAATACGACTCACTATAGGCAAAAACC</i> CCCT <u>GGAC</u> GCTTTATTTTGACCTGTGTTA GGGACCAAAAACGAATAAAAGATAGC CTCCC	+
	OligoT7M10F	3960-4028	5' <i>GTAATACGACTCACTATAGGCAAAAACC</i> CCCAGGTCGCTTTATTTTGACCTGT <u>GAA</u> AGGGACCAAAAACGAATAAAAGATAG CCTCCC	+
	Oligo0001	1- 19	5'GGTAATCTGC AAATCCCTG	+
	OligoT7RdR	14-76	5' <i>GTAATACGACTCACTATAGGGGCACUAG</i> UUCACUUGUUAGCCUUUGACA <u>UACUAG</u> UGAAGCAUGGGATAGGCGGGTGCCAGG GA	-
	OligoT7H4R	14-76	5' <i>GTAATACGACTCACTATAGGTTTTTGGT</i> CCCTAACACAGGTCAAATAAAGCGAC TGGGGGTTTTGTAGGCGGGTGCCAGGA	-
	OligoT7L2R	14-76	5' <i>GTAATACGACTCACTATAGGTTTTTGGT</i> CCCTAACACAGGTCA <u>TATATA</u> AAGCGACC TGGGGGTTTTGTAGGCGGGTGCCAGGA	-
	OligoT7M5R	14-76	5' <i>GTAATACGACTCACTATAGGTTTTTGGT</i> CCCTAACACAGGTCAAATAAAGCG <u>CTC</u> <u>AGGGGG</u> TTTTGTAGGCGGGTGCCAGGA	-
	OligoT7M1R	14-76	5' <i>GTAATACGACTCACTATAGGTTTTTGGT</i> CCCT <u>TTT</u> CACAGGTCAAATAAAGCGACC TGGGGGTTTTGTAGGCGGGTGCCAGGA	-

<sup>a</sup> Coordinates corresponds to those of TCV, except OligoT7C5 and satC178R, which correspond to satC. Positions 3870 to 3901 are deleted in OligoH4del and OligoCxdelH4.

<sup>b</sup> Bases in italics indicate T7 RNA polymerase promoter sequence. Bold residues denote bases changed to generate a *Bam*HI site. Mutant bases are underlined.

<sup>c</sup> “+” and “-” polarities refer to homology and complementarity with TCV plus-strands, respectively.



**Figure 4.2** Flow chart for the construction procedure of M1. pTCV66 is pUC19 that contains TCV full-length plus-strand sequence (Oh et al., 1995) inserted into the *SmaI* site of the plasmid. The primers used for PCR are shown in Table 4.1. See the text for details.

ligase, and used as template to amplify the ligation product by PCR using end primers Oligo3241F and Oligo4005R. The resulting PCR products were digested with *MscI* and *SpeI*, gel purified and used to replace the *MscI*–*SpeI* segment of pTCV66. Other constructs (from M2 to M17) were generated as with M1, except that OligoH4del was replaced with primers OligoC3891G, OligoG3876C, OligoG3876C, Oligo3873, Oligo3873, Oligo3873, Oligo3873, OligoL2UUR, OligoPA5MUT, OligoRA5mut, Oligo3892(+), OligoPA5MUT, Oligo3869, Oligo3869, Oligo3869, respectively, and Oligo3911(-) was replaced with Oligo3892(-), Oligo3877(-), Oligo3877/3905, OligoTAat, OligoatTA, OligoTATA, OligoGtGt, OligoaCaC, OligoL2UUF, Oligo3892(-), Oligo3911(-), Oligo3911(-), OligoPRA5mut, OligoCC, OligoGG, OligoGC, respectively. All clones were confirmed by sequencing.

Construct TT, TT-H4 and TT-H4<sub>M11</sub> were obtained as follows. The 5' PCR fragments were generated using primers OligoT70001 and OligoTCV180R with pTCV66 as template. The 3' PCR fragment of TT was obtained using primers OligoCxdelH4 and Oligo8 with M1 as template. The 3' fragments of TT-H4 and TT-H4<sub>M11</sub> were obtained using primers OligoCXgg and Oligo8, with pTCV66 and M11 as templates, respectively. Both the 5' and 3' fragments were treated with *Bam*HI, gel purified, ligated together, and cloned into *Sma*I-digested pUC19. Construct CT, CT-H4 and CT-H4<sub>M11</sub> were generated as with TT and its derivatives except that the 5' PCR fragments were obtained using primers OligoT7C5 and OligosatC178R with pT7C(+) as template (Song and Simon, 1994). All clones were confirmed by sequencing.

Pr-link1-Rd1, Pr-link1-H4(+), Pr-link1-H4(+)<sub>M5</sub>, Pr-link1-H4(+)<sub>M10</sub>, and Pr-link1-H4(+)<sub>M11</sub> were obtained by PCR using pTCV66 as template with 3' end primer Oligo8



and one of the following 5' primers: OligoT7RdF, OligoT7H4F, OligoT7M5F, OligoT7M10F, and OligoT7L2mF. CCS-link2-Rd2, CCS-link2-H4(-), CCS-link2-H4(-)<sub>M5</sub>, CCS-link2-H4(-)<sub>M10</sub>, and CCS-link2-H4(-)<sub>M11</sub> were obtained by PCR using pTCV66 as template with 3' end primer Oligo0001 and one of the following 5' primers: OligoT7RdR, OligoT7H4R, OligoT7M5R, OligoT7M10R, and OligoT7L2mR.

### **Protoplast inoculation**

Protoplasts ( $5 \times 10^6$ ) prepared from callus cultures of *Arabidopsis thaliana* ecotype Col-0 were inoculated with 20  $\mu\text{g}$  of plus-strand TCV transcripts synthesized from *Sma*I-linearized pTCV66 or its derivatives using T7 polymerases, with or without 2  $\mu\text{g}$  of subviral RNA transcripts (TT, CT or their derivatives). Total RNAs were extracted from protoplasts at 40 hpi as described in Chapter II.

### **Northern blot hybridization**

The Northern blot hybridization was performed as described in Chapter II. Plus-strand RNAs were probed with [ $\gamma$ -<sup>32</sup>P] ATP-labeled Oligo13 complementary to positions 3950 to 3970 of TCV genomic RNA. Minus-strand RNAs were probed with [ $\alpha$ -<sup>32</sup>P] UTP-labeled oligonucleotide complementary to positions 3501 to 3664 of TCV minus-strands.

### **In vitro RNA translation using wheat germ extracts**

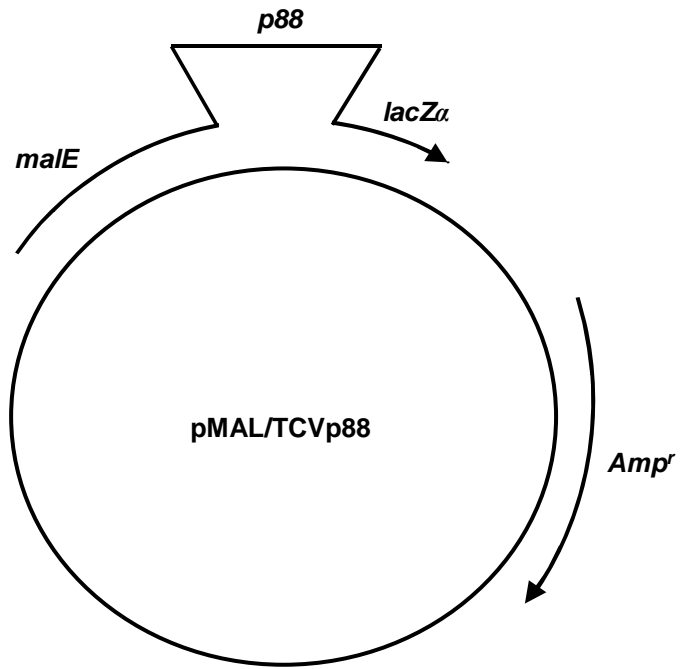
TCV RNA templates were translated in wheat germ extracts (Wheat Germ Extract Plus; Promega) according to the manufacturer's instructions. The wheat germ extracts

were thawed on ice and gently mixed several times by pipetting. One microgram of purified RNA templates were added to a 10- $\mu$ l reaction containing 6  $\mu$ l of wheat germ extracts and 10  $\mu$ Ci [ $^{35}$ S] methionine (Amersham). The reaction was incubated at 25°C for 2 hours, mixed with SDS sample buffer (1x SDS sample buffer contains 50mM Tris-HCl (pH 6.8), 2% SDS, 0.1% bromophenol blue, 10% glycerol, and 100mM dithiothreitol), and heated at 100°C for 2 minutes to denature the proteins. The denatured proteins were analyzed by 12% SDS-polyacrylamide gel followed by autoradiography.

### **Purification of p88 from *E. coli***

TCV p88 with the p28 termination codon (UAG) in the open reading frame changed to a tyrosine codon (UAC) was expressed in *E. coli* as a heterologous protein fused to a maltose-binding protein (MBP). The plasmid (MAL/TCVp88; Figure 4.3) expressing the recombinant p88 was a generous gift of Dr. P. D. Nagy (U of Kentucky). The plasmid was transformed by Dr. G. Zhang into Rosetta (DE3)-pLacI competent cells (Novagen) that enable efficient high-level expression of the MBP-p88 fusion proteins, which are comparable with the RdRp preparation purified from infected plants in transcription of TCV-associated small RNAs (Rajendran et al., 2002).

Purification of the recombinant p88 from *E. coli* was carried out as previously described (Rajendran et al., 2002). Briefly, the transformed bacterial cells were cultured overnight in 3 ml of rich growth medium (MB) containing 50  $\mu$ g/ml ampicillin at 37°C with shaking at 250 rpm. 1 liter of MB contains 10 g of tryptone, 5 g of yeast extract, and 5 g of NaCl. The overnight cultures were diluted to 1:200 in fresh MB containing 0.2% glucose and 50  $\mu$ g/ml ampicillin, incubated at 37°C until the optical density (OD<sub>600</sub>)



**Figure 4.3** The pMAL/TCVp88 plasmid. Arrows indicate the direction of transcription. TCV p88 is fused to the *malE* gene that encodes MBP. The insertion of p88 sequence interrupts a *lacZ $\alpha$*  gene that allows a blue-to-white screen for inserts on X-gal.

reached 0.5. Protein expression was then induced at 14°C with 0.3 mM IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) for 10 hours with shaking at 250 rpm. The induced cells were harvested by centrifugation in a Sorvall GSA rotor at 5,000 rpm for 10 minutes at 4°C and resuspended in 4 ml (for cells from 25 ml culture) of ice-cold column buffer (20 mM Tris-HCl, pH 7.4, 25 mM NaCl, 1 mM EDTA, 10 mM  $\beta$ -mercaptoethanol, 10% glycerol). The solution was sonicated 8 times (10 seconds per time) on ice with intervals of 2 minutes to disrupt the cells. The sonicated solution was then centrifuged at 13,000 rpm for 5 minutes at 4°C and the supernatant was loaded onto an equilibrated 0.8 cm X 4 cm amylose resin column (New England Biosciences). The column was washed thoroughly with 12 column volumes of ice-cold column buffer, and eluted with 5 ml of ice-cold column buffer containing 10 mM maltose. Fractions (0.5 ml) were collected. All purification steps were carried out at 4°C or on ice.

The concentration of purified recombinant proteins was measured using the Bio-Rad protein assay that is based on the Bradford method. High concentration of proteins usually eluted in the 2<sup>nd</sup> and 3<sup>rd</sup> fractions (0.8 ug/ul and 1.3 ug/ul, respectively). The 3<sup>rd</sup> fraction was used in the RNA transcription and binding with protein concentration modified to 1 ug/ul. The proteins were aliquoted into Eppendorf tubes, frozen over liquid nitrogen, and stored at -80°C.

### **In vitro RNA transcription using purified recombinant TCV p88**

PCR-amplified DNAs were used as templates for transcription by T7 polymerase. After phenol/chloroform extraction, unincorporated nucleotides were removed by repeated ammonium acetate/isopropanol precipitation (Nagy et al., 1997; Song and

Simon, 1994). In vitro RdRp assays were also performed as previously described (Nagy et al., 1999; 2001). Briefly, 3 µg of purified RNA template was added to a 25 µl reaction mixture containing 50 mM Tris-HCl (pH 8.2), 100 mM potassium glutamate, 10 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 1 mM each of ATP, CTP, GTP, 0.01 mM UTP, 10 µCi [ $\alpha$ -<sup>32</sup>P] UTP (Amersham) and 2 µg of p88. The assay conditions for assays were modified to keep the molar ratios of RNA templates to p88 identical among all samples. After incubation at 20°C for 90 minutes, 1 µg of yeast tRNA was added and the mixture was subjected to phenol-chloroform extraction and ammonium acetate-isopropanol precipitation. For some of the reactions, half of the RdRp products were treated with S1 nuclease at 37°C for 1 hour (Nagy et al., 1998). Products synthesized using CCS-link2-H4(-) were either treated with S1 nuclease or heated to 90°C, slowly cooled to 25°C, and then treated with S1. Radiolabeled products were analyzed by denaturing 8M urea-5% polyacrylamide gel electrophoresis followed by autoradiography. Gels were stained with ethidium bromide, photographed, and dried, followed by analysis with a phosphorimager (Nagy et al., 1997). Data were normalized based on the number of template-directed radioactive UTP incorporated into the RdRp products and the molar amount of template RNA used in the RdRp reaction.

### **p88 binding assays**

T7 RNA polymerase was used to transcribe *Sma*I-linearized pT7D(-), which contains a T7 RNA polymerase promoter upstream of satD full-length minus-strand sequence (Song and Simon, 1994), in the presence of 5 mM each of ATP, CTP, GTP, 0.05 mM UTP, and 10 µCi [ $\alpha$ -<sup>32</sup>P] UTP (Amersham). Approximately 10 ng of labeled

RNAs, together with unlabeled competitor RNAs, were mixed with 1 µg of p88 in binding buffer (50 mM Tris-HCl, pH 8.2, 10 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 10% glycerol, and 2 U of RNase inhibitor [Invitrogen]) (Rajendran et al., 2002). After incubating for 25 minutes at 25°C, the samples were analyzed by electrophoresis on native 1% agarose gels. Electrophoresis conditions were 100 V for 2 h at 4°C in 0.5x TBE buffer. The gels were dried and visualized by autoradiography.

## Results

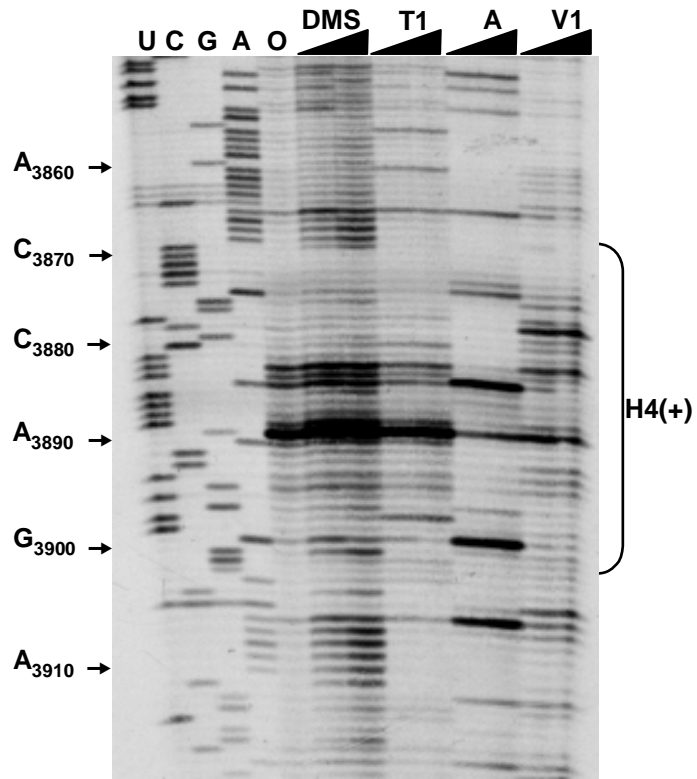
### **Solution structure probing of the H4 region in TCV plus and minus strands**

The sequence and structural conservation of TCV and CCFV H4 suggests that H4(+) and/or H4(-) exist as stem-loop structures. To determine if the H4(+) hairpin is present on plus-strands, RNA solution structural probing was performed by subjecting TCV full-length plus-strand transcripts to chemical and enzymatic probing in vitro. Transcripts were subjected to partial treatments with DMS (methylates the N1 and N3 positions of unpaired adenylates and cytidylates, respectively), RNase T1 (cleaves at single-stranded guanylates), RNase A (cleaves at single-stranded pyrimidines) and RNase V1 (cleaves at double-stranded and stacked residues). The location of cleaved or modified bases was determined following primer extension and electrophoresis through 6% sequencing gels (Figure 4.4).

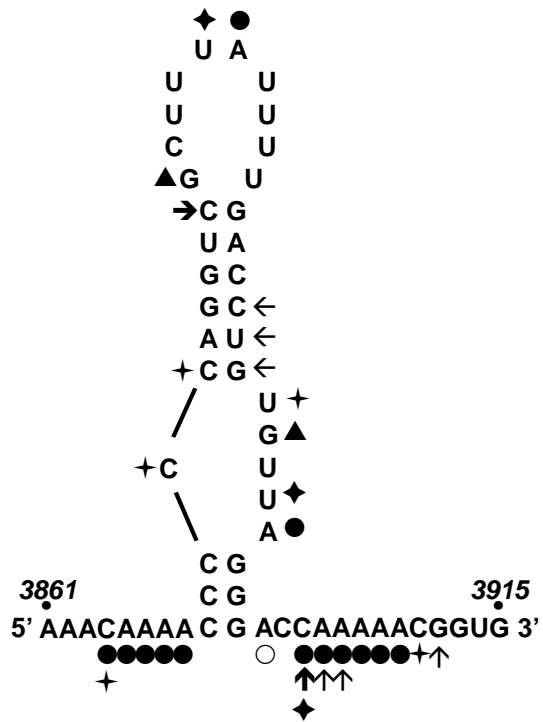
For H4(+), five of the six bases comprising the internal asymmetric loop were

**Figure 4.4** Chemical and enzymatic probing of H4(+). TCV plus-strand transcripts were treated with DMS for 10 or 20 minutes or with RNase T1, RNase A and RNase V1 for 5 or 10 minutes. The modified or cleaved RNAs were subjected to primer extension using an oligonucleotide complementary to positions 3950 to 3970 of TCV genomic RNA. (A) Representative gel showing H4(+) and surrounding sequence. The region corresponding to H4 is indicated. The sequencing ladder comprises the first four lanes. Bases corresponding to specific nucleotides are indicated at left. “0” indicates sample that was not treated with reagents prior to primer extension. Solid triangles above lanes indicate increasing incubation time. (B) Summary of solution structure probing. Low or high sensitivity to reagents is indicated by open or solid symbols, respectively, or by light or bold arrows, respectively. Circles, DMS; triangles, RNase T1; stars, RNase A; arrows, RNase V1.

**A**



**B**



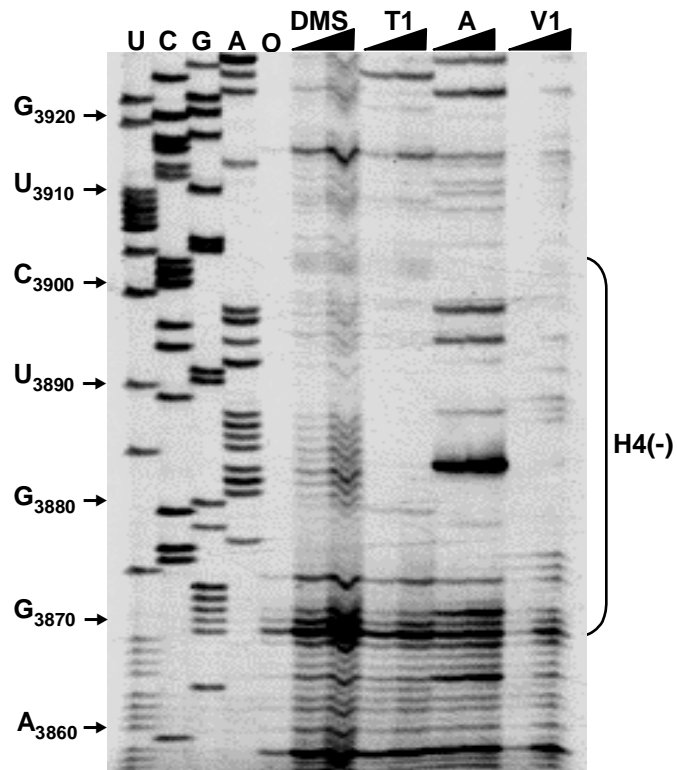


recognized by various single-strand-specific reagents, strongly suggesting that this loop is unpaired. In the terminal loop, G3879, U3883, and A3884 were accessible to single-strand-specific reagents, while five of the remaining six positions (3881-3882, 3886-3888) had premature polymerase termination sites in the absence of treatment and thus could not be evaluated. The predicted upper stem contained one strong and three weak RNase V1 signals, while C3873 adjacent to the interior asymmetric loop was susceptible to RNase A, suggesting breathing at the base of the stem. The bases flanking both sides of H4(+) were strongly susceptible to single-strand-specific reagents, with positions 3905-3907 and 3912 also cleaved by RNase V1. This suggests that these residues can assume more than one alternative or tertiary RNA structure. All together, these results are consistent with the computer-predicted and phylogenetically conserved H4(+) structure.

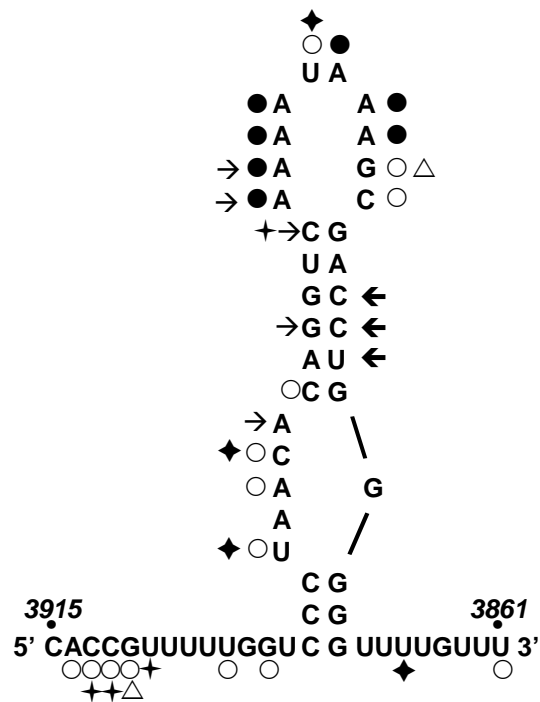
Solution structure probing was also performed on TCV full-length minus-strands to examine if H4(-) also exists as a hairpin. As shown in Figure 4.5, most residues in the loop region were recognized by DMS, RNase T1, or RNase A, and thus are likely single-stranded. One adenylylate in the internal loop and two adenylylates in the terminal loop were recognized by RNase V1, suggesting that these bases are stacked or may pair with another region. The upper stem was recognized by RNase V1, producing three strong and two weak signals, while C3889 and C3894, which are adjacent to the terminal and internal loop, respectively, were susceptible to single-stranded specific reagents. In this experiment, some uridylylates and guanylylates were weakly susceptible to DMS for unknown reasons. Taken together, the probing data for TCV minus-strands was consistent with the predicted H4(-) structure, suggesting that the sequence forms a hairpin

**Figure 4.5** Chemical and enzymatic probing of H4(-). TCV minus-strand transcripts were treated as described in in the legend to Figure 4.1. The modified or cleaved RNAs were subjected to primer extension using an oligonucleotide complementary to positions 3773 to 3789 of TCV minus-strand RNA. (A) Representative gel showing H4(-) and surrounding sequence. (B) Summary of solution structure probing. Symbols are as described in the legend to Figure 4.1.

**A**



**B**



in both strands.

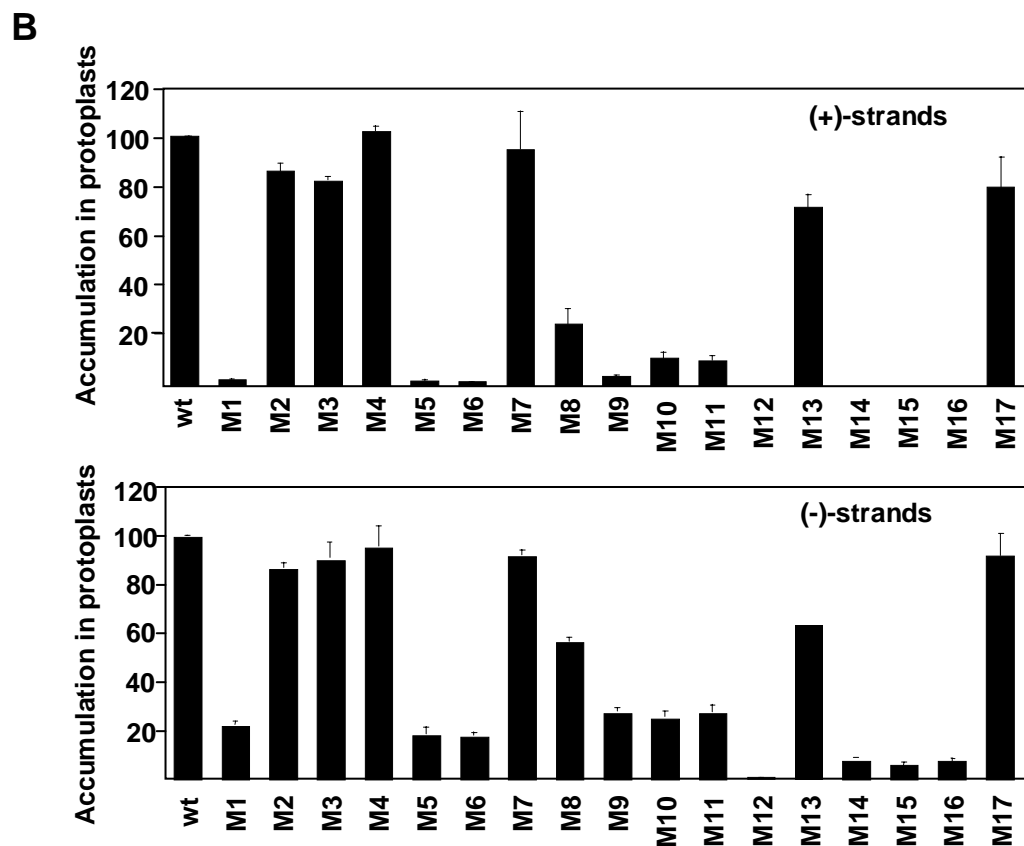
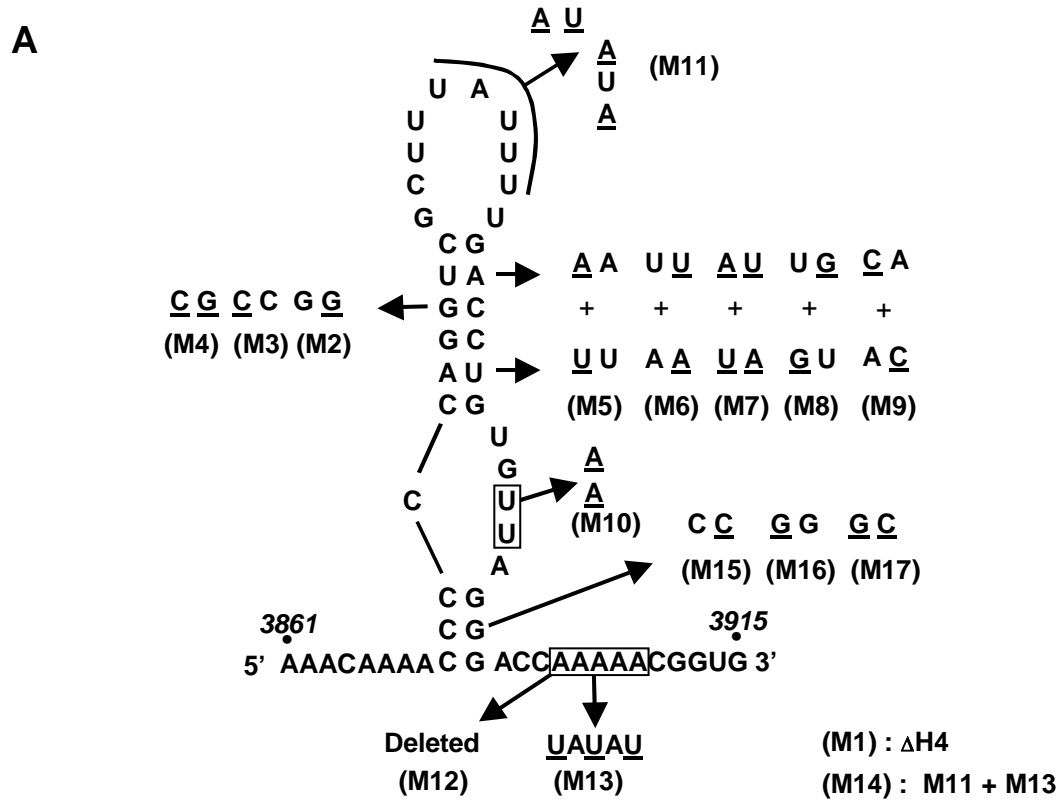
#### **H4 is important for efficient accumulation of TCV in vivo**

H4(-) is a hot spot for reinitiation of transcription by the RdRp during the process of RNA recombination, which results in the joining of full-length or nearly full-length satD to the 3' region of TCV (Carpenter et al., 1995). Since recombination hotspots attract RdRp to the acceptor strand and have enhancer activity (Cheng et al., 2005; Nagy et al., 1999), it seemed likely that TCV H4(-) serves as an enhancer of TCV replication. However, the location of H4(+) just upstream from the plus-sense 3' proximal hairpins H4a/H4b/H5/Pr suggests that H4(+) may perform a related role in minus-strand synthesis.

To examine how H4 participates in viral accumulation, H4 was deleted or subjected to site-specific mutations (Figure 4.6A). Transcripts of TCV H4 mutants were inoculated into protoplasts and total RNA extracted at 40 hpi was analyzed by RNA gel blots (Figure 4.6B). Deletion of H4 (construct M1) reduced plus-strand RNA accumulation to 2% of wt TCV levels indicating that H4 is important for efficient accumulation of TCV. To confirm the importance of the hairpin structure, a middle position of the upper six base stem (G3876-C3891) was altered to a G G mismatch (construct M2) or a C C mismatch (construct M3). M2 and M3 showed only slight reductions in accumulation (to 82-90% of wt TCV), while a compensatory G-C to C-G exchange at this position (construct M4) increased plus-strand accumulation to wt levels.

Since it was possible that the six base H4 stem might not be fully disrupted by eliminating a canonical base-pair in a central position, two additional constructs were

**Figure 4.6** Mutational analysis of H4. (A) Locations of mutations introduced into the H4 region. Names of constructs are given in parentheses. Nucleotide substitutions are underlined. (B) Accumulation of wt and mutant viral RNAs in protoplasts. Arabidopsis protoplasts were inoculated with transcripts of wt TCV and the H4 mutants. Total RNA was extracted at 40 hpi and subjected to RNA gel blot analysis using an oligonucleotide probe complementary to either plus or minus-strands.



enerated that simultaneously disrupted two base-pairs in the stem, A3874-U3893 and U3877-A3890. Disrupting these pairs by converting A3874 to U and U3877 to A (construct M5) or A3890 to U and U3893 to A (construct M6) decreased plus-strand accumulation to 2% of wt TCV levels. When both pairings were re-established with compensatory changes (construct M7), accumulation of plus-strands was restored to wt TCV levels. These results establish the importance of the larger H4 stem, which was likely not substantially affected by disruption of a single central position. This is similar to a previous finding with H5, where disruption of the central C-G pairing in the three base stem of satC H5 was detrimental, while disruption of the same base-pair in satC with H5 of TCV, which contains a five base stem in this location, did not affect satC accumulation (Zhang and Simon, 2005).

To evaluate the importance of the three-base lower stem, the center position (C3870-G3901) was altered to a C C mismatch (construct M15) or a G G mismatch (construct M16). M15 and M16 plus strands did not accumulate to detectable levels, while a compensatory C-G to G-C exchange at this position (construct M17) restored plus-strand accumulation to 80% of wt levels. All together, these results confirm the RNA solution structure assays and mFold structural predictions for the H4(+) structure depicted in Figure 4.6A.

I next examined the importance of the H4 terminal and interior loops as well as H4 3' flanking sequences on TCV accumulation. Altering two residues in the asymmetric internal loop (construct M10) or four base changes in the terminal loop (construct M11) reduced levels of plus-strands to about 10% of wt levels, indicating important functions for the H4 single-stranded regions. Deleting five consecutive adenylates flanking the 3'

side of H4(+) (construct M12) was strongly inhibitory with no TCV detected. To determine whether the missing bases were sequence specific or required as a spacer between elements, three of the adenylates were converted to uridylates (construct M13). In addition, to explore the possibility that these residues might participate in a pseudoknot with uridylates in the terminal loop, M13 mutations and M11 mutations were combined into a single construct (M14), which would re-establish putative pairing between the two locations. M13 accumulated to 72% of wt while M14 did not reach detectable levels. These results suggest that the spacing between H4(+) and downstream elements, which was altered in M12, and not the identity of the five adenylates, is critical for TCV accumulation.

For nearly all constructs with mutations that significantly reduced levels of TCV plus-strands (M1, M5, M6, M10, M11, M14, M15, M16), accumulation of minus-strands exceeded the accumulation of plus-strands (when compared with wt levels) (Figure 4.6B, bottom). For example, deletion of H4 (construct M1) resulted in plus-strand levels that reached only 2% of wt while minus-strand levels were 22% of wt, an 11-fold difference. This differential accumulation in plus- and minus-strands was not reflected in construct M13, which had alterations in sequence flanking H4. While reduced accumulation of plus-strands compared with minus-strands is generally interpreted as an indication of a minus-sense element functioning in plus-strand synthesis, recent findings when altering plus-sense hairpins H5 and H4b in satC indicated similar asymmetric effects on accumulation of plus and minus-strands (Zhang et al., 2005; 2006a). Since H5 and the related Tombusvirus hairpin SL3 have been proposed to be sites of replicase organization (McCormack and Simon, 2004; Panaviene et al., 2004; 2005), this led to the suggestion



that altering these elements disrupts the interacting replicase, which was more consequential for plus-strand synthesis (Zhang et al., 2005).

To gain further insights into whether one or both hairpin orientations function in TCV accumulation, A3874 and A3890 in the upper stem were replaced with guanylates to allow for G-U pairings at these positions in plus-strands, and C A mismatches in minus-strands (construct M8). M8 plus-strands accumulated to 25% of wt TCV levels while minus-strands reached 56% of wt TCV levels. When U3877 and U3893 were replaced by cytidylates, allowing for C A mismatches in plus-strands and G-U pairings in minus-strands, (construct M9), TCV plus- and minus-strand accumulation was reduced to 4% and 27% of wt, respectively. These results indicate that disruption of H4(+) was more consequential to RNA accumulation than H4(-). However, mutations that allowed for maintenance of the H4(-) structure but not H4(+) (construct M9) were less detrimental than disrupting both orientations (M5 and M6). Therefore, these results suggest that both orientations of H4 contribute to TCV accumulation.

#### **H4 enhances replication of TCV RNAs**

While site-specific mutagenesis and deletion analyses indicated that H4 is important for efficient accumulation of TCV *in vivo*, these experiments did not address if H4 is involved in replication or translation. To begin addressing this question, we made use of the ability of TCV to support the replication of non-coding subviral RNAs, which are frequently used as models to study replication-specific elements (Fabian et al., 2003; Nagy et al., 1999; 2001; Ray and White, 2003). Since H4 is not a component of any natural TCV subviral RNA, it was inserted into the central portion of two artificial non-

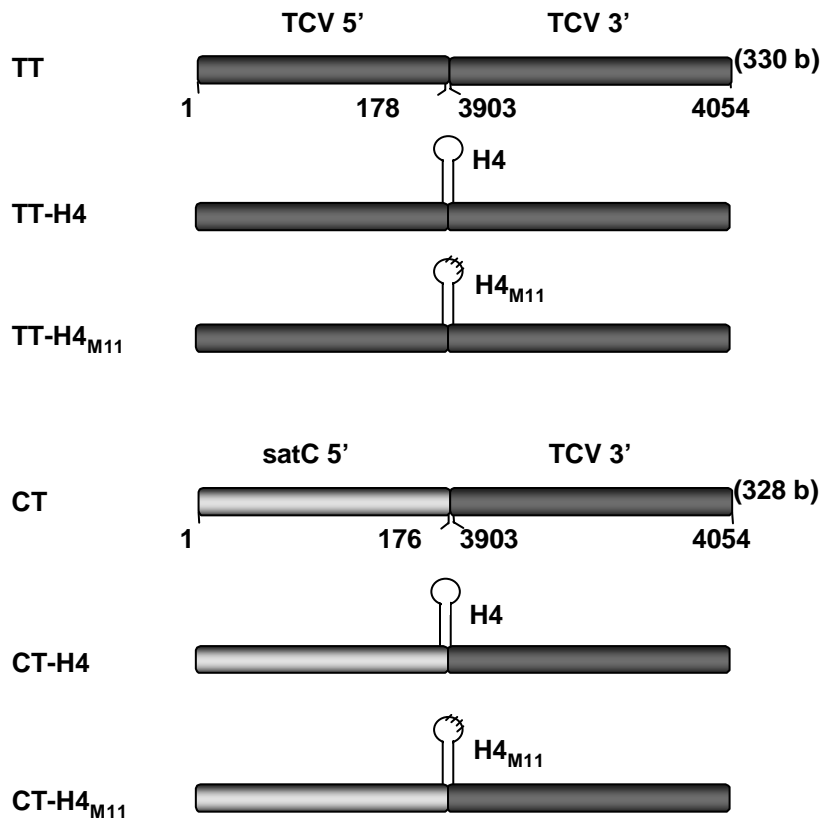
coding RNAs, TT and CT (Figure 4.7A). TT was constructed by joining the 5' end region of TCV (positions 1 to 178) to the 3' 152 bases of TCV (sequence downstream from H4). CT contained the same 3' segment joined to a satC 5' fragment (positions 1 to 176). Wt H4 and H4 with the terminal loop mutations from construct M11 were inserted into the central region of both constructs, producing TT-H4, TT-H4<sub>M11</sub>, CT-H4 and CT-H4<sub>M11</sub> (Figure 4.7A).

Transcripts of all constructs were inoculated into protoplasts with TCV variant CPmT (Wang and Simon, 1999). CPmT contains alterations at the CP translation initiation site that eliminate translation of CP, which allows for enhanced accumulation of subviral RNAs with TCV-related Pr elements (Kong et al., 1995; 1997). Total RNAs were extracted at 40 hpi and analyzed by RNA gel blots (Figure 4.7B). No TT construct accumulated to detectable levels while construct CT generated both monomers and dimers, similar to satC. CT-H4 monomers accumulated to 2-fold higher levels than CT monomers, indicating that H4 enhances accumulation of this artificial construct. Alterations in the H4 terminal loop that reduced accumulation of TCV also decreased levels of CT-H4 by nearly 60% (construct CT-H4<sub>M11</sub>). These results indicate that H4 enhances replication of CT, and by analogy, TCV. In addition, it is interesting that the presence of H4 had a negative effect on levels of dimers, similar to previous observations for satC M1H.

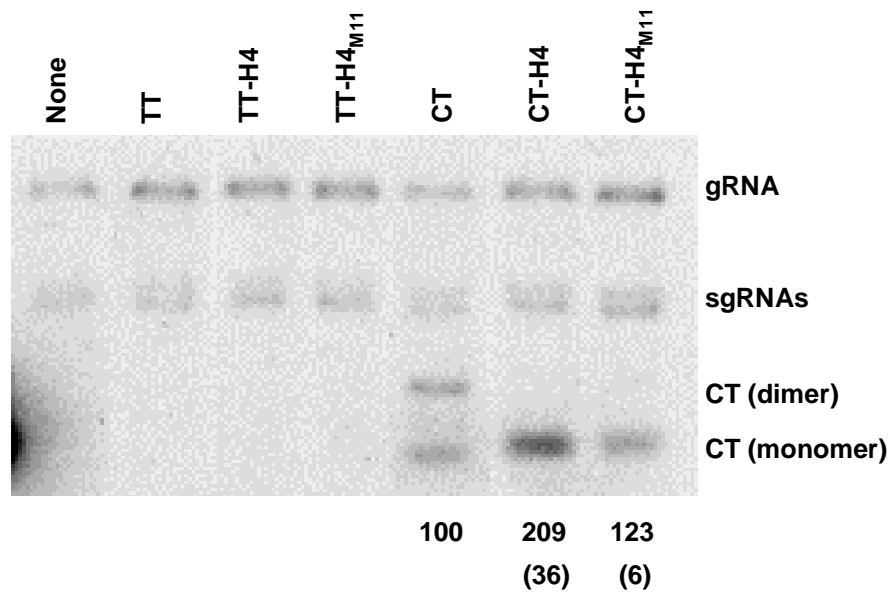
To determine if H4 contributes to viral translation, wt TCV, M1, M11 (Figure 4.6A), and  $\Delta 3'$ UTR that contains a complete deletion of the 3' UTR were subjected to in vitro translation reaction containing wheat germ extracts and [<sup>35</sup>S]-labeled methionine. H<sub>2</sub>O was used as a negative control in this assay. The experiment was carried out at 25°C

**Figure 4.7** Effect of H4 on replication of artificial non-coding subviral RNAs. (A) Schematic representation of two artificial subviral RNAs and their derivatives generated by insertion of H4 or H4M11. Numbers denote boundaries of segments derived from TCV or satC. (B) Accumulation of the non-coding RNAs in protoplasts. Total RNA was extracted at 40 hpi and subjected to RNA gel blot analyses using an oligonucleotide probe complementary to the 3' end of TCV. Averaged levels of the accumulating CT RNAs from three independent experiments are given below each lane. Numbers in parentheses denote standard deviations. gRNA, TCV genomic RNA; sgRNAs, TCV subgenomic RNAs.

**A**



**B**



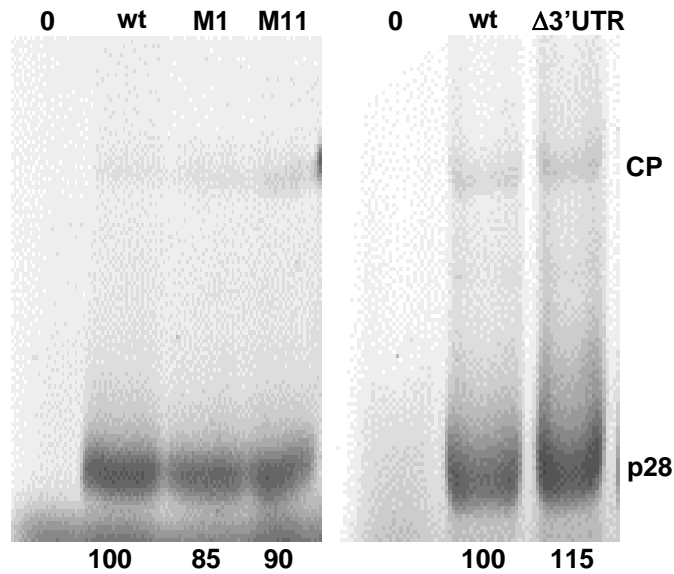
for 2 hours and the translation products were separated by SDS-PAGE gel and detected by autoradiography.

As shown in Figure 4.8, *in vitro* translation of the viral templates resulted in efficient expression of p28. The CP, which is translated from the 1.7 kb subgenomic RNA *in vivo*, was poorly expressed from the genomic RNA under these conditions. Deletion of H4 (construct M1) and four base changes in the terminal loop (construct M11) decreased p28 expression to 85% and 90% of the wt level, respectively. However, deletion of the 3' UTR increased p28 expression to 115% of the wt level, whereas it is not consistent with previous finding that the TCV 3' UTR is necessary for efficient translation in cucumber protoplasts (Qu and Morris, 2000). Because of the non-responsiveness to the loss of the 3' UTR, the current results cannot definitely determine if H4 is necessary for viral translation.

#### **H4(+) and H4(-) have enhancer activity *in vitro***

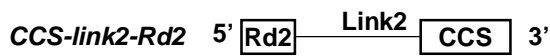
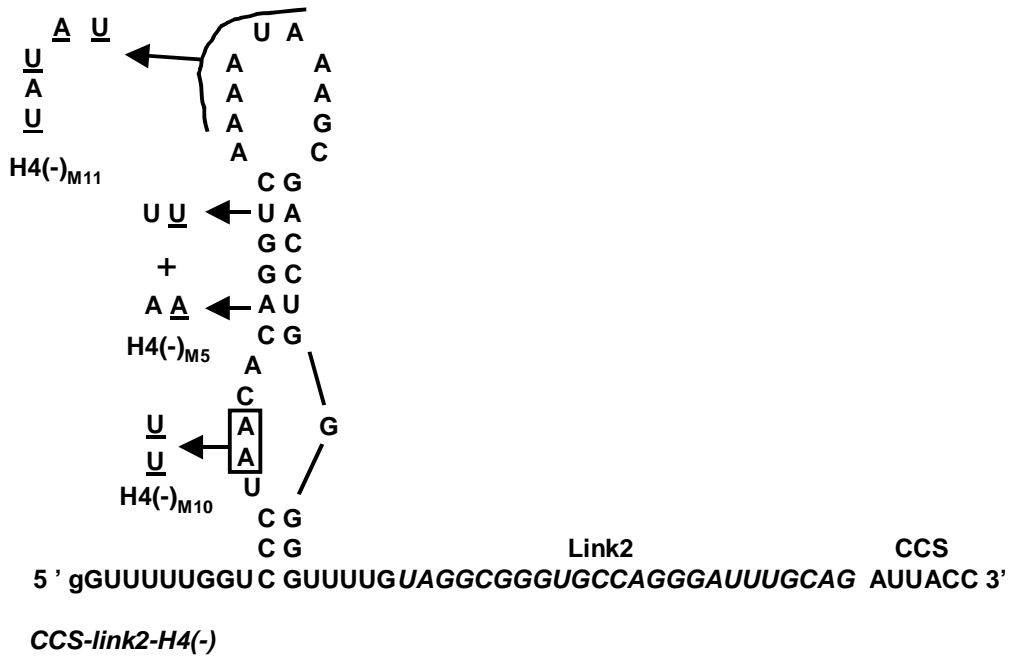
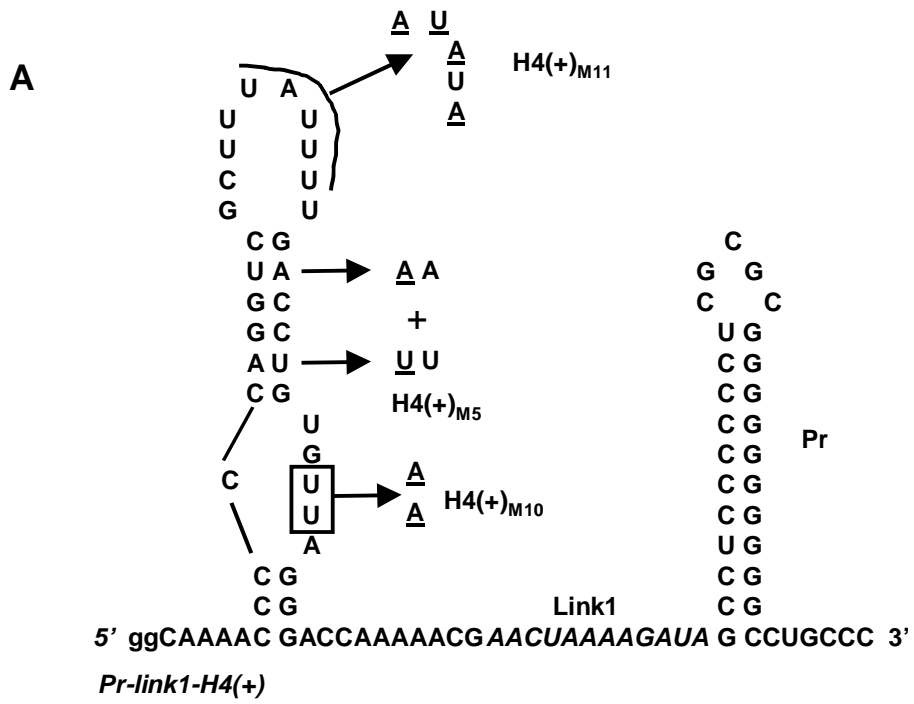
*In vitro* transcription by the TCV p88 RdRp was used to determine if both H4(+) and H4(-) can enhance the activity of TCV core promoters. The starting constructs contained either the plus-strand Pr core promoter (Song and Simon, 1995) or the minus-strand 3' terminal Carmovirus Consensus Sequences (CCS; C<sub>2-3</sub>A/U<sub>3-7</sub>) (Guan et al., 2000b). The promoters were flanked by their natural sequences (12 bases for Pr [link1] and 24 bases for CCS [link2]) joined to either wt or mutant H4(+), H4(-) or randomized H4(+) or H4(-) sequence (Rd1 or Rd2, respectively; Figure 4.9A).

H4(+) was a weak enhancer of the plus-strand Pr promoter, augmenting

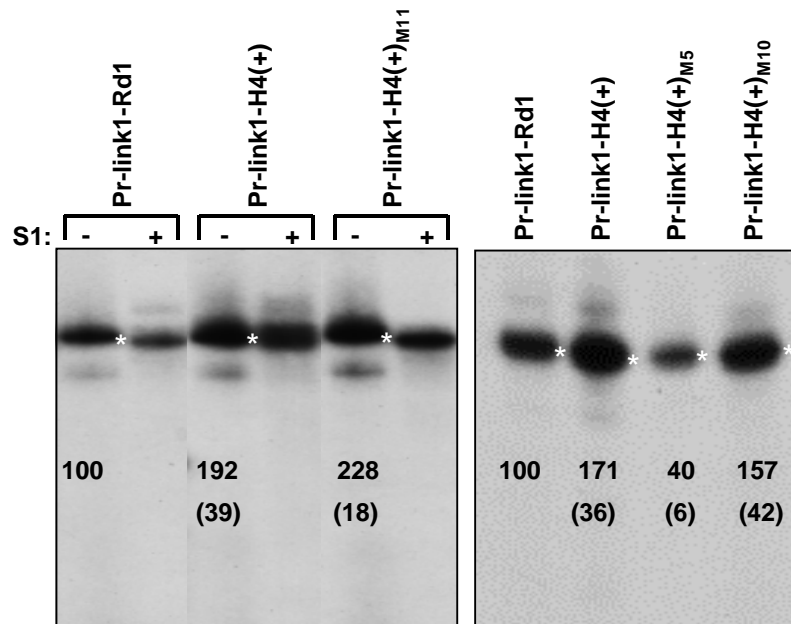
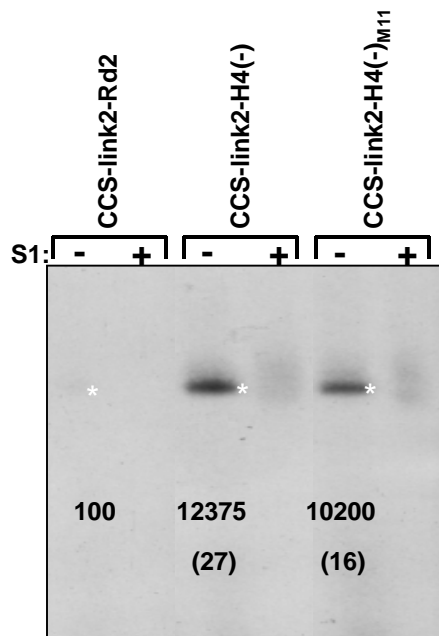
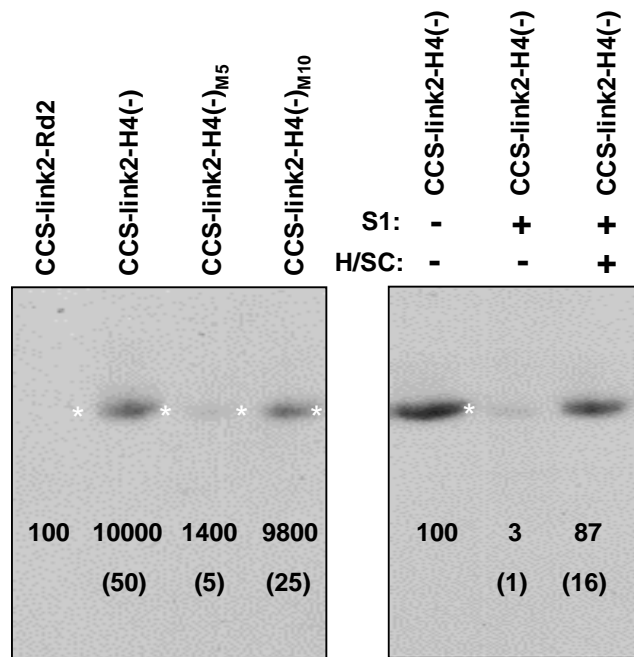


**Figure 4.8** Analysis of in vitro translation products of viral RNA templates in wheat germ extracts. Equal amounts of RNA templates were used. “0” indicates the negative control reaction where equal volume of H<sub>2</sub>O was used to replace viral RNA; wt, wt TCV; M1, TCV with deletion of H4 (Figure 4.3); M11, TCV containing base changes in the H4 terminal loop (Figure 4.3); D3’UTR, TCV with deletion of the 3’ UTR. Translation products were separated on 12% SDS-polyacrylamide gels and detected by autoradiography. The level of wt TCV was assigned a value of 100.

**Figure 4.9** Effect of H4 on transcription in vitro using recombinant p88. (A) Composition of RNA constructs. Rd1 and Rd2 denote randomized H4(+) and H4(-) sequences, respectively. Bases in lower case were included for efficient transcription by T7 RNA polymerase. Names of mutant constructs containing alterations in H4 are shown as subscripts. (B) In vitro transcription of Pr-containing constructs. Products in the left panel were either untreated (-) or treated (+) with S1 nuclease. (C) In vitro transcription of CCS-containing constructs. Products in the left panel were either untreated or treated with S1 nuclease. (D) Determination that products synthesized from the CCS promoter are single-stranded. After in vitro transcription of CCS-link2-H4(-) with p88, the reaction mix was either untreated, treated with S1 nuclease, or subjected to heating and slow cooling (H/SC) to anneal any de novo synthesized product with template followed by treatment with S1 nuclease. A weak band was visible in the CCS-link2-Rd2 lane in the absence of S1 nuclease treatment in the original autoradiogram. All constructs were tested in three independent assays. White asterisks indicate the template-sized products that are not treated with S1.





**B****C****D**

transcription by nearly 2-fold (Figure 4.9B). Enhancer activity of H4(+) was not affected by mutations in either the terminal loop [Pr-link1-H4(+)<sub>M11</sub>] or the internal loop [Pr-link1-H4(+)<sub>M10</sub>], however, disruption of the upper H4(+) stem eliminated enhancer activity [Pr-link1-H4(+)<sub>M5</sub>]. In contrast, H4(-) was a very strong enhancer of the minus-strand CCS promoter, enhancing transcription over 100-fold compared with randomized sequence [compare CCS-link2-H4(-) with CCS-link2-Rd2, Figure 4.9C]. As with H4(+), mutations in the terminal and internal loops did not affect H4(-) enhancer activity while disruption of the stem substantially reduced enhancer activity. These results indicate that both H4(+) and H4(-) are able to enhance activity of core TCV promoters. Furthermore, the single-stranded sequences in H4, which when disrupted in the context of full-length TCV (Figure 4.6) or CT constructs (Figure 4.7) affect accumulation, do not impact on the ability of H4(+) and H4(-) to enhance the activity of core promoters in vitro.

Initially, transcripts synthesized by p88 were either not further treated or treated with single-stranded specific S1 nuclease to determine the nature (single-stranded or double-stranded) of the products (Figure 4.9B and C, left panels). Treatment of three different products transcribed from the Pr promoter with S1 nuclease only slightly reduced transcript levels, indicating that products were mainly double-stranded as previously found for all RdRp products tested using satC promoters (Nagy et al., 2001). However, CCS-generated products were S1-sensitive, suggesting that these products might be single-stranded. To eliminate the possibility that the RdRp was serving as a terminal transferase and adding radiolabeled nucleotides to the 3' end of the template (which would remain single-stranded and thus degraded by S1 nuclease), template CCS-link2-H4(-) and products were heated and slow cooled to promote annealing prior to S1

nuclease treatment. If de novo synthesized full-length products had been synthesized by the RdRp in the reaction, then this treatment should anneal template and products into S1 nuclease-resistant double-stranded RNAs (since only a small percentage of available templates are transcribed by the RdRp in vitro, all products should theoretically be able to pair with templates). However, if radiolabeled nucleotides had been added to the template by terminal transferase activity, the template-derived products should remain single-stranded and be S1 nuclease sensitive. As shown in Figure 4.9D, heating and slow-cooling the reaction mix resulted in 87% of the products becoming S1 nuclease resistant. These results indicate that products transcribed from the CCS promoter were synthesized by de novo initiation and were single-stranded.

#### **p88 binds to H4(+) and H4(-)**

The finding that H4(+) and H4(-) can enhance transcription from core promoters suggests that these hairpins might function by binding the RdRp and thus attracting the replicase to the template. To test the validity of this hypothesis, I used a competition assay programmed with radiolabeled satD minus-strands [satD(-)], which are efficiently bound by TCV p88 (Rajendran et al., 2002). To test whether p88 binds to H4(+) and/or H4(-), the constructs described in Figure 4.9A were added in one, ten and 100-fold molar excess to fixed levels of the satD(-) probe and purified p88. Competitiveness for binding was determined using gel mobility shift assays. Yeast tRNA, a poor competitor in previous competition experiments (Rajendran et al., 2002), was used as a control.

One hundred-fold molar excess of Pr-link1-Rd1 only reduced satD(-) binding by about 30%, indicating that the TCV Pr is a poor competitor for RdRp binding compared

with the satD(-) probe (Figure 4.10). In contrast, 10-fold molar excess of Pr-link1-H4(+) reduced satD(-) binding by approximately 50%, indicating that H4(+) binds the RdRp more strongly than the Pr core promoter. Mutations in the H4(+) interior or terminal loops [Pr-link1-H4(+)<sub>M10</sub> and Pr-link1-H4(+)<sub>M11</sub>] did not reduce RdRp binding to H4(+), which supports the RdRp transcription results indicating these regions do not affect H4(+) enhancer activity (Figure 4.9B). Pr-link1-H4(+)<sub>M5</sub> was less effective at reducing satD(-) binding compared with Pr-link1-Rd1, consistent with its weaker activity in the *in vitro* transcription assays (Figure 4.9B).

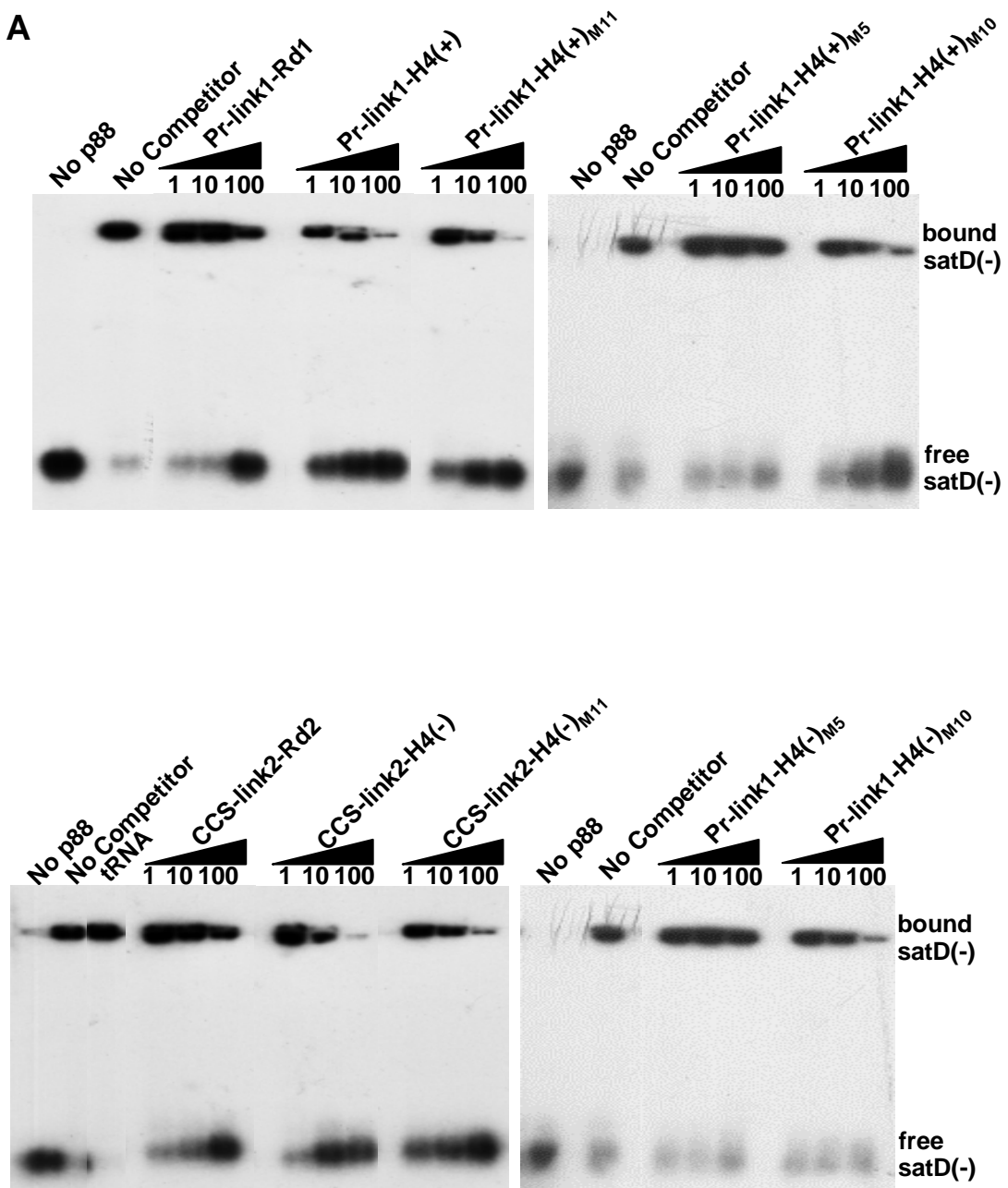
Results with the CCS promoter constructs were very similar to the Pr constructs. CCS was also a weak competitor for RdRp binding (CCS-link2-Rd2; Figure 4.10) as was CCS-link2-H4(-)<sub>M5</sub>. Constructs containing wt H4(-) competed as effectively as constructs containing H4(+), and terminal or interior loop mutations (CCS-link2-M<sub>10</sub> and CCS-link2-M<sub>11</sub>) had no apparent effect on binding. These results indicate that H4(+) and H4(-) have similar affinities for the RdRp *in vitro*. In addition, alternations in the loop sequences do not affect RdRp binding and thus must disrupt an ancillary function of the hairpin.

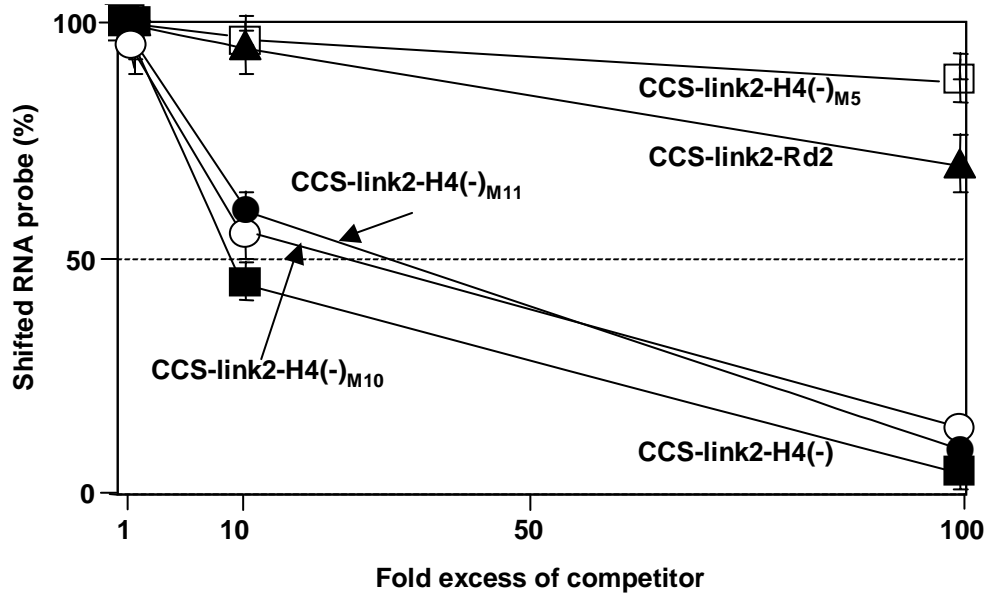
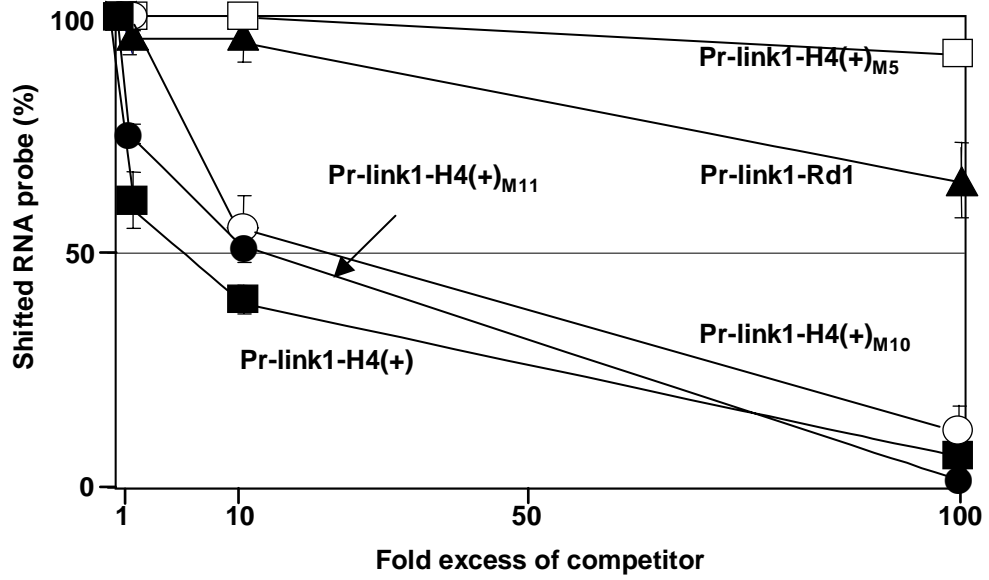
## Discussion

In this Chapter, I examined the function of TCV hairpin H4, which is structurally similar to the satC enhancer M1H and located in a similar position relative to the 3' end (Nagy et al., 1999). Solution structure analyses suggested that H4 exists as stem-loop structures in both plus and minus strands *in vitro*. H4 deletion and site-specific alterations

**Figure 4.10** Preferential binding of p88 to H4(+) and H4(-). (A) Representative gel mobility shift gels of <sup>32</sup>P-labeled satD(-) RNA probe bound to p88 in the absence or presence of unlabeled competitor RNAs described in Figure 4.9A. No p88, without added p88; no competitor, <sup>32</sup>P-labeled satD(-) probe alone. One hundred fold-excess of yeast tRNA was used as a non-specific binding control as shown in the lower panel. (B) Graphic presentation of data obtained in panel A and one additional independent experiment. The relative levels of the shifted probes were averaged and shown as percentage of the probe level shifted by p88 in the absence of competitor.

**A**



**B**

confirmed that H4 participates in accumulation of TCV, and demonstrated the importance of the terminal loop, internal asymmetric loop, and upper and lower stems in H4 function. The sequence that links H4 with H4a was also critical as demonstrated by elimination of detectable TCV accumulation when five consecutive adenylates were deleted. Since accumulation was partially restored when three adenylates were converted to uridylates, this suggests that the region may be important to spatially position H4 for correct function.

Insertion of H4 into a poorly replicating, artificial subviral RNA construct (CT) led to enhanced accumulation of subviral RNA monomers, suggesting that H4 functions during TCV replication, although a translational role cannot be ruled out. CT with wt or mutant H4 did not generate detectable levels of dimers (Figure 4.7B), which is consistent with previous studies of the satC M1H enhancer, whose presence also correlated with a substantial decrease in the level of satC dimers (Nagy et al., 1999). Dimers are formed from reinitiation of synthesis by the RdRp before release of the newly synthesized strand (Carpenter et al., 1991). The mechanism underlying the involvement of H4 and M1H in dimer accumulation remains unknown. Some deletions in the satC 5' region also greatly increased dimer levels while substantially reducing levels of monomers (Carpenter et al., 1991; Simon et al., 1988). This led to the suggestion that monomers and dimers accumulate independently of each other and may not share the same *cis*-requirements for replication. H4 (and M1H) inhibition of dimer accumulation could therefore reflect either a reduction in initial dimer formation or a suppression of dimer replication.

The 5' and 3' UTRs of *Hibiscus chlorotic ringspot virus*, another member of the genus *Carmovirus*, have been determined to synergistically contribute to translation in



wheat germ extracts (Koh et al., 2003). However, my *in vitro* translation results did not show the synergistic effect of TCV UTRs, which has been determined *in vivo* (Qu and Morris, 2002). There are several possible explanations for this inconsistency. For example, the TCV genomic RNA in wheat germ extracts may fold into secondary and tertiary structures that are not identical to those in host cells, which may prevent the interaction between the 5' and 3' UTRs or between the RNA template and cellular translational factors.

Previous studies indicated that TCV H4(-) is a recombination hotspot (in the absence of an adjacent hairpin) leading to the suggestion that H4 might serve as a *cis*-replication element primarily in its minus-sense orientation during plus-strand synthesis (Carpenter et al., 1995). My current findings that H4(-) is bound by p88 *in vitro* and can function as an enhancer *in vitro* provides additional support for H4(-) contributing to TCV accumulation by helping to attract the RdRp to minus-strands. How a 5' proximal element (on minus-strands) might enhance transcription from the distal 3' end was recently elucidated for the minus-sense, dual hairpin enhancer/RdRp binding element [SL1-III(-) and SL2-III(-)] of viruses in the Tombusvirus genus (Panavas and Nagy, 2005). A sequence linking the two hairpins acts as a bridge to the 3' promoter by pairing with sequence near the 3' end.

Based on strand-specific disruptions in Tombusviral RNAs assayed *in vivo* (Ray and White, 2003) and ability of only minus-sense SL1-III and SL2-III to enhance transcription by the RdRp from a minus-sense promoter *in vitro* (Panavas and Nagy, 2003; 2005), SL1-III and SL2-III were proposed to function in their minus-sense orientations. However, inverting the hairpins did not decrease accumulation of constructs

in vivo compared with constructs containing the hairpins in their forward orientations (Ray and White, 2003) and both plus- and minus-sense stem-loops bound RdRp in vitro, although affinity was higher for the minus-sense enhancers (Panavas and Nagy, 2005). These results suggest that the Tombusvirus enhancers may also function in their plus-sense orientation. H4(+) is just upstream from four 3' terminal hairpins conserved in satC, TCV, CCFV, and the related carmovirus *Japanese iris necrosis virus*, which are important for initiation of satC minus-strand synthesis (Zhang et al., 2004a; 2006a) and which appear to play similar roles in TCV accumulation (J. McCormack and A. E. Simon, unpublished data). My results indicate an important role for H4(+) since maintaining the structure of H4(+) (construct M8) resulted in greater TCV accumulation than maintaining the H4(-) structure (construct M9). However, preserving H4(-) while disrupting H4(+) led to greater TCV accumulation when compared with simultaneous disruptions of both structures (constructs M5 and M6), suggesting that both H4(+) and H4(-) function in TCV genomic RNA replication.

Both H4 orientations had similar binding affinities for p88, which were substantially greater than the binding affinities of the two TCV core promoters tested. This suggests that H4 may function in the initial binding of the RdRp to plus- and minus-strands during viral replication. Alternatively, it is possible that the weak interaction of the Pr (and CCS) with the RdRp reflects a requirement for additional upstream elements for efficient promoter function (Zhang et al., 2006a). Strong RdRp binding to H4(+) and H4(-) was reflected in enhanced p88 transcription of constructs containing core promoters and either H4(-) or H4(+) in vitro. Putative communication between H4 and other elements is suggested by the negative effect of mutations in the H4 terminal and

internal loops on TCV accumulation that was not reflected in either enhancer activity or RdRp binding. The detrimental effects caused by disruption of the upper stem suggest that correct H4 folding is essential for enhancer function and the related ability to bind RdRp.

Interestingly, H4(+) was a much weaker enhancer of Pr activity compared with the effect of H4(-) on the linear minus-sense CCS promoter. This may indicate that correct function of H4(+) in the genomic RNA requires the downstream H4a/H4b/H5 elements. The satC 3' region consisting of H4a/H4b/H5 and Pr was recently found to undergo a conformational switch from a pre-active structure that does not apparently contain these hairpins, to an active structure where these hairpins are needed to direct minus-strand synthesis (Zhang et al., 2006a). Since TCV genomic RNA is translated, a similar conformational switch could convert a translation-active form of the template to one that is replication-active. The location of H4(+) proximal to the 3' hairpins leads to the following proposal for H4(+) function in minus-strand synthesis: an interaction between the H4(+) terminal loop and a downstream sequence helps to maintain the translation-active structure. RdRp binding to H4(+) disrupts this interaction leading to a conformational switch to a replication-active form. The role of the RdRp in mediating such a switch may be similar to the binding of poliovirus-encoded 3CD to a cloverleaf structure near the 5' end of the poliovirus genome that causes translation to cease and replication to commence (Gamarnik and Andino, 1998). In *Alfalfa mosaic virus*, CP binding to 3' viral elements also causes a switch from the translation to the replication form of the RNA (Olsthoorn et al., 1999). The structural and functional similarity between TCV H4 and satC M1H suggests that M1H assumes the role of H4 in attracting

the RdRp to plus- and minus-strand satellite templates and may also help mediate the plus-strand conformational switch.

## CONCLUSIONS

In this dissertation, I studied the roles of RNA elements in TCV replication and virion accumulation. M1H of satC is a recombination hotspot *in vivo* and a replication enhancer functioning primarily on minus strands. My analyses of M1H by 10-base SELEX, together with mutational assays, suggested that M1H functions on plus strands by bringing together flanking CA-rich sequences that are involved in both satC replication and virion assembly. These results improve our understanding of the function of M1H(+) significantly, because previous evidence that the M1H region of satC plus strands may be involved in virion assembly was based on only a single SELEX winner containing a 7-base M1H replacement sequence (Zhang and Simon, 2003b). In addition to its role in virion assembly, M1H(+) shows weak enhancer activity during viral promoter-directed *in vitro* RNA synthesis (Nagy et al., 2001). Altogether, it can be concluded that M1H is involved in at least three viral processes: recombination, replication, and virion assembly. During the recombination and replication, it functions by attracting the RdRp to the template (Nagy et al., 1998; 1999; Nagy and Simon, 1998).

Further studies of the M1H-flanking CA-rich sequences revealed three short sequence-specific elements (CCCA, CAAAA, and CGGCGG). Protoplast assays determined that CCCA and CGGCGG are involved in satC replication while CCCA and CAAAA are involved in the satC-mediated virion interference. While the detailed mechanism is not known, a possible *trans*-acting role of CCCA that explains how it is involved in both satC replication and virion repression was presented in Chapter III. RNA elements involved in both replication and virion accumulation have been identified for a

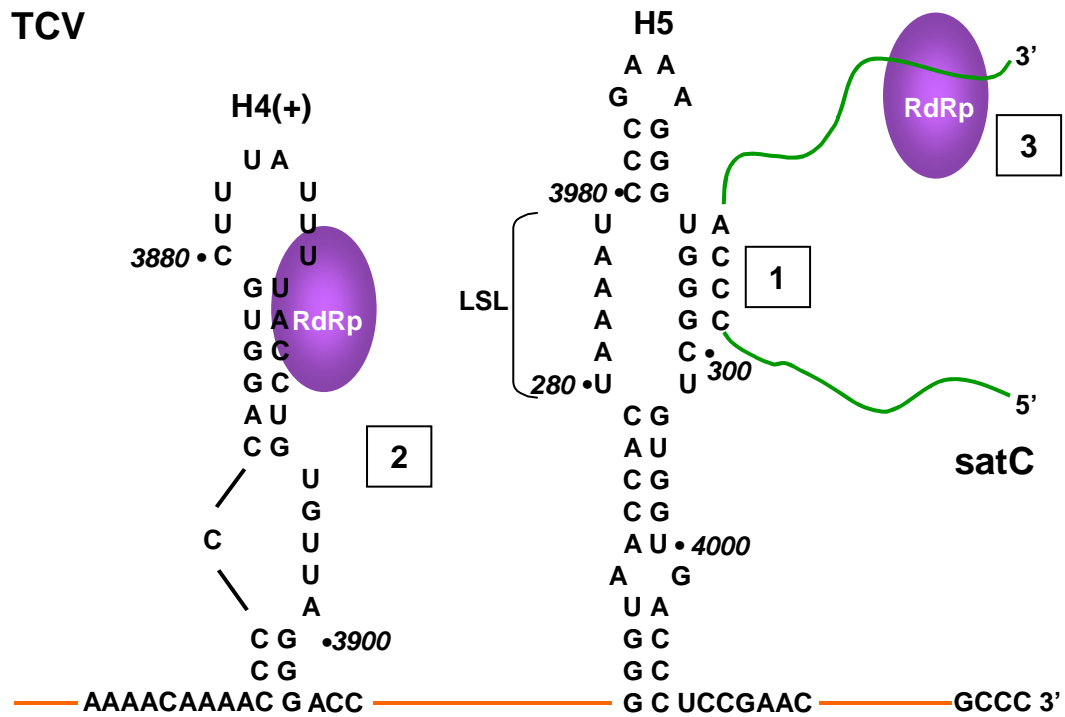
number of other positive-strand RNA viruses. For example, the 34-nt element of RCNMV RNA-2, which acts as a *trans*-activator of the subgenomic RNA synthesis (Sit et al., 1998), has recently been found to mediate co-packaging of the two genomic RNAs (Basnayake et al., 2006). The 3'-terminal TLS of BMV is also required for both minus-strand synthesis (Chapman and Kao, 1999; Dreher and Hall, 1988) and virion assembly (Choi et al., 2002).

CAAAA is located just 2 bases downstream of CCCA. It is not known how CAAAA affects virion assembly or if its function is related to CCCA. As described in Chapter III, in the absence of CCCA and CAAAA, satC was able to repress virion accumulation by 36%, suggesting that additional sequences may affect virion accumulation. Some viruses have been found to contain more than one packaging element. For example, three different RNA elements are required for virion assembly of *Ross river virus* (RRV; family *Togaviridae*, genus *Alphavirus*) (Frolova et al., 1997). In BMV, an internal 187-base sequence located within the movement protein open reading frame, in addition to the 3'-terminal TLS mentioned above, has recently been identified as required for virion assembly (Choi et al., 2002; Choi and Rao, 2003).

My *in vivo* SELEX and protoplast assays indicated that CGGCGG is necessary for efficient replication of satC. This is consistent with recent *in vitro* experimental results suggesting that the CGGCGG participates in the 3'-terminal conformational switch regulating the minus-strand synthesis (Zhang et al., unpublished data). Short RNA elements participating in 3'-terminal switch have been identified for other positive-strand RNA viruses such as AMV, BYDV, and TBSV (Olsthoorn et al. 1999; Koev et al. 2002; Pogany et al. 2003).

As described in Chapter I, H4 is a hairpin located within the TCV 3' UTR, which is similar to satC M1H in several aspects. The presence of H4 hairpin structures in both polarities was supported by RNA solution structural probing in full-length TCV. In vivo and in vitro studies of H4 suggested that it functions in TCV replication by enhancing the synthesis of both strands through attracting the RdRp to the template. So far, three replication enhancers have been identified for viruses included in the family of *Tombusviridae*: RIII(-) of TBSV (described in Chapter I), satC M1H, and TCV H4. At least four features are shared among these replication enhancers: (i) they contain one or two hairpins flanked by short single-stranded sequences that are also important for replication; (ii) they are positioned more than 100 nt upstream of the related core promoters; (iii) they function by attracting the RdRp to the template; (iv) they have stronger enhancer activity on minus strands than on plus strands during viral promoter-directed in vitro RNA syntheses (Nagy et al., 1999; 2001; Panavas and Nagy, 2003; 2005; Ray and White, 1999; 2003; Figure 4.9).

As summarized above, identification and characterization of these viral RNA elements provide new insights into the mechanisms of viral RNA replication and virion assembly. One model is presented in Figure 5.1 that can explain how TCV and satC RNA elements regulate RNA replication and virion assembly. It is proposed that satC RNA interacts in *trans* with TCV RNA through the CCCA element that base pairs with the TCV H5 loop sequence (UGGG) (Figure 5.1). This interaction inhibits virion formation by mechanisms that are discussed in Chapter III. Before or after this interaction, the RdRp binds to TCV H4, which facilitates the recruitment of the RdRp to the TCV template as suggested by the in vitro experimental results shown in Chapter IV. It is



**Figure 5.1** Model for the connection between virion repression and satC replication. Orange lines, TCV sequence; green lines, satC sequence. In step 1, satC RNA interacts in *trans* with TCV RNA through the CCCA element. This interaction inhibits virion formation. In step 2, RdRp binds to TCV H4. In step 3, the satC-TCV interaction in step 1 leads to transfer of RdRp from TCV to satC. Step 2 is not necessary to occur prior to step 1, while step 3 is proposed to occur after the other two steps. Numbering is from the 5' end of the TCV genomic RNA.



possible that the satC-TCV interaction leads to transfer of RdRp from the TCV to satC, which facilitates recruitment of the RdRp to the satC template. More experiments should be carried out to test this model. In addition, since M1H is structurally and functionally similar to H4, in vitro binding of M1H by p88 should be carried out to obtain direct evidence supporting its role in replication and recombination.

## REFERENCES

- Albarino C.G., Eckerle, L.D., Ball, L.A., 2003. The cis-acting replication signal at the 3' end of Flock House virus RNA2 is RNA3-dependent. *Virology* 311, 181-191.
- Alvarez D.E., Lodeiro, M.F., Luduena, S.J., Pietrasanta, L.I., Gamarnik, A.V., 2005. Long-range RNA-RNA interactions circularize the dengue virus genome. *J Virol.* 79, 6631-6643.
- Andino, R., Rieckhof, G.E., Achacoso, P.L., Baltimore, D., 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5' - end of viral RNA. *EMBO J.* 12, 3587–3598.
- Andino, R., Rieckhof, G.E., Baltimore, D., 1990. A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* 63, 369–380.
- Arnold, J.J., Ghosh, S.K., Cameron, C.E., 1999. Poliovirus RNA-dependent RNA polymerase (3D<sup>pol</sup>). Divalent cation modulation of primer, template, and nucleotide selection. *J. Biol. Chem.* 274, 37060-37069.
- Baker, T.A., Bell, S.P., 1998. Polymerases and the replisome: machines within machines. *Cell* 92, 295-305.

Barrera, I., Schuppli, D., Sogo, J.M., Weber, H., 1993. Different mechanisms of recognition of bacteriophage Q $\beta$  plus and minus- strand RNAs by Q $\beta$  replicase. *J. Mol. Biol.* 232, 512-521.

Barton, D.J., O'Donnell, B.J., Flanagan, J.B., 2001. 5' cloveleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. *EMBO J.* 20, 1439-1448.

Basnayake, V.R., Sit, T.L., Lommel, S.A., 2006. The genomic RNA packaging scheme of Red clover necrotic mosaic virus. *Virology* 345, 532 – 539.

Baulcombe, D., 2002. Viral suppression of systemic silencing. *Trends Microbiol.* 10, 306-308.

Baumstark, T., Ahlquist, P., 2001. The brome mosaic virus RNA3 intergenic replication enhancer folds to mimic a tRNA T $\Psi$ C-stem loop and is modified in vivo. *RNA* 7, 1652–1670.

Bol, J.F., 1999. Alfalfa mosaic virus and ilarviruses: involvement of coat protein in multiple steps of the replication cycle. *J Gen. Virol.* 80, 1089–1102.

Bolten, R., Egger, D., Gosert, R., Schaub, G., Landmann, L., Bienz, K., 1998. Intracellular localization of poliovirus plus- and minus-strand RNA visualized by strand-specific fluorescent In situ hybridization. *J Virol.* 72, 8578-8585.

Brown, D.M., Cornell, C.T., Tran, G.P., Nguyen, J.H., Semler, B.L., 2005. An authentic 3' noncoding region is necessary for efficient poliovirus replication. *J Virol.* 79, 11962-11973.

Bruenn, J.A., 2003. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Research* 31, 1821-1829.

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

Carpenter, C.D., Cascone, P.J., Simon, A.E., 1991. Formation of multimers of linear satellite RNAs. *Virology* 183, 586-594.

Carpenter, C.D., Oh, J.-W., Zhang, C., 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.

Carpenter, C.D., Simon, A.E., 1996. In vivo restoration of biologically active 3' ends of virus-associated RNAs by nonhomologous RNA recombination and replacement of a terminal motif. *J. Virol.* 70, 478-486.

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

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

Chapman, R. M., 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.

Chen, J., Noueir, A., Ahlquist, P., 2001. Brome mosaic virus Protein 1a recruits viral RNA2 to RNA replication through a 5' proximal RNA2 signal. *J. Virol.* 75, 3207-3219.

Cheng, C.P., Panavas T., Luo, G., Nagy, P.D., 2005. Heterologous RNA replication enhancer stimulates in vitro RNA synthesis and template-switching by the carmovirus, but not by the tombusvirus, RNA-dependent RNA polymerase: implication for modular evolution of RNA viruses. *Virology* 341, 107-121.

Choi, Y.G., Dreher, T.W., Rao, A.L., 2002. tRNA elements mediate the assembly of an

icosahedral RNA virus. Proc. Natl. Acad. Sci. USA. 99, 655-660.

Choi, Y.G., Rao, A.L., 2003. Packaging of brome mosaic virus RNA3 is mediated through a bipartite signal. J Virol. 77, 9750–9757.

Citovsky, V., Knorr, D., Schuster, G., Zambryski, P., 1990. The P30 movement protein of tobacco mosaic virus is a single-strand nucleic acid binding protein. Cell 60, 637–647.

Corver, J., Lenches, E., Smith, K., Robison, R.A., Sando, T., Strauss, E.G., Strauss, J.H., 2003. Fine mapping of a *cis*-acting sequence element in yellow fever virus RNA that is required for RNA replication and cyclization. J. Virol. 77, 2265–2270.

de Graaff, M., Jaspars, E.M.J., 1994. Plant viral RNA synthesis in cell-free systems. Annu. Rev. Phytopathol. 32, 311-335.

Dreher, T.W., 1999. Functions of the 3'-untranslated regions of positive strand RNA viral genomes. Annu. Rev. Phytopathol. 37, 151-174.

Dreher, T.W., Bujarski, J.J., Hall, T.C., 1984. Mutant viral RNAs synthesized in vitro show altered aminoacylation and replicase template activities. Nature 311, 171-175.

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

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

Eckerle, L.D., Albarino, C.G., Ball, L.A., 2003. Flock House virus subgenomic RNA3 is replicated and its replication correlates with transactivation of RNA2. *Virology* 317, 95-108.

Eckerle, L.D., Ball, L.A., 2002. Replication of the RNA segments of a bipartite viral genome is coordinated by a transactivating subgenomic RNA. *Virology* 296, 165-176.

Fabian, M.R., Na, H., Ray, D., White, K.A., 2003. 3'-Terminal RNA secondary structures are important for accumulation of Tomato bushy virus DI RNAs. *Virology* 313, 567-80.

French, R., Ahlquist, P., 1987. Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus RNA3. *J. Virol.* 61, 1457-1465.

Frolova, E., Frolov, I., Schlesinger, S., 1997. Packaging signals in alphaviruses. *J Virol.* 71, 248–258.

Gallagher, W.H., Lauffer, M.A., 1983. Calcium ion binding by isolated tobacco mosaic virus coat protein. *J Mol Biol.* 170, 921-929.

Gamarnik, A.V., Andino, R., 1997. Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA. *RNA* 3, 882–892.

Gamarnik, A.V., Andino, R., 1998. Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev.* 12, 2293-2304.

Gerber, K., Wimmer, E., Paul, A.V., 2001. Biochemical and genetic studies of the initiation of human rhinovirus 2 RNA replication: identification of a *cis*-replicating element in the coding sequence of 2A(pro). *J. Virol.* 75, 10979–10990.

Goebel, S.J., Hsue, B., Dombrowski, T.F., Masters, P.S., 2004. Characterization of the RNA components of a putative molecular switch in the 3' untranslated region of the murine coronavirus genome. *J. Virol.* 78, 669-682.

Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J.W., Barclay, W., Evans, D.J., 2000. Identification of a *cis*-acting replication element within the poliovirus coding region. *J. Virol.* 74, 4590–4600.

Guan, H., Carpenter, C.D., Simon, A.E., 2000a. Requirement of a 5'-proximal linear sequence on minus strands for plus-strand synthesis of a satellite RNA associated with



TCV. *Virology* 268, 355-363.

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

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

Guarino, L.A., Ghosh, A., Dasmahapatra, B., Dasgupta, R., Kaesberg, P., 1984. Sequence of the black beetle virus subgenomic RNA and its location in the viral genome. *Virology* 139, 199-203.

Guenther, R.H., Sit, T.L., Gracz, H.S., Dolan, M.A., Townsend, H.L., Liu, G., Newman, W.H., Agris, P.F., Lommel, S.A., 2004. Structural characterization of an intermolecular RNA-RNA interaction involved in the transcription regulation element of a bipartite plant virus. *Nucleic Acids Res.* 32, 2819-2828.

Hacker, D.L., Petty, I.T.D., Wei, N., Morris, T.J., 1992. Turnip crinkle virus genes required for RNA replication and virus movement. *Virology* 186, 1-8.

Hahn, C.S., Hahn, Y.S., Rice, C.M., Lee, E., Dalgarno, L., Strauss, E.G., Strauss, J.H., 1987. Conserved elements in the 3'-untranslated region of flavivirus RNAs and potential cyclization sequences. *J. Mol. Biol.* 198, 33–41.

Hardy R.W., 2006. The role of the 3' terminus of the Sindbis virus genome in minus-strand initiation site selection. *Virology* 345, 520-531.

Herold, J., Andino, R., 2001. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol. Cell* 7, 581-591.

Holstege, F.C., Fiedler, U., Timmers, H.T., 1997. Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.* 16,7468-7480.

Houser-Scott, F., Baer, M.L., Liem, K.F., Cai, J.M., Gehrke, L., 1994. Nucleotide sequence and structural determinants of specific binding of coat protein or coat protein peptides to the 3' untranslated region of alfalfa mosaic virus RNA 4. *J. Virol.* 68, 2194–2205.

Hsue, B., Hartshorne, T., Masters, P.S., 2000. Characterization of an essential RNA secondary structure in the 3' untranslated region of the murine coronavirus genome. *J. Virol.* 74, 6911–6921.

Hsue, B., Masters, P.S., 1997. A bulged stem-loop structure in the 3' untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. *J. Virol.* 71, 7567–7578.

Isken, O., Grassmann, C.W., Sarisky, R.T., Kann, M., Zhang, S., Grosse, F., Kao, P.N., Behrens, S.E., 2003. Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus. *EMBO J.* 22, 5655-5665.

Jaspars, E.M.J., 1999. Genome activation in alfamo- and ilarviruses. *Arch. Virol.* 144, 843–863.

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

Khromykh, A.A., Meka, H., Guyatt, K.J., Westaway, E.G., 2001. Essential role of cyclization sequences in flavivirus RNA replication. *J. Virol.* 75, 6719-6728.

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

Koev, G., Liu, S., Beckett, R., Miller, W.A., 2002. The 3'-terminal structure required for replication of barley yellow dwarf virus RNA contains an embedded 3' end. *Virology* 292, 114-126.

Koh, D.C.Y., Wong, S.M, Liu, D.X., 2003. Synergism of the 3'-untranslated region and an internal ribosome entry site differentially enhances the translation of a plant virus coat protein. *J. Bio. Chem.* 278, 20565–20573.

Kong, Q., Oh, J.-W., Carpenter, C.D., Simon, A.E., 1997. The coat protein of Turnip crinkle virus is involved in subviral RNA-mediated symptom modulation and accumulation. *Virology* 238, 478-485.

Kong, Q., Oh, J.-W., Simon, A.E., 1995. Symptom attenuation by a normally virulent satellite RNA of Turnip crinkle virus is associated with the coat protein open reading frame. *Plant Cell* 7, 1625-1634.

Kong, Q., Wang, J, Simon, A.E., 1997. Satellite RNA-mediated resistance to turnip crinkle virus in *Arabidopsis* involves a reduction in virus movement. *Plant Cell* 9, 2051-2063.

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

Levis, R., Schlesinger, S., Huang, H.V., 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. *J. Virol.* 64, 1726-1733.

Levis, R., Weiss, B.G., Tsiang, M., Huang, H., Schlesinger, S., 1986. Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. *Cell* 44,137-145.

Li, W.-Z., Qu, F., Morris, T.J., 1998. Cell-to-cell movement of turnip crinkle virus is controlled by two small open reading frames that function in trans. *Virology* 244, 405-416.

Li, W.X., Ding, S.W., 2001. Viral suppressors of RNA silencing. *Curr. Opinions Biotech.* 12, 150-154.

Li, X.H., Heaton, L., Morris, T.J. and Simon, A.E. 1989. Defective interfering RNAs of turnip crinkle virus intensify viral symptoms and are generated de novo. *Proc. Natl. Acad. Sci. USA* 86, 9173-9177.

Li, X.H., Simon, A.E., 1990. Symptom intensification on cruciferous hosts by the virulent sat-RNA of turnip crinkle virus. *Phytopathology* 80, 238-242.

Lobert, P.E., Escriou, N., Ruelle, J., Michiels, T., 1999. A coding RNA sequence acts as a replication signal in cardioviruses. *Proc. Natl. Acad. Sci. USA* 96, 11560–11565.

Marsh, L.E., Hall, T.C., 1987. Evidence implicating a tRNA heritage for the promoters of positive-strand RNA synthesis in brome mosaic and related viruses. Cold Spring Harbor Symp. Quant. Biol. 52, 331–341.

Mason, P.W., Bezborodova, S.V., Henry, T.M., 2002. Identification and characterization of a *cis*-acting replication element (*cre*) adjacent to the internal ribosome entry site of foot-and-mouth disease virus. J. Virol. 76, 9686–9694.

Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. 288, 911-940.

McCormack, J., Simon, A.E., 2004. Biased hypermutagenesis of an RNA virus associated with mutations in an untranslated hairpin. J. Virol. 78, 7813-7817.

McKnight, K.L., Lemon, S.M., 1996. Capsid coding sequence is required for efficient replication of human rhinovirus 14 RNA. J. Virol. 70, 1941–1952.

Monkewich, S., Lin, H.-X., Fabian, M.R., Xu, W., Na, H., Ray, D., Chernysheva, O.A., Nagy, P., White, K.A., 2005. The p92 polymerase coding region contains an internal RNA element required at an early step in tombusvirus genome replication. J. Virol. 79, 4848-4858.

Morasco, B.J., Sharma, N., Parilla, J., Flanagan, J.B., 2003. Poliovirus cre(2C)-dependent synthesis of VPgpUpU is required for positive- but not negative-strand RNA synthesis. *J. Virol.* 77, 5136–5144.

Murray, K.E., Barton, D.J., 2003. Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. *J. Virol.* 77, 4739–4750.

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

Nagy, P.D., Carpenter, C.D., Simon, A.E., 1997. A novel 3'-end repair mechanism in an RNA virus. *Proc. Natl. Acad. Sci. USA* 94, 1113-1118.

Nagy, P.D., Pogany, J., 2000. Partial purification and characterization of Cucumber necrosis virus and Tomato bushy stunt virus RNA-dependent RNA polymerases: similarities and differences in template usage between Tombusvirus and Carmovirus RNA-dependent RNA polymerases. *Virology* 276, 279–288.

Nagy, P.D., Pogany, J., 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 J.* 18, 5653-5665.

Nagy, P.D., Pogany, J., Simon, A.E., 2001. In vivo and in vitro characterization of an RNA replication enhancer in a satellite RNA associated with turnip crinkle virus. *Virology* 288, 315-324.

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

Nowakowski, J., Tinoco, I.Jr., 1997. RNA structure and stability. *Sem. Virol.* 8, 153-165.

O'Reilly E.K., 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.

Oh, J.-W., Kong, Q., Song, C., Carpenter, C.D., Simon, A.E., 1995. Open reading frames of turnip crinkle virus involved in satellite symptom expression and incompatibility with *Arabidopsis* ecotype Dijon. *Mol. Plant-Microbe Interact.* 8, 979-987.

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



Olsthoorn, R.C., Mertens, S., Brederode, F.T., Bol, J.F., 1999. A conformational switch at the 3' end of a plant virus RNA regulates viral replication. *EMBO J.* 18, 4856-4864.

Osman, T.A., Buck, K.W., 2003. Identification of a region of the tobacco mosaic virus 126- and 183-kilodalton replication proteins which binds specifically to the viral 3'-terminal tRNA-like structure. *J. Virol.* 77, 8669–8675.

Osman, T.A., Hemenway, C.L., Buck, K.W., 2000. Role of the 3' tRNA-like structure in tobacco mosaic virus minus-strand RNA synthesis by the viral RNA-dependent RNA polymerase in vitro. *J Virol.* 74, 11671-80.

Panavas, T., Nagy, P.D., 2003. The RNA replication enhancer element of tombusvirus contains two interchangeable hairpins that are functional during plus-strand synthesis. *J. Virol.* 77, 258-269.

Panavas, T., Nagy, P.D., 2005. Mechanism of stimulation of plus-strand synthesis by an RNA replication enhancer in a tombusvirus. *J. Virol.* 79, 9777-9785.

Panavas, T., Pogany, J., Nagy, P.D., 2002. Analysis of minimal promoter sequences for plus-strand synthesis by the cucumber necrosis virus RNA-dependent RNA polymerase. *Virology* 296, 263–274.

Panaviene, Z., Panavas, T., Nagy, P.D., 2005. Role of an internal and two 3'-terminal RNA elements in assembly of tombusvirus replicase, *J. Virol.* 79, 10608-10618.

Parsley, T.B., Towner, J.S., Blyn, L.B., Ehrenfeld, E., Semler, B.L., 1997. Poly(rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA* 3, 1124–1134.

Paul, A.V., Rieder, E., Kim, D.W., van Boom, J.H., Wimmer, E., 2000. Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. *J. Virol.* 74, 10359-10370.

Paul, A.V., van Boom, J.H., Filippov, D., Wimmer, E., 1998. Protein-primed RNA synthesis by purified RNA polymerase. *Nature* 393, 280–284.

Petrillo, J.E., Rocheleau, G., Kelley-Clarke, B., Gehrke, L., 2005. Evaluation of the conformational switch model for alfalfa mosaic virus RNA replication. *J Virol.* 79, 5743–5751.

Plasterek, R.H., 2002. RNA silencing: the genome's immune system. *Science* 296, 1263-1265.

Pogany, J., Fabian, M.R., White, K.A., Nagy, P.D., 2003. A replication silencer element in a plus-strand RNA virus. *EMBO J.* 22, 5602-5611.

Pogue, G.P., Marsh, L.E., Connell, J.P., Hall, T.C., 1992. Requirements for ICR-like sequences in the replication of brome mosaic virus genome RNA. *Virology* 188, 742–753.

Qu, F., Morris, T.J., 1997. Encapsidation of turnip crinkle virus is defined by a specific packaging signal and RNA size. *J. Virol.* 71, 1428-1435.

Qu, F., Morris, T.J., 2000. Cap-independent translational enhancement of turnip crinkle virus genomic and subgenomic RNAs. *J. Virol.* 74, 1085–1093.

Qu, F., Ren, T., Morris, T.J., 2003. The coat protein of turnip crinkle virus suppresses posttranscriptional gene silencing at an early initiation step. *J. Virol.* 77, 511-522.

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

Rajendran, K.S., Pogany, J., Nagy, P.D., 2002. Comparison of Turnip crinkle virus RNA-dependent polymerase preparations expressed in *Escherichia coli* or derived from infected plants. *J. Virol.* 76, 1707-1717.

Ranjith-Kumar, C.T., Zhang, X., Kao, C.C., 2003. Enhancer-like activity of a brome mosaic virus RNA promoter. *J. Virol.* 77, 1830-1839.

Ray, D., White, K.A., 1999. Enhancer-like properties of an RNA element that modulates Tombusvirus RNA accumulation. *Virology* 256, 162-171.

Ray, D., White, K.A., 2003. An internally located RNA hairpin enhances replication of Tomato bushy stunt virus RNAs. *J. Virol.* 77, 245-257.

Restrepo-Hartwig, M.A., 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, 8908-8916.

Reusken, C.B.E.M., Bol, J.F., 1996. Structural elements of the 3'-terminal coat protein binding site in alfalfa mosaic virus RNAs. *Nucleic Acids Res.* 24, 2660–2665.

Rieder, E., Paul, A.V., Kim, D.W., van Boom, J.H., Wimmer, E., 2000. Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylylation. *J. Virol.* 74, 10371–10380.

Sarnow, P., 1989. Role of 3'-end sequences in infectivity of poliovirus transcripts made in vitro. *J. Virol.* 63, 467-470.

Schuppli, D., Georgijevic, J., Weber, H., 2000. Synergism of mutations in bacteriophage Q $\beta$  RNA affecting host factor dependence of Q $\beta$  replicase. *J. Mol. Biol.* 295,149-154.

Shen, R., Miller, W.A., 2004. Subgenomic RNA as a riboregulator: negative regulation of RNA replication by Barley yellow dwarf virus subgenomic RNA 2. *Virology* 327, 196–205.

Silvera, D., Gamarnik, A.V., Andino, R., 1999. The N-terminal K homology domain of the poly(rC)-binding protein is a major determinant for binding to the poliovirus 59-untranslated region and acts as an inhibitor of viral translation. *J. Biol. Chem.* 274, 38163–38170.

Silvestri, L.S., Parilla, J.M., Morasco, B.J., Ogram, S.A., Flanagan, J.B., 2006. Relationship between poliovirus negative-strand RNA synthesis and the length of the 3' poly(A) tail. *Virology* 345, 509-519.

Simon, A.E., 1999. Replication, recombination and symptom-modulation properties of the satellite RNAs of Turnip crinkle virus. In: *Cur. Top. in Microbiol. and Immunol.* P. K. Vogt and A. O. Jackson, Eds. Springer-Verlag.

Simon, A.E., Engel, H., Johnson, R.P., Howell, S.H. 1988. Identification of regions affecting virulence, RNA processing and infectivity in the virulent satellite of Turnip crinkle virus. *EMBO J.* 7, 2645-2651.

Simon, A.E., Howell, S.H., 1986. The virulent satellite RNA of Turnip crinkle virus has a major domain homologous to the 3'-end of the helper virus genome. *EMBO J.* 5, 3423-3428.

Simon, A.E., Roossinck, M.J., and Havelda, Z., 2004. Plant virus satellite and defective interfering RNAs: New paradigms for a new century. *Annu. Rev. Phytopathol.* 42, 415-447.

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

Sivakumaran, K., Bao, Y., Roossinck, M.J., Kao, C., 2000. Recognition of the core RNA promoter for minus-strand RNA synthesis by the replicase of *Brome mosaic virus* and *Cucumber mosaic virus*. *J. Virol.* 74, 10323–10331.

Sivakumaran, K., Kim, C.H., Tayon, R.Jr., Kao, C., 1999. RNA sequence and secondary structural determinants in a minimal viral promoter that directs replicase recognition and initiation of genomic plus-strand RNA synthesis. *J. Mol. Biol.* 294, 667–682.

Smirnyagina, E., Hsu, Y.H., Chua, N., Ahlquist, P., 1994. Second-site mutations in the brome mosaic virus RNA3 intercistronic region partially suppress a defect in coat protein mRNA transcription. *Virology* 198, 427–436.

Song, C., Simon, A.E., 1994. RNA-dependent RNA polymerase from plants infected with turnip crinkle virus can transcribe (+)- and (-)-strands of virus-associated RNAs. Proc. Natl. Acad. Sci. USA. 91, 8792-8796.

Song, C., 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., 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., Ahlquist, P., 1997. Cis-acting signals in bromovirus RNA replication and gene expression: Networking with viral proteins and host factors. Semin. Virol. 8, 221–230.

Sullivan, M.L., Ahlquist, P., 1999. A brome mosaic virus intergenic RNA3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA in vivo. J. Virol. 73, 2622–2632.

Tabor, S., Richardson, C.C., 1989. Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. Proc. Natl. Acad. Sci. USA 86:4076–4080.

Thomas, C.L., Leh, V., Lederer, C., Maule, A.J., 2003. Turnip crinkle virus coat protein mediates suppression of RNA silencing in *Nicotiana benthamiana*. *Virology* 306, 33-41.

Todd, S., Towner, J.S., Brown, D.M., Semler, B.L., 1997. Replication-competent picornaviruses with complete genomic RNA 3' non-coding region deletions. *J. Virol.* 71, 8868-8874.

Tretheway, D.M., Yoshinari, S., Dreher, T.W., 2001. Autonomous role of 3'-terminal CCCA in directing transcription of RNAs by Q $\beta$  replicase. *J. Virol.* 75, 11373-11383.

Turner, R.L., Buck, K.W., 1999. Mutational analysis of cis-acting sequences in the 3'- and 5'-untranslated regions of RNA2 of red clover necrotic mosaic virus. *Virology* 253, 115-124.

van Rossum, C.M.A., Reusken, C.B.E.M., Brederode, F.T., Bol, J.F., 1997. The 3' untranslated region of alfalfa mosaic virus RNA3 contains a core promoter for minus-strand synthesis and an enhancer element. *J. Gen. Virol.* 78, 3045-3049.

Voinnet, O., Pinto, Y.M., Baulcombe, D.C., 1999. Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc. Natl. Acad. Sci. USA* 96, 14147-14152.



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

Wang, J., Simon, A.E., 1999. Symptom attenuation by a satellite RNA in vivo is dependent on reduced levels of virus coat protein. *Virology* 259, 234-245.

Williams, G.D., Chang, R.Y., and Brian, D.A., 1999. A phylogenetically conserved hairpin-type 3' untranslated region pseudoknot functions in coronavirus RNA replication. *J. Virol.* 73, 8349–8355.

Wilson, D.S., Szostak, J.W., 1999. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* 68, 611–647.

Wu, B., Vanti, W.B., White, K.A., 2001. An RNA domain within the 5 untranslated region of the Tomato bushy stunt virus genome modulates viral RNA replication. *J. Mol. Biol.* 305, 741–756.

Xiang, W.K., Paul, A.V., Wimmer, E., 1997. RNA signals in entero- and rhinovirus genome replication. *Semin. Virol.* 8, 256–273.

Yang, Y., Rijnbrand, R., McKnight, K.L., Wimmer, E., Paul, A., Martin, A., Lemon, S.M., 2002. Sequence requirements for viral RNA replication and VPg uridylylation

directed by the internal *cis*-acting replication element (*cre*) of human rhinovirus type 14. J. Virol. 76, 7485–7494.

Yoshinari S., Nagy, P.D., Simon, A.E., 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.

Yoshinari, S., Dreher, T.W., 2000. Internal and 3' RNA initiation by Q $\beta$  replicase directed by CCA boxes. Virology 271, 363-370.

You, S., Falgout, B., Markoff L., Padmanabhan, R., 2001. In vitro RNA synthesis from exogenous Dengue viral RNA templates requires long-range interactions between 5'- and 3'- terminal regions that influence RNA structure. J. Biol. Chem. 276, 15581-15591.

Zhang, F., Simon, A.E., 2003a. Enhanced viral pathogenesis associated with a virulent mutant virus or a virulent satellite RNA correlates with reduced virion accumulation and abundance of free coat protein. Virology 312, 8-13.

Zhang, G., Simon, A.E., 2003b. A multifunctional turnip crinkle virus replication enhancer revealed by in vivo functional SELEX. J. Mol. Biol. 326, 35-48.

Zhang, G., Zhang, J., George, A.T., Baumstark T., Simon, A.E, 2006a. Conformational changes involved in initiation of minus-strand synthesis of a virus-associated RNA. *RNA* 12, 147-162.

Zhang, G., Zhang, J., Simon, A.E., 2004a. Repression and derepression of minus-strand synthesis in a plus-strand RNA virus replicon. *J. Virol.* 78, 7619-7633.

Zhang, J., Simon, A.E., 2005. Importance of sequence and structural elements within a viral replication repressor. *Virology* 333, 301-315.

Zhang, J., Stuntz, R., Simon, A.E., 2004b. Analysis of a viral replication repressor: Sequence requirements in the large symmetrical loop. *Virology* 326, 90-102.

Zhang, J., Zhang, G., McCormack, J.C., Simon, A.E. 2006b. Evolution of Virus-Derived Sequences for High Level Replication of a subviral RNA. *Virology*, in press.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406-3415.