ABSTRACT

Title of dissertation: DIFFERENT REPLICATION REQUIREMENTS IN THE HOMOLOGOUS 3' ENDS OF A POSITIVE STRAND RNA VIRUS AND ITS SUBVIRAL RNA

Rong Guo, Doctor of Philosophy, 2010

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

SatC is a noncoding subviral RNA associated with *Turnip Crinkle Virus* (TCV), a small (4054 nt) single-stranded (+)-strand RNA virus belonging to the Carmovirus genus. Because of its small size (356 nt) and TCV-derived 3' end, satC has been successfully used as a model to elucidate sequence and structural requirements for TCV RNA replication.

Although satC is considered a model to identify cis-acting elements required for TCV replication, recent findings indicate distinct differences in structures and functions of these related sequences. RNA2D3D predicts that part of the TCV 3' end (H5, H4a, H4b and two pseudoknots) folds into an internal T-shaped structure (TSS) that binds to 60S ribosomal subunits and is required for translation. SatC contains a similar 3' end with 6 nt differences in the 100 nt TSS region. RNA2D3D did not predict a similar structure for satC TSS region, and satC did not bind yeast ribosomes. satC nucleotides were changed into TCV TSS bases to determine which base differences are responsible for the loss of the TSS in satC. Changing these bases all increased ribosome binding but

surprisingly none of them had an effect on satC accumulation in protoplasts and plants. Therefore satC may need these and other 3' end base differences for its required conformational switch for efficient replication, and not to inhibit ribosome binding.

In vivo genetic selection (SELEX) of the linker sequence between H5 and the Pr showed the conservation of UCC, which led to the discovery of Ψ_2. Ψ_2 is required for both viral and satC accumulation in protoplasts. H5-Pr linker had no significant structural change after RdRp binding in satC, which is different with TCV H5-Pr linker. TCV H5-Pr linker had a major structural change upon RdRp binding, and is proposed to be involved in a conformational switch.

Replacement of satC H4a with randomized sequence and scoring for fitness in plants by SELEX resulted in winning sequences that contain an H4a-like stem-loop. SELEX of H4a/H4b in satC generated two different structures: wt H4a/H4b-like structure and a single hairpin structure. Two highly distinct RNA conformations in the H4a and H4b region can mediate satC fitness in protoplasts.

With the protection of CP, satC can form higher amount of dimers that have additional nucleotides at the junction sites in the absence of TCV. The extra nucleotides are not necessarily associated with an active TCV RdRp.

DIFFERENT REPLICATION REQUIREMENTS IN THE HOMOLOGOUS 3' ENDS OF A POSITIVE STRAND RNA VIRUS AND ITS SUBVIRAL RNA

by

Rong Guo

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
2010

Advisory Committee:

Professor Anne E. Simon (Chair)
Professor Jonathan Dinman
Professor James N. Culver
Assistant Professor Brenda Fredericksen
Associate Professor Douglas Julin (Dean's Rep)

ACKNOWLEDGEMENTS

The first person I would like to acknowledge is my advisor, Dr. Anne E. Simon, for her willingness to accept random unknown students, her enthusiasm in science, her endless ideas for our projects, her ability in maintaining funding. She encouraged us to give talks at national conferences. To prepare for my first conference talk, she sat with me for four hours. She taught me how to write, and corrected my writings word by word.

To my committee members, Dr. Jonathan D. Dinman, Dr. James N. Culver, Dr. Brenda Fredericksen and Dr. Douglas Julin, I thank them for their support and constructive criticism of my research in the past six years. I would like to thank Dr. David Kushner from Dickinson College, for his interest in satC evolution and collaborating with us in H4a and H4b SELEX. His students performed the first two rounds of the H4a and H4b SELEX during a laboratory course, and I thank Wai Lin at Dickinson College finished the H4a/H4b SELEX. I would like to express my gratitude to Dr. Arturas Meskauskas and Dr. Jonathan D. Dinman, for performing satC ribosome binding. I thank Dr. Vera A. Stupina, Dr. John McCormack, Dr. Jiuchun Zhang, Dr. Fengli Zhang, Dr. Xiaoping Sun, and Dr. Guohua Zhang, for their advice, help and patience during my first year in the lab. I would like to acknowledge Megan Young, whose brain knows everything, who bakes delicious cakes on everybody's birthday. I also thank her for proofreading everything I wrote. I thank Dr. Xuefeng Yuan for his precious advice on my research.

To my past and present colleagues, Dr. Vera A. Stupina, Dr. John McCormack, Dr. Jiuchun Zhang, Dr. Guohua Zhang, Dr. Fengli Zhang, Dr. Xiaoping Sun, Dr. Alicia J.

Manfre, and Dr. Xuefeng Yuan, Dr. Feng Gao, Megan Young, Maitreyi Chattopadhyay, Micki Kuhlmann, My Le, I thank them for their discussions, friendship, and reagents.

My husband Shengjun Liu and my lovely daughters Julia and Jayde are my biggest support. I thank them for giving me a happy family. I also want to thank my parents, for their love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

%: Percent

°C: Degrees in Celsius

2,4 D: 2,4-Dichlorophenoxyacetic acid

3'PE: 3'-Proximal element

5'PE: 5'-Proximal element

ATP: Adenosine triphosphate

b: Base(s)

BMV: *Brome mosaic virus*

BaMV: *Bamboo mosaic virus*

BBSV: *Beet black scorch virus*

bp: Basepair(s)

BSA: Bovine serum albumin

BVDV: *Bovine viral diarrhea virus*

BYDV: *Barley yellow dwarf virus*

CarMV: *Carnation mottle virus*

CCFV: *Cardamine chlorotic fleck virus*

CCS: Carmovirus consensus sequence

CIRV: *Carnation Italian ringspot virus*

CMV: *Cucumber mosaic virus*

CNV: *Cucumber necrosis virus*

CP: Coat protein

CPMoV: *Cowpea mottle virus*

CTP: Cytidine triphosphate

CyRSV: *Cymbidium ringspot virus*

D: Dimer

DI RNA: Defective interfering RNA

diG: Defective interfering RNA G

DNA: Deoxyribonucleic acid

DNA: Deoxyribonucleic acid

dNTP: Deoxynucleotide triphosphate

dpi: Days post inoculation

DR: Derepressor element in satC

DTT: Dithiothreitol

DV: *Dengue virus*

E. coli: Escherichia coli

EAV: *Equine arteritis virus*

EDTA: Ethylene diamine tetraacetic acid, disodium, dihydrate

eEF1A: Eukaryotic elongation factor 1A

eIF(iso)4E: Plant isoform of eIF4E

eIF(iso)4F: Plant isoform of eIF4F

eIF(iso)4G: Plant isoform of eIF4G

eIF: Eukaryotic initiation factor

ER: Endoplasmic reticulum

EMSA: Electrophoretic mobility shift assay

FHV: *Flock house virus*

g: Gram

GaMV: *Galingsoga mosaic virus*

GAPDH: Glyceraldehyde 3-phosphate dehydrogenase

GDD: Gly-Asp-Asp domain

gRNA: Genomic RNA

GTP: Guanosine triphosphate

GRV: *Groundnut rosette virus*

H4: Hairpin 4

H4a: Hairpin 4a

H4b: Hairpin 4b

H5: Hairpin 5

HCRSV: *Hibiscus chlorotic virus*

HCV: *Hepatitis C virus*

HDV: *Hepatitis delta virus*

HIV: Human immunodeficiency virus

hpi: Hours post inoculation

hr: Hour(s)

IGR: Intergenic region

JINRV: Japanese iris necrosis virus

K_d: Dissociation constant

L: Liter

LB: Lysogeny broth

LS: Lower stem

LSL: Large symmetrical internal loop

M1H: Motif 1 hairpin

M3H: Motif 3 hairpin

MDV: Qβ bacteriophage-associated midivariant RNA

mg: Milligram

min: Minute(s)

ml: Milliliter

mM: Millimolarity

MNSV: *Melon necrotic spot virus*

MP: Movement protein

mRNA: Messenger RNA

MS salts: Murashige and Skoog basal salt mixture

N: Normality

ng: Nanogram

nt: Nucleotide(s)

OD: Optical density

ORF: Open reading frame

PABP: Poly(A) binding protein

PCR: Polymerase chain reaction

PEG: Polyethylene glycol

PEMV: *Pea enation mosaic virus*

pH: Measure of acidity or alkalinity of a solution

PIM: Protoplast isolation medium

pmol: Picomole

PMV: *Panicum mosaic virus*

PNK: Polynucleotide kinase

Pr: Core promoter for negative-strand initiation

PTGS: Post transcriptional gene silencing

RCNMV: *Red clover necrotic mosaic virus*

RdRp: RNA-dependent RNA polymerase

RE: Replication enhancer

RNA: Ribonucleic acid

rRNA: Ribosomal RNA

RT: Reverse transcription

SARS: *Severe acute respiratory syndrome*

satC: Satellite RNA C

satRNAs: Satellite RNAs

SDS: Sodium dodecyl sulfate

SELEX: Systematic Evolution of Ligands by Exponential Enrichment

SL: Stem loop

sgRNA: Subgenomic RNA

STNV: Tobacco necrosis virus satellite RNA

TAV: T*omato aspermy virus*

TBSV: *Tomato bushy stunt virus*

TCV: *Turnip crinkle virus*

TGGE: Temperature-gradient gel electrophoresis

TMV: *Tobacco mosaic virus*

Tris: Tris[hydroxymethyl]aminomethane

TSS: T-shaped structure

TuMV: *Turnip mosaic virus*

TYMV: *Turnip yellow mosaic virus*

UTR: Untranslated region

UV: Ultraviolet

wt: Wild-type

Δ: Deletion

Ψ_1: Pseudoknot 1

$\Psi_{2'}$: Pseudoknot 2’

Ψ_3: Pseudoknot 3

Ψ_4: Pseudoknot 4

α-^{32}P: Alpha phosphorus-32

γ-^{32}P: Gamma phosphorus-32

μg: Microgram

μL: Microliter

μM: Micromolarity

μmol: Micromole

CHAPTER I

ELEMENTS INVOLVED IN REPLICATION OF POSITIVE STRAND RNA VIRUSES

Introduction

Positive strand RNA viruses comprise over one-third of known virus genera (van Regenmortel et al., 2000), and more than 75% of plant viruses. Crop losses caused by (+)-strand RNA viruses have large economic influences worldwide. Positive strand RNA viruses, like *Poliovirus*, *West nile virus*, and *Severe acute respiratory syndrome coronavirus* (SARS), are major threats to humans and animals. To control the diseases caused by (+)-strand RNA viruses, we need to understand the replication of these viruses.

Positive strand RNA viruses have a wide range of genome configurations. They can be either nonsegmented or segmented, have subgenomic RNAs, or can be associated with satellite RNAs (satRNA) or defective interfering RNAs (DI RNA). Positive strand RNA viruses replicate through multiple steps involving the viral RNA-dependent RNA polymerase (RdRp), other viral proteins, and the host system (Lai, 1998; Mackenzie, 2005).

As a (+)-strand RNA virus progresses through its life cycle, it takes various forms or conformations. Infection of animal viruses is initiated by receptor-mediated endocytosis, whereas plant viruses are transmitted by insects or mechanical abrasion. Upon infection, the virus uncoats from the virion and releases the viral (+)-strand RNA

into the cytoplasm. pH changes, membrane receptor binding, or protease activity may trigger this conformational change (Flint et al, 2004). The viral RdRp is then translated using the cellular translation machinery. At some point, translation stops and the same genomic RNA acts as the template for (-)-strand synthesis. Since translation (5' to 3' end) and transcription (3' to 5' end) cannot occur simultaneously, the initial viral (+)-strand RNA needs to switch from a translation-competent template to a replication-competent template allowing (-)-strand RNAs to be transcribed. Newly synthesized (-)-strands then serve as templates for nascent (+)-strand RNA synthesis. Progeny (+)-strand genomes have been proposed to fold into a conformation that is not accessible to the RdRp, so the newly formed (+)-strand may not be a template for further (-)-strand synthesis (Zhang et al., 2006a; Zhang et al., 2006b). (+)-strand and (-)-strand synthesis can be asymmetrical with up to 1000 (+)-strands synthesized for every one (-)-strand (Buck, 1996).

Minus-strands also serve as templates for subgenomic RNA (sgRNA) synthesis for viruses that have 3' coterminal subgenomic RNAs. sgRNAs, which code for structural and/or movement proteins, are produced by three possible mechanisms: premature termination of transcription by the RdRp during (-)-strand synthesis (Sit, Vaewhongs, and Lommel, 1998; Tatsuta et al., 2005; Wu et al., 2010); internal initiation on (-)-strands during (+)-strand synthesis (Levis, Schlesinger, and Huang, 1990; Miller, Dreher, and Hall, 1985); or discontinuous transcription (Jeong and Makino, 1994).

In the life cycle of RNA viruses, packaging, or encapsidation of the viral genome by structural proteins is an essential step for transmission. Plant RNA viruses have different mechanisms for packaging. For monopartite Tymoviruses, the genomic RNA and sgRNA are packaged into separate virions. In the genus Luteovirus, only the genomic

RNA is packaged. The two RNA molecules from bipartite RNA viruses are packaged either separately into distinct virions (Comovirus and Nepovirus) or into the same virion (Dianthovirus). In the tripartite genera Bromovirus and Cucumovirus, the largest two genomic RNAs are packaged individually into virions and the third genomic RNA and its sgRNA are co-packaged into a third virion. In the genus Alfamovirus, the three genomic RNAs and sgRNA are packaged individually into four distinct virions (Rao, 2006). sgRNA, satRNA, and DI RNAs are not required to initiate an infection, and are dependent on the parental virus for replication and packaging (Simon, Roossinck, and Havelda, 2004). Co-packaging of two or more RNA molecules requires either that they all contain a packaging signal (Annamalai and Rao, 2005) or that RNAs without a packaging signal interact with a signal-containing RNA (Basnayake, Sit, and Lommel, 2006).

Replication and transcription of (+)-strand RNA viruses are conducted by the RdRp. RdRp, along with the other three types of polymerases (DNA-dependent DNA polymerase, DNA-dependent RNA polymerase, RNA-dependent DNA polymerase), all have a right hand shape consisting of a palm, fingers and thumb domains, with the active site of the enzyme located in the palm subdomain (Baker and Bell, 1998). In spite of their wide variation in genomic conformation and sequences, the four types of polymerases all have four common motifs. One critical motif is a highly conserved Gly-Asp-Asp (GDD) sequence located in the catalytic site of the enzyme (Jablonski and Morrow, 1995; Letzel, Mundt, and Gorbalenya, 2007; O'Reilly and Kao, 1998; Vázquez, Alonso, and Parra, 2000; Wang and Gillam, 2001; Wang et al., 2007).

Unlike DNA replication that requires a primer, RNA transcription by most (+)-strand RNA virus polymerases occurs *de novo* (i.e., no primer needed), starting with the 3' end. RdRp from some viruses can use DNA as a template to synthesize RNA, though it is about 10% as efficient as RNA synthesis with RNA templates (Siegel et al., 1999).

Viruses are considered to be molecular genetic parasites that utilize cellular systems for their own replication (Villarreal, 2005). They depend entirely on the host cells to reproduce their genome and form infectious progeny. For all (+)-strand RNA viruses investigated so far, the viral replication complex co-purifies with cellular membrane extracts. Viral RNA synthesis takes place in virus-induced membrane compartments that are derived from different organelles (Table 1.1), with viral proteins modifying the structure of intracellular membranes. Different viruses induce diverse but specific cellular structures, and can also use different membranes in the absence of the cognate organelle membrane. These virus-induced membrane structures often contain the viral replication complex and are thought to be the replication sites (Laliberté and Sanfaçon, 2010; Miller and Krijnse-Locker, 2008). The membranes provide a scaffold for anchoring the replication complex, which increases the local concentration of components required for viral replication, and protects the viral RNA from being degraded by the host defense system. Although membrane association is important for (+)-strand RNA virus replication, the mechanism of replication and the formation of these membrane structures are still poorly understood.

In addition to viral replication proteins, cellular proteins have been found within virus-induced membrane structures including several proteins involved in protein synthesis and folding. eIF(iso)4E, eEF1A, polyA-binding protein (PABP) and Hsp70 are

localized in vesicles induced by *Turnip mosaic virus* (TuMV; *Potyviridae*, *Potyvirus* genus) and are involved in viral RNA synthesis (Beauchemin, Boutet, and Laliberte, 2007; Beauchemin and Laliberte, 2007; Dufresne et al., 2008; Thivierge et al., 2008). Hsp70 and eEF1A are also found in *Tomato bushy stunt virus* (TBSV; *Tombusviridae*, *Tombusvirus* genus)-induced vesicles (Li et al., 2009; Wang, Stork, and Nagy, 2009). DnaJ and eIF3 are involved in *Brome mosaic virus* (BMV; *Bromoviridae*, *Bromovirus* genus) RNA replication (Quadt et al., 1993; Tomita et al., 2003). The mechanism and function of these cellular proteins in viral RNA replication is not clear, with possibilities including stabilizing viral replication proteins and controlling their activity.

Replication site	Family	Genus	Virus	Reference
ER membrane	*Potyviridae*	*Potyvirus*	*Tobacco etch virus*	(Schaad, Jensen, and Carrington, 1997)
			Zucchini yellow mosaic virus	(Zechmann, Müller, and Zellnig, 2003)
	Comoviridae	*Nepovirus*	*Grapevine fanleaf virus*	(Ritzenthaler et al., 2002)
			Tomato ringspot virus	(Han and Sanfacon, 2003)
			Cowpea mosaic virus	(Carette et al., 2000)
	Alphaflexiviridae	*Potexvirus*	*Potato virus X*	(Bamunusinghe et al., 2009)
	Bromoviridae	*Bromovirus*	*Brome mosaic virus*	(Restrepo-Hartwig and Ahlquist, 1996; Schwartz et al., 2002)
	Tombusviridae	*Dianthovirus*	*Red clover necrotic mosaic virus*	(Turner et al., 2004)
	unclassified	*Tobamovirus*	*Tobacco mosaic virus*	(Kawakami, Watanabe, and Beachy, 2004; Más and Beachy, 1999; Nishikiori et al., 2006; Reichel and Beachy, 1998)
mitochondrial membrane	*Tombusviridae*	*Carmovirus*	*Melon necrotic spot virus*	(Mochizuki et al., 2009)
		Tombusvirus	*Carnation Italian ringspot*	(Weber-Lotfi et al., 2002)
peroxisomal membrane	*Tombusviridae*	*Tombusvirus*	*Tomato bushy stunt virus*	(McCartney et al., 2005b)
			Cymbidium ringspot virus	(Navarro et al., 2006)
			Cucumber necrosis virus	(Panavas et al., 2005; Russo, Di Franco, and Martelli, 1983)
chloroplast membrane	*Potyviridae*	*Potyvirus*	*Turnip mosaic virus*	(Prod'homme et al., 2003; Prod'homme et al., 2001; Wei et al., 2010)
	Tymoviridae	*Tymovirus*	*Turnip yellow mosaic virus*	(Hatta, Bullivant, and Matthews, 1973),

Table 1.1 Cellular membranes associated with plant (+)-strand RNA viruses

Subviral RNAs Associated With Positive Strand RNA Viruses

Viruses can be associated with subviral RNAs (DI RNAs or satellite RNAs) that are dependent on the helper virus for replication, movement, and encapsidation (Simon, Roossinck, and Havelda, 2004). DI RNAs are derived mainly or completely from the genome of the helper virus, while by definition satRNAs have sequences either mostly or completely unrelated to any large contiguous segment of the helper virus genome. Plant RNA viruses are more frequently associated with satellite RNAs. SatC from *Turnip crinkle virus* (TCV; *Tombusviridae, Carmovirus* genus) and *Bamboo mosaic virus* (BaMV; *Flexiviridae, Potexvirus* genus) satRNA are among the best studied. DI RNAs are commonly associated with animal RNA viruses, and have been found in a subset of plant viruses including carmoviruses, tombusviruses, bromoviruses, furoviruses, potexviruses, hordeiviruses, rhabdoviruses, and bunyavirus (reviewed by Simon and Bujarski, 1994). The best characterized plant virus DI RNAs are DI RNAs from TBSV and TCV.

Origins of subviral RNAs

Animal and plant RNA viruses can form DI RNAs *de novo* upon high multiplicity passage of DI RNA-free isolates, and are considered to be mistakes generated by the error-prone replicase during replication (Burgyan, Rubino, and Russo, 1991; Knorr et al., 1991; Li et al., 1989; Marsh et al., 1991; Resende et al., 1991). Upon infection with *in vitro* transcribed TCV (DI RNA-free), new low molecular weight RNAs accumulated that hybridized to TCV-specific probes. These new DI RNAs (e.g., DI1) were collinear

deletion mutants of TCV containing the exact TCV 5' and 3' ends and an internal segment (Li et al., 1989). diG (346 nt), a DI RNA associated with TCV isolate TCV-B, is a mosaic molecule composed of 21 nucleotides of unknown origin at the 5' end, TCV 5' portion in the middle and TCV 3' terminus with a centrally repeated block at the 3' end (Li et al., 1989). There are two different mechanisms proposed for DI RNA generation; enzyme cutting and ligation and replicase-driven copy choice (Simon and Bujarski, 1994). Recombination between TCV subviral RNAs support the replicase-driven copy choice model for the formation of DI RNAs (Cascone et al., 1990). During replication, when the replicase copies the viral (-)-strand, it can dissociate from the template prematurely. Before the newly made (+)-strand is released from the replicase complex, the replicase reinitiates (+)-strand synthesis either on the same template or on a different viral (-)-strand. Besides generation of DI RNAs, RNA recombination plays a pivotal role in virus evolution (reviewed by Simon and Bujarski, 1994).

Since satRNAs share little sequence similarity with their helper viruses, their origins remain unknown (Hu, Hsu and Lin, 2009). At least three sources are the possible origins of satRNAs: the genomes of the helper viruses, the host plants, and transmission vectors. While the origin of satRNAs and the mechanism by which they are generated remain unknown, this does not limit their usefulness as tools for investigating properties shared with their helper viruses.

Relationship between subviral RNAs and their helper viruses

In the past, subviral RNAs were considered to be parasites because of their interference with helper virus replication. They are usually not required by their helper

viruses, with the exception of the satRNA associated with *Groundnut rosette virus* (GRV; *Umbravirus*). This satRNA is essential for both encapsidation of GRV genomic RNA (Robinson et al., 1999) and aphid transmission of GRV (Murant, 1990). What has not been generally considered is that maintenance of non-essential subviral RNAs by helper viruses, even those that interfere with helper virus replication, requires that the subviral RNA confers some selective advantage (Simon et al., 2004). *Pea enation mosaic virus* (PEMV; *Umbravirus*) satRNA may stabilize virions, which is advantageous to the two genomic RNAs (de Zoeten and Skaf, 2001). The relationship between satC and its helper virus TCV is mutualistic, where both participants gain fitness from the interaction. By interfering with virion accumulation, satC indirectly enhances movement of TCV (Zhang and Simon, 2003a) by supporting the ability of TCV to overcome plants' defense system, post transcriptional gene silencing (PTGS). Reduced accumulation of virions by satC results in enhanced levels of free CP, which is the TCV silencing suppressor (Qu, Ren, and Morris, 2003; Thomas et al., 2003).

Subviral RNAs require the presence of a helper virus to provide *trans*-acting factors necessary for replication and often accumulate at the expense of the helper virus. Interference with the helper virus frequently results in substantial symptom attenuation. Most DI RNAs can reduce viral genome RNA accumulation and attenuate helper virus symptoms in a host-specific manner (Simon, Roossinck, and Havelda, 2004), while diG associated with TCV is responsible for intensified symptoms when co-inoculated with TCV on cruciferous plants (Li et al., 1989).

SatRNAs usually have little sequence similarity with the helper virus, and have varied effects on the helper virus. Different satRNAs associated with the same helper

virus, or the same satRNAs on different host plants, can cause different effects on host phenotypes and on helper virus accumulation. Most *Cucumber mosaic virus* (CMV; *Bromoviridae, Cucumovirus* genus) satRNAs reduce the accumulation of CMV genomic RNA and attenuate symptoms, while pathogenic satRNAs from the same virus induce lethal necrosis in tobacco plants, or programmed cell death on different plants (Simon, Roossinck, and Havelda, 2004; Xu and Roossinck, 2000). SatRNAs from TBSV (Célix, Burgyán, and Rodríguez-Cerezo, 1999), *Peanut stunt virus* (*Bromoviridae, Cucumovirus* genus) (Militão et al., 1998), and BaMV (Hsu et al., 2006) have varied effects on viral RNA replication or symptom attenuation on the same host plants. Two satRNA clones BSL6 and BSF4 from BaMV had very different effects on helper virus accumulation. BSL6 reduced BaMV accumulation and attenuated BaMV-induced symptoms in co-inoculated plants, while BSF4 had only a negligible effect on helper virus accumulation (Hsu et al., 2006).

Some satRNAs, like satC from TCV and satRNA from *Panicum mosaic virus* (PMV; *Tombusviridae, Panicovirus* genus) can intensify symptoms caused by their helper viruses. PMV satRNA enhances viral accumulation in infected plants (Scholthof, Jones, and Jackson, 1999), while satC interferes with its genomic RNA accumulation (Li and Simon, 1990; Simon and Howell, 1986).

The mechanism of subviral RNA symptom attenuation was reviewed in Simon et al. (2004). For DIs that attenuate symptoms, competition with the helper virus for limited *trans*-acting replication factors, specific interaction with viral encoded protein p19 (a strong silencing suppressor), or activation of PTGS may be involved. Several different mechanisms may result in down-regulations of helper virus accumulation by satRNAs. 5'

end sequences determine GRV attenutation by satRNA (Taliansky and Robinson, 1997). Down regulation of BaMV replication was mapped to specific sequences and the secondary structure of a hairpin in the 5' UTR (Chen, Hsu, and Lin, 2007; Hsu et al., 2006). SatC-mediated reduction of TCV levels requires the CP (but not necessarily the silencing suppressor activity of the CP), either cognate or from a related carmovirus, and is independent of CP interaction with the RNAs (Manfre and Simon, 2008). CMV symptom attenuation caused by associated satRNAs was proposed to be due to competition between helper virus and satRNA for RdRp (Wu and Kaper, 1995). Another CMV satRNA can enhance viral accumulation. The pathogenicity of CMV satRNA was mapped to a few specific nucleotides at the 3' terminus, and a hairpin stemloop within the 5' half of the (-)-strand was responsible for enhanced symptoms in infected leaves (Xu and Roossinck, 2000).

Changes in genomic RNA accumulation caused by subviral RNAs are not always correlated with symptom variation. High levels of genomic RNAs were detected in the presence of TBSV pepper isolate (TBSV-P) DI-5 RNA, but the plants were protected from necrosis. While inoculation of TBSV-P and DI RNAs resulted in reduced viral genome RNA accumulation and high levels of DI RNA, the infected plants were completely necrotic (Havelda, Szittya, and Burgyán, 1998). TCV levels were reduced in plants co-inoculated with satC and TCV, while symptoms were more severe than when inoculated with TCV only (Li and Simon, 1990).

Subviral RNAs as model systems

Subviral RNAs have sequence and structural requirements for replication of their RNAs but do not encode their own replicase and thus replication is solely dependent on their helper viruses. DI RNAs share substantial sequence similarity with their helper viruses and some satellite RNAs possess structural conservation with their helper viruses. They all contain *cis*-acting RNA elements that allow for efficient amplification when helper virus replicase is provided in *trans*. To separate replication from translation on the genomic RNA, the small, non-translated subviral RNAs are frequently used as models to identify and analyze the *cis*-acting elements required for genomic RNA replication (Chernysheva and White, 2005; Fabian et al., 2003).

BaMV is a single-stranded (+)-strand monopartite RNA virus containing a 5' cap and 3' poly(A) tail. The 3' untranslated region (UTR) folds into four stem-loop hairpins (domains A to D) and a pseudoknot structure (domain E) (Cheng and Tsai, 1999). BaMV replication requires both structure and sequences in the 3' end *in vivo* (Chen et al., 2003; Cheng and Tsai, 1999) and *in vitro* (Cheng et al., 2001). The satRNA associated with BaMV is a messenger type RNA but does not encode its own CP. The 3' end of satBaMV folds into three stem-loops (SLA, SLB and SLC) that are structurally similar to BaMV 3' end domains B, C, and D, although they share no significant sequence similarity (Huang et al., 2009). Like BaMV, the 3' end structure is required for satBaMV replication in protoplasts and plants (Huang et al., 2009). Both mFold (Zuker, 1994) and enzyme structure probing showed that the 5' end of BaMV (Chen, Desprez, and Olsthoorn, 2009) and satBaMV (Annamalai et al., 2003; Chen, Hsu, and Lin, 2007) fold into a large stem-loop structure with an asymmetrical internal loop consisting of either C:C or C:A mismatches and an A bulge at identical positions.

Despite the sequence, structural or functional similarities between subviral RNAs and their helper viruses, some discrepancies in function do occur. SatBaMV with 5' UTR, 3' UTR, or both regions from BaMV accumulated to less than 10% of wt satBaMV, although BaMV and satBaMV UTRs share high structural similarity (Huang et al., 2009). This implies that subviral RNAs have evolved distinct features for their own replication in the absence of the genomic RNA sequence.

Basic Aspects of Virus Replication

Viral replicases, which comprise the virus-encoded RdRp together with host proteins and/or additional viral proteins (Lai, 1998), must recognize their cognate RNAs through direct or indirect interaction with specific sequence or structural elements (*cis*-acting elements) to initiate complementary strand synthesis. Unlike DNA-dependent RNA polymerases, RdRp require no general conserved sequence or structure specificity for binding to RNA templates. There are various *cis*-acting elements in both (+)- and (-)-strand (the replication intermediate) RNAs to direct transcription. Specific template recognition by the replicase complex is essential for faithful genome replication. The *cis*-acting elements include core promoters and various regulatory RNA elements such as replication enhancers, silencers that down regulate RNA replication, template recruitment elements, and RNA elements that are required for genome circularization. These RNA elements are usually located in 3' UTRs. tRNA-like structures (TLS) are found on the genomes of tymoviruses, furo-like viruses, bromoviruses, cucumoviruses, hordeiviruses, tobamoviruses, and tobraviruses. TLSs have important roles in viral RNA replication,

translation and encapsidation (Dreher, 1999; Dreher, 2009). These viral TLSs can be either aminoacylated by specific aminoacyl-tRNA synthetases (*Turnip yellow mosaic virus*, valine aminoacylation; BMV, tyrosine aminoacylation; *Tobacco mosaic virus*, histidine aminoacylation), or serve as substrates for 3' terminal adenylation by tRNA nucleotidyltransferase (tobravirus and tymovirus). Some plant (+)-strand RNA viruses, including potyviruses, potexviruses, comoviruses, capilloviruses, carlaviruses, and *beet necrotic yellow vein virus*, terminate their genomes in poly(A) tails, similar to cellular mRNAs. Viruses from genera *Sobemovirus, Luteovirus, Tombusvirus, Carmovirus, Dianthovirus, Umbravirus,* and *Closterovirus* have neither a TLS nor a poly(A) tail at the 3'end, but terminate their genome with hairpins, pseudoknots, or short primary sequences (Dreher, 1999).

Core promoters are the minimal sequences required for complementary strand synthesis. The promoter for (-)-strand synthesis usually requires both sequence and structure specificity, while the promoter for (+)-strand initiation does not appear to have such complicated requirements. The promoter for (+)-strand initiation is thought to be stronger than the (-)-strand promoter, accounting for the asymmetric synthesis of (+)- and (-)-strand RNAs. Replication enhancers can either stimulate the core promoter activity, or stabilize and recruit RNA templates to replication sites. Enhancers can (i) increase the basal levels of RNA synthesis; (ii) function with no orientation; (iii) and function independently of position.

5' end elements or internal elements have been found to bind RdRp, with RNA bridges or RNA-protein interactions directing the RdRp to the 3' end (Alvarez et al., 2005; Klovins and van Duin, 1999; Miller and White, 2006; Pogany, White, and Nagy, 2005;

Wu, Vanti, and White, 2001; Yoshinari and Dreher, 2000; Yuan et al., 2009). The most common RNA-RNA interaction is the pseudoknot (either H-type or H-H type kissing loop). Pseudoknots play crucial roles in viral gene expression, either in *cis* or in *trans* (Brierley, Pennell, and Gilbert, 2007). Pseudoknots are reported to regulate *Barley yellow dwarf virus* (BYDV; *Luteoviridae, Luteovirus* genus) and TBSV translation initiation (Fabian and White, 2004; Guo, Allen, and Miller, 2001); TCV and Qβ replication (Klovins and van Duin, 1999; McCormack et al., 2008; Zhang, Zhang, and Simon, 2004; Zhang et al., 2006b); TBSV, *Red clover necrotic mosaic virus* (RCNMV; *Tombusviridae*, *Dianthovirus* genus) and *Flock house virus* (FHV; *Nodaviridae, Alphanodavirus* genus) subgenomic RNA synthesis (Choi and White, 2002; Guenther et al., 2004; Lin and White, 2004; Lindenbach, Sgro, and Ahlquist, 2002; Sit, Vaewhongs, and Lommel, 1998); RCNMV virion packaging (Basnayake, Sit, and Lommel, 2006); and *Human immunodeficiency virus-1* (HIV-1) viral genome dimerization (Mujeeb et al., 1998).

RNA-RNA base-pairing interactions can be protein-mediated. The RNA/RNA/protein model has been suggested for translation initiation of BYDV (Guo, Allen, and Miller, 2001), Satellite *Tobacco necrosis virus* (STNV) (Gazo et al., 2004), *Cucumber leafspot virus* (CLSV) (Xu and White, 2009), TBSV (Fabian and White, 2006), and *Carnation Italian ringspot virus* (CIRV; *Tombusviridae, Tombusvirus* genus) (Nicholson and White, 2008). The 3' end of *Melon necrotic spot virus* (MNeSV; *Tombusviridae, Carmovirus* genus) (3' cap-independent translation enhancer; 3'CITE) can form a complex RNA structure that binds to eIF4F. The eIF4F-bound 3'CITE complex interacts with the 5' UTR of MNeSV through RNA-RNA interaction. The 5'UTR/3'CITE/eIF4F complex functions similar to the cap-dependent mechanism and

mediates translation initiation (Nicholson et al., 2010).

Model systems for studying virus replication

TBSV (*Tomato bushy stunt virus*) replication

Tombusviruses have been excellent model systems for understanding molecular aspects of (+)-strand RNA virus replication. TBSV, the prototype member of Tombusviruses, terminates its genome with a 5-nt stem and a stable tetraloop upstream. TBSV's DI RNA has four noncontiguous regions (I to IV) derived from the viral genome (White, 1996). The *cis*-acting RNA elements necessary for TBSV replication are contained in these four regions, so the TBSV DI RNA has been used as a model to study *cis*-acting elements required for TBSV accumulation both *in vivo* and *in vitro* (White and Nagy, 2004). Region I is from the 5' UTR, and has a T-shaped domain (TSD). Mutation of this TSD can reduce DI RNA accumulation in protoplasts by 20-fold (Wu, Vanti, and White, 2001), thus it is important, but not essential, for efficient DI RNA replication (Wu and White, 1998). Regions II and III are from internal positions, and region IV is from the 3' end of the genomic RNA. Region III functions in the (-)-strand and acts to promote efficient (+)-strand synthesis and is dispensable for DI RNA accumulation in plants (Chang et al., 1995). The enhancing activity was mapped to a 35-nt segment (Ray and White, 1999; Ray and White, 2003). Region II and region IV are critical in DI RNA (Fabian et al., 2003) and TBSV genome replication (Monkewich et al., 2005).

The core promoter for TBSV (-)-strand synthesis was mapped to the 3'-terminal 19 nucleotides (5'-CAUUGCAGAAAUGCAGCCC) by *in vitro* transcription (Panavas,

Pogany, and Nagy, 2002), and to three 3' terminal stem loops *in vivo* (Fabian et al., 2003). For (+)-strand synthesis, deletion and mutagenesis analyses showed that the minimal promoter sequence for a DI RNA (+)-strand replicon is the 3'-terminal 11 nucleotides (3'-CCUUUAAGAGG), which are highly conserved among tombusviruses and their associated subviral RNAs (Panavas, Pogany, and Nagy, 2002). *In vitro* competition experiments using purified *Cucumber necrosis virus* (CNV; *Tombusviridae*, *Tombusvirus* genus) RdRp showed that the (+)-strand core promoter is a stronger promoter than the (-)-strand core promoter. Surprisingly, the CNV RdRp preparation was able to synthesize RNA from a Carmovirus (satC associated with TCV) core promoter, but not an artificial promoter. The core promoters of TBSV DI RNA and satC are 50% different in sequence, yet share some common features, such as the 3'-terminal cytidylates and a stretch of purich-rich sequence (Panavas, Pogany, and Nagy, 2002).

TBSV replicase can internally bind to its RNA template. TBSV replicase protein p33 binds to RII on the (+)-strand (Pogany, White, and Nagy, 2005) and RIII on the (-)-strand (Panavas and Nagy, 2003; Ray and White, 2003).

One cytopathological feature upon TBSV infection is the presence of peroxisome-derived multivesicular bodies. The incorporation of tritiated uridine indicates that these vesicles are sites of TBSV replication (Appiano et al., 1981; Appiano et al., 1986). TBSV replicase proteins p33 and p92 are membrane bound proteins (Scholthof, Scholthof, and Jackson, 1995), and co-localize with endogenous peroxisomal catalase (McCartney et al., 2005a). p33 contains the peroxisomal membrane targeting signal and by itself can cause the peroxisomal boundary membrane vesiculation (McCartney et al., 2005a).

Several cellular proteins are found associated with the TBSV replicase complex. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), a host metabolic enzyme, is the only cellular protein that has a clear function in viral RNA replication. During TBSV replication in yeast, GAPDH moves from the cytosol to the peroxisomal membrane surface, selectively binding the (-)-strand RNA through an AU pentamer sequence. GAPDH retains the (-)-strand RNA within the replication complex, which is proposed to promote asymmetric synthesis of (+)-strand RNA (Wang and Nagy, 2008).

BMV (*Brome mosaic virus*) replication

BMV, a member of alphavirus-like superfamily, has a tripartite genome. Genomic RNA1 encodes protein 1a, which is a key factor for replicase assembly; RNA2 encodes the RdRp 2a; RNA3 contains the open reading frame (ORF) for the movement protein and CP. The three BMV genomic RNAs contain a tRNA-like structure that is sufficient to initiate (-)-strand RNA synthesis *in vitro* (Dreher, Bujarski, and Hall, 1984; Miller et al., 1986). Using a template competition assay, BMV stem C (3' proximal 136-160 nt) was identified as the determinant bound by the RdRp (Chapman and Kao, 1999). Although the 200-nt 3' end of the three genomic RNAs share high sequence and structure similarity, they cannot be interchanged (Duggal, Rao, and Hall, 1992).

There are major differences in (+)- and (-)-strand RNA synthesis for BMV. BMV (+)-strand RNA initiation requires a 26-nt region from the 3' end of (-)-strand template (Sivakumaran et al., 1999) and a 3' nontemplated nucleotide (-1 position) for *in vitro* RdRp-directed synthesis (Sivakumaran and Kao, 1999). Unlike (-)-strand synthesis, which requires a highly structured promoter, the 27-nt region on the three genomic RNAs

required for (+)-strand synthesis can adopt different secondary structures. The -1, +1, and +2 nucleotides determine RdRp binding to the template (Sivakumaran et al., 1999). The 3' nontemplated nucleotide that is added by the RdRp during (-)-strand synthesis (Siegel, Adkins, and Kao, 1997) is required for (+)-strand synthesis, but the identity of this nucleotide is not crucial. BMV RdRp prefers the initiating cytidylate to be at the penultimate position (Sivakumaran and Kao, 1999), with more than one nontemplated nucleotide dramatically reducing RNA synthesis.

The BMV tRNA-like 3' end functions as the core promoter for (-)-strand synthesis, but correct (-)-strand initiation may require a 5' end sequence *in vivo* (Sullivan and Ahlquist, 1997). The 250 nt intergenic region (IGR) of BMV RNA3 contains sequences that can stimulate (-)-strand synthesis by about 100-fold in yeast (Quadt et al., 1995). The BMV 1a protein has putative helicase and capping domains. The expression of 1a protein alone dramatically stabilizes RNA3 in yeast, and this 1a induced stabilization was mapped to IGR (Sullivan and Ahlquist, 1999). The BMV 1a protein recognizes RNA3 through this region and recruits it from translation to replication. This IGR stimulates RNA3 (-)-strand synthesis by 50 to 100-fold in protoplasts (French and Ahlquist, 1987) and yeast (Sullivan and Ahlquist, 1999) and is designated RE (replication enhancer).

Template competition assays showed that BMV RdRp recognizes RNA1 and RNA2 by similar mechanisms while RNA3 was recognized in a different way. BMV RdRp internal binding occurred only with RNA3 (Choi et al., 2004). BMV RNA3 has ten replicase binding sites, with 7 of 10 at the core promoters for (+)-strand, (-)-strand or subgenomic RNA synthesis (Choi et al., 2004).

Under high-resolution immunofluorescence confocal microscopy, BMV replication-required proteins 1a and 2a accumulated in cytoplasmic patches that often assumed a vesicular appearance in barley protoplasts (Restrepo-Hartwig and Ahlquist, 1996) and yeast (Restrepo-Hartwig and Ahlquist, 1999). The co-localization of 1a, 2a, as well as the nascent RNA containing bromouridine (BrU) in the presence of actinomycin D and ER markers indicates tha the ER is the primary location for BMV RNA synthesis (Restrepo-Hartwig and Ahlquist, 1996). Protein 1a, which is responsible for ER localization and retention of the RNA replication complex, localizes to the outer perinuclear ER membranes when expressed in the absence of other viral factors (Restrepo-Hartwig and Ahlquist, 1999). The ER luminal membrane displayed 50-70 nm vesicular invaginations or spherules, with necks connecting them to the cytoplasm in yeast (Schwartz et al., 2002). By interacting with RE on BMV genomic RNA template RNA3, 1a sequesters viral (+)-strand template in a nuclease-resistant, detergent-susceptible state. Polymerase 2a, which co-localizes in these spherules, is proposed to initiate (-)-strand RNA synthesis, and retains the (-)-strand RNA for (+)-strand RNA synthesis (Schwartz et al., 2002).

TCV (*Turnip Crinkle Virus*) replication

At 4054 nt, TCV is one of the smallest and simplest of the (+)-strand RNA viruses. It infects plants by flea beetle or mechanical abrasion and replicates on mitochondrial membranes after uncoating and translation. The single genomic RNA contains five ORFs. The 5' ORF is translated from the genomic RNA and encodes the replication-required protein, p28, and the longer p88 RdRp, which is translated by

ribosomal readthrough of the p28 amber termination codon (Figure 1.1; Carrington et al., 1987; Hacker et al., 1992). p88 contains the GDD motif, and by itself is able to direct complementary strand synthesis of exogenously added (+)- and (-)- strand TCV subviral RNAs *in vitro* (Rajendran, Pogany, and Nagy, 2002). TCV is also associated with two subgenomic RNAs: the 1.7 kb subgenomic RNA is the mRNA for p8 and p9, which are required for virus movement (Li et al., 1989), and the 1.45 kb subgenomic RNA contains a single ORF for the viral CP (Hacker et al., 1992). In addition to its monopartite (+)-sense genome, TCV is associated with at least three non-essential satRNAs: satC, satD and satF. The 356 nt satC is a unique chimeric RNA composed of a true satRNA, satD (194 nt), which shares little contiguous sequence similarity with the genomic RNA, and two regions from the 3' terminus of TCV genomic RNA (Simon and Howell, 1986). The 190 nucleotides from the 5' end of satC share 88% sequence similarity with satD, and the 166 nucleotides from the 3' end share 94% similarity with TCV 3' end regions (Simon and Howell, 1986). A replicase-driven copy choice model requiring two recombination events between satD and the two TCV fragments can explain how satC was generated (Cascone et al., 1990). The original formation of satD, which has little sequence similarity with TCV, remains unknown. SatC, the unique 5' sat/3' DI structure RNA intensifies TCV symptoms on different hosts (Li and Simon, 1990; Simon and Howell, 1986), interferes with TCV accumulation both in protoplasts and in plants (Kong, Oh, and Simon, 1995; Li and Simon, 1990), and represses virion formation (Zhang and Simon, 2003a). TCV satRNAs have no coding capability and their replication and encapsidation are dependent on TCV-encoded RdRp and CP. Because satC and TCV share similar 3' ends (3' co-terminal 200-nt), satC has proven useful in identifying sequences and

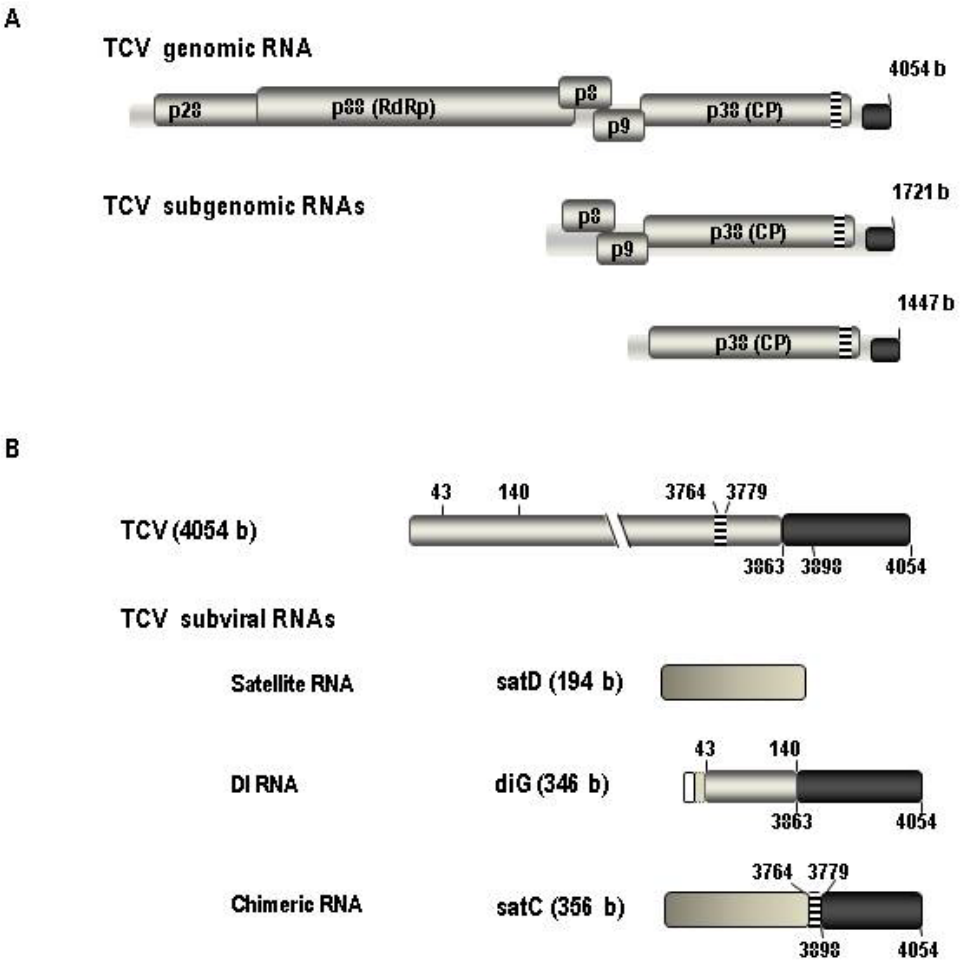


Figure 1.1 Genome organization of TCV: TCV genomic and subgenomic RNAs. The single (+)-strand genome of TCV and five open reading frames are shown. p28 and the readthrough protein p88 (the RdRp) are required for replication. p8 and p9 are required for cell-to-cell virus movement. (B) TCV genomic RNA and associated subviral RNAs. Similar regions are shaded alike. Numbers refer to positions of sequences in the TCV genome.

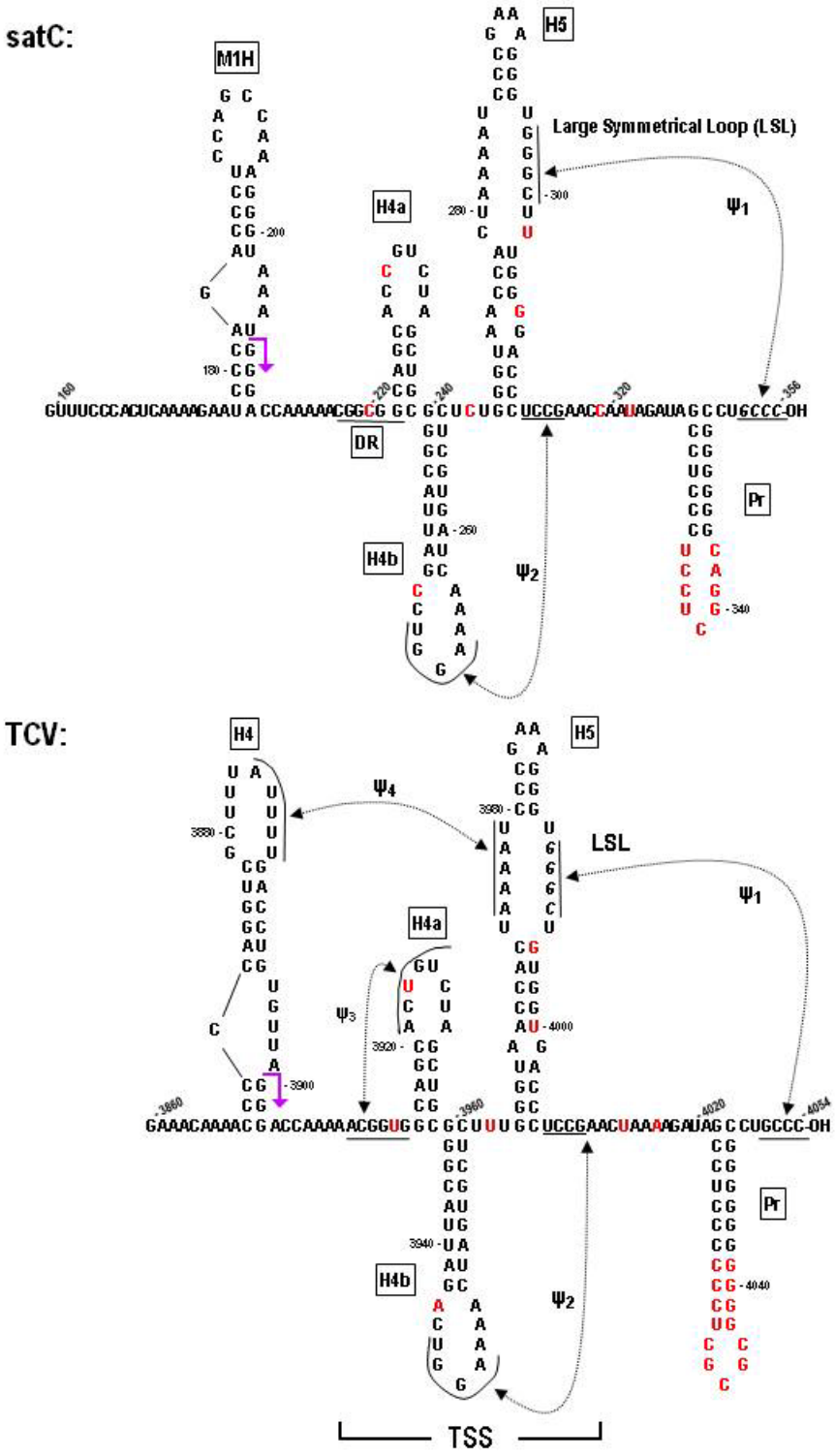

satC:
M1H
H5
H4a
H4b
DR
Pr
Large Symmetrical Loop (LSL)
Ψ1
Ψ2
TCV:
H4
H5
H4a
H4b
Pr
Ψ4
Ψ1
Ψ2
Ψ3
LSL
TSS

Figure 1.2 MPGAfold-predicted TCV and satC 3' terminal structures. Hairpins are described in the text. The region the folds into TSS in TCV is indicated. Ψ_1 and Ψ_4 are not compatible with efficient ribosome binding to the TSS and are likely found in the replication-competent form of TCV. Sequence differences between satC and TCV are shown in red. Although the satC structure is shown in a similar configuration as TCV for ease of comparison, none of the hairpins or Ψ_1 are present in the satC structure assumed by transcripts synthesized *in vitro*, while H5 and Ψ1 are present in the satC replication-active structure (Zhang et al., 2004). Arrowheads in H4 and M1H denote 5' end of shared 3' terminal sequences.

structures required for viral replication.

Structural comparison between satC and TCV

Carmoviruses have no 5' cap or 3' poly(A) tail, but terminate their genomes with short conserved single-stranded tails ending with hydroxyl groups (CCUGCCC-OH). Deletion of 6-13 nt at the 3' end of satD was repaired *in vivo* by recombination events (Carpenter and Simon, 1996a; Carpenter and Simon, 1996b). Up to a 6-nt deletion at the satC 3' end is repaired *in vivo* using TCV RdRp-generated abortive products as RNA primers (Nagy, Carpenter, and Simon, 1997).

Sequence in the 3' region shared by satC and TCV (94% sequence identity) is predicted by the RNA structure program MPGAfold (Shapiro et al., 2006; Shapiro and Wu, 1997) to form four hairpins (Figure 1.2) that are phylogenetically conserved among the more closely related carmoviruses and have been confirmed genetically (McCormack et al., 2008). The 3' terminal hairpin (Pr), found in 15 of 16 carmoviruses, is the core promoter for (-)-strand synthesis in satC (Song and Simon, 1995) and in TCV (Sun and Simon, 2006). Because of their sequence similarity, TCV-associated satC was used to identify *cis*-acting elements required for TCV replication. *In vitro* experiments were carried out using purified TCV RdRp, which transcribes (+)- and (-)-strand viral and subviral RNAs (Rajendran, Pogany, and Nagy, 2002; Song and Simon, 1994). Deletion assays and site-directed mutagenesis showed that the core promoter for satC maps to a 3'-terminal hairpin (5'-GCCUCCCUCCUCGGACGGGGGGC with underlined sequences base paired) both *in vitro* (Song and Simon, 1995) and *in vivo* (Stupina and Simon, 1997). When this hairpin was fused to a heterologous RNA, it was able to direct complementary

strand synthesis by TCV RdRp (Song and Simon, 1995). The importance of maintaining the stem-loop hairpin structure of satC core promoter was confirmed by *in vivo* genetic selection (Carpenter and Simon, 1998). *In vivo* genetic selection also showed that the hairpin has some variability in stem-loop size and composition (Carpenter and Simon, 1998).

In vitro RdRp assays suggested that satC (-)-strands had two promoter sequences, one (the 3' proximal promoter sequence, 3'PE, 3'-UCCCAAAGUAU) is located at the 3' end at position 11-21, while the other one (the 5' proximal promoter sequence, 5'PE, 3'-AACCCCUGGGAGGC) is at the 5' end and is contained within positions 302-315 (Guan, Song, and Simon, 1997). The 5'PE was necessary for satC replication in *vitro* and *in vivo*. *In vivo* genetic selection of the 5'PE revealed a strict sequence requirement for 10 out of the 14 nucleotides (Guan, Carpenter, and Simon, 2000b). It is now known that the 5'PE occupies a very important region in the (+)-strand, which may account for the conservation as it is part of a hairpin stem. Mutations introduced into the 3'PE had little or no effect on satC accumulation *in vivo* (Guan, Carpenter, and Simon, 2000a). *In vivo* genetic selection of the (-)-strand 3' terminal 21-nt sequence (including 3'PE) revealed that all satC species recovered from plants contained a sequence that is conserved in all carmoviruses 3'-$CC_{1\text{-}2}$(A/U)(A/U)(A/U). This Carmovirus consensus sequence (CCS) is required for satC replication in protoplasts and in plants, but it did not serve as a promoter *in vitro* (Guan, Carpenter, and Simon, 2000a).

Although satC and TCV share 94% sequence identity at the 3' terminus, satC with TCV 3' terminal 100 bases accumulates very poorly in plants and protoplasts (Wang and Simon, 2000). The weak accumulation was due to the differences in the Pr region (Zhang

et al., 2006b; Zhang et al., 2006c). *In vitro* experiments showed that the satC Pr is a much stronger promoter compared to the TCV Pr, and functions efficiently by itself *in vitro* (Zhang et al., 2006b). SatC Pr can tolerate a wide range of different sequences in its Pr loop (Carpenter and Simon, 1998; Song and Simon, 1995; Stupina and Simon, 1997), while the TCV Pr loop has more strict sequence requirement, and appears to interact with upstream elements in the 3' UTR (Yuan et al., 2010). Evolutionary adaptations appear to have allowed the satC Pr to function independently without any additional upstream elements thus accounting for its template strength *in vitro*.

The upstream hairpin H5 is conserved in 15 out of 16 carmoviruses. H5 contains a large symmetrical loop (LSL) that pairs with 3' terminal bases forming Ψ_1 (between the 3' side of the LSL GGGC and the 3' terminal GCCC-OH), with a similar pairing present in all carmoviruses. Ψ_1 is required for normal viral RNA accumulation *in vivo* (Zhang et al., 2006c) and satRNA transcription *in vitro* (Zhang, Zhang, and Simon, 2004). Deletion of the three cytidylates at the 3' terminus of satC resulted in a substantial increase in transcription of both full-length and aberrantly initiated products (Zhang, Zhang, and Simon, 2004). Sequences in the H5 LSL are required for satC accumulation in both protoplasts and plants (Zhang, Stuntz, and Simon, 2004). SatC with H5 sequences from the closely related carmoviruses *Japanese iris necrosis virus* (JINRV) or *Cardamine chlorotic fleck virus* (CCFV) did not replicate in protoplasts even though Ψ_1 was maintained and JINRV has identical H5 LSL sequence (Zhang and Simon, 2005). *In vivo* SELEX showed that the H5 upper stem has sequence plasticity, while the lower stem has both stringent sequence and structure requirement (Zhang and Simon, 2005). H5 is

thought to act as a 'RNA chaperone' in TCV to nucleate the formation of an active RdRp complex and help the RdRp fold correctly (McCormack and Simon, 2004).

H4b is conserved in some carmoviruses. H4b loop sequence UGGA was shown genetically to interact with sequence at the base of H5 (UCCG) forming Ψ_2. Ψ_2 is required for both viral (McCormack et al., 2008) and satC (Zhang et al., 2006b) accumulation in protoplasts. H4a is only conserved in the most related carmovirus, CCFV. In TCV, H4a loop sequence ACUGU can pair with H4a upstream flanking sequence ACGGU, forming Ψ_3, which is required for TCV accumulation (McCormack et al., 2008). As I will show in Chapter III, although a similar pseudoknot is predicted in satC by MPGAfold, genetic assays did not support Ψ_3 in satC (Guo et al., 2009).

In addition to a role in suppressing virion accumulation, H4a upstream flanking sequence 217 CGGCGG (DR) is involved in many functions such as fitness of satC in plants; high-level accumulation in protoplasts; and efficient transcription of satC transcripts by the TCV RdRp *in vitro* (Sun, Zhang, and Simon, 2005; Zhang et al., 2006a; Zhang et al., 2006b). The DR is also important for a conformational switch in satC (discussed below).

TCV H5 and the two upstream hairpins (H4a, H4b) along with two associated pseudoknots (Ψ_2 and Ψ_3) were predicted by RNA structural modeling protocol RNA2D3D to fold into a tRNA-shaped structure (TSS) (McCormack et al., 2008), which was experimentally confirmed by NMR-small angle X-ray scattering (Zuo et al., 2010). The TSS binds to the ribosome P-site through the 60S subunit in the absence of translation factors and serves as part of the TCV translational enhancer (Stupina et al., 2008). The TSS is also a stable scaffold for interactions with the surrounding sequences

through the H5 LSL (Yuan et al., 2009). SatC contains a similar 3' end with 6-nt differences in the 100-nt TSS region. RNA2D3D did not predict a similar structure for the satC TSS analogous region (Stupina et al., 2008).

Just upstream of the shared regions of TCV and satC, is motif 1 hairpin (M1H) in satC, and H4 in TCV. M1H consists of sequences derived from satD and TCV, and is required for satC accumulation (Nagy and Simon, 1998). M1H is a hot spot for recombination between satC and satD during (+)-strand synthesis *in vivo* (Cascone, Haydar, and Simon, 1993). In its (-)-sense orientation, M1H is also a replication enhancer both *in vivo* and *in vitro* (Nagy, Pogany, and Simon, 1999; Nagy, Pogany, and Simon, 2001). *In vivo* SELEX assays in which M1H sequences were randomized and selected by TCV RdRp in plants showed that most recovered satC progeny contained short motifs in (-)-strands that show strong similarity to sequences found in TCV and satC promoter elements (TCV 3' CCS or satC 5' PE) (Sun and Simon, 2003; Zhang and Simon, 2003b). M1H, which cannot function in locations other than its wt position (Nagy, Pogany, and Simon, 1999), has no sequence specificity but the structure is required for satC fitness in plants (Sun and Simon, 2003; Zhang and Simon, 2003b). Sequence non-specific M1H in the wt location appears to function by bridging the flanking sequences. These flanking sequences contain three motifs, revealed by *in vivo* SELEX: 164 CCCA, 211 CAAAA and 217 CGGCGG (Sun, Zhang, and Simon, 2005). By 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, Ren, and Morris, 2003; Zhang and Simon, 2003a).

In TCV, the DR analogous sequence (CGGUGG) interacts with H4a to form Ψ_3, which structurally mimics a tRNA amino-acid acceptor stem. The Ψ_3-H4a region is important for ribosome binding to TCV TSS (Stupina et al., 2008). Disruption of the TSS by altering Ψ_3 reduced translation and virus accumulation *in vivo* (McCormack et al., 2008; Stupina et al., 2008). The DR region might be a RdRp binding site in TCV based on the loss of residue flexibility by in-line probing upon RdRp binding (Yuan et al., 2009).

Conformational changes

As mentioned above, translation (5' to 3' end) and replication (3' to 5' end) are mutually exclusive processes that cannot occur simultaneously on the same strand (Barton, Morasco, and Flanegan, 1999; Gamarnik and Andino, 1998). It has been proposed that (+)-strand RNA must switch from a conformation that favors translation to a conformation that favors replication. RNA conformational changes are known to control switches between translation and replication (Olsthoorn et al., 1999), transcription and replication, and (+)- strand and (-)- strand synthesis (Goebel et al., 2004; Koev et al., 2002; Na and White, 2006; Pogany et al., 2003; Zhang et al., 2006a; Zhang, Zhang, and Simon, 2004).

The in-line cleavage pattern of TCV 3' end fragments changes significantly after RdRp binding, causing both enhanced and reduced cleavages at TCV 3' end. The altered susceptibility to in-line cleavage in the presence of the TCV RdRp indicates that RdRp interaction with the TCV 3' end caused a dramatic structural change. RdRp binding to the 3' end disrupts structures required for 60S ribosome subunit binding, likely prohibiting

ribosomes access to the RNA. TCV replication is thus proposed to be initiated by a structural switch mediated by RdRp binding to an upstream hairpin (H4), causing loss of translation-favored TSS conformation, freeing H5 to form replication-required Ψ_1 (Yuan et al., 2009).

Since satC encodes no ORF, there is no need to switch between translation and transcription conformations. However, satC still has at least two conformations (Simon and Gehrke, 2009). Hairpin exchanges with CCFV and *in vivo* functional selection of satC with various regions randomized showed that satC can form TCV-like hairpin structures within the TCV-derived region (Carpenter and Simon, 1998; Guo et al., 2009; Zhang and Simon, 2005; Zhang, Stuntz, and Simon, 2004; Zhang et al., 2006b). However, structure probing of satC did not support the presence of Ψ_1 or any of the TCV-related hairpins, which indicated that the initial structure of satC transcripts was different from the structure that was required *in vivo* (Zhang, Zhang, and Simon, 2004). These transcripts used for structure probing were not metastable misfolded, kinetically-trapped intermediates because the RNAs were treated using different mono/divalent ion concentrations and different heating/cooling methods and still folded into one detectable conformation (Zhang et al., 2006a). This structure, termed the pre-active structure, contains Ψ_2, an alternative Pr and is not an active template for transcription. Releasing the 3' end by mutating several regions of satC H5 or deleting 3' end caused structural rearrangements in the Pr, H4a and DR that correlated with enhanced transcription *in vitro* (Zhang, Zhang, and Simon, 2004). This structure, termed the active structure, contains the conserved H5, Pr and is active for (-)-strand synthesis. It has been proposed that newly transcribed RNAs need to switch from the pre-active structure to the active

structure to start complementary strand synthesis. Mutations in the DR prevented the activation of satC templates, decreased satC accumulation *in vivo* and substantially reduced transcription *in vitro* (Zhang, Zhang, and Simon, 2004). However, mutations in the DR did not inhibit satC transcription when the DR mutations were combined with alterations that shifted the structure to the active form *in vitro* (Zhang et al., 2006a). DR mutations also had a reduced effect when combined with alterations that stabilized the active form *in vivo* (Zhang et al., 2006b). This suggests that the DR is important for implementing the conformational switch from the pre-active to the active structure, but with reduced importance if RNAs initially assume the active structure.

Thesis plan

This thesis investigates differences in *cis*-acting elements involved in satC replication compared to TCV replication. Chapter II discusses the differences between satC and TCV in the TSS region, and the reasons for satC evolving away from the TCV sequence. In Chapter III, the importance of the linker sequence between H5 and Pr is described. The structure of H4a and H4b in satC and their relationship will be discussed in Chapter IV. Chapter V shows the occurence of satC dimers in the absence of TCV RdRp, and the possible mechanism of dimer formation is discussed.

CHAPTER II

The difference between satC and TCV in the TSS region

Introduction

DI RNAs and other subviral RNAs that share substantial sequence similarity with their helper viruses are frequently used as models with results extrapolated to their helper viruses. However, subtle sequence differences in satC appear to reflect a need to alter the function of genomic RNA elements removed from helpful but not essential interacting sequences and to generate a local pre-active structure that differs from one that may incorporate translational activity. Both satC with TCV 3' 100 bases and TCV with satC 3' 100 bases accumulated poorly in plants and protoplasts (Wang and Simon, 2000) with the heterologous bases that caused the down regulation of replication mapped to the Pr (Zhang et al., 2006c). *In vitro* experiments demonstrated that the satC Pr is stronger than the TCV Pr, which implies that satC evolved a better Pr in the absence of upstream TCV sequences (Zhang et al., 2006c). Besides the Pr, we wanted to determine what other structural and functional differences satC evolved to better use the TCV-derived region in the absence of the rest of the TCV sequence.

As mentioned in Chapter I, in TCV, H5 and the two upstream hairpins (H4a, H4b) along with two associated pseudoknots (Ψ_2 and Ψ_3) were predicted by RNA structural modeling protocol RNA2D3D to fold into TSS (McCormack et al., 2008) and the computer prediction was experimentally confirmed by NMR-small angle X-ray scattering

(Zuo et al., 2010). The TSS binds to the ribosome P-site through the 60S subunit in the absence of translation factors and serves as part of the TCV translational enhancer (Stupina et al., 2008). The TSS is also a stable scaffold for interactions with the surrounding sequences (Yuan et al., 2009).

SatC contains 6-nt differences within the comparable TSS region: two uridylate-to-cytidylate changes in the Ψ_3–H4a region (position 220 and 229 in satC), one adenylate-to-cytidylate change in H4b loop (position 249 in satC), one uridylate-to-cytidylate change in the linker sequence between H4b and H5 (position 268 in satC), and one guanylate-to-uridylate change and one uridylate-to-guanylate change in H5 (position 302 and 306 in satC). Structure modeling did not support a TSS in satC with these 6-nt differences (Stupina et al., 2008). Ribosome binding to isolated satC TSS was not detected (Stupina et al., 2008) and binding was reduced by 27-fold when assaying full-length satC compared with the TCV TSS. Here I analyzed how the 6-nt differences between satC and TCV in the TSS region affect ribosome binding in satC. Stepwise mutations in the satC TSS analogous region that substantially enhanced ribosome binding surprisingly had little effect on satRNA replication, suggesting that satC evolved to eliminate the TSS for reasons other than reducing ribosome binding.

Materials and Methods

Construction of satC mutants

To generate plasmid C220U, C229U, C220U/C229U and C249A, PCR reactions were performed with a common 5' primer (T7C5') and analogous 3' primers using

template pT7C+ (pUC19 containing full-length satC cDNA downstream of T7 promoter) (see Table 2.1 and 2.2). SpeI and NcoI digested PCR products were ligated into similarly digested pT7C+, replacing the endogenous fragment.

To construct plasmid C220U/C229U/C249A, PCR products using template pC220U/C229U with primers T7C5' and C249A were digested with SpeI and NcoI, then inserted into similarly treated analogous region of pT7C+. To construct plasmid C249A/C268U and C220U/C229U/C249A/C268U, PCR reactions were performed with primers T7C5' and C249A using template pT7C+ or pC220U/C229U. SpeI and NcoI digested PCR products were inserted to analogous region of plasmid C268U. All constructs were identified by DNA sequencing.

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

Linearized DNA template (8 μg) or PCR products were mixed with 6 μL of dithiothreitol (DTT), 12 μL of 5 mM each of ATP, GTP, CTP, UTP, 12 μL of 5X T7 buffer (125 mM NaCl, 40 mM $MgCl_2$, 2mM spermidine, 40 mM Tris-HCl, pH8.0), 60 units of RNase OUT (RNase inhibitor from Invitrogen, for genomic RNA only), 80 units of T7 RNA polymerase, and distilled H_2O to give a final volume of 60 μl. The mixture was incubated at 37^{o}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_4OAc (pH5.3)/isopropanol (1:5), and washed with 70% ethonal.

Plus-strand transcripts of TCV were synthesized using T7 RNA polymerase from *Sma*I-linearized pT7TCVms, which contains a T7 RNA polymerase promoter upstream of TCV full-length plus-strand sequence (Oh et al., 1995; Table 2.1). 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).

Small-scale plasmid DNA isolation (Alkaline Lysis)

Bacterial cells (*E. coli*) from 1.5 mL of overnight culture were collected in an eppendorf tube by centrifugation at 13,000 rpm for 1 min. The pellet was resuspended in 1 volume (200 µL) Solution I (0.9 % glucose, 25 mM Tris [pH 8], 10 mM EDTA [pH 8], 0.01% RNaseA). After adding 1 volume of solution II (2-4 % SDS/0.1 M NaOH), gently invert the tube several times to mix, and incubate the tube at room temperature for 5 min. Add 0.75 volume of solution III (3 M KOAc, pH 4.8) to the tube and vigorously invert the tubes for several time, then spin the tube at 13,000 rpm for 5 min at room temperature. The supernatant was transferred to a new eppendorf tube, extracted with equal volume of phenol/chloroform. The aqueous layer was precipitated by adding 1/10 volume of 3M NaOAc (pH 5.3) and 2.5 volume of 100% ethanol, incubated at −80°C for 1 h our and centrifuged at 13,000 rpm for 15 min at 4°C. The pellet was washed with 70% ethanol, air dried, re-suspended in 50 µl of sterile distilled water.

Small-scale plasmid DNA isolation (STET)

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 µL 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 min, and centrifuged at 13,000 for 14 min at room temperature. The pellet of cell debris was removed with a toothpick. The remaining solution was mixed with 140 µl of isopropyl alcohol, incubated on ice for 5 min, 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 µL of distilled H_2O.

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 MS agar plates (Murashige-Skoog salts, 4.3g/L, 3% sucrose, 1X vitamins/glycine, 0.5 mg/L kinetin, 0.5 mg/L 2,4-D, 1% agar, pH 5.8). 1X vitamins/glycine contains 1 mg/L nicotinic acid, 10 mg/L thiamine HCl, 1 mg/L pyridoxine, 0.1 mg/L myoinositol, and 4 mg/L glycine. The calli were incubated in a growth chamber at 20°C with a cycle of 16-hour light and 8-hour dark. Cultures were passaged every 21 days.

To prepare protoplasts, calli in the 4^{th}, 5^{th} or 6^{th} passage were collected and soaked in 40 ml of 0.6 M mannitol at room temperature 25°C for 20 min with shaking. The calli were recovered by centrifugation (Beckman GPR-type swinging bucket) at 2000 rpm for 5 min at room temperature, and then suspended in 50 ml of freshly prepared 0.5 g of cellulase (11,900 U/g) and 0.1 g of pectinase (3,140 U/g) (Calbiochem, La Jolla, CA) dissolved in protoplast isolation medium (PIM, 4.3 g/L MS salts, 0.1 M sucrose, 3mM MES, 0.5M mannitol, 5mM $CaCl_2$, 1 mg/L thiamine-HCl, 0.5 mg/L pyridoxine-HCl, 0.5 mg/L nicotinic acid, and 0.1 g/L myo-inositol, 0.2 mg/L 2,4-D, 0.2 mg/L kinetin, and 25

mM KOH, pH 5.8). The calli/PIM/enzyme mixture was incubated at 26°C in dark for 4 hours with shaking at 100 rpm.

The turbid solution was filtered through a 53-µm nylon mesh (Small Parts), followed by centrifugation at 1000 rpm for 5 min at 4°C. The precipitated protoplasts were washed 3 times with 20 ml of cold 0.6 M mannitol. The number of cells was calculated using a microscope and hemocytometer. Protoplasts were divided into 50 ml centrifuge tubes so that each tube contained 5 x 10^6 cells, and collected by centrifugation at 1000 rpm for 5 min at 4°C.

Protoplasts (in a volume of about 100 µL) were swirled well with 20 µL of TCV genomic RNA transcripts (1µg/µL), 2 µl of satRNAs (1µg/µL), 8 µl of 1 M $CaCl_2$, 400 µL of distilled H_2O, 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 sec, followed by addition of cold 0.6 M mannitol/1 mM $CaCl_2$ and incubation on ice for 15 minutes. The protoplasts were collected by centrifugation at 1000 rpm for 5 minutes at 4°C. After washing 3 times with 20 mL of cold 0.6 M mannitol/1 mM $CaCl_2$, protoplasts were resuspended in protoplast culture medium (PCM, 4.3 g/L MS salts, 0.1 M sucrose, 3mM MES, 0.4M mannitol, 1 mg/L thiamine-HCl, 0.5 mg/L pyridoxine-HCl, 0.5 mg/L nicotinic acid, and 0.1 g/L myo-inositol, 0.2 mg/L 2,4-D, 0.2 mg/L kinetin, and 25 mM KOH, pH 5.8) and incubated at 25°C for 40 hours in the dark.

Extraction of total RNA from Arabidopsis protoplasts

1 ml of Protoplasts (~1.67 x 10^6 cells) at 40 hours post inoculation (hpi) were collected in an eppendorf tube by centrifugation at 3000 rpm for 3 min at room temperature, and re-suspended 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 min 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 re-suspended in 20 μL of nuclease-free water.

Northern blotting using RNA gels

4 μg of total RNA extracted from protoplasts was combined with an equal amount of 2X formamide loading buffer (prepared by mixing 800 μL of formamide with 200 μL of 10X formaldehyde gel-loading buffer (50% (v/v) glycerol, 1 mM EDTA, pH 8.0, 0.25% bromophenol blue, 0.25% xylene cyanol FF)). The mixture was heat denatured at 65°C for 5 min, quenched on ice, and subjected to electrophoresis through a freshly prepared 1.2% nondenaturing agarose gel. The gel was then soaked completely in 6% formaldehyde solution with gentle shaking for 1 hour to denature any double stranded region. The formaldehyde was decanted and the gel was soaked in 10X SSC containing 0.15 M NaCl and 0.015 M sodium citrate for 25 minutes. RNA was transferred to a Nitrocellulose membrane for overnight by traditional paper towel method (the capillary transfer method). The membrane was then placed, face-down, on an ultraviolet light box (310 nm, Fotodyne Inc.) for 2 min and dried at 80°C for 5 min.

The membrane was pre-hybridized for at least 1 hr at 42°C using a 30% (v/v) formamide prehybridization buffer. Pre-hybridization buffer contains 5X SSPE (20X SSPE [pH 7.4]: 3 M NaCl, 0.2M NaH_2PO_4, 0.02 M EDTA), 10X Denhardt's reagent (50X Denhardt's reagent: 1% Ficoll, 1% Polyvinylpyrrolidone, 1% Bovine Serum

Albumin), 0.2% SDS, 0.2 mg/mL freshly denatured salmon sperm DNA, and 50% formamide (every 1% of formamide reduces the Tm by 0.7 °C).

During the prehybridization incubation, oligonucleotides were labeled with [γ-^{32}P] ATP using T4 polynucleotide kinase (10,000 U/mL; New England Biolabs) according to the manufacturer's protocol. The kinase reaction was incubated for 30 min at 37°C. The radiolabeled oligonucleotide was added to the prehybridization buffer and incubated for at least 2 hours at 42°C. After hybridization, the blot was washed in a high salt solution containing 6X SSPE and 0.1% SDS for 10 minutes, then washed in a low salt solution containing 0.1X SSPE and 0.1% SDS for 15 minutes at 42°C. The membrane was dried at room temperature for 5 min, covered with a plastic wrap, and subjected to autoradiography performed at -80°C for at least 12 hours.

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) and satC transcripts (0.015 μg/μL). The inoculation buffer contains 0.05 M glycine, 0.03 M K_2HPO_4, pH 9.2, 1% cleaned bentonite, and 1% Celite.

Competition in Plants

5 μg of T7 RNA polymerase-generated *in vitro* transcripts of wt satC, mutants C249A, C268U, C220U/C229U, C220U/C229U/C249A/C268U and 2 μg of wt TCV transcripts were co-inoculated onto 6 turnip seedlings. Total RNA was extracted from un-inoculated leaves at 21 dpi. Full-length satC amplified by reverse transcription-PCR (RT-

PCR) with primers T7C5' and oligo 7 were cloned into the SmaI site of pUC19 and then sequenced.

Small-scale viral RNA extraction from infected turnip plants

Total RNA was isolated from the un-inoculated leaves 21 dpi. Briefly, 0.5 g leaf tissue was ground to a fine powder in liquid nitrogen, then each sample was transferred to a 1.5 mL Eppendorf tube and extracted with 0.55 mL of RNA extraction buffer [25 mM EDTA, 0.4 M LiCl, 1% (w/v) SDS, 0.2 M Tris-HCl, pH 9.0] and 0.55 mL of H_2O-saturated phenol/chloroform (1:1 v/v) by vigorously vortexing. The sample was centrifuged at 13,000 rpm for 5 min at 4°C. The aqueous layer was collected into a fresh microfuge tube and re-extracted with equal volume of phenol/chloroform. Total RNA was precipitated by adding 1/10 volume of 3M NaOAc (pH 5.3) and 2.5 volume of 100% ethanol. The sample was centrifuged at 13,000 rpm for 15 min at 4°C and the pellet was re-suspended in 0.3 mL of 2 M LiCl. Following centrifugation at 13,000 rpm for 5 min at 4°C, the pellet was dissolved in 0.3 mL distilled H_2O and precipitated with 1/10 volume of 3M NaOAc (pH 5.3) and 2.5 volume of 100% ethanol again. The tubes were kept at −80°C for 2 h or overnight and centrifuged at 13,000 rpm for 15 min at 4°C. The RNA pellet was washed with 70% ethanol, air dried, re-suspended in 50 µL of sterile distilled water, and stored at −80°C. RNA concentration was estimated by measuring the absorbance of the sample at 260 nm with a spectrophotometer.

Cloning of viral progeny into pUC19

Total RNA was reverse transcribed by M-MLV reverse transcriptase (New England Biolabs). To generate first strand cDNA, 5 µg of total RNA was mixed with 10 pmol of oligo 7 (Table 2.2), the reverse transcription buffer provided by the manufacturer,

dNTP, and DTT to a final concentration of 0.5 mM and 10 mM respectively. The mixture was heated at 75°C for 10 min and cooled briefly on bench top for primer and RNA annealing. The solution was then incubated at 37°C for 1 hour with 1 μL of M-MLV reverse transcriptase (USB, 200 U/μL). The reaction was stopped by heating at 75°C for 15 min and placed immediately on ice.

SatC molecules were amplified by PCR in the presence of Pyrostase polymerase (Molecular Genetic Resources, Tampa, FL). PCR was performed using 8 μL of cDNA synthesized as described above with primers T7C5' and oligo7 in a 100 μL reaction. Immediately after PCR, 1 μL Klenow was added to the reaction with 10 min incubation at room temperature to blunt-end treat the PCR products. The PCR mixture was extracted with equal volume of phenol/chloroform. The aqueous layer was precipitated by adding 1/10 volume of 3M NaOAC (pH 5.3) and 2.5 volume of 100% ethanol, incubated at −80°C for 1 hour and centrifuged at 13,000 rpm for 15 min at 4 °C. The pellet was washed with 70% ethanol, air dried, re-suspended in desired volume of sterile distilled water. The purified PCR products were treated with T4 Polynucleotide kinase in the presence of T4 DNA ligase buffer, then precipitated with 2.5 volume of NaOAC (pH 5.3)/ethanol (3:50). The pellet was washed with 70% ethanol, air dried, re-suspended in desired volume of sterile distilled water, and ligated with SmaI-linearized and CIP treated pUC19 vector, transformed into *E. coli* DH5α, plated into LB plates containing ampcillin and X-gal. The white colonies were randomly picked, extracted, confirmed with PVUII digestion, and subjected to sequencing.

RNA in-line probing

In-line probing was performed essentially as previously described (Yuan et al., 2009).

Briefly, full length wt satC and mutant C249A, C268U, C220U/C229U/C249A/C268U transcripts were synthesized using T7 RNA polymerase, and then purified by agarose gel electrophoresis. RNAs were either treated with Antarctic phosphatase, then T4 PNK in the presence of [γ-^{32}P] ATP, or ligated with [α-^{32}P] pCp and T4 RNA ligase. 5% polyacrylamide gel-purified radiolabeled RNAs were heated to 95°C for 5 min and snap cooled on ice for 3 min. RNA (20 pmol) was then incubated at room temperature in 50 mM Tris-HCl (pH 8.5), 20 mM $MgCl_2$, 100 mM NaCl for 14 h. Reactions were then inactivated and precipitated using inactivation buffer from Ambion. The pellets were then dissolved in 10 μL RNA loading buffer (Ambion). 1 μL samples were heat denatured at 95°C for 2 min, and subjected to electrophoresis through sequencing length 8% denaturing polyacrylamide gels containing 8 M urea.

Table 2.1 SatC constructs used in Chapter II

Name	Description
pT7C(+)	T7 promoter and full length satC cloned into pUC19
pTCVms	Full length TCV cloned into pUC19
G219C	G to C change at position 219
C228G	C to G change at position 228
G219C/C228G	G to C change at position 219 and a C to G change at position 228
Δ216-220	SatC with a 5-nt ACGGC deletion at positions 216-220
C229G/Δ216-220	C to G change at position 229 and a 5-nt ACGGC deletion at positions 216-220
C229G/M216-219	C to G change at position 229 and a GCGG to UGCC change at positions 216-219
C220U	C to U change at position 220
C229U	C to U change at position 229
C220U/C229U	C to U change at position 220 and a C to U change at position 229
C249A	C to A change at position 249
C268U	C to U change at position 268
C220U/C229U/C249A/C268U	C to U change at position 220,C to U change at position 229, C to A change at position 249, and C to U change at position 268

Table 2.2 Oligonucleotides used in Chapter II

Name	Position	Sequence	Polarity
T7C5'	1-19	GTAATACGACTCACTATA GGGATAACTAAGGGTTTCA	+
Oligo7	338-356	GGGCAGGCCCCCCGTCCGA	-
Oligo13	249-269	GTTACCCAAAGAGCACTAGTT	-
G219C	207-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACGGTG CTGCCGGCGTTTTTGGTCC	-
C228G	214-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACGCTG CTGCCGCCGTTT	-
G219C/C228G	207-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACGCTG CTGCCGGCGTTTTTGGTCC	-
Δ216-220	200-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACGGTG CTGCCTTTTGGTCCCATTTAC	-
C229G/Δ216-220	200-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACCGTG CTGCCTTTTGGTCCCATTTAC	-
C229G/M216-219	207-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACCGTG CTGGGCACGTTTTTGGTCC	-
C220U	205-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACGGTG CTGCCACCGTTTTTGGTCCCA	-
C229U	214-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACAGTG CTGCCGCCGTTT	-
C220U/C229U	205-264	GC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTAGACAGTG CTGCCACCGTTTTTGGTCCCA	-
C249A	235-264	GC*ACTAGT*TTTCCAGTCTAATGCCCGCAGC	-

Underlined nucleotides indicate the mutations.

Italics are restriction enzyme recognition sites.

Results

Ψ_3 is not required for satC accumulation in protoplasts.

TCV TSS is composed of three hairpins and two pseudoknots, one of which, Ψ_3, is critical for TCV translation (McCormack et al., 2008) and also involved in RdRp binding to the TCV 3' end *in vitro* (Young and Simon, unpublished). Whereas RNA2D3D combined with molecular modeling did not support the formation of Ψ_3 in satC (Stupina et al., 2008), MPGAfold predicted the existence of a pseudoknot that differed slightly from TCV Ψ_3 and formed from the terminal loop of an H4a-like hairpin and upstream sequences. The H4a-like hairpin contained a 3 bp stem capped by a 9-nt loop, with four loop residues predicted to pair with upstream DR sequences. Genetic analysis did not support this pseudoknot (J. Zhang, doctoral dissertation, 2006). Visual inspection also revealed a second possible conformation of the pseudoknot that was more similar to TCV Ψ_3, with a TCV-like H4a forming a 4-nt interaction with the DR sequence. RNA biochemical structure mapping of the DR region in full-length satC transcripts revealed only faint susceptibility to single-stranded specific enzymes, suggesting that the DR is located in a paired or stacked configuration (Zhang et al., 2006a), which is consistent with both structures.

To examine the second possibility, single and potential partner exchanges were introduced into the hairpin and upstream flanking sequence. The accumulation levels of mutant satC were compared with that of wt satC at 40 hpi in *Arabidopsis* protoplasts. One set of single and dual exchanges were introduced into the region (J. Zhang, doctoral

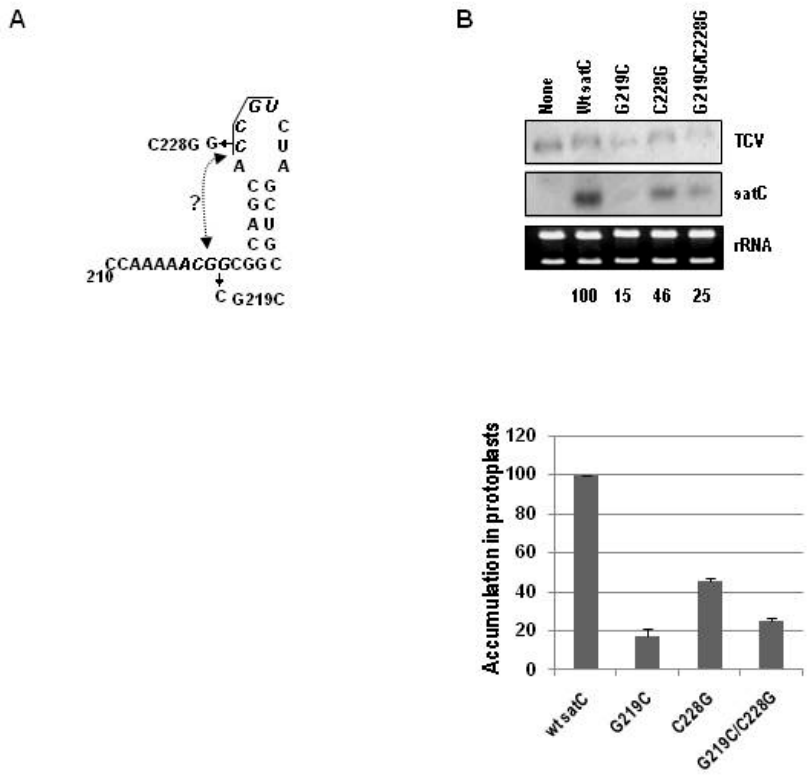


Figure 2.1 Analysis of a predicted pseudoknot between H4a and its 5' side flanking sequence (DR, CGGCGG). (A) Structures of H4a and its 5' side flanking sequence predicted by MPGAfold computer programs. Mutations generated in putative base-paired partners are shown. (B) RNA gel blot of viral RNAs accumulating in protoplasts at 40 hpi. TCV, TCV genomic RNA (+)-strand. satC, satC (+)-strand monomer. None, no satC in the inoculum. wt C, wild-type satC. Ethidium bromide staining of the gel before blotting shows rRNA loading control (panel below the blot). Values given below the blots are the averages of at least two independent experiments, with the satC level arbitrarily assigned a value of 100. None, no satC in the inoculum. wt C, wild-type satC.

dissertation, 2006). Alteration of position 218 (G218C) reduced satC accumulation to 28% of wt. The putative partner mutation, C229G, did not substantially affect satC levels (87% of wt) in protoplasts. Interestingly, combining the two alterations (G218C/C229G) restored satC accumulation to 96% of wt levels, suggesting that C229G was able to compensate for the low accumulation of G218C (J. Zhang, doctoral dissertation, 2006). To rule out a TCV-like Ψ_3, I made another set of mutants (Figure 2.1A). G219C reduced satC accumulation to 15% of wt while the putative partner, C228G, reduced satC accumulation to 45% of wt. SatC containing the combined alterations accumulated to 25% of wt (Figure 2.1B). Therefore, I conclude either that the TCV-like Ψ_3 does not exist in satC or that sequence specificity is important in the region.

DR has some flexibility in its location.

One possible interpretation of the compensatory nature of C229G when combined with G218C was that the negative effect of the G218C alteration in the DR region was compensated by a fortuitous generation of a sequence with properties of the DR sequence (CGG) in the loop of H4a when C229 was converted to a guanylate. To test this interpretation, I assayed if the C229G alteration could compensate for other modifications in the DR that together would not be compensatory for re-establishment of a pseudoknot (Figure 2.2). The DR modifications tested were an alteration of 4 nt (m216-219; ACGG to UGCC), which reduced satC accumulation to 5%, and a construct with a deletion of 5 nt (ACGGC, Δ216-220), which reduced satC accumulation to 22% of wt. C229G restored accumulation to 45% of wt when combined with m216-219 and 75% of wt when

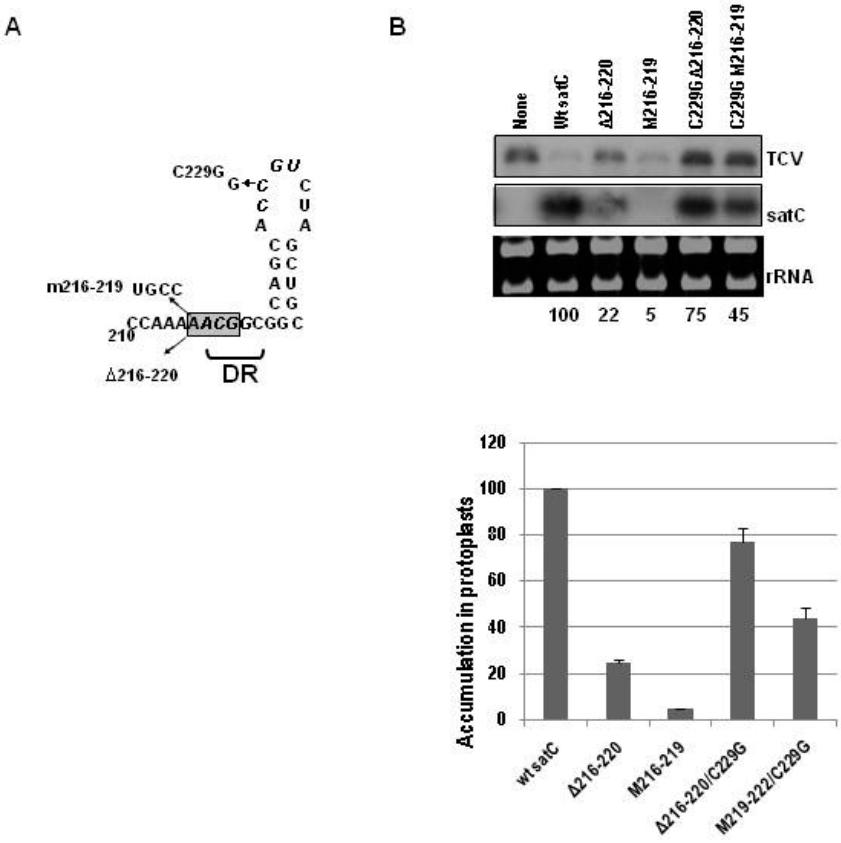


Figure 2.2 Analysis of the position flexibility of DR. (A) Structures of H4a and its 5' side flanking sequence predicted by MPGAfold computer programs. Mutations generated in DR and combined with C229G were shown. Italic letters indicate these 4 nt were altered. The bracket indicates the location of the 5-base deletion. (B) RNA gel blot of viral RNAs accumulating in protoplasts at 40 hpi. TCV, TCV genomic RNA (+)-strand. satC, satC (+)-strand monomer. None, no satC in the inoculum. wt C, wild-type satC. Ethidium bromide staining of the gel before blotting shows rRNA loading control (panel below the blot). Values given below the blots are the averages of at least two independent experiments, with the satC level arbitrarily assigned a value of 100. None, no satC in the inoculum. wt C, wild-type satC.

combined with Δ216-220. These results strongly suggest that generation of a new, partially functional DR element downstream from the original element can compensate in part for DR alterations, indicating that the DR has some flexibility in its location.

Altogether, these data do not support an important pseudoknot in this location in satC. Mutations in the DR that reduced satC accumulation had strong negative effects on transcription of full-length satC by the TCV RdRp *in vitro*, indicating its importance for activation of the pre-active conformation of the satRNA (Zhang et al., 2004a). This information, combined with our results above indicating that the DR is not base paired in Ψ_3 as is the comparable sequence in TCV, indicates that the DR is likely interacting in an alterative location in the satC pre-active structure.

Ribosome binding was enhanced in satC mutants.

The TCV TSS functionally mimics a tRNA by binding to yeast 80S ribosomes with a K_d of 0.45×10^{-6}M. Binding is to the P-site with no requirement for translation initiation factors (McCormack et al., 2008; Stupina et al., 2008). Ribosome binding was correlated with enhancement of translation. Disruption of the TSS by altering Ψ_3 reduced translation and virus accumulation in vivo (Stupina et al., 2008). The satC TSS analogous region (positions 216-315) was an inefficient ribosome-binding template ($K_d=37\times10^{-6}$M) (Stupina et al., 2008), while full length satC bound ribosomes with a K_d of 12.14×10^{-6}M, about 27-fold weaker than ribosome binding to the TCV TSS (Figure 2.3B). Since full-length satC gave detectable levels of binding, all subsequent binding assays were conducted using full-length satC. Ribosome binding was performed by Dr. Arturas Meskauskas.

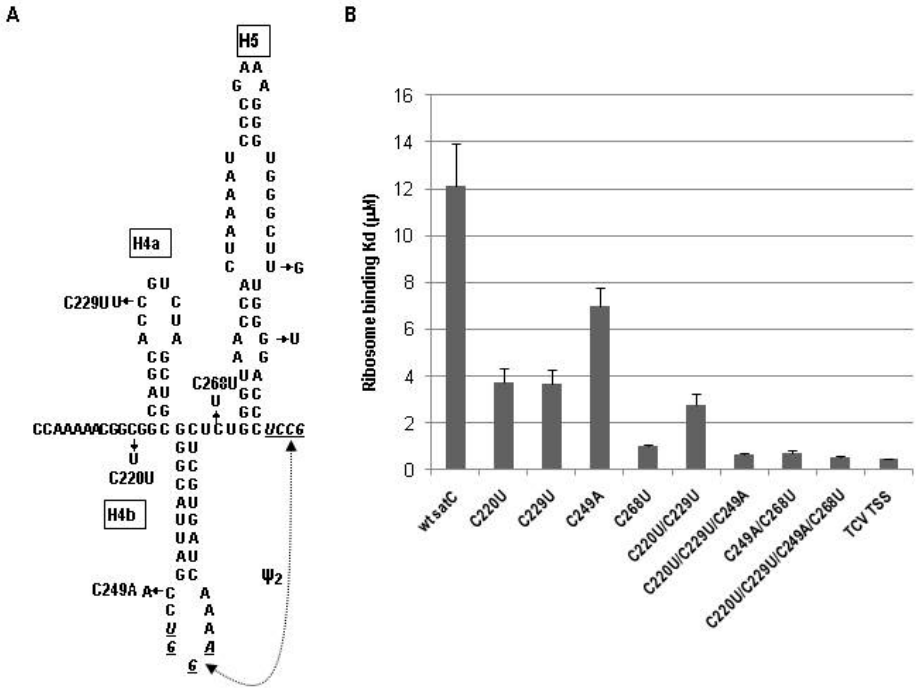


Figure 2.3 Analysis of satC mutants containing TCV TSS sequence. (A) SatC TSS analogous region with arrows showing the nucleotides that differ from TCV. The four nucleotides outside of H5 were step-wise changed into TCV analogous sequences. (B) SatC mutants were subjected to 80S ribosome binding using a filter binding assay.

Of the 6-nt differences between TCV and satC in the TSS region (Fig. 2.3A), four were postulated to have a greater role in disrupting the TSS and reducing ribosome binding: C229U and C220U in the equivalent Ψ_3 region; C268U in the H4b-H5 linker; and C249A in the loop of H4b. The two alterations in H5 were thought less likely to substantially affect binding since ribosomes can bind to the TSS in the absence of H5 (Stupina et al., 2008). To determine which base changes are responsible for reduced ribosome binding, the four differences outside of H5 were converted either individually or in combination to their TCV counterparts (Figure 2.3A). Ψ_3–H4a in TCV TSS correlates to the amino-acceptor arm of a tRNA, and is critical for ribosome binding (Stupina et al., 2008). Altering individual residues in the Ψ_3 region in satC (C220U and C229U) enhanced ribosome binding by 3.2-fold (Figure 2.3B). Unexpectedly, combining the two alterations to generate TCV-equivalent H4a and Ψ_3 reduced ribosome binding by 1.9-fold compared with the individual changes, although binding improved 1.7-fold over wt satC level (Table 2.3). These results indicate that the conversion to TCV-identical sequences in the Ψ_3 region does not necessarily result in a substantial improvement in ribosome binding.

Surprisingly, a major enhancement in ribosome binding (12-fold) occurred when satC contained the single change in the H4b loop (C249A) (K_d=0.99x10^{-6}M). Combining the C249A alteration with C220U/C229U further improved ribosome binding (K_d=0.73x10^{-6}M) while all four exchanges improved binding to near TCV TSS level (K_d=0.55×10^{-6}M; TCV TSS: K_d=0.45×10^{-6}M). The next major enhancement among satC with single changes was at position 268, a cytidylate-to-uridylate change (K_d=2.75x10^{-6}M)

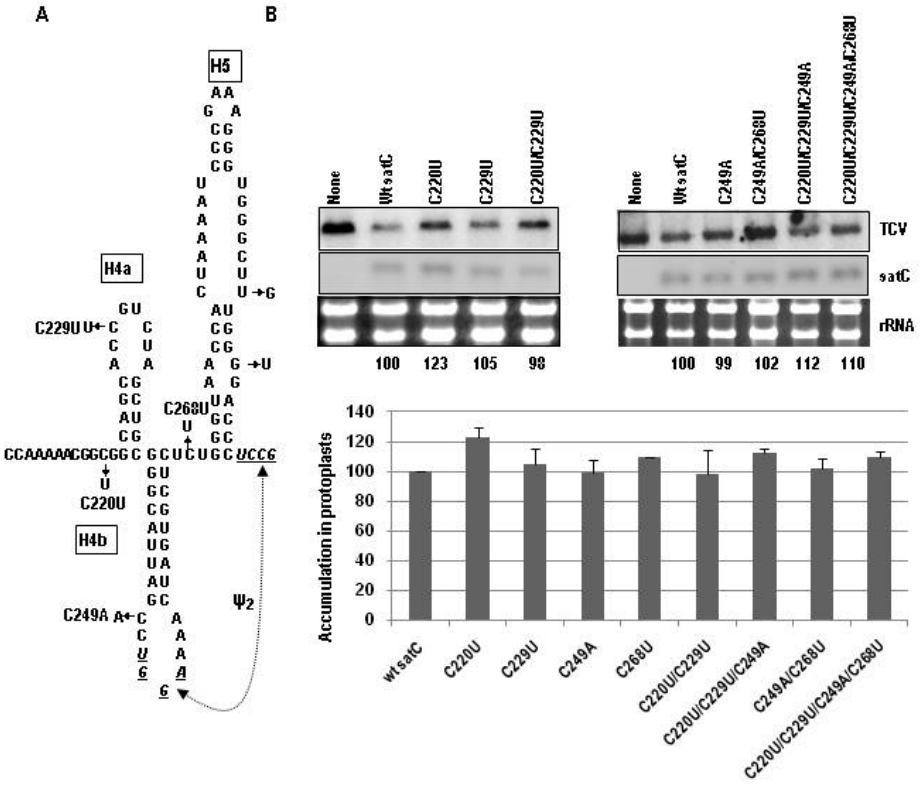


Figure 2.4 Replication of satC mutants which contain TCV TSS sequence. (A) SatC TSS analogous region showing the nucleotides that differ from TCV. The four nucleotides outside of H5 were step-wise converted into TCV analogous sequences. (B) Accumulation at 40 hpi of satRNA with various mutations in 5 x 10^6 Arabidopsis thaliana protoplasts prepared from callus cultures. Denatured RNA was hybridized with a [γ-^{32}P]ATP-labeled oligonucleotide probe, which is complementary to both satC and TCV. Data from three independent experiments were normalized to wt satC levels, arbitrarily assigned a value of 100.

Table 2.3 Ribosome binding and accumulation in protoplasts of satC mutants

	Kd[a,e] ($\times 10^{-6}$ M)	Accumulation[b] (%)
TCV TSS only	0.45 ±0.05	-
satC TSS analogous region only	-[c]	-
wt satC[d]	12.14 ± 1.80	100
C220U[d]	3.73 ± 0.57	123 ± 6
C229U[d]	3.70 ± 0.57	105 ± 10
C249A[d]	0.99 ± 0.10	99 ± 8
C268U[d]	2.75 ± 0.47	105 ± 0
C220U/C229U[d]	6.98 ± 0.80	98 ± 16
C249A/C268U[d]	0.73 ± 0.10	102 ± 6
C220U/C229U/C249A[d]	0.67 ± 0.07	112 ± 3
C220U/C229U/C249A/C268U[d]	0.55 ± 0.06	110 ± 3

[a] Binding to yeast 80S ribosomes. Standard error is indicated
[b] 40 hours after inoculation of Arabidopsis protoplasts.
Standard deviation is indicated.
[c] No ribosome binding detected
[d] Ribosome binding using full length satC
[e] Ribosome binding was done by Arturas Meskauskas.

in the H4b-H5-linker, whose counterpart in TCV TSS mimics the D-loop in a tRNA.

Enhanced ribosome binding to satC does not impact satRNA replication.

To determine if enhanced ability to bind ribosomes interferes with satC replication, satC containing the single and multiple base changes were inoculated into protoplasts and accumulation was assayed at 40 hpi. All satC mutants accumulated to at least wt levels, ranging from 98% to 123% of wt (Figure 2.4; value for C268U from Zhang et al., 2006c). The four-base altered mutant accumulated to higher than wt level in protoplasts although this mutant bound ribosomes at near wt TCV TSS levels. The other two alterations in H5 (U302G/G306U) also accumulated to higher than wt level in protoplasts (Zhang et al., 2006c). TCV levels were comparable, which means TCV replication was not affected by the presence of different satRNAs. SatC mutants with enhanced ribosome binding did not negatively impact TCV translation since the accumulation of both TCV and satC were not affected (Figure 2.4B). These results suggest either that ribosome binding is not incompatible with replication or that the satRNA is not exposed to ribosomes in the cell.

Enhanced ribosome binding to satC affects satRNA competition *in planta*.

Unlike replication in protoplasts, accumulation in plants involves virus replication, movement, and interaction with the host silencing system. While these satC mutants with TCV TSS sequences accumulated to at least wild type satC levels in protoplasts, we wanted to address if the TCV residues reduced satC fitness in plants. Among these satC mutants, C220U/C229U/C249A/C268U has the highest ribosome binding ability. C249A

and C268U bound ribosomes well among the single mutants, while C220U/C229U could potentially form TCV H4a and Ψ_3 which is important for TCV translation. These four mutants were individually inoculated onto turnip seedlings along with TCV genomic RNA. SatC fragments were cloned from total RNA that was extracted from un-inoculated upper leaves at 21 dpi. SatC containing all of the original mutations were recovered, which correlates with their near wt level of replication in protoplasts.

When wt satC and the four mutants were co-inoculated onto turnip seedlings with helper virus TCV, 15 different progeny were identified (Table 2.4). The original input sequences including wt satC, C249A, C220U/C229U and C220U/C229U/C249A/C268U were recovered, with wt sequence as the majority (31 out of 69). C249A was highly competitive among the mutants (19 out of 69) even though C249A has the highest ribosome binding among the single mutants. This demonstrates that higher ribosome binding did not interfere with satC accumulation in plants. C268U was not recovered, either because our sample size was too limited (69 sequences) or because satRNA prefers the adapted satC cytidylate instead of the TCV uridylate in the linker between H4b and H5.

The error-prone nature of viral RdRp can cause second site changes that can improve fitness of mutant viruses. In this experiment, several new alterations were recovered. G142, G157 and C190 were recovered in a wt background. U159C/C220U/C229U/C249A/C268U and U159C/C220U/C229U/C249A/U273C have a second site change in the satD region near G157. An adenylate-to-guanylate change at 216, which is next to the DR sequence, may affect the structural change involving DR. With a uridylate-to-cytidylate change at 312, Ψ_2 was slightly weakened, which may

Table 2.4 SatC sequences obtained from the competition among the satC mutants in plants

Mutant sequence	Recovered sequences	Numbers recovered
Competition (wt, C249A, C268U, C220U/C229U, C220U/C229U/C249A/C268U)	wt	31
	G142A	1
	G157A	2
	C190U	1
	A216G	2
	C220U	1
	C249A	19
	U312C	1
	C220U/C229U	2
	A51G/C220U/C229U	1
	C220U/C229U/C249A	1
	C220U/C229U/C249A/C268U	1
	U159C/C220U/C229U/C249A/C268U	1
	U159C/C220U/C229U/C249A/**U273C**	1
	U115 insertion/C239U/C249A	3
	C249A/**C264G/U265G/C266A**	1

Second site changes are in bold.

enhance the ability of satC to adopt its active conformation. When wt satC was inoculated by itself onto plants, no second site changes were found (Zhang, Stuntz, and Simon, 2004).

SatC with TCV TSS sequence cause structural changes in Ψ_2 and H4a loop.

Full length wt satC and the mutants with TCV sequence in the TSS-equivilent region were subjected to in-line probing to explore possible structural differences that improve ribosome binding. In-line structure probing is an assay that monitors RNA phosphodiester cleavage. The mechanism involves a nucleophilic attack of the 2' oxygen on the phosphate backbone when the 2' oxygen, the phosphate and the oxyanion leaving group adopt a linear configuration, which occurs if the base is flexible. The assay can be used to monitor alternative cleavage sites when the RNA is bound by a ligand (Mandal and Breaker, 2004), or to identify the ligand binding site. Bases susceptible to in-line cleavage usually do not participate in a higher-order structure. C249A, C268U and C220U/C229U/C249A/C268U together with wt satC were subjected to in-line probing. SatC RNAs were either phosphorylated with [γ-^{32}P] ATP at the 5' end or ligated with [α-^{32}P] pCp at the 3' end. When the RNAs were 5' end labeled, nucleotide flexibility of mutant RNAs in the 5' region was identical to wt indicating that the mutations did not disrupt the 5' half of satC which is the satD-derived region (Figure 2.5).

The TSS analogous region was monitored when the RNAs were 3' end labeled with [α-^{32}P] pCp. The wt satC 3' end in-line cleavage pattern was similar to the cleavage pattern of TCV 3' end fragment F4, which includes the TSS, Pr and a hairpin similar to satC M1H (Figure 2A from McCormack et al, 2008). The terminal and internal loops of

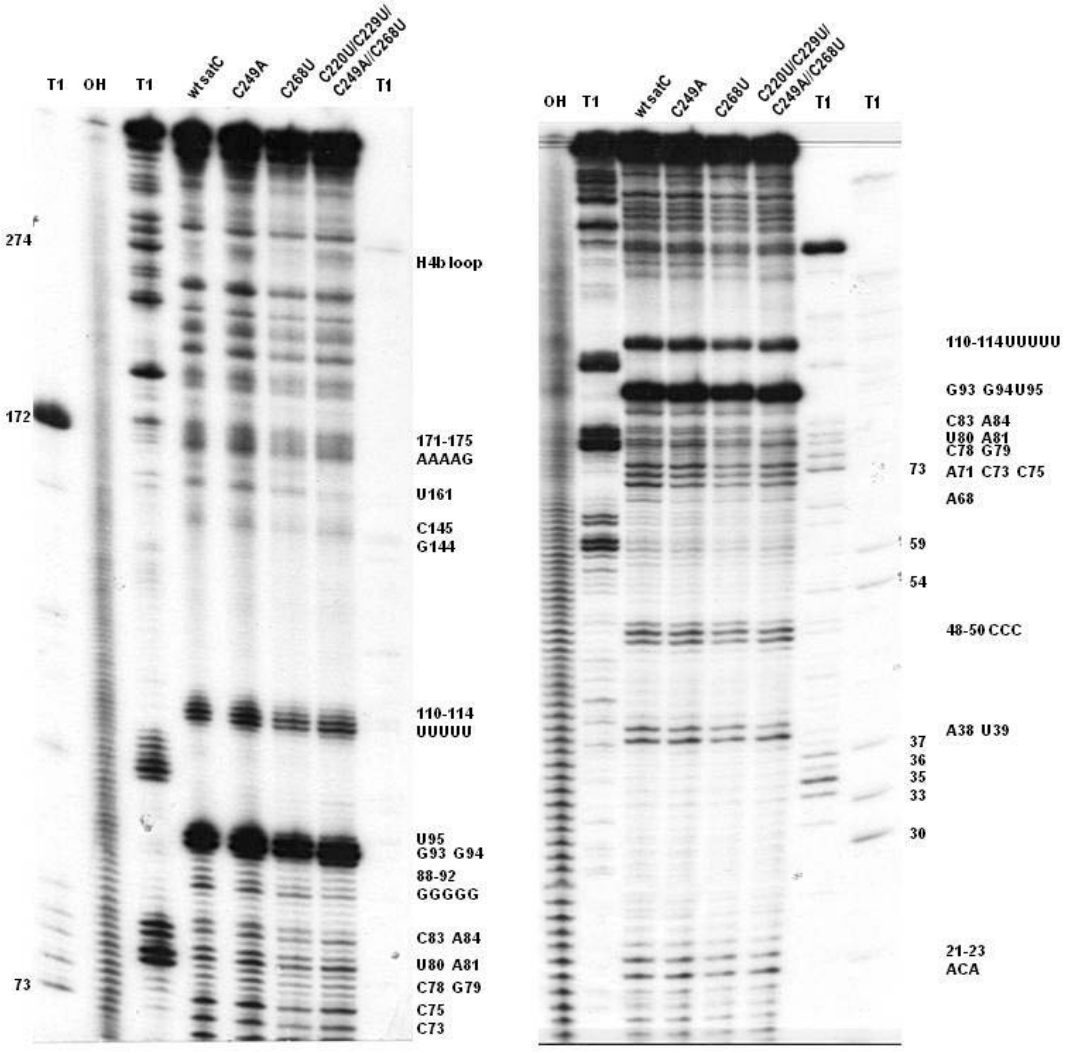


Figure 2.5 In-line probing of the 5' half of wt satC and various mutants. RNAs were 5' end labeled and incubated at 25ºC for 14 hours followed by 8% polyacrylamide gel electrophoresis. Due to non-specific digestion by RNase T1 of the 5' end of satC (lane 3 from left panel and lane 2 from right panel), T1 ladders were generated by partial RNase T1 digestion of two RNA fragments. Locations of the susceptible nucleotides are indicated.

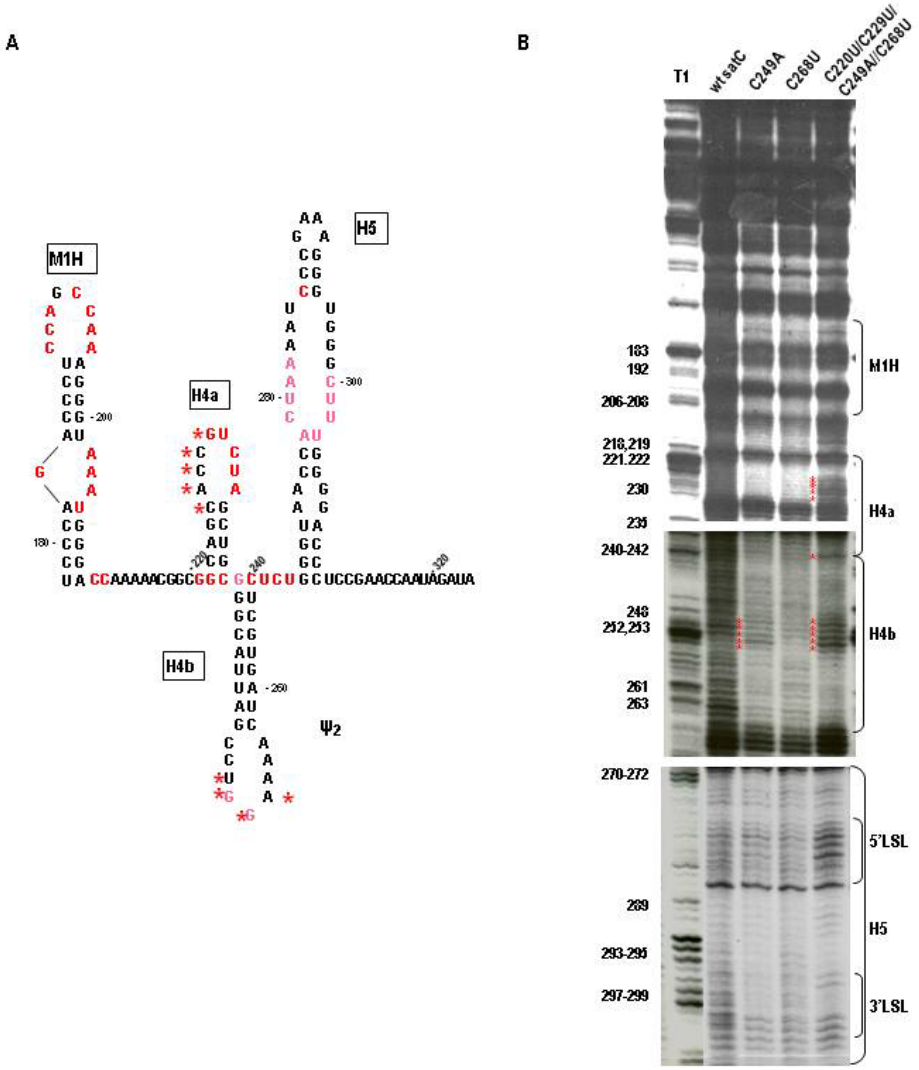


Figure 2.6 In-line probing of the 3' half of wt satC and various mutants. (A) Position of residues cleaved by in-line probing in the satC 3' half are shown. Nucleotides in red are susceptible to in-line cleavage. Purple nucleotides indicate a lower level of cleavages relative to other cleavages. Stared nucleotides are more susceptible in mutant C220U/C229U/C249A/C268U relative to wt satC. (B) In-line cleavage of satC RNAs. T1, ladder generated by partial RNase T1 digestion using partially denatured RNA. Locations of the hairpins are indicated. Stars indicate the differences between mutant C249A, C220U/C229U/C249A/C268U and wt satC.

M1H (H4 in TCV), part of the H4a loop (GUCUA), and the short linker sequence between H4b and H5 (UCU) showed strong in-line cleavages (Figure 2.6B). The triple uridylates in the TCV comparable region also showed strong in-line cleavage. The cytidylate at position 286 in the H5 upper stem was also highly susceptible to in-line cleavage.

RNA containing the two single mutations C249A and C268U had similar cleavage patterns with wt satC (Figure 2.6B). The only exception was in H4b loop between wt and C249A. wt satC showed faint cleavage of the two guanylates at positions 252 and 253 in H4b loop. The cleavage at C249A in the H4b loop was extended to the two adjacent nucleotides. The four-base mutant C220U/C229U/C249A/C268U also had stronger cleavages at the UGGA in the H4b loop which are involved in Ψ_2, thus the mutants with higher ribosome binding had more disrupted Ψ_2. Besides the stronger cleavage in the H4b loop, C220U/C229U/C249A/C268U also had other significant differences compared with wt satC. H4a loop sequence CACCC and the cytidylate at position 239 at the bottom of H4a stem were more susceptible to in-line cleavage. In TCV, H4a loop sequence is part of Ψ_3. Ψ_3–H4a region is important for ribosome binding to TCV fragments *in vitro* and is part of the translational enhancer (Stupina et al., 2008).

Discussion

Our report that a 100-nt region near the 3' end of TCV folds into TSS that binds ribosomes and serves as the TCV translational enhancer (McCormack et al., 2008;

Stupina et al., 2008) provoked the question of whether the related sequence in the TCV-associated satRNA satC forms a similar structure despite not being translated. Previous data showed that the six differences in satC's TSS-analogous region prevent formation of a tRNA-like structure in the satRNA according to the RNA structure program RNA2D3D and molecular modeling (Stupina et al, 2008). The structural predictions correlate with ribosome filter binding assays, showing that isolated satC TSS does not bind detectably to ribosomes (Stupina et al, 2008).

C249A had the highest ribosome binding among the single mutants (12.3-fold greater than wt satC). The modeling protocol RNA2D3D program predicted that C249 in satC is a primary contributor to destabilizing the TSS, which correlates with the high ribosome binding of C249A. While the location of C249 in the 3D pre-active structure of satC is not currently known, H4a and/or H4b exchanges with the related Carmovirus CCFV indicated that sequences within these elements form a single functional domain. Altering C249 disrupts its pre-active conformation allowing formation of TCV-like H4a and H4b, which would promote a more TSS-like structure in the region.

Altering C268 in the short linker between H4b and H5 to the TCV-specific uridylate led to the next most significant enhancement in ribosome binding (4.4-fold). Interestingly, a second untranslated subviral RNA whose sequence is derived mainly from TCV (diG) contains the identical uridylate to cytidylate difference at this location (Li et al., 1989). DiG has only one other base variation with TCV in the TSS region, a cytidylate at TCV position 4000 that enlarges the H5 lower symmetrical loop.

The two differences between TCV and satC in the DR and H4a loop regions (C220 and C229, respectively) likely contribute in part to loss of Ψ_3 in satC. Ψ_3 is

important for both efficient ribosome binding to the TCV TSS, and for translational enhancement (Stupina et al., 2008), and is required for TCV accumulation (McCormack et al., 2008), but not for satC replication in protoplasts (Figure 2.1; Guo et al., 2009). By not requiring formation of Ψ_3 for satC biological activity when satC was initially generated, these regions could have evolved to participate in additional functions. While the loop region of H4a can be altered without a substantial decrease in satC accumulation (Chapter IV; Guo et al., 2009), the upstream sequence flanking H4a (the DR) was apparently able to evolve a separate function. Since satC with C220U/C229U had reduced ribosome binding compared with satC containing the individual changes, Ψ_3 may not have re-established despite sequence identity with TCV in the region. The C220U alteration appeared to enhance DR function, as satC with C220U accumulated 23% better than wt (Figure 2.4).

These experiments revealed that alteration of the H4a loop could complement mutations in the upstream DR element. The DR was originally identified as an important sequence-specific element (Sun et al., 2005). Fitness of satC in plants, high-level accumulation in protoplasts, and efficient transcription of satC transcripts by the TCV RdRp *in vitro* all correlated with wt DR sequences (Sun et al., 2005; Zhang et al., 2006a; 2006b). DR is also involved in the conformational switch between satC pre-active and active forms (Zhang et al., 2006a; Zhang et al., 2006b). Our lab has recently determined that mutations in the DR-equivalent sequence in TCV disrupt single-site RdRp binding to the region (X. Yuan, M. Young, and A. E. Simon, unpublished data). One possibility is that the DR sequence is also important for RdRp binding to satC. Such an interaction

could be restored in mutant C229G, where the CCG in the H4a loop is converted to CGG, which is repeated twice in the DR (CGGCGG).

In addition to the higher than wt levels achieved by satC with C220U, accumulation of satC with C220U/C229U/C249A and C220U/C229U/C249A/C268U also reproducibly reached levels higher than wt. Previous results also showed that satC with H5 of TCV reproducibly accumulated to greater than wt levels in protoplasts (115%; Zhang et al., 2006c). An interesting question is why satC evolved to reduce its accumulation efficiency? One possibility is that accumulation in single cells reflects only the stability of the satRNA and its replication ability while systemic accumulation in host plants requires further interplay with the helper virus to support efficient movement. It is possible that the selective advantage conferred by the satRNA's ability to repress virion levels and thus enhance monomer CP suppression of RNA silencing is reduced by higher satC levels that negatively interfere with some other TCV requirement. Interestingly, increased ribosome binding to full-length satC did not interfere with satC accumulation, suggesting that the RdRp can displace a bound ribosome or that satC is shielded from ribosomes within cells.

Though satC and TCV share 90% sequence identity at the 3' terminus, satC with TCV 3' terminal 100 bases accumulates very poorly in plants and protoplasts (Wang and Simon, 2000). The weak accumulation was due to the differences in the Pr region (Zhang et al., 2006a), not because of the differences at the analogous TSS region (Zhang et al., 2006a; and this chapter). *In vitro* experiments showed that the satC core promoter is a stronger promoter compared to the weak TCV core promoter (Zhang et al., 2006a). TCV has more strict sequence requirements at the Pr loop than satC does (Yuan et al., 2010).

The replication elements are normally affected by the surrounding sequences and structural context. Evolutionary adaptations have allowed the satC Pr to function independently without any additional upstream elements. SatC evolved the differences at Pr for high level accumulation, while the differences in the analogous TSS region evolved for new conformational changes in satC, or for interaction with its 5' end sequences, or for satC fitness in plants. *Tobacco mosaic virus* (TMV) defective RNAs also have different replication requirements compared to genomic RNA replication, in that they require smaller 3' UTR and more sequence specificity. Part of the reason is because TMV is replicated in *cis* by a coupled translation-replication mechanism, while the defective RNAs are replicated in *trans* (Chandrika et al., 2000).

These findings suggests that the 6-nt differences with TCV in the region did not arise with the principal goal of eliminating satRNA ribosome binding, but rather that evolution of these base changes were needed to reduce satC accumulation, or to enhance a beneficial function not yet identified. In addition to the 6 nt differences in the TSS region, satC has two additional base changes in the linker sequence between H5 and Pr and several base changes and deletions within the Pr that together improved satC accumulation by ~ 3-fold (Zhang et al., 2006c). By not requiring maintenance of the entire TSS for ribosome binding, satC could evolve residues in this region to both establish its pre-active structure and maintain an optimum level of accumulation for its mutualistic association with TCV.

CHAPTER III

The importance of a flexible linker region in satC replication

Introduction

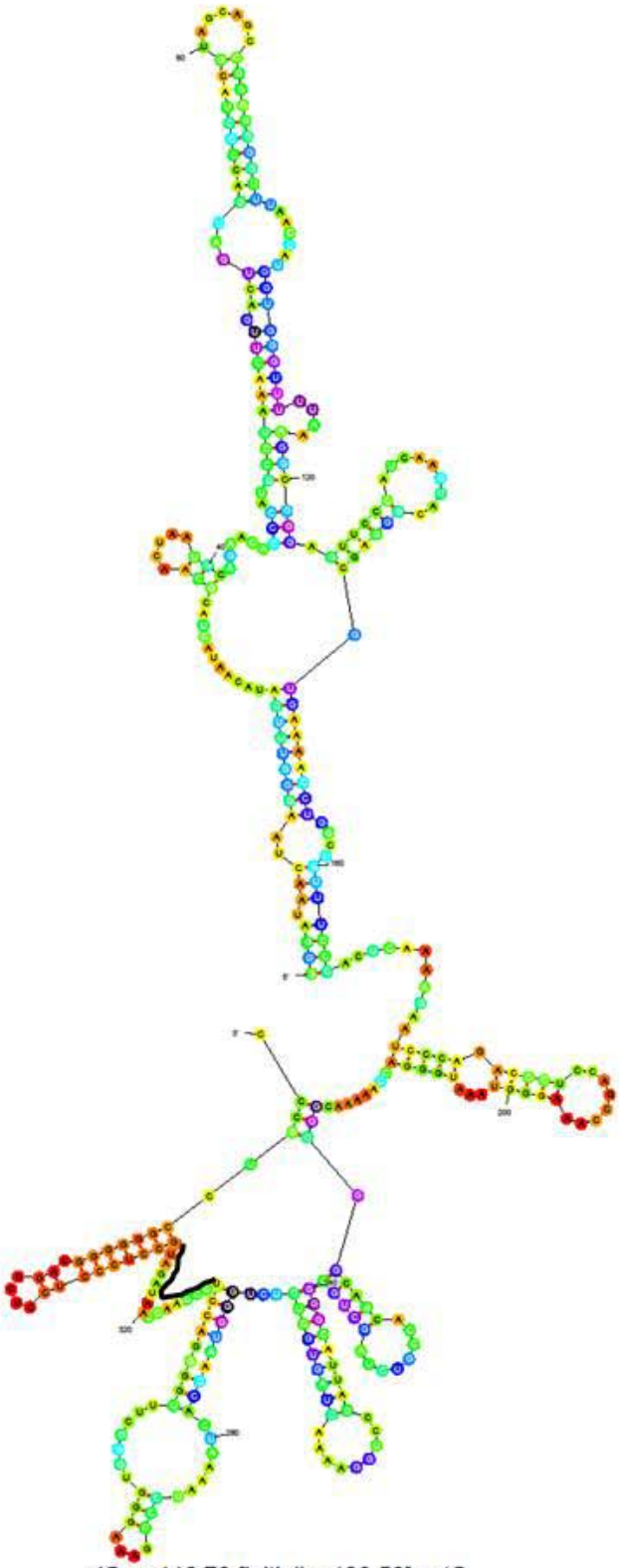
The structure adopted by *in vitro* synthesized (+)-strands of satC, which has biological relevance *in vivo* (Zhang et al., 2006a), does not contain the four 3' terminal hairpins (H4a, H4b, H5 and Pr) or Ψ_1 and has limited transcriptional activity in vitro. The regulation of many RNA-associated cellular processes including transcription termination, translation, and RNA cleavage can involve changes in the conformation of the RNA (Brantl, 2004; Nagel and Pleij, 2002). RNA conformational switches have also been implicated in the mutually exclusive processes of translation and replication of RNA viral templates. Plus-stranded genomes must initially assume a conformation that is recognized by cellular ribosomes for reiterative translation of products required for replication, such as the viral RdRp. When sufficient product accumulation has been achieved, the RNA must be induced to switch conformations to a form that is both inaccessible to ribosomes and contains the *cis*-acting elements necessary to attract the newly translated RdRp (Dreher 1999). Following reiterative transcription of (-)-strand intermediates, (+)-strands are synthesized that may not be templates for further (-)-strand synthesis (Chao, Rang, and Wong, 2002; Zhang and Simon, 2005). This would imply that *de novo* synthesized (+)-strands adopt a conformation that is not recognized by the RdRp. For viral genomic RNAs, this alternative conformation may be the initial translation-competent form.

However, noncoding subviral RNAs derived from the viral genome, such as DI RNAs and chimeric satRNAs, might need to evolve their genomes to produce an alternative inactive conformation. The pre-active structure in satC could be converted to a transcriptionally active structure by alterations that either disrupt a stabilizing pseudoknot or release 3' terminal bases from their initial paired configuration (Zhang et al., 2006a). How satC was converted from the pre-active structure to the active structure and which structure is recognized by the RdRp are not known.

SatC phylogenetically-derived structure shows a single-stranded linker region (Figure 1.2; position 312-327, 5' UCCGAACCAAUAGAUA) between H5 and the Pr, which is conserved in most of the carmoviruses (Table 3.1), but with no obvious sequence similarity. Sequences in satC(-)-strands that can serve as a promoter *in vitro* was mapped to a sequence 41 nt from the (-)-strand 5' end (3' AACCCCUGGGAGGC) (Guan, Song, and Simon, 1997). The complementary sequence is between positions 302 and 315, part of which is located in the single-stranded linker between H5 and Pr (Figure 1.2). *In vivo* SELEX involving this region recovered the sequences "UCC" in position 312-314 back (Guan, Carpenter, and Simon, 2000b; Guan, Song, and Simon, 1997), which suggested a role for at least a portion of the linker sequence in satC replication.

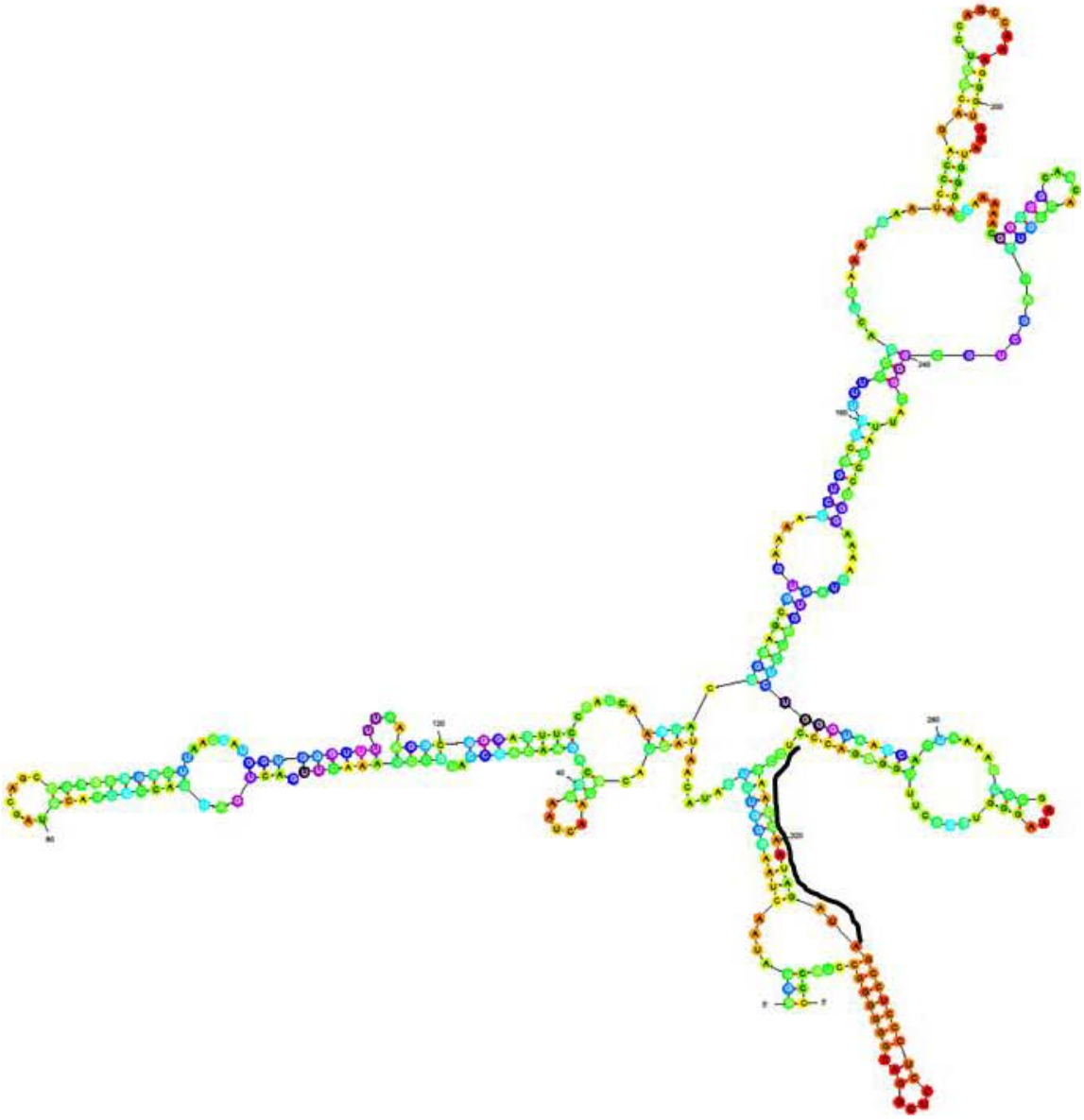
mFold-predicted (Zuker, 1989; Zuker, 1994) satC structures have three conformations for the H5-Pr-linker (Figure 3.1). Figure 3.1A shows that the linker region is single-stranded, the same as the phylogenetically conserved structure. The majority of

A



dG = -116.70 [initially -120.50] satC

B



dG = -108.62 [initially -122.40] satC

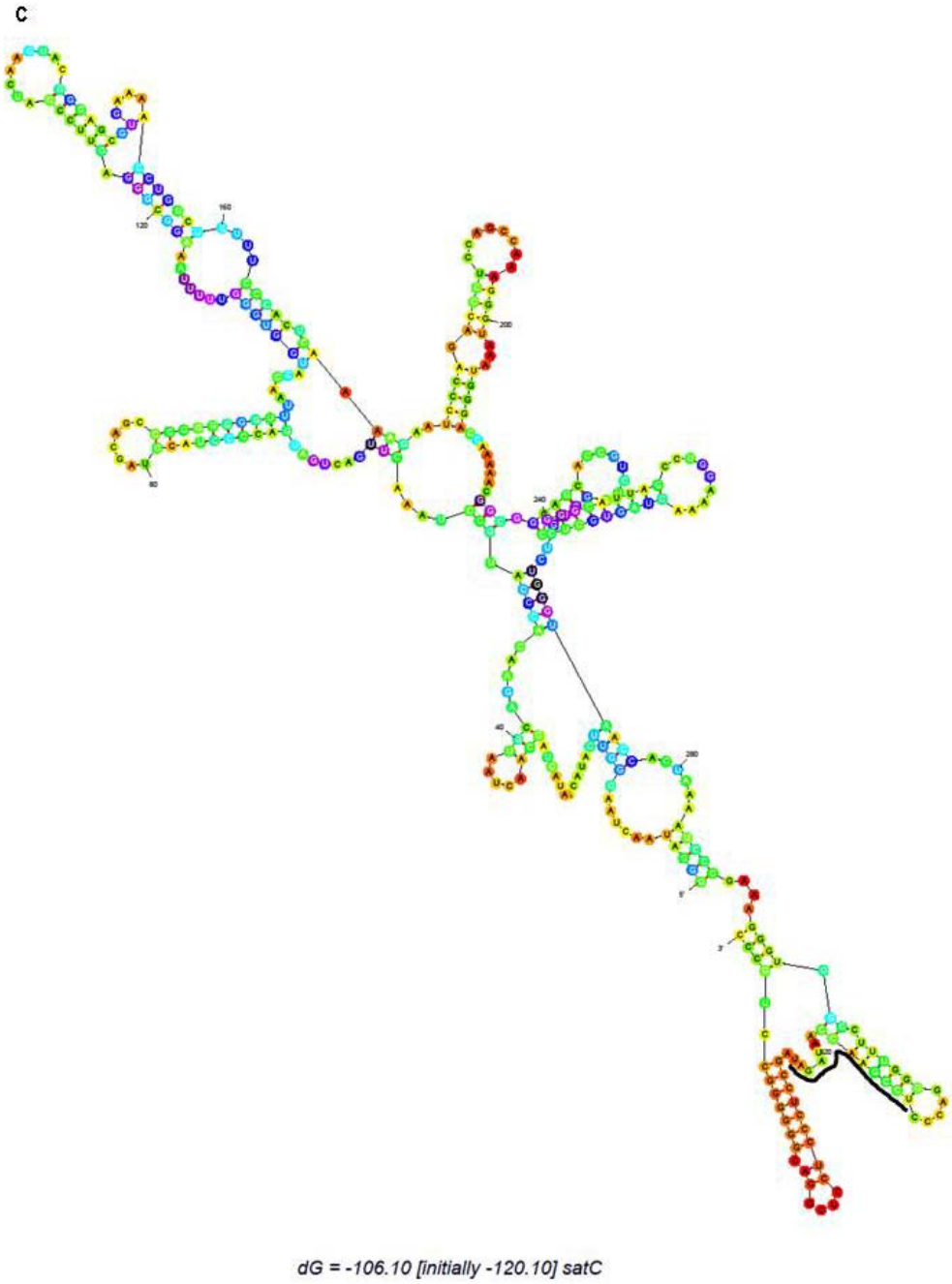


Figure 3.1 mFold-predicted satC structures showing three different conformations for H5-Pr linker. The lines indicate the H5-Pr linker (positions 312-327).

mFold-predicted satC structures have linker region base-pairing with the 5' end, as shown in Figure 3.1B. The linker sequence can also interact with the H5 lower stem (Figure 3.1C).

To determine if the linker region between H5 and the Pr has any sequence or structure specificity, the linker was subjected to *in vivo* functional selection (*in vivo* SELEX; systematic evolution of ligands by exponential enrichment). The methodology of *in vitro* SELEX was initially developed to isolate individual RNAs from a large RNA pool that could bind to a specific target (Ellington and Szostak, 1990; Tuerk and Gold, 1990). The principles of in vitro SELEX can be applied to *in vivo* evolution of viral RNAs by randomizing a portion of viral sequence, introducing a pool of these randomized viral RNAs into host cells, and subsequently isolating replication-competent RNAs. *In vivo* SELEX of satC is performed by inoculating a satC population containing randomized sequences in the element of interest together with helper virus TCV onto host plants for fitness selection through multiple rounds of competition. The mutualistic relationship between TCV and satC makes *in vivo* SELEX of satC possible. By interfering with virion accumulation, satC indirectly contributes to movement of TCV (Zhang and Simon, 2003a) by enhancing the ability of TCV to overcome PTGS, an endogenous antiviral defense system. Reduced accumulation of virions by satC results in enhanced levels of free CP, which is the TCV silencing suppressor (Qu, Ren, and Morris, 2003; Thomas et al., 2003). Therefore, TCV that is isolated from systemic leaves is more likely to be associated with a functional satRNA. Like all other RNA polymerases, the high error rate (10^{-4}) of TCV RdRp makes the satC sequence selection feasible.

Table 3.1 Carmovirus H5-Pr linker sequences.

virus		H5-Pr linker
satC		UCCGAACCAAUAGAUA
TCV	*Turnip crinkle virus*	UCCGAACUAAAAGAUA
CCFV-BL	*Cardamine chlorotic fleck virus*	ACCAAAUGAAGAAAAUA
JINRV	*Japanese iris necrotic ring virus*	GGCCGACACUGCGUAAUGCAAAACUA
CarMV	*Carnation mottle virus*	CCACUGUUUAUUAAUUA
CPMoV	*Cowpea mottle virus*	AACAUAAUCAGAUAAUGAUA
SCV	*Saguaro cactus virus*	CUUAUUAGUUGAAAAUUA
MNSV	*Melon necrotic spot virus*	AUAGTCAUGUAUGUUUGA
PFBV	*Pelargonium flower break virus*	CUAACACGUUCCUGAAGGAAGA
HCRSV	*Hibiscus chlorotic ringspot virus*	ACUUGUUUGAAAUAC
PSNV	*Pea stem necrosis virus*	AGUAACACCAAAUUUAUAAACUA
CbMV	*Calibrachoa mottle virus*	AACUGCACUAAAC
PLPV	*Pelargonium line pattern virus*	CAUCCUUAAUGAAGAA
PCRPV	*Pelargonium chlorotic ring pattern cirus*	GAUGGUGCGAAGUAAAUUAGA
GaMV	*Galinsoga mosaic virus*	No H5- or Pr-equivalent
SYMV	*Soybean yellow mottle mosaic virus*	AGCAAAUCUUGACCUGAUA
ANFB	*Angelonia flower break virus*	UCCAAGUUUCCUUGUGAAAGAAA
Nootka	*Nootka lupine vein-clearing virus*	AACCCUGUUUUUAAAUUA

Materials and Methods

In vivo SELEX

In vivo functional selection (Sun and Simon, 2003; Zhang and Simon, 2003b) was performed to produce full-length satC RNAs with random sequence replacing the 16-base linker region between H5 and Pr. PCR was carried out using pC(+) (pUC19 containing full-length satC cDNA; Table 3.2) as the template and primers T7C5' and C327S or L3C327S (Table 3.3). For the first round, 5 μg of satC transcripts containing randomized sequences were directly synthesized using T7 RNA polymerase and inoculated onto each of 30 turnip seedlings along with 2 μg of TCV genomic RNA transcripts. Immediately before inoculation, the RNA was mixed with an equal amount of 2x infection buffer and rubbed onto two leaves of each seedling (10 μl each leaf) using gloved hands. Total RNA was extracted from uninoculated leaves at 21 dpi. Full-length satC was amplified by reverse transcription-PCR using primers T7C5' and oligo7 and then cloned into the SmaI site of pUC19 and sequenced. For the second round of SELEX, equal amounts of total RNA from each first round plant were pooled and inoculated onto six new turnip seedlings at a concentration of about 5 μg per plant. Total RNA was extracted at 21 dpi and full length satRNA was cloned into pUC19 as described above.

Construction of satC mutants

To generate plasmids, two PCR reactions were performed with primers 5’GGUU and oligo7, T7C5'and 3’AACC using template pT7C+. Immediately after PCR, 1 μl of Klenow was added to the PCR reaction to remove the 3’ overhang A generated by Taq polymerase. PCR products were agarose gel purified, treated with T4 polynucleotide

kinase (10,000 U/ml; New England Biolabs) according to the manufacturer's protocol, ligated to SmaI-digested, CIP-treated pUC19 vector.

To generate plasmid G253C and G254U, PCR reactions were performed with a common 5' primer T7C5' and analogous 3' primers using template pT7C+. SpeI and NcoI digested PCR products were ligated into similarly digested pT7C+, replacing the endogenous fragment.

To generate plasmid C313G and U312A, PCR reactions were performed with T7C5' and analogous 3' primers using template pT7C+. PCR products were treated with klenow, agarose gel purified, kinased with T4 polynucleotide kinase, then ligated to SmaI-digested, CIP-treated pUC19 vector.

To generate plasmid G253C/C313G and G254U/U312A, plasmid G253C and G254U were digested with SpeI and NcoI. The small fragment was agarose gel-purified and ligated into similarly digested plasmid C313G and U312A, replacing the endogenous fragment. All constructs were confirmed by DNA sequencing.

***In vitro* transcription, inoculation of Arabidopsis protoplasts, and Northern blots**

TCV genomic and satC RNA transcripts with precise 5' and 3' ends were synthesized using T7 RNA polymerase and SmaI-digested plasmids. Protoplasts (5×10^6), prepared from callus cultures of Arabidopsis thaliana ecotype Col-0, were inoculated with 20 μg of genomic RNA transcripts with and without 2 μg of satC transcripts using PEG-$CaCl_2$, as previously described. Total RNA was isolated from protoplasts at 40 hpi, denatured with formamide and separated on nondenaturing agarose gels as previously described. Subsequently denatured RNA was hybridized with a [γ-^{32}P]ATP-labeled oligonucleotide probe oligo 13 (GTTACCCAAAGAGCACTAGTT), which is

complementary to both satC and TCV sequence.

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) 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). The transformed bacterial cells were cultured overnight in 3 ml of rich growth medium (MB) containing 50 μg/ml ampicillin and 34 μg/ml chloramphenicol at 37°C with shaking at 250 rpm. One 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 μg/ml ampicillin and 34 μg/ml chloramphenicol, shaken at 37°C until the optical density (OD600) reached 0.5–0.6. Cool the flask on ice for 10 min. Protein expression was then induced at 14°C with 0.3 mM IPTG (isopropyl-β-D-thiogalactopyranoside) with shaking at 250 rpm for overnight. The induced cells were harvested by centrifugation in a Sorvall GSA rotor at 5,000 rpm for 10 min at 4°C and resuspended in 70 ml (for cells from 1 L culture) of ice-cold column buffer (20 mM Tris-HCl, pH 7.4, 25 mM NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol, 10% glycerol). The solution was sonicated at least 8 times (15 seconds per time) on ice

with intervals of 2 min to disrupt the cells. After each sonication, PMSF was added to the solution to help prevent protein degradation. The sonicated solution was then centrifuged at 13,000 rpm for 10 min at 4°C for two times. 1.5 mL of amylose resin (New England Biosciences) was loaded into a 0.8 cm x 4 cm column and the storage alcohol was drained. The column was washed two times with 8 column volumes of ice-cold column buffer. The supernatant was loaded onto the equilibrated amylose resin column. The column was washed with 12 mL of ice-cold column buffer, then the protein was eluted with 5 ml of ice-cold column buffer containing 10 mM maltose and fractions (0.5 mL) collected. All purification steps were carried out at 4°C. The concentration of purified recombinant proteins was measured using the Bio-Rad protein assay, based on the Bradford method, containing 200 µL Biorad protein assay solution, 790 µL water, and 10 µL protein. The OD was read at a wavelength of 595. A high concentration of proteins usually elutes in the 2^{nd} and 3^{rd} fractions (0.8 µg/µL and 1.3 µg/µL, respectively). The proteins from these two fractions were aliquoted into Eppendorf tubes (20 µL each), frozen with liquid nitrogen, and stored at –80°C.

RNA DMS modification and primer extension

Polyacrylamide gel-purified full length wt satC transcripts were heated to 65°C and either allowed to slow cool over 2 hours to room temperature, or snap cool on ice for 3 min. RNA (20 pmol) was then incubated at room temperature in 50 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 100mM NaCl for 10 min with or without 20 pmol of purified p88 RdRp. Five microliters of 1:10 ethanol-diluted DMS was added to the RNAs for 10 min at room temperature. 100% ethanol was added to the control tube. Reactions were ethanol precipitated, dissolved in 12 µL of 0.5X TE buffer. 1 pmol of [γ-^{32}P]ATP labeled

oligonucleotide probe oligo 7 was annealed to the RNA by incubating at 65°C for 5 min and cooled to 55°C. Primer extension reactions were performed with Superscript III reverse transcriptase at 55°C for 30 min, quenched by adding 1 μl of 4 M NaOH and denaturing at 95°C for 5 min. Samples with RNA loading buffer (85% formamide, 0.5X TBE, 50 mM EDTA, pH 8.0, containing bromophenol blue and xylene cyanol tracking dyes) added were heat denatured at 95°C for 5 min, and subjected to electrophoresis through sequencing length 8% denaturing polyacrylamide gels containing 8 M urea. Sequencing reactions were preformed with 1 μL purified wt satC transcripts and individual ddNTP.

Filter binding assay

In vitro transcripts of satC fragments were agarose gel purified, treated with Antarctic Phosphatase, then kinased with [γ-^{32}P]ATP by T4 PNK. The reactions all contained a constant low level of RdRp (5 pmole) and increasing amounts of labeled RNA (from 5 pmole to 120 pmole). The RNA and RdRp were combined in RdRp binding buffer (50 mM Tris, pH 8.2; 10 mM $MgCl_2$; 10 mM DTT; and 10% glycerol) and allowed to bind for 30 min at room temperature. When filtering, the filters were moistened with binding buffer (2 x 1 mL) before the RNA/protein mixture was applied. 500 μL binding buffer was added to the reaction tube and the entire tube's contents were transferred to the filter with vacuum applied to bring down the reaction through the filter. After filtration, the filters were washed two times with 3 mL of binding buffer. The filters were placed into scintillation vials with 5 ml scintillation fluid. The vials were placed in the scintillation counter and the counts for each of the filters recorded. Non-linear regression was performed using GraphPad Prism4.

Table 3.2 Constructs used in Chapter III

Name	Description
pC(+)	Full length satC cloned into pUC19
pT7C(+)	T7 promoter and full length satC cloned into pUC19
pTCVms	Full length TCV cloned into pUC19
5'GGUU	GGUU to CCAA change at positions 13-16
3'AACC	AACC to UUGG change at positions 316-319
GGUU-AACC	GGUU to CCAA change at positions 13-16 and a AACC to UUGG change at positions 316-319
G253C	G to C change at position 253
C313G	C to G change at position 313
G253C/C313G	G to C change at position 253 and a C to G change at position 313

Table 3.3 Oligonucleotides used in Chapter III

Name	Position	Sequence	Polarity
T7C5’	1-19	GTAATACGACTCACTATA GGGATAACTAAGGGTTTCA	+
Oligo7	338-356	GGGCAGGCCCCCCGTCCGA	-
C327S	300-356	GGGCAGGCCCCCCGTCCGAGGAGGGAGGCNNNNNNNNNNNNNNNNGGGTCCCCAAAGC N: ATCG=1:1:1:1	-
L3C327S	309-356	GGGCAGGCCCCCCGTCCGAGGAGGGAGGCNNNNNNNNNNNNNNNGGATACTTAAAGC N: ATCG=1:1:1:1	-
T7bob1 (5’GGUU)	1-15	GTAATACGACTCACTATAGGGATAACTAAGCCAATCATACAATACTAC	+
Bob2 (3’AACC)	299-356	GGGCAGGCCCCCCGTCCGAGGAGGGAGGCTATCTATTCCAACGGAGGGTCCCCAAAGC	-
G253C	235-272	CCCAGAGAGC*ACTAGT*TTTGCAGGCTAATGCCCGCAGC	-
C313G	297-356	GGGCAGGCCCCCCGTCCGAGGAGGGAGGCTATCTATTGGTTCGCAGGGTCCCCAAAGCCC	-

Underlined nucleotides indicate the mutations.

Italics are restriction enzyme recognition sites.

Results

The H5-Pr linker has sequence flexibility.

The H5-Pr linker was randomized (Figure 3.2A; positions 312 to 327), and transcripts containing the randomized sequences were inoculated with TCV onto 30 turnip seedlings as described in Materials and Methods. SatC capable of replication are also able to increase the rate of TCV movement and therefore TCV associated with viable satRNA are present in higher levels in uninoculated tissue. Total RNA extracted from uninoculated leaves of all plants at 21 dpi contained satRNA detectable on ethidium bromide-stained agarose gels (Figure 3.3A). Viable satC from 9 randomly chosen plants were reverse transcribed using primers T7C5' and oligo7 (Figure 3.3B), and cloned into the SmaI site of pUC19. Sequencing of first round progeny revealed 13 different sequences (Table 3.4). These sequences showed some conservations and variations. Seven out of the 13 sequences have UCC (positions 312-314) recovered as in the previous SELEX (Guan, Carpenter, and Simon, 2000b). Others also had AACC or AAC conserved.

Total RNA isolated from the 30 plants was pooled and inoculated onto six new turnip seedlings for direct competition between first round "winners". The six plants showed severe satC crinkle symptoms (Figure 3.3C). Total RNA was again extracted from uninoculated leaves at 21 dpi and satC species were cloned. Four second-round sequences were recovered from the six plants including the wild type sequence. All recovered clones contained the wt UCC sequence and wt AACC sequence (Table 3.4). Besides the two conserved sequences, the rest of the linker sequence was variable.

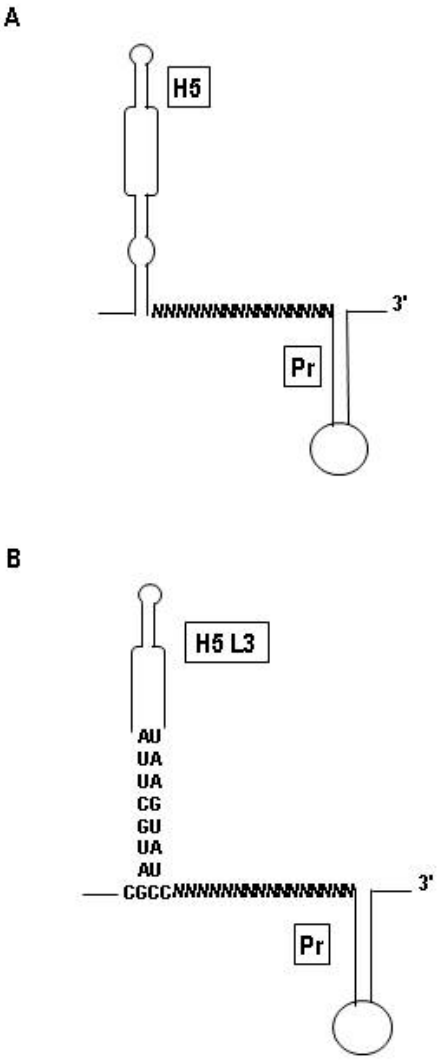


Figure 3.2 *In vivo* selection of H5-Pr linker region. (A) Phylogenetically inferred structure of H5 and Pr region. *N* means randomized nucleotide. (B) Structure of H5 and Pr region from third round winner of H5 lower stem *in vivo* selection. *N* means randomized nucleotide.

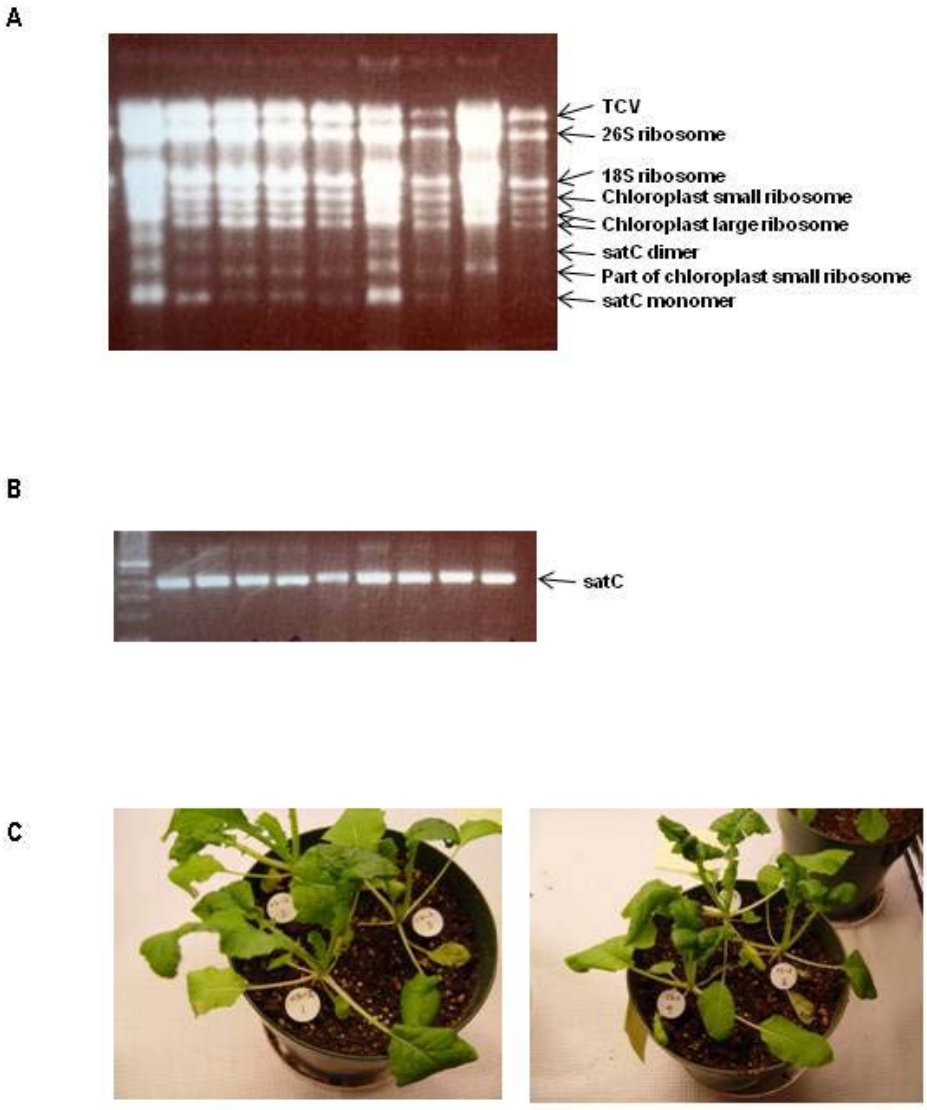


Figure 3.3 Analysis of H5-Pr linker region SELEX. (A) Agarose gel showing part of the total RNA from first round. (B) SatC RT-PCR products from 9 randomly chosen first round plants. (C) Turnip seedlings infected with total RNA from first round SELEX.

Table 3.4 Summary of in vivo SELEX of H5 and Pr linker

		Sequence
	wt	UCCGAACCAAUAGAUA
wt H5	First round	UCCAGAACCUGCAAAA
		UCCGUCUAACCUGUUC
		UACCAACUAUGAAUCC
		UCUGCGCCAUGAACCA
		UCCUAUAUAUUAAACA
		UCCAAACCUGUAAGAU
		UGCCAACGUAAUACUA
		CCAGGAUUUACCGUUA
		UCCAAACCAAUAGAUA
		GUCCAACGUAAUACUA
		UUCUCUGUGUACACCA
		CCGACACAUUUUCUUA
		UCCGCCACGCGUACCC
	Second round	UCCGAACCAAUAGAUA
		UCCACAACCACAAUUA
		UCCGCACCAUGAACCA
		UCCAAACCUGUAAGAC
L3 H5	First round	UCCGAACCAAUAGAUA
		UCCUCGUAUGUUUUC
		UCCGCAUUAUACUACA
		UCUAAUAACCCCGGAA
		UCUAAAAACUAAAUCC
		AUCUGAACCUGAUACA
		UCCAAACAUAUAUCA
		AUCCAACUUAAUCCUA
		UCUACCGCAUCUAACC
		UCCGAUCCAAUAGAUA

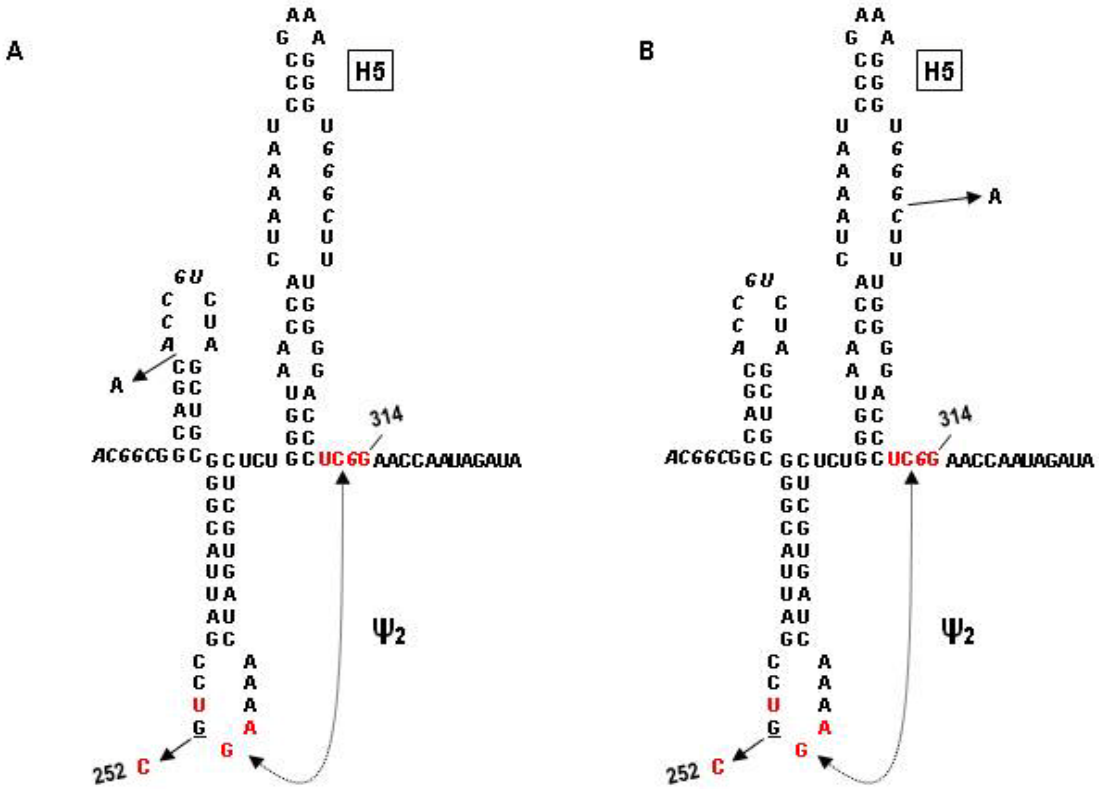
UCC is conserved in the linker region, and forms Ψ_2.

Randomization of the linker region and selection for satC fitness revealed that all satC recovered in the initially inoculated plants contained sequence between positions 312 to 314 that could maintain pairing with at least the GGA in the H4b loop and form Ψ_2.

To provide further evidence for Ψ_2, transcripts of satC mutant C314G, which accumulated to 2% of wild-type levels in protoplasts, were inoculated onto 6 turnip seedlings with helper virus TCV. At 21 dpi, total RNA was isolated from uninoculated leaves. Gel electrophoresis showed all plants contained detectable satC. SatC was cloned and sequenced and 3 different progeny were recovered (Figure 3.4A, B, C, D). Most of the progeny had a guanylate-to-uridylate conversion at position 314. One clone maintained the original cytidylate-to-guanylate change at position 314, but with a new guanylate-to-cytidylate change at position 252 in the H4b loop; the other clone had 314 guanylate changed to wild type cytidylate, and 312 uridylate changed to cytidylate (Figure 3.4). All mutations can re-establish the interaction between UCCG and UGGA in the H4b loop, thus supporting the existance of Ψ_2 (UCCG in the linker and UGGA in the H4b loop), which is required for efficient satC accumulation in protoplasts (Zhang et al., 2006c).

The conserved AACC does not base pair with the 5' end.

The linker *in vivo* genetic selection revealed conservation of the sequence "AACC". The 5' end of satC contains "GGUU" that might potentially interact with this "AACC" sequence, according to mfold. I made two mutations that swapped the AACC


A
H5
314
Ψ2
252
B
H5
314
Ψ2
252

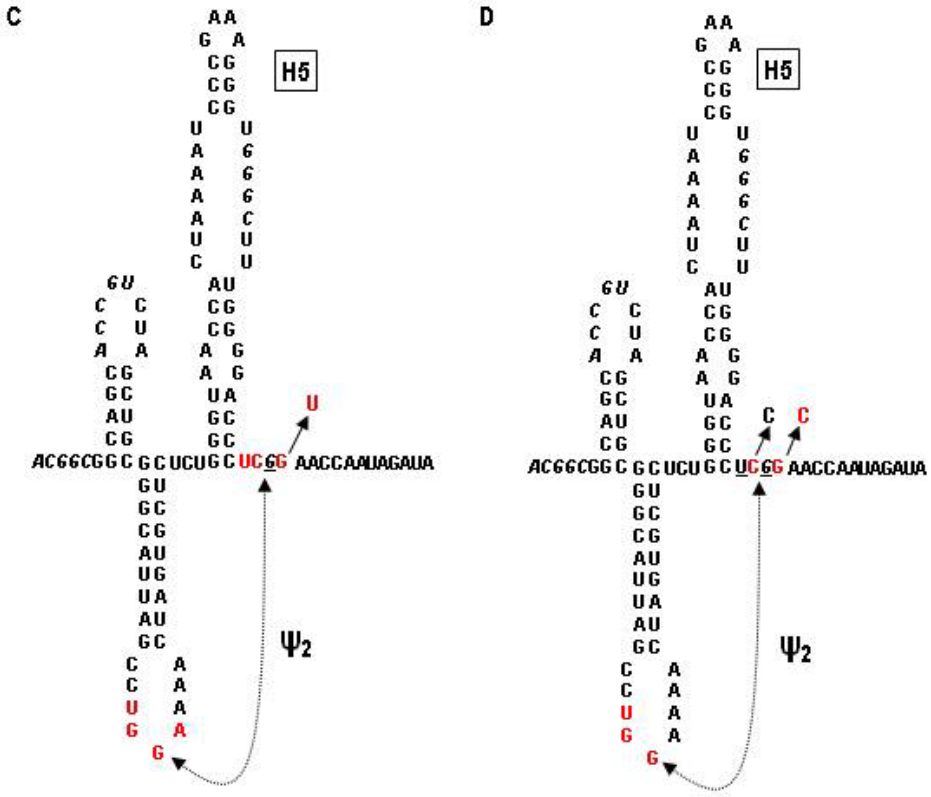


Figure 3.4 UCC is involved in Ψ_2. SatC C314G mutant, which disrupts Ψ_2, was inoculated with TCV genomic RNA onto 6 plants. Functional satC were recovered 3 weeks later. (A-D), Sequences recovered from C314G-infected plants. (A) and (B) each have an adenylate insertion. The sequence in italics is the mutated guanylate from the original cytidylate at position 314. Underlined sequences are sequences that have second site mutations. Arrowheads are the recovered new sequences. Nucleotides in red show the potential base pairs.

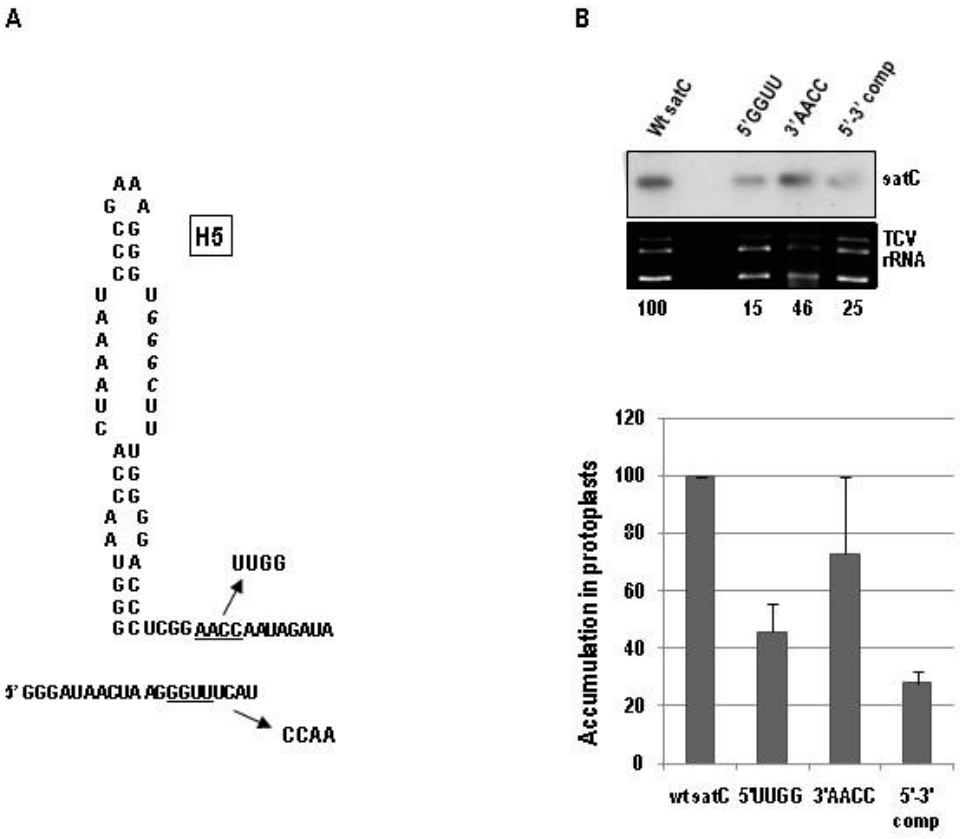


Figure 3.5 Conserved AACC is not interacting with the 5' end. (A) Conserved sequences AACC from H5-Pr linker selection have potential base pairs with GGUU from 5' end. (B) Accumulation at 40 hpi of satRNA with various mutations in 5 x 10^6 *Arabidopsis thaliana* protoplasts prepared from callus cultures. Denatured RNA was hybridized with a [γ-^{32}P]ATP-labeled oligonucleotide probe, oligo7. Mutations are described in the left panel. Data from two independent experiments were normalized to wt satC levels, arbitrarily assigned a value of 100. Bars denote standard deviation.

and GGUU, and one double mutation that combined both 4 base alterations (Figure 3.5A). Transcripts synthesized by T7 RNA polymerase were inoculated into protoplasts. RNA was extracted at 40 hpi and subjected to electrophoresis. SatC was visualized by Northern blot using the satC-specific probe, oligo7. The 5' end GGUU to CCAA mutant accumulated to 46% of wt satC, and the 3' end AACC to UUGG mutant accumulated to 73%. The double mutant only accumulated to 28% of wt levels (Figure 3.5B). These results do not support the conserved AACC sequence base pairing with the 5' end. Since mutating the AACC in the linker region did not have a big effect on satC accumulation in protoplasts, the conserved AACC sequence may be involved in an event other than replication, such as helping satC to move in plants, or suppressing TCV virion formation.

The linker region does not directly interact with the H5 lower stem.

Some structures predicted by mfold have the linker sequence base paired with the H5 lower stem (Figure 3.1C). To determine if the linker region has any interaction with the H5 lower stem, I randomized the linker sequence together with L3, a SELEX winner recovered from a previous SELEX of the satC H5 lower stem (Zhang and Simon, 2005) (Figure 3.2B). Ten first-round winning sequences were recovered (Table 3.4). The sequences were similar to the sequences recovered from the wt background. They also contained conserved UCC sequence and AACC sequence. There were no obvious base pairs between H5 LS and the recovered sequences. The linker region, therefore, appears to have no direct interaction with the H5 lower stem.

Structure of satC H5-Pr linker remains unchanged upon RdRp binding.

To identify the structure of the H5-Pr linker, T7 RNA polymerase transcribed full length satC transcripts were subjected to DMS methylation in the absence or presence of the RdRp, then primer extention by reverse transcriptase with oligo7, which is complementary to the 3’ 21 nt. DMS modifies N1 of adenosines, N7 of guanosines and N3 positions of cytidines. Methylation of these three nucleotides can be detected by primer extension (Tijerina, Mohr, and Russell, 2007). The primer extention products were separated by gel electrophoresis.

In the linker region, UCCG was not modified which means they are either double stranded or buried in a structure (Figure 3.6). The rest of the linker region was detected as single stranded except the two uridines that would not be modified by DMS and the G at position 324. The in vitro structure confirms that the UCCG is involved in base pairs and the rest of the linker region is single stranded.

In the presence of purified recombinant TCV RdRp, the H5-Pr linker region remained the same except for U322 and U326 (Figure 3.6B). These two Us were detected as single stranded in the presence of the RdRp. Because DMS does not modify U, these two stops are not specific. Thus, the linker region did not show major structural changes upon RdRp binding.

In TCV, the H5-Pr linker (4006UCCGAACUAAAAGAUA4021) had a major structure change after RdRp binding (Yuan et al., 2009). TCV fragment F4, which contains TCV 3' end 195 nucleotides, was subjected to in-line probing in the presence and absence of TCV RdRp. RdRp binding caused both enhanced and reduced cleavages

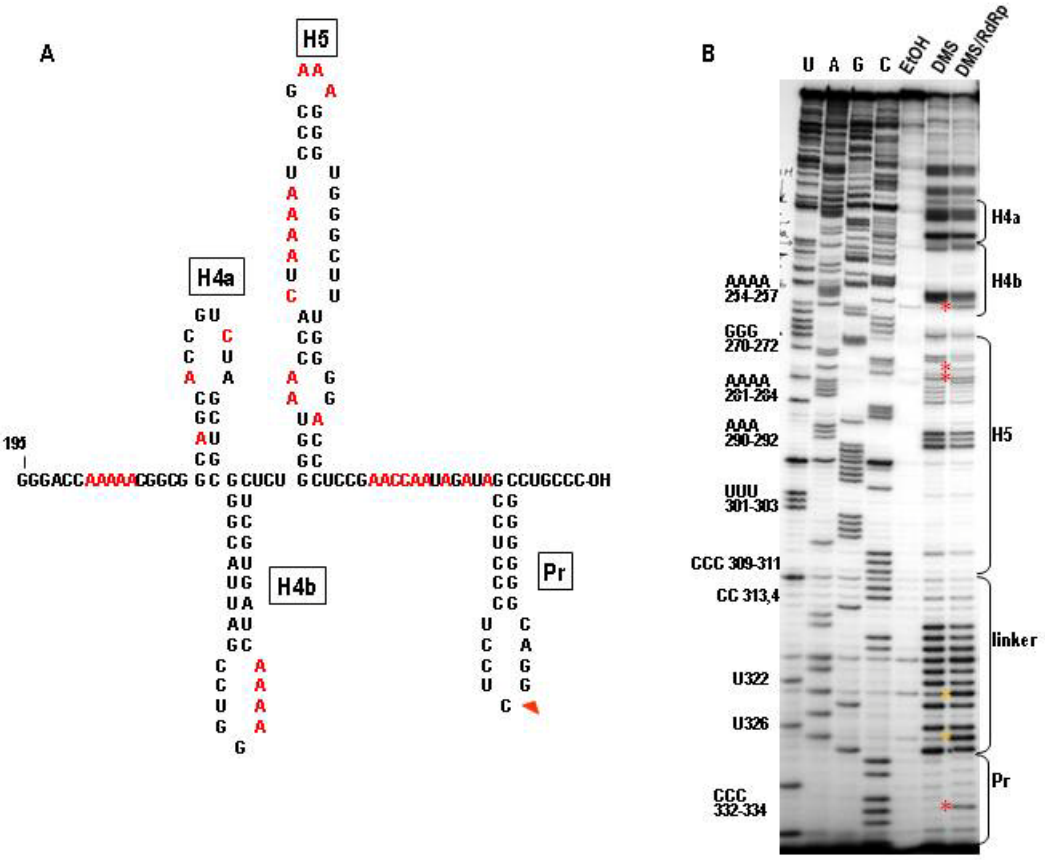


Figure 3.6 DMS structure probing of in vitro transcribed satC with or without TCV RdRp. (A) Phylogenetically conserved satC structure with nucleotides in red showing DMS modification. Due to the nature of different structure probing methods, the flexible nucleotides differ with Figure 2.4. (B) wt satC transcripts were heat denatured and slow cooled to room temperature and subjected to DMS treatment for 10 min followed by primer extension. EtOH is the control. For DMS/RdRp lane, heat denatured and slow-cooled satC were incubated with 1:1 ratio of TCV RdRp at room temperature for 20 min, then subjected to DMS modification followed by primer extension. The hairpins are denoted by brackets at right. Red stars indicate the structural changes after RdRp binding. Orange stars indicate non-specific modifications.

at locations throughout the fragment. Among these structure changes, the H5-Pr linker had altered cleavages in seven nucleotides. U4006, C4007, A4011, and A4014 had increased susceptibility to in-line cleavage after RdRp binding; positions 4018-4020 (GAU) had reduced cleavage (Yuan et al., 2009).

In vitro filter binding assay showed the importance of the Pr and the DR for RdRp binding.

Besides the difference in the H5-Pr linker, TCV also has many other structural changes upon RdRp binding, including reduced flexibility in the H4 asymmetrical loop, the DR analogous sequence (CGGUGG), H4b-H5 linker, H5 lower stem, and the Pr (Yuan et al., 2009). Residues with reduced susceptibility to cleavage after RdRp binding may either have become base paired or the bases were protected by direct interaction with the RdRp. Since structure probing did not reveal major structural changes in satC upon RdRp binding, a filter binding assay was used to identify residues important for RdRp binding.

Radiolabeled satC fragments (see Table 3.5 for fragment sizes and filter binding results) were incubated with TCV RdRp, and then applied to nitrocellulose filters. SatC TSS analogous region binds RdRp with similar affinity as the Pr fragment *in vitro* (TSS analogous region, $3.10 \pm 2.20 \times 10^{-6}$ M; Pr, $3.46 \pm 1.21 \times 10^{-6}$ M). A fragment containing Pr and the TSS analogous region had RdRp binding ability ($2.20 \pm 0.50 \times 10^{-6}$ M) similar to the 3' end 200 nucleotides ($2.30 \pm 1.50 \times 10^{-6}$ M). Both the Pr and the TSS analogous region were beneficial for RdRp binding. When DR sequence was changed, G218C had reduced RdRp binding ability ($5.91 \pm 3.12 \times 10^{-6}$ M) than the full length satC (2.74 ± 1.06

$\times 10^{-6}$ M). The DR may be required for RdRp binding. These results are similar to RdRp filter binding results for the 3' end of TCV (M. Young, A. E. Simon, unpublished).

Table 3.5 RdRp binding of satC fragments

satC fragment	Position	Length (nt)	Kd[a] ($\times 10^{-6}$ M)	100mM NaCl[c]
3'	156-356	201	2.30 ± 1.50	-
Pr	312-356	45	3.46 ± 1.21	-
M1H	156-216	63	-[b]	-
Pr + TSS region	207-356	150	2.20 ± 0.50	-
TSS analogous region	207-323	117	3.10 ± 2.20	-
wt satC	full length	356	2.74 ± 1.06	√
G218C	full length	356	5.91 ± 3.12	√
G218C/C229G	full length	356	8.12 ± 3.68	√

[a] Binding to TCV RdRp. Standard error is indicated
[b] No RdRp binding detected
[c].√ indicates binding performed in the presence of 100mM NaCl, - indicates no NaCl present in the binding reactions.

Discussion

The phylogenetically-derived single-stranded region between H5 and the Pr shared by satC and TCV is structurally conserved in carmoviruses with the exception of *Galinsoga mosaic virus*. In TCV, H5-Pr linker is involved in the conformational change and undergoes a major structural change upon RdRp binding (Yuan et al., 2009).

In this study, satC H5-Pr linker was subjected to *in vivo* SELEX. *In vivo* SELEX

has been successfully applied to satC structure and sequence identification. Our lab has applied *in vivo* SELEX to determine structural and sequence requirements of satC Pr (Guan, Carpenter, and Simon, 2000a; Guan, Carpenter, and Simon, 2000b; Guan and Simon, 2000), structural and sequence requirements of H5 (Zhang and Simon, 2005; Zhang, Stuntz, and Simon, 2004), and structural requirements of M1H with sequence specificity in its flanking regions (Sun and Simon, 2003; Sun, Zhang, and Simon, 2005; Zhang and Simon, 2003b).

Randomization of positions 312 to 327 and selection for satC fitness in plants revealed that all satC recovered in the initially inoculated plants contained some sequence conservation. Sequences between positions 312 to 316 could maintain pairing with at least the UGG in the H4b loop and up to two additional flanking bases. Second site changes recovered from mutation C314G showed changes either at H4b loop or the sequences flanking the 3' side of H5 to maintain the base pairing between H4b loop and the linker region. Genetic assays confirmed the existence of this base paring (Zhang et al., 2006b). Besides the conserved UCC, the H5-Pr linker also had AACC recovered in most progeny (6/15 had AACC and 5/15 had AAC in the first round; by the second round, all had AACC), with unknown function. Except for the conserved UCC and AACC, the remaining sequence was apparently random. The linker sequence did not appear to pair with the 5’ end or the H5 lower stem, as suggested by mFold.

The discovery of this long distance interaction between the linker and H4b loop (Ψ_2) is intriguing. SatC conformational changes and the structure of TCV were studied later based on the discovery of Ψ_2. *In vitro* transcribed satC transcripts do not contain any of the phylogenetically conserved hairpins, but they do contain Ψ_2, which stabilizes the

pre-active structure (Zhang et al., 2006b). The pre-active structure is necessary for robust accumulation in protoplasts, and allows newly synthesized satC (+)-strand RNAs to keep their 3' ends sequestered and promoters unavailable to the RdRp. Thus only the initially infected (+)-strand will be transcribed by the RdRp, reducing potential deleterious mutations. Disruption of Ψ_2 is not sufficient, but appears necessary for satC conversion to the active structure (Zhang et al., 2006b).

DMS structure probing showed the H5-Pr linker is single-stranded and does not undergo significant changes upon RdRp binding which is different with the linker in TCV. As mentioned in Chapter I, TCV and satC have different needs for conformational changes. The altered susceptibility to in-line cleavage in the presence of the TCV RdRp indicates that RdRp interaction with the TCV 3' end causes a dramatic structural change, which likely prohibits ribosomes access to the RNA, freeing the 3' terminal sequences to form replication-required Ψ_1, and initiating replication (Yuan et al., 2009). Apparently, satC H5-Pr linker is not involved in such function.

Besides the H5-Pr linker, our lab has recently determined that mutations in the DR-equivalent sequence in TCV disrupt single-site RdRp binding to the region (X. Yuan, M. Young, and A. E. Simon, unpublished data). The DR sequence may also be important for RdRp binding to satC. *In vitro* structure probing using DMS did not show any major structural changes in the DR region upon RdRp binding, while filter binding of different satC fragments confirmed the importance of DR in RdRp binding. The lack of any structural changes with DMS strcture probing may be due to the inability of DMS to access the DR region. Specific template recognition by RdRp is essential for faithful genome replication. Future work is required to identify the RdRp binding site in satC.

Based on current knowledge, how satC converts from the pre-active structure to its active structure to initiate (-)-strand synthesis remains unknown. The process might require a *trans*-acting element, like RdRp or CP binding.

CHAPTER IV

The relationship between H4a and H4b in satC

Introduction

RNA viruses contain elements in both their 5' and 3' UTRs that participate in translation, a function not required by small subviral RNAs, suggest that subviral RNAs might evolve to differentially use genomic-derived sequences that are no longer required for helper virus-related functions. Elucidation of such functional differences in the utilization of shared sequences could therefore lead to important insights into the relationship between viral and associated subviral RNAs. Recent identification of a 3'-proximal translational enhancer partially contained within the shared region that is necessary for gene expression of TCV, but not satC, has complicated assignment of element function and/or importance based on sequence and structural conservation.

Based on biochemical structure mapping, the pre-active structure of satC does not contain hairpins H4a, H4b, H5, or Pr or Ψ_1, but does contain Ψ_2 (Zhang et al., 2006a). However, *in vivo* genetic selection (SELEX) revealed that the structures and some sequences within Pr and H5 are critical for satC accumulation in plants and protoplasts (Carpenter and Simon, 1998; Zhang et al., 2006b; Zhang et al., 2006c). In addition, fragment exchanges with the related carmovirus CCFV suggested that H4a and H4b form a functional unit in satC (Zhang et al., 2006b).

SatC and TCV have different conformational switches. In TCV, the RdRp-

mediated conformational switch could restrict translation and promote replication (Yuan et al., 2009). Untranslated satC has two conformations: pre-active and active structures (Simon and Gehrke, 2009). This pre-active structure needs to undergo a conformational switch to convert to an active configuration and initiate (-)-strand synthesis. The sequence just upstream of H4a, the DR, appears to be important for the conformational switch between pre-active and active structures. Identifying the DR location in the satC pre-active structure will help us elucidate the pre-active and active structures in satC.

SatC evolved to eliminate the TSS, although not exclusively for reducing ribosome binding. Does satC need to maintain the rest of the TCV sequence for its own replication? *In vitro* structure probing could not detect either of these TCV-derived hairpins for satC, while genetic assay proved the importance of the hairpins. Deletion of H4a or H4b accumulated to low levels in protoplasts (6% and less than 1% of wt level, respectively). SatC with reverse complement of H4a or H4b accumulated to 42% and 3% respectively. SatC with analogous H4a and H4b from CCFV accumulated to higher level than satC with individual hairpins from CCFV, which indicates H4a and H4b function as a single unit (Zhang et al., 2006b). To determine the sequence and structure specificity in satC replication, H4a and H4b within the TSS region were subjected to *in vivo* functional selection. The results discussed in this chapter show: (i) in the presence of wt H4b, sequence in the H4a region is selected to form an H4a-like stem-loop; (ii) satC H4a and H4b are functional either as two adjacent stemloops (as is the case in wt satC) or as a single hairpin, with little sequence similarity to wt satC required. This suggests that given the opportunity, viral RNAs can rapidly evolve topologically distinct elements to perform similar functions.

Materials and Methods

Construction of satC mutants

For the generation of plasmids G219C, C228G, G219C/C228G, Δ216-220, C229G/m216-219, C229G/Δ216-220, C226U, C243U, and C243U/G263A, PCR reactions were performed with a common 5' primer (T7C5') homologous to the 5' end of template pT7C(+) (see Table 4.1 and 4.2). SpeI (or BstEII for C243U/G263A) and NcoI digested PCR products were ligated into similarly digested pT7C(+), replacing the endogenous fragment. To obtain pG263A and pC226U/G263A, pT7C(+) and pC226U and were treated with NcoI and SpeI, the small fragments purified, then ligated into NcoI/SpeI digested pC243U/G263A vector backbone. To construct pC226U/C243U/G263A, PCR reaction was performed using template pC226U and primers T7C5’ and C243U/G263A. The PCR product was digested with NcoI and BstEII, and ligated into similarly digested pT7C(+) vector backbone. All constructs were confirmed by DNA sequencing.

In vivo SELEX (H4a/H4b)

In vivo genetic selection was performed as previously described. To generate the template for in vitro transcription of satC with random sequence in place of the 18-nt H4a region (Figure 4.1; positions 222 to 239), two fragments were generated by separate PCRs with pC(+) as a template. The 5' fragment was produced by using primers T7C5', which contains a T7 polymerase promoter at its 5' end, and H4aL3'. The 3' fragment was

generated by using primers BstE2R5' and oligo7, which is complementary to the 19 nt at the 3' end of satC. To generate satC with a randomized H4a/H4b region (positions 222 to 266), separate PCRs using pC(+) and either oligos T7C5' and H4aH4bL3' or oligos BstE25' and oligo7 were performed. PCR products were subjected to electrophoresis, purified using QIAQuick MinElute columns (Qiagen, Valencia, CA), digested with BstEII (all enzymes from New England Biolabs, Ipswich, MA, except where noted), phenol-chloroform extracted, and ligated together to produce full-length satC cDNA. These satC cDNAs with randomized H4a or H4a/H4b were in vitro transcribed using T7 RNA polymerase; TCV genomic RNA was *in vitro* transcribed from SmaI-linearized pT7TCVms. Both satC and TCV genomic RNA transcripts contain precise 3' and 5' ends. For the first round of selection, 2 µg of wt TCV transcripts and 5 µg of satC transcripts with specific randomized sequences were inoculated onto each of 30 turnip seedlings (Turnip Hybrid Just Right; Gurney's, Greendale, IN). Total RNA was isolated from uninoculated leaves after 21 days and added to six new turnip seedlings for an additional 21-day infection; this was repeated for a total of five rounds. For rounds 1, 2, 3, and 5, satC RNA was reverse transcribed, amplified by PCR in the presence of Pyrostase polymerase (Molecular Genetic Resources, Tampa, FL) using primers T7C5' and oligo7, cloned into SmaI-linearized pUC19, and subjected to sequencing. Selected round 3 H4a/H4b "winner" sequences were transcribed from their pUC19-based plasmids and self-evolved in plants (in the presence of TCV genomic RNA) for six rounds; RNA was reverse transcribed, cloned, and sequenced after rounds 1 and 6. Selected round 3 H4a/H4b winner sequences and selected winners from the sixth round of self-evolution also were subjected to direct competition in six turnip seedlings with RNA reverse

transcribed, cloned, and sequenced 21 days later. Competition experiments between wt satC and SELEX control experiments were simultaneously performed in which individual plants were coinfected with TCV RNA and RNA of each satC SELEX winner, to verify that the satC transcripts were functional and able to be cloned in the absence of competition with wt satC. To obtain clones of satC with randomized H4a or H4a/H4b for use as controls in protoplast experiments, ligated satC cDNAs (described above) were directly cloned into the SmaI site of pUC19.

Four way *in vivo* SELEX

In vivo functional selections with random sequence replacing the DR (CGGCGG), the H4b loop (UGGA), H5 lower stem and flanking sequence (CCC, GGGUCCG) were performed as described above. Two PCRs were carried out using pC(+) as the template and primers T7C5', 3'L3wayS, and 5'R3wayS and oligo7. The PCR products were purified by agarose gel, digested with SpeI, and ligated together to make full length satC DNA. For the first round, five micrograms of satC transcripts containing randomized sequences were directly synthesized using T7 RNA polymerase and inoculated onto each of 30 turnip seedlings along with 2 µg of TCV genomic RNA transcripts. Immediately before inoculation, the RNA was mixed with equal amount of 2x infection buffer and rubbed onto two leaves of each seedling (10 µL each leaf) using gloved hands. Total RNA was extracted from uninoculated leaves at 21 dpi. Full-length satC was amplified by reverse transcription-PCR using primers T7C5' and oligo7 and then cloned into the SmaI site of pUC19 and sequenced. For the second round of SELEX, equal amounts of total RNA from each first round plant were pooled and inoculated onto six new turnip

seedlings at a concentration of about 5 µg per plant. Total RNA was extracted 21 dpi and full length satRNA was cloned into pUC19 as described above.

Accumulation of viral RNAs in protoplasts

TCV genomic RNA and satC transcripts were in vitro transcribed with T7 RNA polymerase using either plasmids pT7TCVms and pT7C(+) that were linearized with SmaI or directly from PCR products. Protoplasts (5×10^6) prepared from callus cultures of Arabidopsis thaliana ecotype Col-0 were inoculated with 20 µg of TCV genomic RNA transcripts with or without 2 µg of satC RNA transcripts using polyethylene glycol (PEG)-$CaCl_2$, as described previously. Total RNA isolated from protoplasts at 40 h postinoculation (hpi) was subjected to RNA gel blot analysis. This 40-h time point was used since accumulation of RNA is still increasing. The RNA was probed with [γ-^{32}P]ATP-labeled oligonucleotide oligo13, which is complementary to positions 3950 to 3970 of TCV genomic RNA and positions 250 to 269 of satC, or [γ-^{32}P]ATP-labeled oligo7, which is complementary to positions 338 to 356 of satC.

Table 4.1 Constructs used in Chapter IV

Name	Description
pC(+)	Full length satC cloned into pUC19
pT7C(+)	T7 promoter and full length satC cloned into pUC19
pTCVms	Full length TCV cloned into pUC19
C226U	C to U change at position 226
C243U	C to U change at position 243
G263A	G to A change at position 263
C226U/G263A	C to U change at position 226 and a G to A change at position 263
C243U/G263A	C to U change at position 243 and a G to A change at position 263

Table 4.2 Oligonucleotides used in Chapter IV

Name	Position	Sequence	Polarity
T7C5’	1-19	GTAATACGACTCACTATAGGGATAACTAAGGGTTTCA	+
Oligo7	338-356	GGGCAGGCCCCCCGTCCGA	-
Oligo13	249-269	GTTACCCAAAGAGCACTAGTT	-
H4a 1mut (C226U)	220-270	CAGAGAGC*ACTAGT*TTTCCAGGCTAATGCCCGCAGCTA GACGGTACTGCCG	-
H4b 2mut (C243U/G263A)	237-290	TCGGGATTTTAGT*GGTTACC*CAGAGAGT*ACTAGT*TTTCC AGGCTAATACCCGCA	-
H4aL3'		T*GGTTACC*CAGAGAGCACTAGTTTTCCAGGCTAATGCC CNNNNNNNNNNNNNNNNNCGCCGTTTTTGG	-
H4aH4bL3'		T*GGTTACC*CAGANNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNCGCCGTTTTTGG	-
BstE2R5'		G*GGTAACC*ACTAAAATCC	+
3’L3wayS		GAGAGCACTAGTTTNNNNGGCTAATGCCCGCAGCTAG ACGGTGCTGNNNNNNTTTTTGGTCCCATTTACCC	-
5’R3wayS		AAACTAGTGCTCTCTNNNTAACCACTAAAATCCCGAA AGGGTGGGCTTTGGGGANNNNNNNAACCAATAGATA GCC	+

Underlined nucleotides indicate the mutations.

Italics are restriction enzyme recognition sites.

Results

SatC H4a SELEX results in retention of a stem-loop.

Our lab's previous work showed that satC has different configurations in vitro. A conformational switch model was proposed (Zhang et al., 2006a). Replacement of satC H4a and H4b with the analogous hairpins of CCFV led to accumulation levels greater than with individual hairpin replacements, which implies that H4a and H4b function as one unit (Zhang et al., 2006b). Structure probing did not confirm that H4a and H4b exist in the satC pre-active structure (Zhang et al., 2006a), while deletion of H4a or H4b from satC markedly impaired satC accumulation in protoplasts (Zhang et al., 2006b).

To further explore the sequence/structural requirements of satC H4a and H4b and their interrelationship, H4a was subjected to in vivo genetic selection. SELEX was done in collaboration with Dr. David Kushner at Dickinson College. Full length satC RNA with 18 randomized bases (Figure 4.1A; positions 222 to 239) replacing H4a was inoculated onto 30 turnip seedlings along with TCV genomic RNA. Total RNA was extracted from top new leaves from each of the 30 plants after 30 days. To allow for further evolution, total RNA from all 30 first round plants were combined and inoculated onto six new turnip seedlings. 21 days later, total RNA was extracted, combined, and re-inoculated onto six additional seedlings. These steps were repeated for a total of five rounds. SatC progeny from round 1, 2, 3, and 5 were cloned and sequenced (Table 4.3).

Sequence J emerged as a SELEX winner, as it comprised nearly one-half of the sequenced clones for round 3. By round 5, only sequence J was isolated, making it the functional winning sequence in this experiment. mFold prediction of full-length satC with

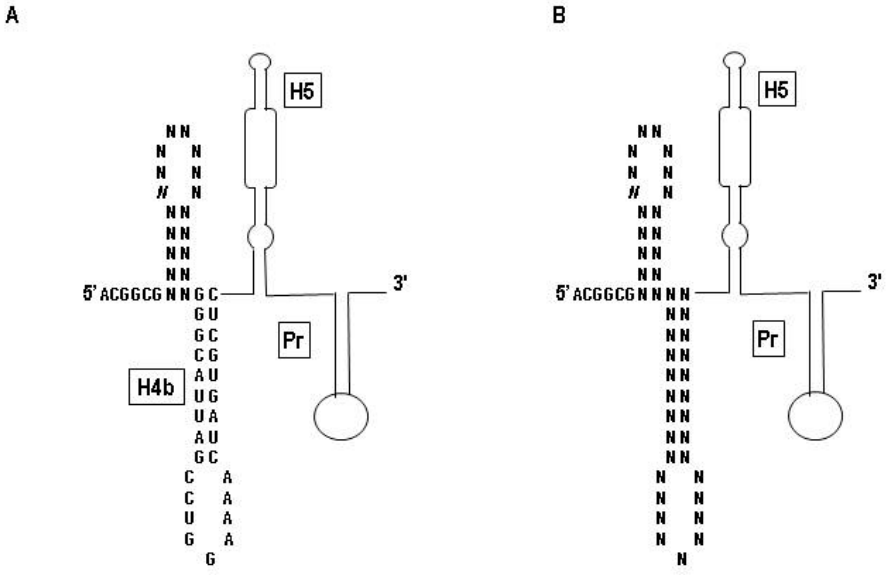


Figure 4.1 H4a and H4a/H4b in vivo SELEX. (A) H4a, (B) H4a and H4b were randomized. Transcripts containing the random sequences were inoculated onto plants. Total RNA was extracted 21 dpi and satC progeny was cloned. N means random nucleotides.

sequence J in place of H4a suggests that sequence J can fold into a stem-loop (figure 4.2A).

The emergence and evolution of sequence J in plant infection could reflect more robust replication or an enhanced ability to support rapid TCV movement (Zhang and Simon, 2003a). To determine if sequence J was capable of enhanced replication in protoplasts compared with related sequences not found in later rounds, wt satC, sequences E, G, and J, and two satCs with random H4a sequences (Rd1 and Rd2) cloned from the initial population of randomized H4a sequences prior to SELEX in plants were individually inoculated into protoplasts along with TCV genomic RNA and assessed 40 h later for (+)-strand accumulation (Table 4.3). Rd1 and Rd2 did not accumulate to detectable levels, indicating the importance of specific sequences/structures in this region for satC amplification (table 4.3). In contrast, satC with H4a sequences E, G, and J accumulated to nearly wt levels. Altogether, these results strongly suggest that a hairpin in this region contributes to replication of satC.

However, lack of detectably enhanced accumulation by winning sequence J compared to related sequences E and G suggests that the two unique alterations in sequence J in the H4a loop (position 229) and upstream H4a-flanking (not SELEXed)

Table 4.3 SatC sequences obtained from H4a SELEX rounds 1, 2, 3, and 5 sequence (position 215) contribute to a different satC attribute that allows for its strong selection by the helper virus in infected plants (Figure 4.2A).

The two new alterations in sequence J that enhance fitness *in planta* compared with progenitor sequence G would allow for a pseudoknot to form that is similar to TCV Ψ_3. Sequence O, which was present in the second-round SELEX, also has the potential to

Table 4.3 SatC sequences obtained from H4a SELEX rounds 1, 2, 3, and 5

		Number per round				Accumulation in protoplasts
		1	2	3	5	
WT	aaaaacggcgGCAGCACCGUCUAGCUGC					100
A	aaaaacggcgGCUUAACAUCCGAACGAA	1				
B	aaaaacggcgGAGCGGUAUCAUGCGACG	1				
C	aaaaacggcgGAGCGGUAUCAUGCGCCG	1				
D	aaaaacggcgGAGCGGUAUCAUGCUCCG		1			
E	aaaaacggcgGAGCGGUAUCAUGCUCCG		1	4		95 ± 4
F	aaaaacggcgGAGCGGUAUCAUGCGCCG			1		
G	aaaaacggcgGUGCGGUAUCAUGCGCCG		1	4		91 ± 7
H	aaaaacggcgGUGCGGUUUCAUGCGCCG			2		
I	aaaaacggcgGUGCGGUAUCAUGCUCCG			1		
J	aaagacggcgGUGCGGUCUCAUGCGCCG			10	20	91 ± 10
K	aaaaacggcgAAAGAAAGACGUUUUAAC	2				
L	aaaaacggcgAGUCUUAAAGAGCCGACC	3				
M	aaaaacggcgAGUCUUAACGAGCCGACC		1			
N	aaaaacggcgCUAGGAUAUGGCCGAAAA	1				
O	aaaaacggcgAACUCCCCGUAAAGGAGU		1			
P	aaa.gcggcg.ACGGACCGGUCGC			1		
Rd1	UAUGUCAAGUCGUUCAAG					0
Rd2	UACAAAGAACAAGACCGA					0

(Nucleotide sequences were from Dr. David Kushner)

Nucleotides in lower case were not subjected to SELEX.

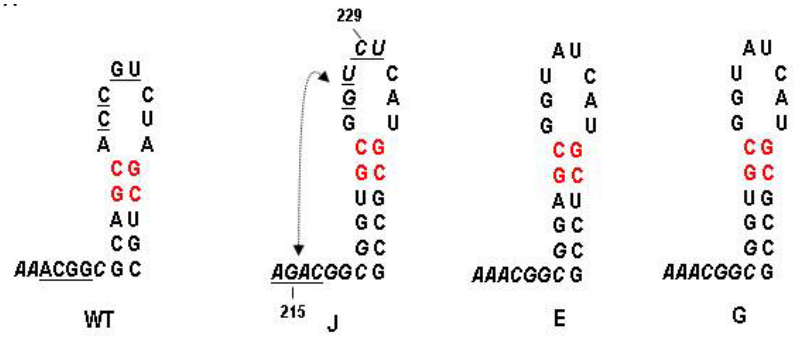


Figure 4.2 H4a *in vivo* SELEX results in retention of a stem-loop. Potential pseudoknots are indicated and complementary sequences are underlined. Red nucleotides indicate identical sequence between wt and respective SELEX sequences.

form a similar TCV-like pseudoknot. Since the importance of Ψ_3 could not be shown in protoplasts as described in Chapter II, Ψ_3 may be required for functions other than replication.

Testing possible interaction between H4a and H4b.

The results of the H4a SELEX suggested that a hairpin is preferred in this location for satC replication in protoplasts, while emergence of sequence J indicated that the region is also involved in an additional effect that led to selection of this particular sequence/structure. As described above, sequence J contained some sequence identity with wt satC in the H4a stem, suggesting possible involvement of specific residues in an alternative conformation. We previously determined that replacing both H4a and H4b with the analogous hairpins from CCFV enhanced accumulation compared with replacing individual hairpins. Analysis of wt H4a and H4b sequence in satC revealed a possible alternative pairing involving sequence from both hairpins (Figure 4.3A; positions 224 to 228 and 261 to 265) that is also present in CCFV H4a and H4b. Interestingly, sequence J also maintains similar possible alternative pairing between the two hairpins.

To test if this alternative pairing is important for satC amplification in protoplasts, mutations were designed that disrupted potential base-pairing in one structure while maintaining the other structure (Figure 4.3B). C226U, which was predicted to not significantly affect the H4a/H4b structure or the alternative structure, had no effect on satC replication. C243U, which was predicted to have little impact on the stem of H4b in the H4a+H4b structure and no obvious effect on the stem of the alternative structure, reduced satC accumulation by 33%. G263A should impact both the stem of H4b in the

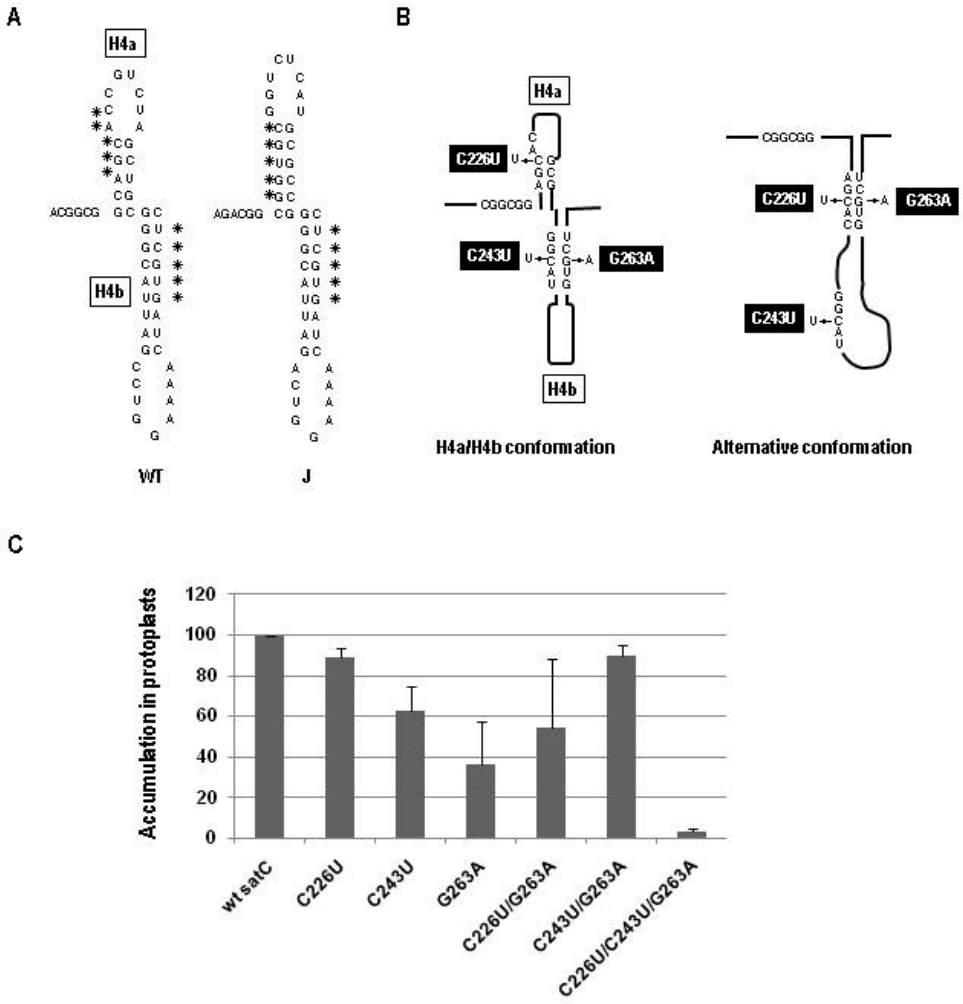


Figure 4.3 Two potential conformations of satC H4a and H4b. (A) H4a and H4b regions of satC. Left, wt sequence; right, H4a SELEX winner sequence J. Asterisks denote possible alternate base-pairing interactions, as depicted for wt satC in (B). (B) Two possible conformations for satC H4a and H4b. Boxes indicate satC mutations generated for analysis of accumulation in protoplasts. (C) Accumulation of satRNA (with various mutations) in protoplasts at 40 hpi. Mutations are indicated in (B). Data from at least three independent experiments were normalized to wt satC levels. Standard deviations are indicated.

H4a/H4b structure, as well as potentially disrupt the stem of the alternate structure. Of the three point mutations, G263A had the greatest impact on satC accumulation in protoplasts, reducing accumulation to 29% of wt. Combining C226U with G263A, which should be compensatory for the alternative structure, did not significantly improve the replication of satC (2 sample t test p value = 0.08 when C226U/G263A and G263A were compared). However, combining C243U with G263A, which was compensatory for the H4a/H4b structure, restored satC accumulation to 69% of wt (2 sample t test p value = 0.015 when C243U/G263A and G263A were compared; 2 sample t test p value = 0.05 when C243U/G263A and C243U were compared). These results support the importance of maintaining the H4b stem for satC accumulation in protoplasts and do not provide evidence for the alternative structure. However, combining C226U, which by itself had no effect on satC accumulation, with C243U/G263A dramatically reduced accumulation to 3% of wt (Figure 4.3C). This result was unexpected and suggested that H4a and H4b is more complex in terms of sequence, structure, and function than previously thought.

Two distinctive functional structures result from *in vivo* SELEX of satC H4a/H4b.

To further explore the relationship between satC H4a and H4b, satC with 45 randomized nucleotides representing the entire H4a and H4b region (Figure 4.1B; positions 222 to 266) was transcribed and RNA was inoculated onto 30 turnip seedlings along with TCV genomic RNA. SELEX was performed by Dr. David Kushner. After 3 weeks, RNA was extracted from new leaves of 30 turnip plants. To allow for possible further evolution, four additional SELEX rounds were completed as described earlier for the H4a SELEX.

Table 4.4 SatC sequences obtained from H4a/H4b SELEX rounds 1,2,3,and 5

		Number per round				Accumulation in protoplasts
	H4a/H4b	1	2	3	5	
WT	GCAGCACCGUCUAGCUGCGGGCAUUAGCCUGGAAAACUAGUGCUC					100
A_{ab}	AAGCGGAAUGUCGGUACCAAGCCUUGCUGUAGCAUCCCACAAACC	8				
B_{ab}	AUUUGACUGCCAAAGCAAAUUGACCGCUCGGGCGCACCAACGCA.	5				
C_{ab}	GGAGCGAGGUAGCUAGAAUCCCUGGCAAGAAGCAGAGAAAGACCG	2				
D_{ab}	GGAGCGAGGUAGCUAGAAUCCCUGGCAAGAAGCAGAGGAAGACCG	2				
E_{ab}	GGAGCGAGGUAGCUAGAAUCCCUGGAAAGAAGCAAAGGAAGACCG	1				
F_{ab}	UUACGUCGGAGUUCAACUGCCAGUAAAACUUUUGAGUGAUAAGA.	2				
G_{ab}	UAUGCUCGAACAAAAAAUGGGGCAGCCGCACACCACUCCCGACUA	1				
H_{ab}	UAUGCUCGAACGAAAAAUGGGGCAGCCGCACACCACUCCCGACUA	1				
I_{ab}	GCCUGUGCAAACUUCUCGUGCAUCAAAGCGGCUAACUGUCAGGAC	2				
J_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAAUACCGGAAAU	1	5	3		
K_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAAAAUACCGGAAAU			1		42 ± 5
L_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAAAAUACCGGAAAC				3	
M_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAAUACCGGAAA.			2		
N_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAAUACCGGAACU			7		
O_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAAUACCGGAAAC			3		
P_{ab}	CUGAUCACUAAAAUUCGAGGUGAAGCGCUCUAGAAUACCGGAAAC			2		
Q_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAGUACCGGAAAC			1		9 ± 1
R_{ab}	GUGAUCGCUAAAUUUUCGUGGGAAUCGCUCGAUAAUAGUCGAGAU	1				
S_{ab}	GCAAACUGUCAUCACUUGCAUACUUUUCCACGUACCGACGCACAC	1				
T_{ab}	UCGAUUCACGGAAGGAACGGUAGUUUAAAAUAUUUACGCGCUUCU	1				
U_{ab}	AUCCAUCCUAUGUCCAAACUCGAAUCCAA.....	1				
V_{ab}	AGACUUAGGUAGACCAACAAAUCUAGUUUCACCAACUCCCGGCGA			3	2	
W_{ab}	AGACUUAGGUAGACCAACAAAUCUAGUUUCACCAACUCCCGGCGC				1	
X_{ab}	AGACUUAGGUAGACUAACAAAUCUGGUUUCGCCAACUCCCGGCGA			1		64 ± 8
Y_{ab}	AGACUUAGGUAGACUAACAAAUCUGGUUUCACCAACUCCCGGCGA				1	
Z_{ab}	AAAUGCAGCCGUCAUCUGUUUCUCGGGCAGACUGACAAGACAGGC			3	4	68 ± 14
AA_{ab}	AAAUUCAGCCGUCAUCUGUUUCUCGGGCAGACUGACAAGACAGGC				1	
$Rd_{ab}1$	UGCUAAUCCCCAUUCGCAUCAGAGUAGCAAUGCGCAAUAUAAAUUA					0
$Rd_{ab}2$	ACGAUAGAGAGCUAAUCCUAUCAUUUCUAUGUUCUUAUAAUUAAC.					0

(Nucleotide sequences were from Dr. David Kushner)

Nine distinct sequences, some with variants, were obtained from 29 clones recovered from round 1 plants (Table 4.4). Unlike the H4a SELEX, no clear winner emerged after five rounds. Of these sequence winners, only related sequences Z_{ab}/AA_{ab} could potentially fold into H4a-like and H4b-like stem-loops, requiring the DR to form the H4a-like stem (Figure 4.4B). In contrast, sequences $J_{ab}Q_{ab}$ and $V_{ab}Y_{ab}$ were predicted to fold into single hairpins that also incorporated the DR into the 5' stem base (Figure 4.4C). mFold prediction of full-length satC with either sequence X_{ab} (exemplifying the hairpin) or Z_{ab} (exemplifying the two stem-loops) in place of H4a/H4b suggests that these SELEXed sequences can fold into the single hairpin structures (Figure 4.4D).

To determine the ability of selected sequences to accumulate in protoplasts, representatives of the three groups of round 3 winners (sequences K_{ab}, Q_{ab}, X_{ab}, and Z_{ab}) were transcribed and RNAs inoculated into protoplasts with TCV genomic RNA. While no round three winner was able to accumulate to wt satC levels, accumulation of all selected sequences was significantly enhanced compared with satC containing either of two random H4a+H4b sequences ($Rd_{ab}1$ and $Rd_{ab}2$).

The poorest-replicating sequence was Q_{ab} (9% of wt satC levels), which differed in sequence from K_{ab} (42% of wt satC levels) by only 3 nt. To determine which nucleotides were responsible for the enhanced accumulation of sequence K_{ab}, point mutations found in K_{ab} were introduced into sequence Q_{ab} (Table 4.5). The C-to-U change of the end of the 45-nt region (Q_{ab}-C45U) improved accumulation in protoplasts by threefold, while G33A enhanced replication to K_{ab} levels. Interestingly, G35A or a combination of G35A and G33A was more beneficial for Q_{ab} replication than when these base changes were combined with C45U. Since none of these changes is within the single

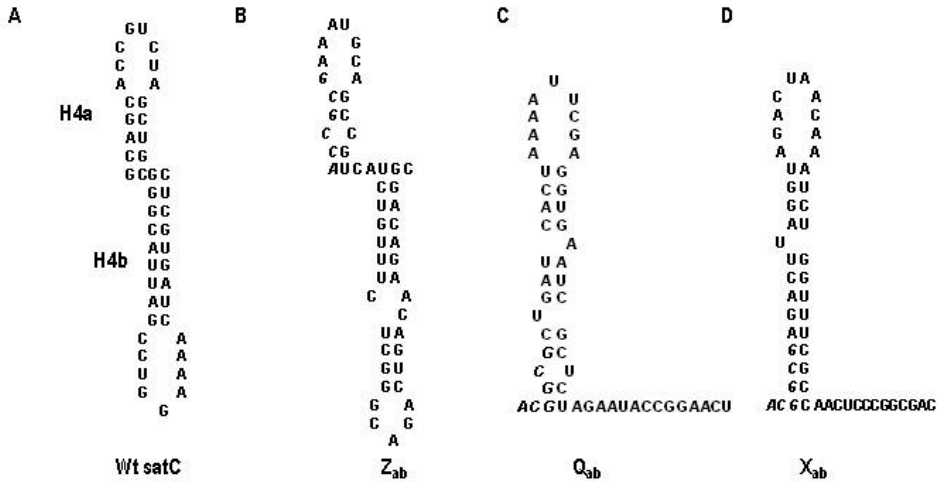


Figure 4.4 Possible structures adopted by the H4a/H4b SELEX winners. (A) Structure of wt H4a and H4b. (B) mFold-predicted structure of Z_{ab} showing a H4a/H4b-like structure. (C) mFold-predicted structure of Q_{ab} and (D) mFold-predicted structure of X_{ab} showing single hairpins.

hairpin or the putative 2 region, this result suggests that these downstream residues are interacting elsewhere in the satRNA, possibly as part of a pre-active structure, promoting replication fitness.

Self-evolution of three H4a/H4b SELEX 3rd winners.

Since K_{ab}, X_{ab}, and Z_{ab} did not replicate to wt satC levels in protoplasts, the three satC species were allowed to further evolve *in planta*. Transcripts from satC clones with sequences K_{ab}, X_{ab}, and Z_{ab} were independently transcribed and separately inoculated onto three turnip seedlings along with TCV genomic RNA. After three weeks, total RNA was extracted and a portion used to infect new plants. This protocol was repeated for a total of six rounds, with RNA from rounds 1 and 6 used to clone satC. After one round, the original K_{ab}, X_{ab}, and Z_{ab} sequences were preferentially cloned along with additional variants for K_{ab} and X_{ab}. After six rounds, no original K_{ab}, X_{ab}, and Z_{ab} sequences were detected, but rather newly evolved variants K_{ab}B/K_{ab}C, X_{ab}B/X_{ab}G, and Z_{ab}A/Z_{ab}D were recovered (Table 4.6). For Z_{ab}A Z_{ab}D, only one change (A37C) was found in the 45-nt region originally subjected to randomization. However, second-site mutations accumulated just 5' of the DR in all but one variant, suggesting that this upstream A-rich

Table 4.5 H4a/H4b SELEX winner Q_{ab} based derivatives sequence supports the function of the H4a/H4b region. Compared to parental K_{ab}, new variants K_{ab}B/K_{ab}C had A33C and A44C changes and K_{ab}C featured a U31C alteration that could have been retained from K_{ab}A after round 1. These changes are all found 3' of the hairpin. New variants X_{ab}B and X_{ab}G differed from parental X_{ab} by containing G44A. In addition, X_{ab}B and X_{ab}F had positions 13 to 16 (5'-ACUA) converted to CCC; these alterations occurred in the loop of

Table 4.5 H4a/H4b SELEX winner Q_{ab} based derivatives

		Accumulation in protoplasts
	H4a/H4b	
WT	GCAGCACCGUCUAGCUGCGGGCAUUAGCCUGGAAAACUAGUGCUC	100
K_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAAAAUACCGGAAAU	42 ± 5
Q_{ab}	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAGUACCGGAAAC	9 ± 1
Q_{ab}-G33A	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAAAGUACCGGAAAC	37 ± 4
Q_{ab}-G35A	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAAUACCGGAAAC	56 ± 5
Q_{ab}-G33A+G35A	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAAAAUACCGGAAAC	68 ± 2
Q_{ab}-C45U	CUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAGAGUACCGGAAAU	27 ± 6
$Rd_{ab}1$	UGCUAAUCCCCAUUCGCAUCAGAGUAGCAAUGCGCAAUAUAAAUUA	0
$Rd_{ab}2$	ACGAUAGAGAGCUAAUCCUAUCAUUUCUAUGUUCUUAUAAUUAAC.	0

(Nucleotide sequences were from Dr. David Kushner)

Table 4.6 SatC sequences obtained from self-evolution of satC H4a/H4b SELEX 3rd round winners Kab and Xabs

		Number per round		Accumulation in protoplasts
	H4a/H4b	1	6	
WT	aaaaacggcgGCAGCACCGUCUAGCUGCGGGCAUUAGCCUGGAAAACUAGUGCUC			100
K_{ab}	aaaaacggcgCUGAUCACUAAAAUUCGAGGUGAAUCGCUCUAAAAUACCGGAAAU	5	0	42 ± 5
$K_{ab}A$	aaaaacggcgCUGAUCACUAAAAUUCGAGGUGAAUCGCUCCAAAAUACCGGAAAU	3	0	
$K_{ab}B$	aaaaacggagCUGAUCACUAAAAUUCGAGGUGAAUCGCUCUACAAUACCGGAACU	0	8	56 ± 7
$K_{ab}C$	aaaaacggagCUGAUCACUAAAAUUCGAGGUGAAUCGCUCCACAAUACCGGAACU	0	1	
X_{ab}	aaaaacggcgAGACUUAGGUAGACUAACAAAUCUGGUUUCGCCAACUCCCGGCGA	4	0	64 ± 8
$X_{ab}A$	aaaaacggcgAGACUUAGGUAGACUAACAA.UCUGGUUUCGCCAACUCCCGGCGA	2	0	
$X_{ab}B$	aaaaacggcgAGACUUAGGUAGCCC.ACAAAUCUGGUUUCGCCAACUCCCGGCAA	0	4	98 ± 2
$X_{ab}C$	aaaaacggcgAGACUUAGAUAGCCC.ACAAAUCUGGUUUCGCCAACUCCCGGCAA	0	2	95 ± 4
$X_{ab}D$	aaaaacggcgAGACUUAGGUAGCCC.ACCAAUCUGGUUUCGCCAACUCCCGGCAA	0	2	96 ± 4
$X_{ab}E$	aaaa.cggcgAGACUUAGGUAGCCC.ACCAAUCUGGUUUCGCCAACUCCCGGCAA	0	1	
$X_{ab}F$	aaaaccggcgAGACUUAGGUAGCCC.AC.AAUCUGGUUUCGCCAACUCCCGGCAA	0	1	
$X_{ab}G$	aaaaacggcgAGACUUAGGU.....CAAAUCUGGUUUCGCCAACUCCCGGCAA	0	1	98 ± 1
Z_{ab}	aaaaacggcgAAAUGCAGCCGUCAUCUGUUUCUCGGGCAGACUGACAAGACAGGC	6	0	68 ± 14
$Z_{ab}A$	aaaaacggcgAAAUGCAGCCGUCAUCUGUUUCUCGGGCAGACUGACCAGACAGGC	0	1	64 ± 11
$Z_{ab}B$	aauaacggcgAAAUGCAGCCGUCAUCUGUUUCUCGGGCAGACUGACCAGACAGGC	0	9	77 ± 8
$Z_{ab}C$	aacaacggcgAAAUGCAGCCGUCAUCUGUUUCUCGGGCAGACUGACCAGACAGGC	0	3	59 ± 10
$Z_{ab}D$	agcaacggcgAAAUGCAGCCGUCAUCUGUUUCUCGGGCAGACUGACCAGACAGGC	0	2	
$Rd_{ab}1$	aaacggggUGCUAAUCCCCAUUCGCAUCAGAGUAGCAAUGCGCAAUAUAAAUUA			0
$Rd_{ab}2$	aaacggggACGAUAGAGAGCUAAUCCUAUCAUUUCUAUGUUCUUAUAAUUAAC.			0

(Nucleotide sequences were from Dr. David Kushner)

Nucleotides in lower case were not subjected to SELEX.

the putative single hairpin. $X_{ab}G$ was distinctive in that it featured a 6-nt deletion, reducing the size of the modeled loop.

To determine if further evolution *in planta* enhanced replication of these satC, two or three evolved sequences from parentals K_{ab}, X_{ab}, and Z_{ab} were assessed for accumulation in protoplasts (Table 4.6). While accumulation of Z_{ab} variants was not enhanced, a variant derived from K_{ab} ($K_{ab}B$) featured enhanced accumulation, and X_{ab} variants $X_{ab}B$ and $X_{ab}G$ accumulated to near-wt levels. This suggested that the changes in the terminal loop, poly(A) upstream region, and/or G44A of X_{ab} enhanced replication leading to selection in plants.

Ψ_2 is confirmed in H4a/H4b SELEX derived two distinctive functional structures.

In wt satC, the H4b loop forms the important pseudoknot (Ψ_2) with sequence adjacent to the 3' side of the H5 stem (5'UCCG). When winning sequences were examined for nucleotides that could maintain this pseudoknot, both $J_{ab}Q_{ab}$ and Z_{ab}/AA_{ab} contained sequence that could maintain this pairing or extend it (3'-AGGC and 3'-GGGCU, respectively). While $V_{ab}Y_{ab}$, which was structurally similar to $J_{ab}Q_{ab}$, did not have sequence that could form Ψ_2, it did contain the similarly situated sequence 3'-CGGC. Examination of the sequence at the base of H5 in the $V_{ab}Y_{ab}$ winners revealed that seven of eight contained a second-site mutation at satC position 312 that converted the Ψ_2 sequence 5'-UCCG to 5'-GCCG, which was capable of reforming the pseudoknot. These results support selection for maintenance of Ψ_2 in the H4a/H4b SELEX winners.

Ψ_2 is not directly interacting with the DR.

Mutations in the DR that reduced satC accumulation had strong, negative effects on transcription of full-length satC by the TCV RdRp in vitro, indicating its importance for activation of the pre-active conformation of the satRNA (Zhang et al., 2004). This information, combined with our results from Chapter II indicating that the DR is not base-paired in Ψ_3 as is the comparable sequence in TCV, indicates that the DR is likely interacting in an alterative location in the satC pre-active structure.

Previous mutations in Ψ_2 partner sequences resulted in strong susceptibility of DR guanylates to single-stranded-specific RNase T1. In addition, mutations in the DR produced more intense double-stranded specific RNase V1 cleavages in the Ψ_2 UCCG sequence (Zhang et al., 2004). Since there are no obvious pairing partners for the DR in this region, one possibility was that the DR participated in a triplet base or other non-canonical interaction with the Ψ_2 region. To determine if alterations in the adjacent H5 lower stem also affect the susceptibility of the DR region to single-and double-stranded specific enzymes, a compensatory exchange was generated in the center of the three GC pairs (G271C/C310G; Fig. 4.5A) and the resultant satC transcripts subjected to RNA structure mapping (Fig. 4.5B). G271C/C310G mutant was generated and subjected to structure probing by Dr. Guohua Zhang. The compensatory changes had pleiotropic effects throughout the H5/H4b/H4a region of the pre-active structure including new susceptibility of cytidylates in the H4b and H4a regions to single-stranded specific RNase A, and increased RNase T1 cleavage of guanylates at the base of H5. These results confirmed our previous findings that H5 and the upstream hairpins do not exist in the conformation assumed by satC transcripts in vitro (Zhang et al., 2004). Importantly, these

alterations in the lower stem of H5 resulted in strong RNase T1 and RNase A cleavages in the DR, indicating that the DR was assuming a more unpaired configuration.

To determine if a direct interaction exists between the DR and the adjoining regions at the base of H5 and Ψ_2, in vivo genetic selection was performed to generate functional sequence variations in Ψ_2, as well as the lower stem of H5 and the DR. The rationale was that if DR residues are hydrogen bonding in this region, then alternative sequences that maintain Ψ_2 in the pre-active structure and the H5 lower stem in the satC active structure would permit sequence variation in the DR that might correlate with particular sequences in the other regions. This was the first time we selected discontinuous regions at the same time. For this assay, the DR (CGGCGG), Ψ_2-partner sequences (H4b loop UGGA, H5 flanking UCCG) and the H5 lower stem (GGG/CCC) were randomized (Figure 4.5C), and transcripts containing the randomized sequences inoculated with TCV onto 30 turnip seedlings. Total RNA extracted from uninoculated leaves of all plants at 21 dpi contained satRNA detectable on ethidium bromide-stained agarose gels. Cloning and sequencing viable satC from 9 plants revealed 13 different first round sequences (Table 4.7). All recovered sequences could reform the 3 bp H5 stem, indicating the importance of maintaining these pairings in the satC active structure. Three of the sequences could additionally form 3 to 5 bp Ψ_2, which for two sequences involved pairing with the non-randomized adenylate just downstream of the randomized sequence flanking H5. A partial DR sequence (CGG) was recovered in four of the selected sequences, with one (clone 1-6) recovering the wt DR (CGGCGG).

Total RNA isolated from the 30 plants was pooled and inoculated onto six new turnip seedlings for direct competition between first round sequences. Total RNA was

again extracted from uninoculated leaves at 21 dpi and satC species were cloned. Twenty-six clones were recovered from the six plants revealing five second-round winning sequences. All recovered clones contained the wt DR sequence; all but one contained sequence that could pair to form a 3 bp lower H5 stem; and all but one could form a 5 bp Ψ_2 (Table 4.7). Neither Ψ_2 nor the H5 lower stem contained sequence that resembled their wt versions.

All second round winners and one satC control with randomly chosen non-selected sequence were assayed for accumulation in protoplasts. Four of the five winners accumulated above background (control) levels with clone 2-2 accumulating to 26% of wt levels. This clone, the most recovered sequence in the 2^{nd} round, had a wt DR, 5 bp Ψ_2 and a CGC/GCG H5 lower stem. Clone 2-2 accumulated nearly 2-fold better than clone 2-4, which had identical DR and H5 sequences, but where a G:U pair replaced a G:C pair in Ψ_2. Clone 2-2 also accumulated nearly 12-fold better than clone 2-5, which contained the same DR and H5 sequences while regaining only a 3 bp Ψ_2. Clone 2-3, which contained identical DR and Ψ_2 as Clone 2-2, had only a 2 bp lower H5 stem, which accumulated to only 17%. Since all recovered second round satC had wt DR sequence and non-wt H5 stem and Ψ_2 sequences, direct interaction between the DR and these regions was not supported. In addition, base-specificity in the H5 and Ψ_2 regions is required for efficient satC accumulation since the top recovered winner still accumulated 4-fold less efficiently than wt satC.

In summary, while compensatory mutation analysis strongly supported base-pairing between the DR analogous sequence and H4a loop sequence in TCV (McCormack et al., 2008), similar analysis of sequences in satC did not support either of

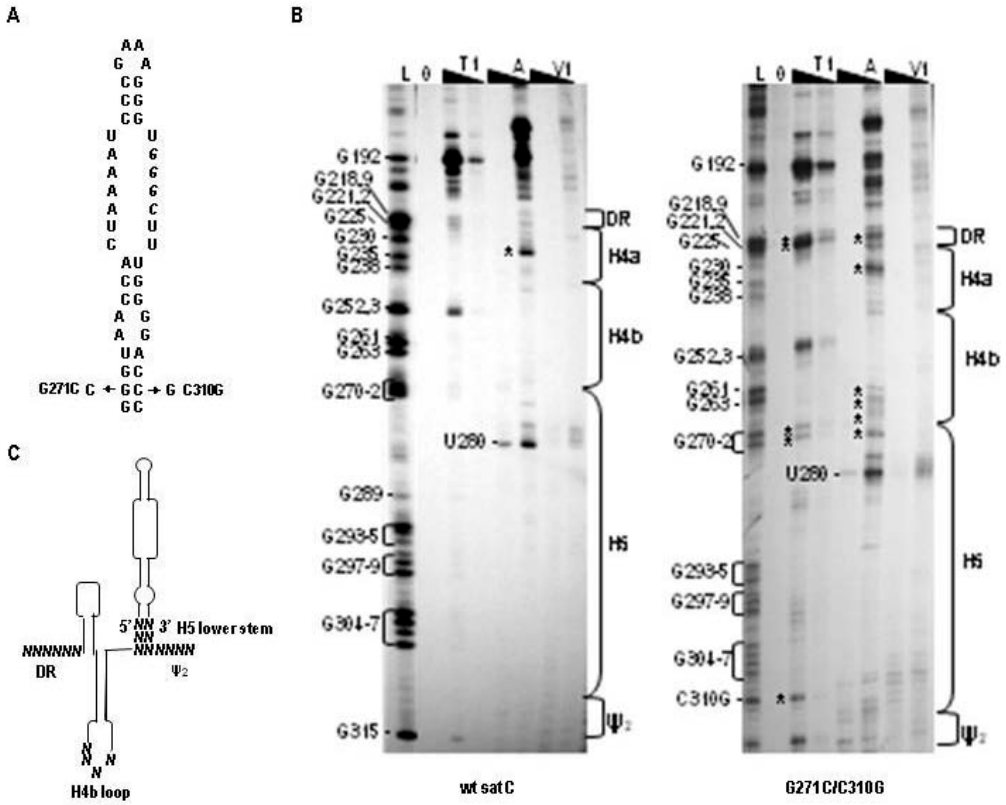


Figure 4.5 Structure probing of satC transcripts with and without compensatory mutations in H5. (A) SatC H5 showing positions of the compensatory alterations G271C and C310G. (B) Structure probing of wt and mutant satC (preformed by Dr. Guohua Zhang). Wt satC and the satC containing G271C/C310G were subjected to limited cleavage with two concentrations each of RNase T1, RNase A and RNase V1. L, RNase T1 ladder; 0, no added enzymes. High and low concentrations of each enzyme are indicated by the extent of the filled triangles above the lanes. Positions of guanylates determined from the T1 ladder reactions are given at left. Residues within the DR, TCV-analogous hairpins and Ψ_2 are denoted by brackets at right. U280G is a strong RNase A cleavage site. Asterisks denote enzyme-specific cleavages that differ between mutant and wt satC. (C) Residues (N) in satC that were randomized for in vivo genetic selection.

Table 4.7 SatC Four Way SELEX

	Clone	Winning sequences[a]					# recovered	Reformed:			Accumulation[b] (%)
		DR	H4b loop	H5 lower stem 5'	H5 lower stem 3'	Ψ_2		DR[c]	Ψ_2	H5 stem	
	wt	CGGCGG	cUGGAaa	GGG	CCC	UCCGaa					100
Round 1	1-1	AAAAACUCGG	cUCCGaa	UGA	UCG	UCGGaa	1	√	√	√	ND[d]
	1-2	AACUCGG	cUCCGaa	UGA	UCG	UCGGaa	1	√	√	√	ND
	1-3	AGAAAG	cUCUCaa	CUG	CAG	AUUCaa	3			√	ND
	1-4	AAGAAAG	cUCUCaa	CUG	CAG	AUUCaa	1			√	ND
	1-5	CUCCAC	[e]Aaa	UUG	CAU	UGGCaa	4			√	ND
	1-6	CGGCGG	cUCUCaa	CGC	GCG	UGAUaa	3	√	√	√	ND
	1-7	UGAUCC	cACUGaa	AGG	CCA	GACGaa	3			√	ND
	1-8	CCGGAC	cCAACaa	GUU	AAC	GGUCaa	3	√		√	ND
	1-9	AGCUGG	cGAACaa	UGU	ACA	CGGUGaa	1			√	ND
	1-10	AGCUGG	cGAACaa	UGU	ACA	CGGCga	1			√	ND
	1-11	AGCUGG	cGAACaa	UGU	ACA	CGGUaa	1			√	ND
	1-12	GCGCGA	cAUAAaa	GUG	CAC	GCGCaa	5			√	ND
Round 2	2-1	CGGCGG	cUAUCaa	CGC	GCG	UGAUaa	13	√	√	√	12.5 ± 3.5
	2-2	CGGCGG	cUGUCaa	CGC	GCG	UGACaa	10	√	√	√	26.5 ± 4.9
	2-3	CGGCGG	cUCUCaa	CGC	GCA	UGAGaa	1	√	√		17.5 ± 3.5
	2-4	CGGCGG	cUGUCaa	CGC	GCG	UGAUaa	1	√	√	√	15 ± 4.2
	2-5	CGGCGG	cUAUCaa	CGC	GCG	UGACaa	1	√	√	√	2.5 ± 0.7
Random		CGACAC	cGUGAaa	CUU	GCC	UAUCaa	-	-	-	-	2

[a]Sequences are shown in 5' to 3' orientation. Only upper case residues were derived from randomized sequence. Underlined sequences are complementary for Ψ_2 formation
[b]40 hours after inoculation of Arabidopsis protoplasts
[c]Based on prior analyses, a minimal DR is "CGG"; "UGG" is the DR analogous sequence in TCV
[d]ND, not determined
[e]Clone 1-5 had a 4 nt deletion

two possible pseudoknots. Mutations that disrupted Ψ_2 displaced the DR from its natural paired configuration (Zhang et al., 2006b) as did mutations in the lower loop of H5. In vivo genetic selection did not support any specific non-canonical interaction with these elements, suggesting that the DR is pairing elsewhere.

The DR may not interact with an upstream satD-derived sequence region X.

The second site changes recovered from the competition between the satC mutants with TCV TSS sequence from Chapter II, G157A, U159C/C220U/C229U/C249A/C268U and U159C/C220U/C229U/C249A/U273C, suggested that the DR may interact with upstream satD-derived region. mFold predicts two possible interactions between the DR region and the satD-derived region. There are potential base pairs between DR sequence AAACGGC (214-220) and sequence GCUGUUU (157-163; referred to as region X; Figure 4.6A), or between GGCGGCAGC (218-226) and GCUGUUUCC (157-165, Figure 4.6C).

Mutations were made to test the potential interaction. C158 was changed to a G, the potential partner sequence G219 was changed to a C, then the two mutations were combined to potentially re-form the interaction (Figure 4.6A). The C to G change at position 158 reduced satRNA accumulation in protoplasts by 30%, while the double mutation (25% of wt) did not compensate for the low accumulation of G219C (15% of wt, Figure 4.6B). From the protoplast data, we rule out one possible DR-region X interaction. The other possible interaction (Figure 4.6C) needs to be tested. There is always the possibility that the DR region has a strict sequence requirement, which would not allow us to see effects from compensatory mutations.

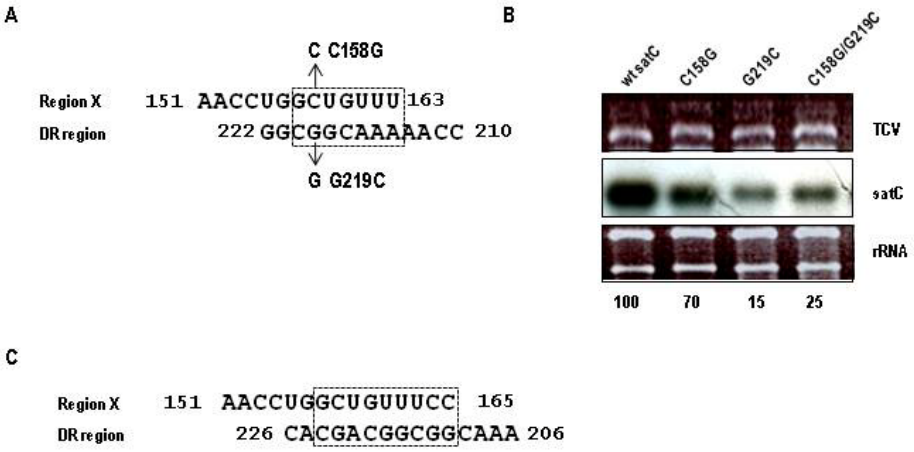


Figure 4.6 Possible interactions between DR and region X. (A) One of the two possible interactions between DR and region X with the box showing the base pairs. A set of compensatory mutations (C158G, G219C, and C158G/G219C) were designed to test this interaction. (B) Accumulation at 40 hpi of satRNA with various mutations in 5 x 10^6 Arabidopsis thaliana protoplasts prepared from callus cultures. Denatured RNA was hybridized with a [γ-^{32}P]ATP-labeled oligonucleotide probe oligo13, which is complementary to both satC and TCV. Data from at least three independent experiments were normalized to wt satC levels. (C) Another possible interaction between region X and the DR region that was not tested further.

Discussion

Although structure probing did not confirm that H4a and H4b exist in the satC pre-active structure (Zhang et al., 2006a), both H4a and H4b are important for satC accumulation in protoplasts (Zhang et al., 2006b). H4a is an important *cis* element for robust satC RNA accumulation in plants; deletion of H4a reduced satC accumulation to 6% of wt, whereas replacing wt H4a with the reverse complement, which alters the loop sequence and maintains the wt stem, reduced accumulation to about 40% of wt levels; Replacing wt H4a with H4a from CCFV, which alters most of the stem and the lower bases in the loop, also reduced accumulation of satC to about 40% of wt levels in protoplasts (Zhang et al., 2006b). H4b is involved in forming Ψ_2 (Chapter III; Zhang et al., 2006b).

H4a SELEX winner J, as well as related intermediate-round sequences G and E, contain selected sequence that can model into a single hairpin, and all sequences support satC accumulation to near wt levels in protoplasts. The presence of sequences G and J in the later rounds and the emergence of sequence J as the SELEX winner suggests that a U:G pairing in the middle of the stem, found in J and G, is preferred over the A:U pairing found in sequence E. One of the two base differences between sequences J and G, located in the loop of the hairpin, enhances loop sequence similarity with wt H4a (5'-GUCU). This new base alteration in sequence J, along with the additional alteration upstream of the H4a region, did not enhance accumulation levels in protoplasts, suggesting that these changes benefit the satRNA or the interaction of the satRNA with TCV in a different process.

Since these two differences in sequence J would permit formation of a TCV-like pseudoknot (Ψ_3 in TCV), Dr. Jiuchun Zhang and I examined the analogous putative pseudoknot in wt satC for a role in accumulation in protoplasts as described in Chapter II. Genetic analyses did not support a TCV-like pseudoknot in this region of wt satC. This pseudoknot in TCV (Ψ_3) is important for TCV accumulation in protoplasts (McCormack et al., 2008), and disruption of the pseudoknot affects the 3' translational enhancer and ribosome binding to this element (Stupina et al., 2008). A similar pseudoknot may be required in satC for functions other than replication in protoplasts, such as movement in plants and silencing suppression.

Interestingly, several of the hairpins selected in the H4a SELEX incorporated either a portion of or the complete DR as part of the stem of the hairpin. Since there is no evidence for formation of H4a in the pre-active structure of satC (Zhang et al., 2006a), it is possible that the DR sequence is involved in RdRp binding.

Replacement of satC H4a and H4b with the analogous hairpins of CCFV led to accumulation levels greater than with individual replacements, which implies that H4a and H4b function as one unit (Zhang et al., 2006b). Because the compensatory mutagenesis studies of H4a and H4b did not definitely resolve the conformation of this region of satC, the entire 45-nt region was subjected to SELEX and two types of winning sequences were obtained. The related sequences Z_{ab}/AA_{ab} are predicted to fold into H4a/H4b-like stem-loops, while $J_{ab}Q_{ab}$ and $V_{ab}Y_{ab}$ selected sequences model into single hairpins. Examination of H4a/H4b-like Z_{ab} revealed a second-site mutation in the DR (G218C) that appeared in round 3 and was retained through round 5 and the six rounds of self-evolution. G218C was also seen in AA_{ab} (round 5). Interestingly, G218C reduces

accumulation of wt satC by 72% and severely abrogates its *in vitro* transcription by the TCV RdRp (Zhang et al., 2004). However, when a chimeric satC with H5 replaced by CCFV H5 was passaged in plants to improve fitness via acquisition of a second-site mutation(s), a more fit satC was recovered possessing G218C (Zhang et al., 2006b). One possibility is that acquisition of G218C allows for alternate control of the switch between pre-active and active structures, which contributes to fitness of Z_{ab}.

The plasticity of the satC H4a and H4b elements, represented by SELEX winners Kab and Xab and sequences evolved from them (especially X_{ab}B due to its wt-level accumulation in protoplasts), is striking because of the size of the region randomized (45 nt), which represents 13% of its genome. Since the H4a/H4b region in TCV contributes to the TSS and binds ribosomes (McCormack et al., 2008; Stupina et al., 2008), plasticity would not be expected for this region in TCV. Notably, structural plasticity may have been possible because satC does not form the TSS and is not translated; therefore satC H4a/H4b sequences obtained in the SELEX only need to function in terms of replication and spread (and not translation).

The DR was originally identified as an important sequence-specific element when *in vivo* selection of sequences flanking both sides of M1H resulted in recovery of wt DR sequence CGGCgg (the terminal two guanylates were not subjected to randomization; Sun et al., 2005). Fitness of satC in plants, high-level accumulation in protoplasts, and efficient transcription of satC transcripts by the TCV RdRp in vitro all correlated with wt DR sequences (Sun et al., 2005; Zhang et al., 2006a; 2006b). While compensatory mutation analysis strongly supported base-pairing between the DR and H4a loop sequence in TCV (McCormack et al., 2008), similar analysis of sequences in satC did not

support the pseudoknot (Chapter II). Mutations that disrupted Ψ_2 displaced the DR from its natural paired configuration (Zhang et al., 2006b) as did mutations in the lower loop of H5 (this chapter). *In vivo* genetic selection did not support any specific non-canonical interaction with these elements, suggesting that the DR is pairing elsewhere. The possible interaction between DR and satD-derived sequence region X is under investigation with collaborator, Dr. David Kushner.

CHAPTER V

The formation of satC dimers in the absence of TCV RdRp

Introduction

Some linear satellite RNAs associated with plant viruses can have both multimeric and monomeric forms, like TCV (Carpenter, Cascone, and Simon, 1991; Simon and Howell, 1986), CMV (Kuroda et al., 1997), *peanut stunt virus* (*Bromoviridae*, *Cucumovirus* genus) (Linthorst and Kaper, 1984), *Cymbidium ringspot virus* (CymRSV; *Tombusviridae, Tombusvirus* genus) (Dalmay and Rubino, 1994), *Beet black scorch virus* (BBSV; *Tombusviridae, Necrovirus* genus) (Guo et al., 2005). Circular satellite RNAs also have multimer forms, which are replication precursors that are cleaved and ligated to form circular monomeric progeny (Diener, 1991). Unlike circular satellite RNA multimers, there is no reported significance for linear satellite RNA multimers.

Formation of linear satellite RNA multimers is considered to be a type of RNA recombination. There are two mechanisms that lead to the generation of RNA recombinants. One is replicase driven template switching (also called copy choice), and the other is nonreplicative RNA breakage and rejoining.

Replicase driven template switching mechanism is the most accepted model for multimer formation (Figure 5.1) (Carpenter, Cascone, and Simon, 1991). The replicase initiates at the 3' end and elongates on the template. The replicase will terminate when it meets a structural obstacle, or at the natural end. Before the nascent strand can detach

from the replicase, the replicase reinitiates either on a new template or on the same template, without release of the nascent strand. The reason this model is generally accepted is that, when dimer junctions are sequenced, there are additional nucleotides frequently found at the junction site which are thought to be added by the replicase before reinitiation in a non-template fashion.

The other mechanism of RNA recombination is non-replicative RNA breakage and rejoining (Chetverin, 1999). Nonreplicative RNA recombination was reported in a cell-free Qβ system (Chetverin et al., 1997; Chetverina et al., 1999), as well as *poliovirus* (Gmyl et al., 1999; Gmyl et al., 2003) and *Bovine viral diarrhea virus* (BVDV) (Gallei et al., 2005). The structure of recombinants suggested that they were generated by an attack of the 3' terminus of the 5' fragment at the phosphodiester bonds within the 3' fragment, so RNA recombination occurs via a transesterification reaction. Additional nucleotides were found, though rare, for bacteriophage QB and *poliovirus*. Both groups are claiming the additional non-viral nucleotides were caused by T7 RNA polymerase generating foreign inserts during their laboratory preparation of transcripts.

In this chapter, I will show the occurence of satC dimers in the absence of the viral RdRp, which contradicts the current replicase-driven copy choice model.

Materials and methods

Construction of T-DNA-based plasmids for agroinfiltration

Full length TCV and TCV with a non-functional RdRp (GDD mutant; GDD active site of enzyme mutated to inactive GAA), satC and satC mutants, and CP ORF

were inserted into a binary vector pCASS4-HDVδ (Padmanaban and Rao, 2007) (Table 5.1). pCASS4-HDVδ has double 35S promoters and hepatitis delta virus ribozyme, and was provided by ALN Rao. To construct full length satC, satC mutants and wt TCV, pCASS4-HDVδ was digested with StuI and NcoI, treated with Mung bean nuclease to create blunt-ended products followed by CIP treatment. Full-length TCV and full-length satC fragmetns were obtained by SmaI digestion of plasmids pTCV6 and pC(+), respectively, and agarose gel purification. The fragments were then inserted into the StuI site and nuclease-treated NcoI site of pCASS4-HDVδ, to get pCASS4-TCV and pCASS4-satC. For construct pCASS4-GDD, pGDD was digested by BamHI and SpeI followed by agarose gel purification. The fragment was then ligated into similarly treated pCASS4-TCV, replacing the endogenous fragment. To construct pCASS4-CP, CP ORF was amplified by PCR with 5' end primer CPF and 3' end primer CPR (Table 5.2), followed by StuI- and NcoI- digestion and agarose gel purification, and was inserted into StuI and NcoI treated pCASS4-HDVδ. To construct satC mutant pCASS4-Δ5'GGGA, satC was amplified by PCR reactions with 5' end primer Δ5G and 3' end primer oligo7. To construct satC mutant pCASS4-Δ3'CCC, satC was amplified by PCR reactions with 5' end primer C5' and 3' end primer Δ3C. Blunt-end PCR products were agarose gel purified and T4 PNK kinased, and then inserted into the StuI site and nuclease-treated NcoI site of pCASS4- HDVδ. To construct pCASS4-Δ22, pCASS-C6 was treated by SnaI and NcoI digestion, followed by Mung bean nuclease treatment, and self-ligation. All constructs contain, in sequential order, a double 35S promoter, gene of interset, hepatitis delta virus ribozyme sequence, and a *Nos* terminator. Both 5' and 3' end sequences cloned into pCASS4- HDVδ contain no vector sequences after in vivo transcription, except for

pCASS4-CP, which has two non-viral nucleotides at the 3' end. All sequences were confirmed by sequencing.

Agroinfiltration

Agroinfiltration was performed as described by Padmanaban and Rao (2006) with minor modifications. T-DNA based plasmids were transformed into the *Agrobacterium* strain EHA105 either by heat-shock or electroporation, grown on LB plates containing 50 μg/ml kanamycin and 30 μg/mL rifampicin, and incubated at 30°C for 48 hours. Individual colonies were grown at 30°C in LB containing 50 μg/mL kanamycin, 20 μg/mL rifampicin, 20 μM acetosyringone, and 10 mM Mes, pH5.7 for 16-20 hours with vigorous shaking. The bacteria were spun down at 3000×*g* for 10 min at room temperature and the pellet was resuspended in 10 mM $MgCl_2$. The resuspended cells were again spun down at 3000×*g* for 10 min at room temperature and the pellet was resuspended in 10 mM $MgCl_2$ with 100 μM of acetosyringone, and the OD_{600} reading was adjusted to 1.0. The bacteria were then kept at room temperature for at least 3 h without shaking. These cultures were infiltrated into abaxial surface of the fully expanded *Nicotiana benthamiana* leaves using 1 ml needle-less syringes.

RNA extraction and Northern blots

Total RNA was isolated from leaf samples 5 dpi as described in Chapter II. Briefly, 0.5 g infiltrated leaf tissue was ground to a fine powder in liquid nitrogen, then each sample was transferred to a 1.5 mL Eppendorf tube and extracted with 0.55 mL of RNA extraction buffer [25 mM EDTA, 0.4 M LiCl, 1% (w/v) SDS, 0.2 M Tris-HCl, pH 9.0] and 0.55 ml of H_2O-saturated phenol by vigorous vortexing. The sample was centrifuged at 13,000 rpm for 5 min at 4 °C. The aqueous layer was collected into a fresh

microcentrifuge tube and re-extracted with with equal volume of phenol/chloroform. Total RNA was precipitated by adding 1/10 volume of 3M NaOAc (pH 5.3) and 2.5 volume of 100% ethanol. The sample was centrifuged at 13,000 rpm for 15 min at 4°C and the pellet was re-suspended in 0.3 ml of 2 M LiCl. Following centrifugation at 13,000 rpm for 5 min at 4°C, the pellet was dissolved in 0.3 mL distilled H_2O and precipitated with 1/10 volume of 3M NaOAc (pH 5.3) and 2.5 volume of 100% ethanol again. The tubes were kept at −80°C for 2 hour and centrifuged at 13,000 rpm for 15 min at 4°C. The RNA pellet was washed with 70% ethanol, air dried, re-suspended in desired volume of sterile distilled water, and stored at−80°C. RNA concentration was estimated by measuring the absorbance of the sample at 260 nm with a spectrophotometer.

RT-PCR and cloning

Total RNA was reverse transcribed by M-MLV reverse transcriptase. To generate first strand cDNA, 5 μg of total RNA was mixed with 10 pmol of oligo 13, the reverse transcription buffer provided by the manufacturer, dNTP, and DTT to a final concentration of 0.5 mM and 10 mM respectively. The mixture was heated at 75°C for 10 min and cooled briefly at room temperature for oligo and RNA annealing. The solution was then incubated at 37°C for 1 hour with 1 μL of M-MLV reverse transcriptase (USB, 200 U/μL). The reaction was stopped by heating at 75°C for 15 min and placed immediately on ice. SatC multimers were amplified by PCR in the presence of Pyrostase polymerase (Molecular Genetic Resources, Tampa, FL). PCR was performed using 8 μL of cDNA synthesized as described above with primers oligo13 and H5F in a 100 μL reaction. PCR products were blunt-end treated with Klenow, kinased with T4

Polynucleotide kinase, and cloned into SmaI-linearized pUC19 vector, and subjected to sequencing.

Protein extraction from plants and Western blotting analysis

Ground leaves into a fine powder under liquid nitrogen, and add 1 mL protein extraction buffer containing 10 mM Tris-HCl (pH 8.3) and 10 mM NaCl to 100 μL ground leaves, then vortex well. Centrifuge for 2 min at 10K rpm at 4°C. Transfer the supernatant into a new eppendorph tube and repeat the centrifugation two or three times, until no solid material remained at the bottom of the tube. Transfer the supernatant to a new eppendorph tube, then quantify the total protein by using the BCA Protein Assay Kit (Pierce). The protein was transfer to PVDF membrane by using Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell (BioRad) at 10V for 30 min.

For Western blot analysis, the membrane was wet in MeOH for 1 min, then blocked in Tris-Buffered Saline (TBST, 20 mM Tris-HCl, 140 mM NaCl, 0.1% Tween-20 v/v, pH 7.5) with 5% nonfat dry milk. After two 15-min washes with TBST, the membrane was incubated in 0.1% nonfat dry milk/TBST and 1:1000 diluted anti-CP antibody, followed by 3 15 min washes with TBST. The membrane was then incubated in 0.1% nonfat dry milk/TBST and 1:2000 diluted secondary antibody, followed by three 15-min washes with TBST. Super Signal West Pico Chemiluminescent Substrate kit (Pierce) was used for chemiluminescent staining according to the manufacturer's instructions. Membranes incubated with the substrate were covered with plastic wrap and exposed to an X-ray film for 5 min before developing the film.

Table 5.1 Constructs used in Chapter V

Name	Description
pCASS4-HDVδ	a binary vector with double S35 promoter for agroinfiltration
pC(+)	Full length satC in the absence of T7 promoter cloned into pUC19
pTCV6	Full length TCV in the absence of T7 promoter cloned into pUC19
pGDD	TCV with GDD enzyme active site mutated to inactive GAA
pCASS4-satC	Full length satC cloned into pCASS-HDVδ
pCASS4-TCV	Full length TCV cloned into pCASS-HDVδ
pCASS4-GDD	Full length TCV GDD mutation cloned into pCASS-HDVδ
pCASS4-CP	CP ORF cloned into pCASS-HDVδ
pCASS4-ΔGGGA	Full length satC with 5'-terminal GGGA deletion cloned into pCASS-HDVδ
pCASS4-Δ3C	Full length satC with 3'-terminal deletion 3 Cs cloned into pCASS-HDVδ
pCASS4-Δ22	Full length satC with 22 nucleotides deletion between SnaBI and NcoI cloned into pCASS-HDVδ

Table 5.2 Oligonucleotides used in Chapter V

Name	Position	Sequence	Polarity
T7C5’	1-19	GTAATACGACTCACTATAGGGATAACTAAGGGTTTCA	+
Oligo7	338-356	GGGCAGGCCCCCCGTCCGA	-
CPF		AAAAGGCCTGAAATCAAACCGATTCACACATC	+
CPR		CATGCCATGGGGGTTTTTGTTTTCTTTTCT	-
Δ5G		TAACTAAGGGTTTCATACAATACTACG	+
C5'		GGGATAACTAAGGGTTTCATACAATAC	+
Δ3C		CAGGCCCCCCGTCCGA	-
Oligo13		GTTACCCAAAGAGCACTAGTT	-
H5F		TCTGGGTAACCACTAA	+

[underlined is NcoI site. Italized is the future non-viral sequence in the construct]

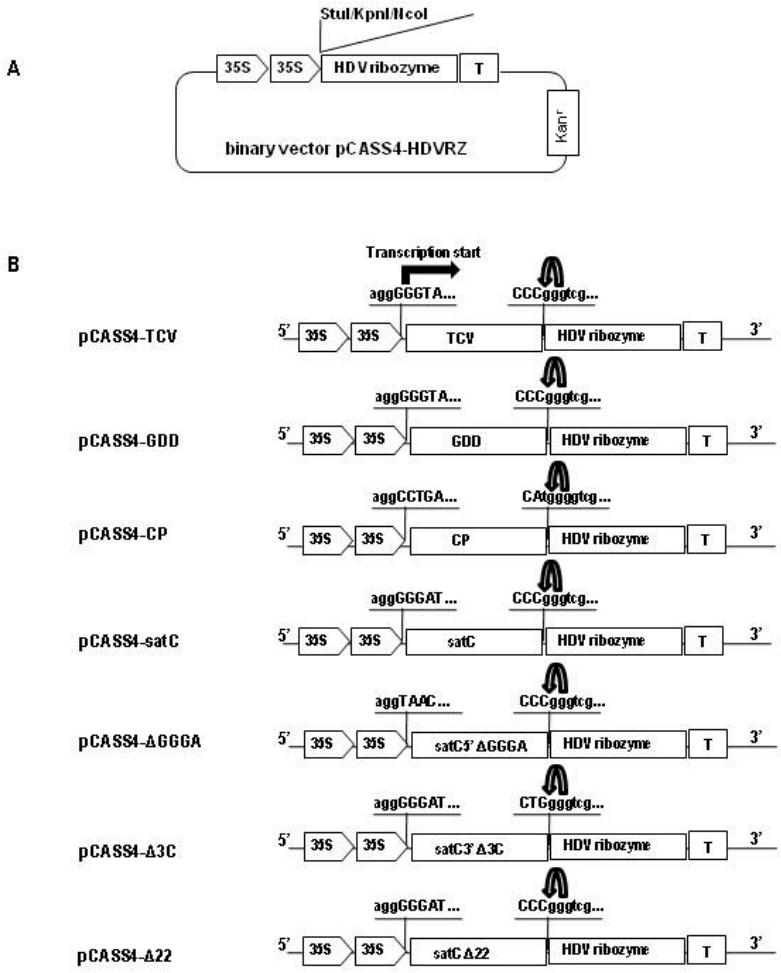

A
StuI/KpnI/NcoI
35S
35S
HDV ribozyme
T
Kan r
binary vector pCASS4-HDVRZ
B
Transcription start
pCASS4-TCV
aggGGGTA...
CCCgggtcg...
5'
35S
35S
TCV
HDV ribozyme
T
3'
pCASS4-GDD
aggGGGTA...
CCCgggtcg...
GDD
pCASS4-CP
aggCCTGA...
CAtggggtcg...
CP
pCASS4-satC
aggGGGAT...
CCCgggtcg...
satC
pCASS4-ΔGGGA
aggTAAC...
CCCgggtcg...
satC3' ΔGGGA
pCASS4-Δ3C
aggGGGAT...
CTGgggtcg...
satC3' Δ3C
pCASS4-Δ22
aggGGGAT...
CCCgggtcg...
satC Δ22

Figure 5.1 T-DNA constructs of TCV and satC RNAs. (A) Schematic representation of T-DNA vector, pCASS4-HDVRz that was used to clone the cDNAs of TCV and satC RNAs. The vector contains a double CaMV 35S promoter (35S) followed by a HDV ribozyme sequence and a Nos terminator (T). A flanking region of three restriction enzyme recognition sites located between the 35S promoter and the ribozyme cassette is shown. Note: Since the StuI site is methylation sensitive, it is necessary to propagate recombinant plasmids in dcm- strains of E. coil (B) Characteristics of T-DNA plasmids harboring TCV, satC and related RNAs used for transient expression in plants. The pCASS4-TCV, pCASS4-GDD, and pCASS4-satC constructs contain full-length cDNA copies of TCV genomic RNA, GDD mutant and satC, respectively. The pCASS4-CP construct contains cDNA copy of CP ORF and the 5' and 3' UTRs with the first two AUG codons deleted in the 5' UTR. The pCASS4-ΔGGGA, pCASS4-Δ3C, pCASS4-CΔ22 constructs contain cDNA copies of satC constructs with 5' end GGGA deletion, 3' end CCC deletion, internal 22 nucleotides between SnaBI and NCoI deletion, respectively. At the 5' junction, nucleotide sequence of the 35S promoter (indicated by lower case) and the 5' sequence of each genomic cDNA (indicated by upper case) are shown. The transcription start site was indicated by a bent arrow. At the 3' end of each construct, the 3' nucleotide sequence from the viral RNA (indicated by upper case) and the non-viral sequences (indicated by lower case) left after self-cleavage by the ribozyme is shown. The predicted self-cleavage site is indicated by a bent arrow.

Results

Successful expression of TCV and satRNAs by agroinfiltration.

Full length TCV, satC monomer, a non-replicating TCV mutant with mutations at the GDD active site, and full length CP with 5' and 3' end UTR were inserted into T-DNA vector pCASS4-HDVRZ downstream of two 35S promoters (Figure 5.1). Immediately after the genes of interest, there is a HDV ribozyme sequence that produces the exact 3' end of TCV and satC by self-cleavage. pCASS4-HDVRZ vector was provided by Dr. A.L.N. Rao (Annamalai and Rao, 2006). The constructs were transformed into agrobacteria, then agroinfiltrated into *Nicotiana Benthamiana*. Total RNA was extracted 5 days after infiltration. Full-length satC and TCV 3' UTR-generated random primers were used to probe both (+)- and (-)-strand RNAs.

TCV, either with or without satC, successfully infects *Nicotiana benthamiana* (Figure 5.2B; lane 1), and TCV-GDD mutant accumulation was not detectable (Figure 5.2B; lane 4). Western blot of total proteins from infected plants by using anti-CP antibody showed that both TCV and CP construct successfully expressed CP, while GDD mutant did not express CP, as expected (Figure 5.2B).

The requirement of CP for satC expression in the absence of active TCV RdRp.

When satC was agroinfiltrated into *Nicotiana Benthamiana* by itself, no RNA was detected using satC-specific oligos (Figure 5.2C; lane 3). SatC was also not detectable when infiltrated with the TCV-GDD mutant to see if a non-replicating TCV could recover satC expression (Figure 5.2C; lane 6). When CP and satC were co-infiltrated,

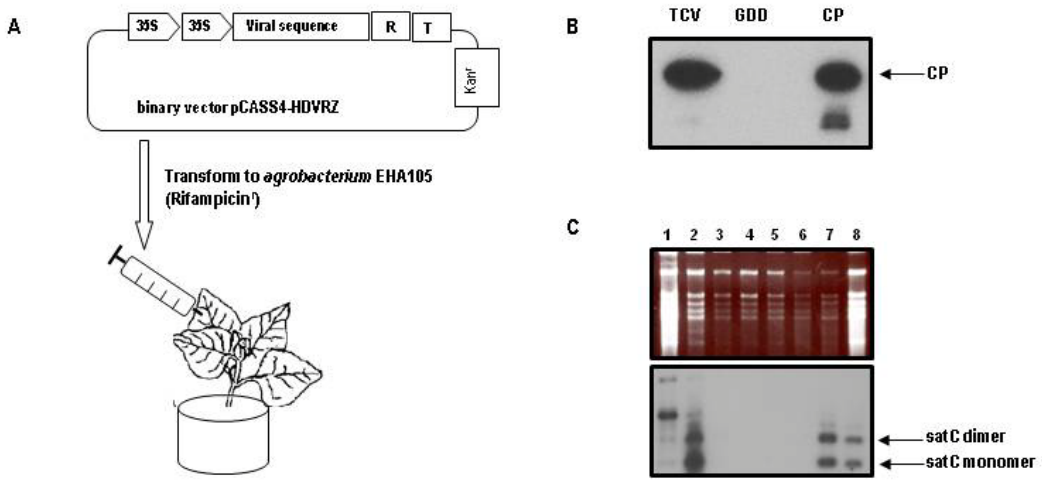


Figure 5.2 (A) The procedure of agroinfiltration. (B) Western blot of total proteins from 35S-TCV, 35S-GDD, and 35S-CP infected plants with CP antibody. Arrowhead shows CP. (C). Total RNA from agroinfiltrated plants. Top panel shows ethidium bromide stained agarose gel. Bottom panel is the same gel hybridized with ^{32}P labeled random primers from full length satC and TCV 3817-4054 fragment. 1, pCASS4-TCV. 2, pCASS4-TCV + pCASS4-satC. 3, pCASS4-satC. 4, pCASS4-GDD. 5, pCASS4-CP. 6, pCASS4-GDD + pCASS4-satC. 7. pCASS4-CP + pCASS4-satC. 8, pCASS4-CP + pCASS4-GDD + pCASS4-satC. Arrows point to satC dimers and monomers.

satC was expressed to detectable levels (Figure 5.2B; lane 7). CP, GDD mutant and satC were then co-infiltrated, and satC expression level was similar to the level of satC with CP (Figure 5.2B; lane 8).

The effects of other silencing suppressors for satC expression in the absence of TCV.

CP is the silencing suppressor of TCV, which targets the plants' defense system (Qu, Ren, and Morris, 2003; Zhang and Simon, 2003a). Other silencing suppressors were infiltrated with satC to determine if they could also enhance satC expression (Figure 5.3B). Protein 2b from *Tomato aspermy virus* (TAV) and CMV were obtained from Dr. Shouwei Ding. Protein p19 was obtained from Dr. Feng Qu. The two 2b were unable to improve satC amplification, while in the presence of p19, satC expressed to a level similar to the level with CP.

Different silencing suppressors target various components in a plants' silencing pathway. TBSV p19 suppresses antiviral response by sequestering viral siRNAs (reviewed by Silhavy and Burgyán, 2004). TCV CP functions as a homodimer to suppress RNA silencing by binding argonaute (AGO1) (Azevedo et al., 2010). CMV 2b suppresses PTGS by interfering with the systemic delivery of the silencing signal (Guo and Ding, 2002).

Compared to TCV CP and TBSV p19, CMV 2b is a relatively weak silencing suppressor as shown by Guo and Ding (2002), CMV 2b was much less effective than TCV CP in preventing GFP silencing (Qu, Ren, and Morris, 2003), which may explain why protein 2b did not protect satC amplification in the absence of TCV in my experiment.

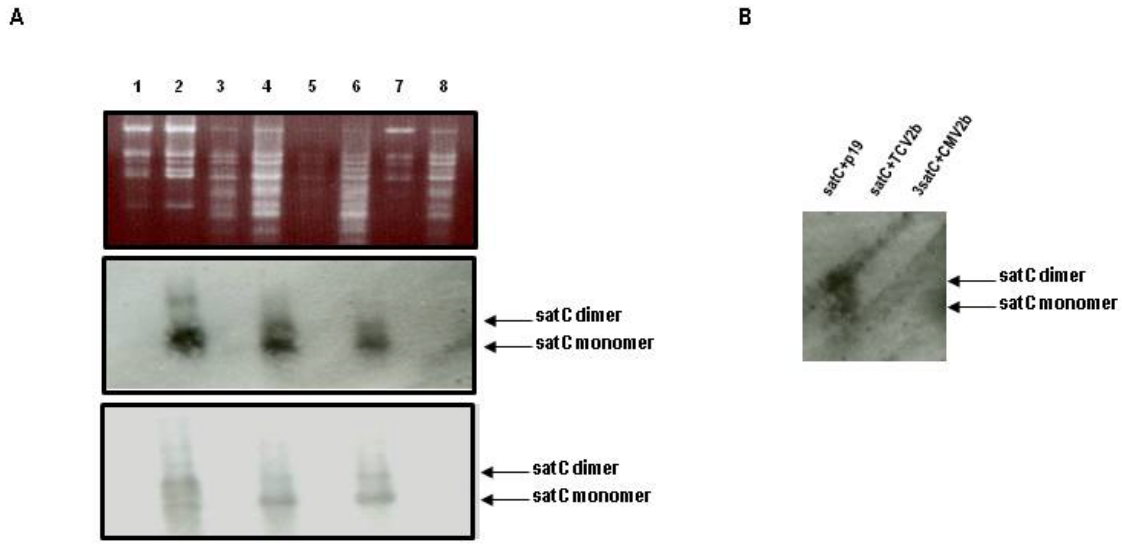


Figure 5.3 SatRNA expression in the presence of CP. (A) Total RNAs from satC mutants infected plants. Top panel shows ethidium bromide stained agarose gel. Middle panel is the same gel hybridized with ^{32}P labeled random primers from full length satC . Bottom panel is the gel hybridized with ^{32}P labeled satC primer oligo7. 1, pCASS4-satC. 2, pCASS4-CP + pCASS4-satC. 3, pCASS4-ΔGGGA. 4, pCASS4-CP + pCASS4-ΔGGGA. 5, pCASS4-Δ3C. 6, pCASS4-CP + pCASS4-Δ3C. 7, pCASS4-Δ22. 8, pCASS4-CP + pCASS4-Δ22. Arrows point to satC dimers and monomers. (B) The effects of other silencing suppressors on satC expression. satC expression in the presence of TBSV p19, CMV 2b and TAV 2b. p19 enhanced satC expression in the absence of TCV, while 2b did not.

SatC can form dimers in the absence of active TCV RdRp.

When monomeric satC was co-agroinfiltrated with CP into *N. benthamiana*, equal amounts of satC dimers and monomers accumulated in total RNA. SatC is found in both monomeric and dimeric forms when expressed in plants or in protoplasts. After infection of plants and protoplasts, gels of extracted RNA showed a monomer to dimer ratio of about 20 to 1. Extra bases are frequently found at the dimer junctions.

Since 5' and 3'-proximal sequences of BBSV satRNA have a direct role in maintaining normal monomer to dimer ratios (Guo et al., 2005), satC 5'-terminal GGGA and 3'-terminal three Cs were deleted to see if these borders have any effect on dimer formation. The dimer to monomer ratio was not changed in the presence of these two satC mutants when both strands were detected (Figure 5.3A middle panel), so the 5' and 3'-terminal sequences are not required to form the dimers. When oligo13 was used to specifically probe (+)-strand RNAs, the dimer to monomer ratios were lower than with wt satC, especially the 5'-terminal GGGA deletion has more monomers than dimers (Figure 5.3A bottom panel). The 5' and 3'-terminal sequences may therefore be required for the formation of (+)-strand dimers.

The observation of preferential accumulation of satC dimers in *N. benthamiana* is similar to results found for satC containing a 22-base segment deletion (position 79-100) (Carpenter et al., 1991). This 22-base deletion mutant was agroinfiltrated into *N. benthamiana* with CP. For unknown reason, the 22-base deletion mutant expressed to below detectable levels (Figure 5.3A).

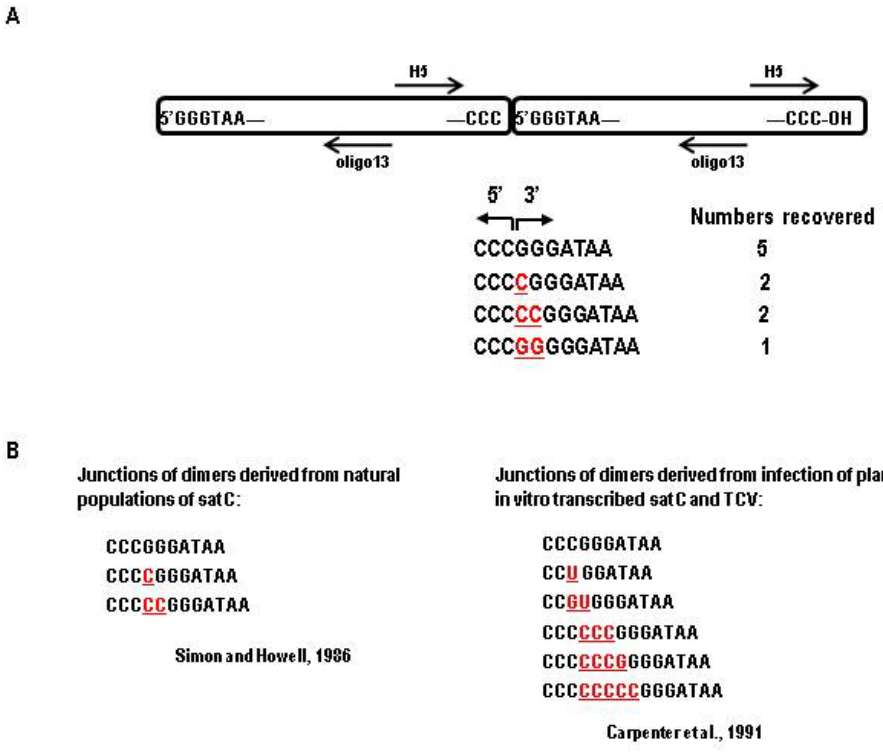


Figure 5.4 Junction sequences of the dimers. (A) Schematic representation of the dimer with the 5'-terminal and 3'-terminal sequences of each monomer shown. The two arrows indicate the two back-to-back primers used to amplify the dimer junctions. Dimer junction sequences are shown with the two arrows indicating the ends from the two monomers. The additional non-viral nucleotides are shown in red and underlined. (B) satC dimer junction sequences cloned from satC natural populations (Simon and Howell, 1986) and from infection of plants with in vitro transcribed RNAs (Carpenter, Cascone, and Simon, 1991). The additional non-viral nucleotides are shown in red and underlined.

Dimer junction sequence.

The junctions of the dimers were cloned by using two back to back primers H5 and oligo13 (Figure 5.4A). Only the junctions of the dimers can be cloned. The dimers are 5' and 3' head-to-tail repeats of satC monomers with different non-viral sequences at the junctions (Figure 5.4A). Two of ten sequenced clones have a cytidylate insertion, two have a CC insertion, one has a GG insertion, and five have no insertion. These results are similar to satC dimers cloned from satC natural populations (Simon and Howell, 1986) and from infection of plants with in vitro transcribed RNAs (Carpenter, Cascone, and Simon, 1991) (Figure 5.4B). Compared with these results, extra bases at multimer junctions are not necessarily associated with a viral RdRp-mediated mechanism. For other satRNA dimers, such as CMV and BBSV, nucleotide deletions often occur at the junction regions in the head-to-tail repeats of satellite RNA, with no additional non-viral sequences (Guo et al., 2005; Kuroda et al., 1997).

Discussion

SatC dimers are thought to be generated by a manner similar to the formation of recombinants and DI RNAs, i.e., replicase-driven copy choice. The replicase detaches from the template because of premature termination or when it reaches the end; then the replicase reinitiates the complementary strand synthesis at specific sequence motifs before it releases the newly made complementary strand (Carpenter, Cascone, and Simon, 1991). The amount of satC dimers is usually one twentieth of the amount of satC monomers when co-inoculated with TCV into either plants or protoplasts.

The observation of higher dimer-to-monomer ratio from agroinfiltrated satC challenges the existing mechanism of satC dimer formation. TCV-GDD mutant does not have an active RdRp (RdRp active site GDD was changed to inactive GAA), so the agroinfiltrated satC with TCV-GDD mutant and CP construct is TCV RdRp-independent. Either TCV RdRp is not involved in satC dimer formation or there is an alternative replicase-independent mechanism.

The subviral RNA dimers can be either replication by-products or double stranded monomer aggregates (Dalmay and Rubino, 1994); replication intermediates that are cleaved to form monomers (Finnen and Rochon, 1995); produced by the autocatalytic reaction of double stranded RNA (Linthorst and Kaper, 1984; Roossinck, Sleat, and Palukaitis, 1992).

In the absence of the TCV RdRp, these satRNA dimers can be generated by three possible ways: ligation product by cellular ligase; replication event by cellular RdRp or other cellular polymerase; nonenzymatic template-directed ligation. RNA polymerase II can transcribe viroids and human *hepatitis delta virus* (HDV), which contains a viroid-like circular RNA and does not encode any protein that has polymerase activity. HDV is transcribed by cellular RNA polymerase II via dual rolling-circle transcription from circular genomic and antigenomic RNAs (Chang et al., 2008; Macnaughton et al., 2002). HDV RNA is the only example of an RNA species that can be amplified by cellular enzymes in mammalian cells. Besides the involvement of host RNA Pol II, little is known about other host factors involved in HDV replication.

Multimeric forms of viroids and HDV RNA are often detected in infected cells, and these RNAs can undergo autocatalytic cleavage into monomer RNAs *in vitro*.

However, the exact role of multimer RNAs in viroid and HDV replication is not clear. It has been shown that the HDV RNA ligation reaction can occur both inter- and intramolecularly (Reid and Lazinski, 2000), these multimers can result from aberrant ligations between monomer species and may represent dead-end products. Ligation of viroid monomer RNAs into multimer RNAs can also occur *in vitro* (Prody et al., 1986). The multimeric forms associated with satC from agroinfiltration may result from similar ligation events.

Arabidopsis possesses six identifiable RdRps. These cellular RdRps play important roles in plant defense system (Voinnet, 2008). They also have important endogenous functions, including the control of chromatin structure and the regulation of cellular gene expression in addition to their role in antiviral defense in plants. Cellular RdRps can be the reason of satC multimer formation in the absence of TCV RdRp.

The satellite RNA of *tobacco ringspot virus* (STobRV RNA) can spontaneously join in a non-enzymic reaction (Buzayan, Hampel, and Bruening, 1986), and form ApG bonds, which are mixtures of 3' to 5' and 2' to 5' forms. As I previously pointed out, non-replicative RNA breakage and rejoining can also happen in cell-free Qβ system (Chetverin et al., 1997; Chetverina et al., 1999), *poliovirus* (Gmyl et al., 1999; Gmyl et al., 2003) and BVDV (Gallei et al., 2005), by 3' terminus attack of the 5' fragment at phosphodiester bonds within the 3' fragment via transesterification reactions.

APPENDIX

SatC co-migrates with TCV in non-denaturing gels.

SatC does not encode any protein, and it is packaged into the virion using the TCV CP. How it is packaged is unknown. The packaging signal for TCV (Qu and Morris, 1997) has no obvious sequence or structure counterpart in satC. Instead of developing independent packaging signals for self-packaging with TCV CP, satC may interact with TCV and become co-packaged with TCV.

As a preliminary experiment to examine the interaction between satC and TCV, I used a gel mobility shift assay (EMSA) to detect if satC and TCV are capable of co-migration. EMSA was modified from Alvarez et al. (2005). SatC was *in vitro* transcribed with T7 RNA polymerase in the presence of [α-^{32}P] UTP. TCV and radiolabeled satC were heat denatured at 65°C for 5 min, and slow cooled to room temperature. In a 30 μL reaction, 0.1 nM of ^{32}P-labeled satC transcripts were mixed with 5mM HEPES (pH7.9), 100mM KCl, 5mM $MgCl_2$, 3.8% glycerol, 2.5 μg yeast tRNA, and 500 nM TCV transcripts. Reactions were subjected to electrophoresis on a 1% nondenaturing agarose gel at 140V for 4-5 hours in an ice bath. The gel was dried at 40°C for 2 hours followed by autoradiography.

A low amount of satC co-migrated at the TCV position (Figure A.1A). This suggests there may be a stable association of RNAs: either satC interacts with TCV or sufficient satC aggregates to migrate at the TCV position in non-denaturing gels.

SatC variant S1-9, a third round winner from a SELEX involving randomizing M1H flanking sequences, contains an important element CCCA. It only replicates to 26%

of wt level in protoplasts; while it dramatically reduces TCV virion level. SatC may interact with TCV through the M1H flanking sequence CCCA. S1-9 also co-migrated with TCV in agarose gel (Figure A.1B).

When CCCA in satC was mutated to GAGA to abolish any potential interactions (mutant LS3), it was still able to interact with TCV *in vitro* (Figure A.1B), which did not support our hypothesis that satC may interact with TCV through the CCCA element. Furthermore, LS3 did not affect satC transcription *in vitro* by using purified TCV RdRp (Sun et al., 2005). Further experiments are needed to identify the possible interaction between satC and TCV.

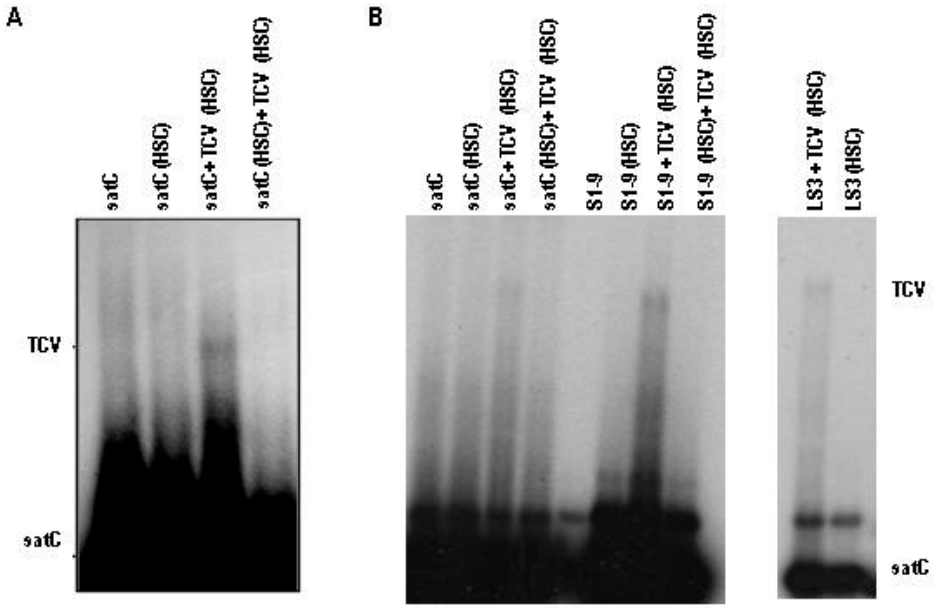


Figure A.1 Putative interaction between satC and TCV. (A) wt satC was uniformly labeled, and then heated and slow cooled (H/SC) either separately or together with full-length TCV. Only when satC and TCV were heated and slow cooled together was satC detected migrating through non-denaturing gels at the position of TCV. (B) Mutant S1-9 and LS3 also migrated at the TCV position.

CONCLUSIONS

Because of the sequence and/or structural similarities with their helper virus and dependence on the helper virus for replication, small nontranslated subviral RNAs are frequently used as models to identify *cis*-acting elements also needed for replication of the much larger helper virus genome. More recent findings that RNA viruses contain elements in both their 5' and 3' UTR that participate in translation, a function not required by small subviral RNAs, suggest that subviral RNAs might evolve to differentially use genomic-derived sequences that are no longer required for helper virus-related functions. This dissertation focuses on the different replication requirements for satC and its helper virus.

Three TCV 3' end hairpins (H4a, H4b, and H5) and two pseudoknots (Ψ_2 and Ψ_3) fold into an internal tRNA-like structure (TSS), which binds to the ribosome P-site through the 60S subunit and is required for translation. SatC was not predicted to adopt a similar structure in its TSS-analogous region although there are only 6-nt differences in this region. Genetic assays showed that Ψ_3 is not required for satC replication in protoplasts (Figure 2.1), but rather one partner, the DR region, is important for satC replication and fitness in plants, and has position flexibility (Figure 2.2). When the different nucleotides in the satC TSS region were changed into TCV sequence, satC mutants all had increased ribosome binding (Figure 2.3), which did not interfere with satRNA replication in protoplasts (Figure 2.4). SatC mutants with ribosome binding affinity similar to the TCV TSS level had structural changes in the H4b loop and the H4a loop, which is the ribosome binding site in TCV (Figure 2.6). These results raise the

question as to why satC evolved to eliminate ribosome binding since enhancing ribosome binding did not affect accumulation in protoplasts. Competition against wt satC in plants showed that wt satC out-competed all the mutants with TCV sequence (Table 2.4). Thus, satC must have evolved to eliminate the TSS structure for better movement *in planta*.

In vivo SELEX of the H5-Pr linker showed that the linker sequence has the conserved sequence UCC (Table 3.4), which is involved in Ψ_2 (Figure 3.4 and 3.5). H5-Pr linker does not appear to interact with the 5' end sequence (Figure 3.5) or with the H5 lower stem (Table 3.4) as predicted by mFold. The linker has a flexible structure since the nucleotides in that region are sensitive to DMS modification. H5-Pr linker connects the Pr and the 3'-terminus to H5 to form replication-required Ψ_1, which at some point has to be open for the replicase access the template 3'-terminus. Having the single-stranded region between these two elements may provide the flexibility necessary for Ψ_1 to form.

With the exception of the six different nucleotides that eliminate the TSS structure, does satC need to maintain the rest of the TCV-derived sequence for its own replication? Replacement of satC H4a with randomized sequence and scoring for fitness in plants by SELEX resulted in winning sequences that contain an H4a-like stem-loop (Table 4.3; Figure 4.2). Replacement of both H4a and H4b with randomized sequence resulted in three distinct groups of winning sequences (Table 4.4). One group formed H4a/H4b-like stem-loop structure, and the other two groups formed a single hairpin (Figure 4.4). Evolution of these single-hairpin SELEX winners in plants resulted in satC that can accumulate to wt levels in protoplasts but remain less fit *in planta* when competed against wt satC (Table 4.6). Thus two highly distinct RNA conformations with various sequences in the H4a and H4b region can mediate satC fitness in protoplasts.

With sequences derived from the TCV 3' end and satD, and without TCV's 5' thousands of nucleotides, satC needed to evolve a stronger Pr (Zhang et al., 2006b), eliminate the ribosome binding in the TSS region (Chapter II), and increase the sequence flexibility of H4a and H4b (Chapter IV). All these data indicate that satC evolved differentially to better use the TCV-derived 3' end for its own replication in the absence of the rest of the TCV sequence.

Although satC evolved some different requirements for its own replication, it can still be used as a model for TCV replication because of its small size and non-coding feature. But we do need to keep in mind that findings in satC may not accurately reflect genomic RNA replication. For satC, the questions that need to be addressed are what is the pre-active structure for satC replication and what triggers the switch from the pre-active structure to the active structure. Delineating the pre-active structure will help us to understand the replication of satC, and extend these rules to TCV genomic RNA replication, or (+)-strand RNA virus replication.

The DR is a very important region in satC. Fitness of satC in plants, high-level accumulation in protoplasts, efficient transcription of satC transcripts by the TCV RdRp *in vitro*, and the satC conformational switch all involve the DR sequence (Table 4.7; Sun et al., 2005; Zhang et al., 2006a; 2006b). Identifying the DR-interacting sequence will help us to identify the DR location in the pre-active structure and understand how satC converts from the pre-active structure to its active structure to initiate (-)-strand synthesis. The DR did not base pair with the 3' terminus to help facilitate RdRp access to the 3' end for the *de novo* initiation (Zhang et al., 2006a); nor did it interact with H4a loop to form Ψ_3 which is essential in TCV replication and translation (Figure 2.1); nor did it have

direct interaction with H5 lower stem or Ψ_2 (Table 4.7). Another possible interaction between the DR and a satD-derived region was not supported by genetic assays (Figure 4.6). Although there is always the possibility that the DR has high sequence specificity which would prohibits sequence alterations, *in vivo* SELEX can be used to identify the DR location.

The DR region might be the possible RdRp binding site in TCV based on the loss of the residue flexibility by in-line reaction upon RdRp binding (Yuan et al., 2009). *In vitro* filter binding of satC fragments using purified TCV RdRp showed that DR may be important for satC RdRp binding (Table 3.5). Unlike TCV, satC did not have significant structural changes after RdRp binding revealed by DMS modification (Figure 2.7). How satC recruits TCV RdRp to the template remains unknown. The Pr, DR region, and H5 LSL are the possible candidates to attract the replicase to satC RNA.

Both replication and packaging of satC are solely dependent on TCV. SatC may interact with TCV genomic RNA for co-packaging or for transfer of the RdRp from TCV genomic RNA replication to satC replication. EMSA provided evidence for TCV-satC interaction by showing that TCV co-migrates with satC *in vitro* (Figure A.1A). Does this interaction have biological relevance? And what are the interacting regions? All these questions need to be addressed. Since CP can package cellular RNAs or form empty virions in the absence of TCV, agroinfiltration is not a good way to discover the interaction between TCV and satC.

A number of economically important plant viruses, such as CMV, have subviral RNAs associated with their genomes. Subviral RNAs that can attenuate helper virus-induced symptoms can be used as biocontrol agents of the helper viruses. Like satC from

TCV, subviral RNAs may develop different strategies to take advantage of the replication/transcription/translation machineries of their helper viruses. Identifying the replication features of subviral RNAs and their helper viruses is an essential step to use subviral RNAs to control pathogenic viruses.

REFERENCES

Alvarez, D., Lodeiro, M., Ludueña, S., Pietrasanta, L., and Gamarnik, A. (2005). Long-range RNA-RNA interactions circularize the dengue virus genome. *J Virol* **79**(11)**,** 6631-43.

Annamalai, P., Hsu, Y., Liu, Y., Tsai, C., and Lin, N. (2003). Structural and mutational analyses of cis-acting sequences in the 5'-untranslated region of satellite RNA of bamboo mosaic potexvirus. *Virology* **311**(1)**,** 229-39.

Annamalai, P., and Rao, A. L. N. (2005). Replication-independent expression of genome components and capsid protein of brome mosaic virus in planta: A functional role for viral replicase in RNA packaging. *Virology* **338**(1)**,** 96-111.

Annamalai, P., and Rao, A. L. N. (2006). Packaging of brome mosaic virus subgenomic RNA is functionally coupled to replication-dependent transcription and translation of coat protein. *Journal of Virology* **80**(20)**,** 10096-10108.

Appiano, A., D'Agostino, G., Redolfi, P., and Pennazio, S. (1981). Sequence of cytological events during the process of local lesion formation in the tomato bushy stunt virus-Gomphrena globosa hypersensitive system. *J Ultrastruct Res* **76**(2)**,** 173-80.

Appiano, A., Dagostino, G., Bassi, M., Barbieri, N., Viale, G., and Dellorto, P. (1986). Origin and function of Tomato bushy stunt virus-induced inclusion-bodies. 2. Quantitative electron-microscope autoradiography and immunogold cyto-chemistry. *Journal of Ultrastructure and Molecular Structure Research* **97**(1-3)**,** 31-38.

Azevedo. J., Garcia. D., Pontier. D., Ohnesorge. S., Yu. A., Garcia. S., Braun. L., Bergdoll. M., Hakimi. M.A., Lagrange. T., and Voinnet. O. (2010) Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. Genes Dev. **24**(9):904-15.

Baker, T., and Bell, S. (1998). Polymerases and the replisome: machines within machines. *Cell* **92**(3)**,** 295-305.

Bamunusinghe, D., Hemenway, C. L., Nelson, R. S., Sanderfoot, A. A., Ye, C. M., Silva, M. A. T., Payton, M., and Verchot-Lubicz, J. (2009). Analysis of potato virus X replicase and TGBp3 subcellular locations. *Virology* **393**(2)**,** 272-285.

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

Basnayake, V., Sit, T., and Lommel, S. (2006). The genomic RNA packaging scheme of Red clover necrotic mosaic virus. *Virology* **345**(2)**,** 532-9.

Beauchemin, C., Boutet, N., and Laliberte, J. F. (2007). Visualization of the interaction between the precursors of VPg, the viral protein linked to the genome of Turnip mosaic virus, and the translation eukaryotic initiation factor iso 4E in planta. *Journal of Virology* **81**(2)**,** 775-782.

Beauchemin, C., and Laliberte, J. F. (2007). The poly(A) binding protein is internalized in virus-induced vesicles or redistributed to the nucleolus during turnip mosaic virus infection. *Journal of Virology* **81,** 10905-10913.

Brantl, S. (2004). Bacterial gene regulation: from transcription attenuation to riboswitches and ribozymes. *Trends in Microbiology* **12**(11)**,** 473-475.

Brierley, I., Pennell, S., and Gilbert, R. (2007). Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nat Rev Microbiol* **5**(8)**,** 598-610.

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

Burgyan, J., Rubino, L., and Russo, M. (1991). De novo generation of cymbidium ringspot virus defective interfering RNA. *J Gen Virol* **72 (Pt 3),** 505-9.

Buzayan, J., Hampel, A., and Bruening, G. (1986). Nucleotide sequence and newly formed phosphodiester bond of spontaneously ligated satellite tobacco ringspot virus RNA. *Nucleic Acids Res* **14**(24)**,** 9729-43.

Carette, J. E., Stuiver, M., Van Lent, J., Wellink, J., and Van Kammen, A. B. (2000). Cowpea mosaic virus infection induces a massive proliferation of endoplasmic reticulum but not Golgi membranes and is dependent on de novo membrane synthesis. *Journal of Virology* **74**(14)**,** 6556-6563.

Carpenter, C. D., Cascone, P. J., and Simon, A. E. (1991). FORMATION OF MULTIMERS OF LINEAR SATELLITE RNAS. *Virology* **183**(2)**,** 586-594.

Carpenter, C. D., and Simon, A. E. (1996a). In vivo repair of 3'-end deletions in a TCV satellite RNA may involve two abortive synthesis and priming events. *Virology* **226**(2)**,** 153-160.

Carpenter, C. D., and Simon, A. E. (1996b). In vivo restoration of biologically active 3' ends of virus-associated RNAs by nonhomologous RNA recombination and replacement of a terminal motif. *Journal of Virology* **70**(1)**,** 478-486.

Carpenter, C. D., and Simon, A. E. (1998). Analysis of sequences and predicted structures required for viral satellite RNA accumulation by in vivo genetic selection. *Nucleic Acids Research* **26**(10)**,** 2426-2432.

Carrington, J., Morris, T., Stockley, P., and Harrison, S. (1987). Structure and assembly of turnip crinkle virus. IV. Analysis of the coat protein gene and implications of the subunit primary structure. *J Mol Biol* **194**(2)**,** 265-76.

Cascone, P. J., Carpenter, C. D., Li, X. H., and Simon, A. E. (1990). Recombination between satellite RNAs of Turnip crinkle virus. *Embo Journal* **9**(6)**,** 1709-1715.

Cascone, P. J., Haydar, T. F., and Simon, A. E. (1993). Sequences and structures for recombination between virus-associated RNAs. *Science* **260**(5109)**,** 801-805.

Chandrika, R., Rabindran, S., Lewandowski, D. J., Manjunath, K. L., and Dawson, W. O. (2000). Full-length tobacco mosaic virus RNAs and defective RNAs have different 3 ' replication signals. *Virology* **273**(1)**,** 198-209.

Chang, J., Nie, X., Chang, H., Han, Z., and Taylor, J. (2008). Transcription of hepatitis delta virus RNA by RNA polymerase II. *J Virol* **82**(3)**,** 1118-27.

Chang, Y., Borja, M., Scholthof, H., Jackson, A., and Morris, T. (1995). Host effects and sequences essential for accumulation of defective interfering RNAs of cucumber necrosis and tomato bushy stunt tombusviruses. *Virology* **210**(1)**,** 41-53.

Chao, L., Rang, C. U., and Wong, L. E. (2002). Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage phi 6. *Journal of Virology* **76**(7)**,** 3276-3281.

Chapman, M., and Kao, C. (1999). A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. *J Mol Biol* **286**(3)**,** 709-20.

Chen, H., Hsu, Y., and Lin, N. (2007). Downregulation of Bamboo mosaic virus replication requires the 5' apical hairpin stem loop structure and sequence of satellite RNA. *Virology* **365**(2)**,** 271-84.

Chen, I., Meng, M., Hsu, Y., and Tsai, C. (2003). Functional analysis of the cloverleaf-like structure in the 3' untranslated region of bamboo mosaic potexvirus RNA revealed dual roles in viral RNA replication and long distance movement. *Virology* **315**(2)**,** 415-24.

Chen, S., Desprez, A., and Olsthoorn, R. (2009). Structural homology between Bamboo mosaic virus and its satellite RNAs in the 5' untranslated region. *J Gen Virol*.

Cheng, C., and Tsai, C. (1999). Structural and functional analysis of the 3' untranslated region of bamboo mosaic potexvirus genomic RNA. *J Mol Biol* **288**(4)**,** 555-65.

Cheng, J., Ding, M., Hsu, Y., and Tsai, C. (2001). The partial purified RNA-dependent RNA polymerases from bamboo mosaic potexvirus and potato virus X infected plants containing the template-dependent activities. *Virus Res* **80**(1-2)**,** 41-52.

Chernysheva, O., and White, K. (2005). Modular arrangement of viral cis-acting RNA domains in a tombusvirus satellite RNA. *Virology* **332**(2)**,** 640-9.

Chetverin, A. B. (1999). The puzzle of RNA recombination. *Febs Letters* **460**(1)**,** 1-5.

Chetverin, A. B., Chetverina, H. V., Demidenko, A. A., and Ugarov, V. I. (1997). Nonhomologous RNA recombination in a cell-free system: Evidence for a transesterification mechanism guided by secondary structure. *Cell* **88**(4)**,** 503-513.

Chetverina, H. V., Demidenko, A. A., Ugarov, V. I., and Chetverin, A. B. (1999). Spontaneous rearrangements in RNA sequences. *Febs Letters* **450**(1-2)**,** 89-94.

Choi, I. R., and White, K. A. (2002). An RNA activator of subgenomic mRNA1 transcription in tomato bushy stunt virus. *Journal of Biological Chemistry* **277**(5)**,** 3760-3766.

Choi, S., Hema, M., Gopinath, K., Santos, J., and Kao, C. (2004). Replicase-binding sites on plus- and minus-strand brome mosaic virus RNAs and their roles in RNA replication in plant cells. *J Virol* **78**(24)**,** 13420-9.

Célix, A., Burgyán, J., and Rodríguez-Cerezo, E. (1999). Interactions between tombusviruses and satellite RNAs of tomato bushy stunt virus: a defect in sat RNA B1 replication maps to ORF1 of a helper virus. *Virology* **262**(1)**,** 129-38.

Dalmay, T., and Rubino, L. (1994). The nature of multimeric forms of cymbidium ringspot tombusvirus satellite RNA. *Arch Virol* **138**(1-2)**,** 161-7.

de Zoeten, G., and Skaf, J. (2001). Pea enation mosaic and the vagaries of a plant virus. *Adv Virus Res* **57,** 323-50.

Diener, T. (1991). Subviral pathogens of plants: viroids and viroidlike satellite RNAs. *FASEB J* **5**(13)**,** 2808-13.

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

Dreher, T. (2009). Role of tRNA-like structures in controlling plant virus replication. *Virus Res* **139**(2)**,** 217-29.

Dreher, T., Bujarski, J., and Hall, T. (1984). Mutant viral RNAs synthesized in vitro show altered aminoacylation and replicase template activities. *Nature* **311**(5982)**,** 171-5.

Dufresne, P. J., Thivierge, K., Cotton, S., Beauchemin, C., Ide, C., Ubalijoro, E., Laliberte, J. F., and Fortin, M. G. (2008). Heat shock 70 protein interaction with Turnip mosaic virus RNA-dependent RNA polymerase within virus-induced membrane vesicles. *Virology* **374**(1)**,** 217-227.

Duggal, R., Rao, A., and Hall, T. (1992). Unique nucleotide differences in the conserved 3' termini of brome mosaic virus RNAs are maintained through their optimization of genome replication. *Virology* **187**(1)**,** 261-70.

Ellington, A., and Szostak, J. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**(6287)**,** 818-22.

Fabian, M., and White, K. (2004). 5'-3' RNA-RNA interaction facilitates cap- and poly(A) tail-independent translation of tomato bushy stunt virus mrna: a potential common mechanism for tombusviridae. *J Biol Chem* **279**(28)**,** 28862-72.

Fabian, M. R., Na, H., Ray, D., and White, K. A. (2003). 3 '-terminal RNA secondary structures are important for accumulation of tomato bushy stunt virus DI RNAs. *Virology* **313**(2)**,** 567-580.

Fabian, M. R., and White, K. A. (2006). Analysis of a 3 '-translation enhancer in a tombusvirus: A dynamic model for RNA-RNA interactions of mRNA termini. *Rna-a Publication of the Rna Society* **12**(7)**,** 1304-1314.

Finnen, R., and Rochon, D. (1995). Characterization and biological activity of DI RNA dimers formed during cucumber necrosis virus coinfections. *Virology* **207**(1)**,** 282-6.

, French, R., and Ahlquist, P. (1987). Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus RNA3. *J Virol* **61**(5)**,** 1457-65.

Gallei, A., Rumenapf, T., Thiel, H. J., and Becher, P. (2005). Characterization of helper virus-independent cytopathogenic classical swine fever virus generated by an in vivo RNA recombination system. *Journal of Virology* **79**(4)**,** 2440-2448.

Gamarnik, A., and Andino, R. (1998). Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev* **12**(15)**,** 2293-304.

Gazo, B., Murphy, P., Gatchel, J., and Browning, K. (2004). A novel interaction of Cap-binding protein complexes eukaryotic initiation factor (eIF) 4F and eIF(iso)4F

with a region in the 3'-untranslated region of satellite tobacco necrosis virus. *J Biol Chem* **279**(14)**,** 13584-92.

Gmyl, A. P., Belousov, E. V., Maslova, S. V., Khitrina, E. V., Chetverin, A. B., and Agol, V. I. (1999). Nonreplicative RNA recombination in poliovirus. *Journal of Virology* **73**(11)**,** 8958-8965.

Gmyl, A. P., Korshenko, S. A., Belousov, E. V., Khitrina, E. V., and Agol, V. I. (2003). Nonreplicative homologous RNA recombination: Promiscuous joining of RNA pieces? *Rna-a Publication of the Rna Society* **9**(10)**,** 1221-1231.

Goebel, S., Hsue, B., Dombrowski, T., and Masters, P. (2004). Characterization of the RNA components of a putative molecular switch in the 3' untranslated region of the murine coronavirus genome. *J Virol* **78**(2)**,** 669-82.

Guan, H. C., Carpenter, C. D., and Simon, A. E. (2000a). Analysis of cis-acting sequences involved in plus-strand synthesis of a turnip crinkle virus-associated satellite RNA identifies a new carmovirus replication element. *Virology* **268**(2)**,** 345-354.

Guan, H. C., Carpenter, C. D., and Simon, A. E. (2000b). Requirement of a 5 '-proximal linear sequence on minus strands for plus-strand synthesis of a satellite RNA associated with turnip crinkle virus. *Virology* **268**(2)**,** 355-363.

Guan, H. C., and Simon, A. E. (2000). Polymerization of nontemplate bases before transcription initiation at the 3 ' ends of templates by an RNA-dependent RNA polymerase: An activity involved in 3 ' end repair of viral RNAs. *Proceedings of the National Academy of Sciences of the United States of America* **97**(23)**,** 12451-12456.

Guan, H. C., Song, C. Z., and Simon, A. E. (1997). RNA promoters located on (-)-strands of a subviral RNA associated with turnip crinkle virus. *Rna-a Publication of the Rna Society* **3**(12)**,** 1401-1412.

Guenther, R., Sit, T., Gracz, H., Dolan, M., Townsend, H., Liu, G., Newman, W., Agris, P., and Lommel, S. (2004). Structural characterization of an intermolecular RNA-RNA interaction involved in the transcription regulation element of a bipartite plant virus. *Nucleic Acids Res* **32**(9)**,** 2819-28.

Guo, H. S., and S. W. Ding. 2002. A viral protein inhibits the long range signaling activity of the gene silencing signal. EMBO J. 21:398-407.

Guo, L., Allen, E., and Miller, W. (2001). Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. *Mol Cell* **7**(5)**,** 1103-9.

Guo, L., Cao, Y., Li, D., Niu, S., Cai, Z., Han, C., Zhai, Y., and Yu, J. (2005). Analysis of nucleotide sequences and multimeric forms of a novel satellite RNA associated with beet black scorch virus. *J Virol* **79**(6)**,** 3664-74.

Guo, R., Lin, W., Zhang, J. C., Simon, A. E., and Kushner, D. B. (2009). Structural Plasticity and Rapid Evolution in a Viral RNA Revealed by In Vivo Genetic Selection. *Journal of Virology* **83**(2)**,** 927-939.

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

Han, S. M., and Sanfacon, H. (2003). Tomato ringspot virus proteins containing the nucleoside triphosphate binding domain are transmembrane proteins that associate with the endoplasmic reticulum and cofractionate with replication complexes. *Journal of Virology* **77**(1)**,** 523-534.

Hatta, T., Bullivant, S., and Matthews, R. (1973). Fine structure of vesicles induced in chloroplasts of Chinese cabbage leaves by infection with turnip yellow mosaic virus. *J Gen Virol* **20**(1)**,** 37-50.

Havelda, Z., Szittya, G., and Burgyán, J. (1998). Characterization of the molecular mechanism of defective interfering RNA-mediated symptom attenuation in tombusvirus-infected plants. *J Virol* **72**(7)**,** 6251-6.

Hsu, Y., Chen, H., Cheng, J., Annamalai, P., Annamali, P., Lin, B., Wu, C., Yeh, W., and Lin, N. (2006). Crucial role of the 5' conserved structure of bamboo mosaic virus satellite RNA in downregulation of helper viral RNA replication. *J Virol* **80**(5)**,** 2566-74.

Hu, C.-C.; Hsu, Y.-H.; Lin, N.-S. (2009). Satellite RNAs and Satellite Viruses of Plants. *Viruses* **1**(3), 1325-1350.

Huang, Y. W., Hu, C. C., Lin, C. A., Liu, Y. P., Tsai, C. H., Lin, N. S., and Hsu, Y. H. (2009). Structural and functional analyses of the 3 ' untranslated region of Bamboo mosaic virus satellite RNA. *Virology* **386**(1)**,** 139-153.

Jablonski, S., and Morrow, C. (1995). Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with altered metal ion requirements for activity. *J Virol* **69**(3)**,** 1532-9.

Jeong, Y., and Makino, S. (1994). Evidence for coronavirus discontinuous transcription. *J Virol* **68**(4)**,** 2615-23.

Kawakami, S., Watanabe, Y., and Beachy, R. (2004). Tobacco mosaic virus infection spreads cell to cell as intact replication complexes. *Proc Natl Acad Sci U S A* **101**(16)**,** 6291-6.

Klovins, J., and van Duin, J. (1999). A long-range pseudoknot in Qbeta RNA is essential for replication. *J Mol Biol* **294**(4)**,** 875-84.

Knorr, D., Mullin, R., Hearne, P., and Morris, T. (1991). De novo generation of defective interfering RNAs of tomato bushy stunt virus by high multiplicity passage. *Virology* **181**(1)**,** 193-202.

Koev, G., Liu, S., Beckett, R., and Miller, W. (2002). The 3prime prime or minute-terminal structure required for replication of Barley yellow dwarf virus RNA contains an embedded 3prime prime or minute end. *Virology* **292**(1)**,** 114-26.

Kong, Q. Z., Oh, J. W., and 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**(10)**,** 1625-1634.

Kuroda, T., T. Natsuaki, W.-Q. Wang, and S. Okuda. (1997). Formation of multimers of cucumber mosaic virus satellite RNA. J. Gen. Virol. 78:941-946.

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

Laliberté, J., and Sanfaçon, H. (2010). Cellular Remodeling during Plant Virus Infection. *Annu Rev Phytopathol*.

Letzel, T., Mundt, E., and Gorbalenya, A. (2007). Evidence for functional significance of the permuted C motif in Co2+-stimulated RNA-dependent RNA polymerase of infectious bursal disease virus. *J Gen Virol* **88**(Pt 10)**,** 2824-33.

Levis, R., Schlesinger, S., and Huang, H. (1990). Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. *J Virol* **64**(4)**,** 1726-33.

Li, X. H., Heaton, L. A., Morris, T. J., and Simon, A. E. (1989). Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated de novo. *Proceedings of the National Academy of Sciences of the United States of America* **86**(23)**,** 9173-9177.

Li, X. H., and Simon, A. E. (1990). Symptom intensification on cruciferous hosts by the virulent satellite RNA of Turnip crinkle virus. *Phytopathology* **80**(3)**,** 238-242.

Li, Z., Pogany, J., Panavas, T., Xu, K., Esposito, A., Kinzy, T., and Nagy, P. (2009). Translation elongation factor 1A is a component of the tombusvirus replicase complex and affects the stability of the p33 replication co-factor. *Virology* **385**(1)**,** 245-60.

Lin, H. X., and White, K. A. (2004). A complex network of RNA-RNA interactions controls subgenomic mRNA transcription in a tombusvirus. *Embo Journal* **23**(16)**,** 3365-3374.

Lindenbach, B., Sgro, J., and Ahlquist, P. (2002). Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. *J Virol* **76**(8)**,** 3905-19.

Linthorst, H., and Kaper, J. (1984). Replication of peanut stunt virus and its associated RNA 5 in cowpea protoplasts. *Virology* **139**(2)**,** 317-29.

Mackenzie, J. (2005). Wrapping things up about virus RNA replication. *Traffic* **6**(11)**,** 967-77.

Macnaughton, T., Shi, S., Modahl, L., and Lai, M. (2002). Rolling circle replication of hepatitis delta virus RNA is carried out by two different cellular RNA polymerases. *J Virol* **76**(8)**,** 3920-7.

Manfre, A. J., and Simon, A. E. (2008). Importance of coat protein and RNA silencing in satellite RNA/virus interactions. *Virology* **379**(1)**,** 161-167.

Marsh, L., Pogue, G., Connell, J., and Hall, T. (1991). Artificial defective interfering RNAs derived from brome mosaic virus. *J Gen Virol* **72 (Pt 8),** 1787-92.

McCartney, A., Greenwood, J., Fabian, M., White, K., and Mullen, R. (2005a). Localization of the tomato bushy stunt virus replication protein p33 reveals a peroxisome-to-endoplasmic reticulum sorting pathway. *Plant Cell* **17**(12)**,** 3513-31.

McCartney, A. W., Greenwood, J. S., Fabian, M. R., White, K. A., and Mullen, R. T. (2005b). Localization of the tomato bushy stunt virus replication protein p33 reveals a peroxisome-to-endoplasmic reticulum sorting pathway. *Plant Cell* **17**(12)**,** 3513-3531.

McCormack, J. C., and Simon, A. E. (2004). Biased hypermutagenesis associated with mutations in an untranslated hairpin of an RNA virus. *Journal of Virology* **78**(14)**,** 7813-7817.

McCormack, J. C., Yuan, X. F., Yingling, Y. G., Kasprzak, W., Zamora, R. E., Shapiro, B. A., and Simon, A. E. (2008). Structural domains within the 3 ' untranslated region of Turnip crinkle virus. *Journal of Virology* **82**(17)**,** 8706-8720.

Militão, V., Moreno, I., Rodríguez-Cerezo, E., and García-Arenal, F. (1998). Differential interactions among isolates of peanut stunt cucumovirus and its satellite RNA. *J Gen Virol* **79 (Pt 1),** 177-84.

Miller, S., and Krijnse-Locker, J. (2008). Modification of intracellular membrane structures for virus replication. *Nature Reviews Microbiology* **6**(5)**,** 363-374.

Miller, W., Bujarski, J., Dreher, T., and Hall, T. (1986). Minus-strand initiation by brome mosaic virus replicase within the 3' tRNA-like structure of native and modified RNA templates. *J Mol Biol* **187**(4)**,** 537-46.

Miller, W., Dreher, T., and Hall, T. (1985). Synthesis of brome mosaic virus subgenomic RNA in vitro by internal initiation on (-)-sense genomic RNA. *Nature* **313**(5997)**,** 68-70.

Miller, W. A., and White, K. A. (2006). Long-distance RNA-RNA interactions in plant virus gene expression and replication. *Annual Review of Phytopathology* **44,** 447-467.

Mochizuki, T., Hirai, K., Kanda, A., Ohnishi, J., Ohki, T., and Tsuda, S. (2009). Induction of necrosis via mitochondrial targeting of Melon necrotic spot virus replication protein p29 by its second transmembrane domain. *Virology* **390**(2)**,** 239-49.

Monkewich, S., Lin, H. X., Fabian, M. R., Xu, W., Na, H., Ray, D., Chernysheva, E. A., Nagy, P. D., and White, K. A. (2005). The p92 polymerase coding region contains an internal RNA element required at an early step in tombusvirus genome replication. *Journal of Virology* **79**(8)**,** 4848-4858.

Mujeeb, A., Clever, J., Billeci, T., James, T., and Parslow, T. (1998). Structure of the dimer initiation complex of HIV-1 genomic RNA. *Nat Struct Biol* **5**(6)**,** 432-6.

Murant, A. (1990). Dependence of groundnut rosette virus on its satellite RNA as well as on groundnut rosette assistor luteovirus for transmission by Aphis craccivora. *J Gen Virol* **71 (Pt 9),** 2163-6.

Más, P., and Beachy, R. (1999). Replication of tobacco mosaic virus on endoplasmic reticulum and role of the cytoskeleton and virus movement protein in intracellular distribution of viral RNA. *J Cell Biol* **147**(5)**,** 945-58.

Na, H., and White, K. A. (2006). Structure and prevalence of replication silencer-3' terminus RNA interactions in Tombusviridae. *Virology* **345**(2)**,** 305-316.

Nagel, J. H. A., and Pleij, C. W. A. (2002). Self-induced structural switches in RNA. *Biochimie* **84**(9)**,** 913-923.

Nagy, P. D., Carpenter, C. D., and Simon, A. E. (1997). A novel 3'-end repair mechanism in an RNA virus. *Proceedings of the National Academy of Sciences of the United States of America* **94**(4)**,** 1113-1118.

Nagy, P. D., Pogany, J., and Simon, A. E. (1999). RNA elements required for RNA recombination function as replication enhancers in vitro and in vivo in a plus-strand RNA virus. *Embo Journal* **18**(20)**,** 5653-5665.

Nagy, P. D., Pogany, J., and 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**(2)**,** 315-324.

Nagy, P. D., and 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**(2)**,** 379-392.

Navarro, B., Russo, M., Pantaleo, V., and Rubino, L. (2006). Cytological analysis of Saccharomyces cerevisiae cells supporting cymbidium ringspot virus defective interfering RNA replication. *J Gen Virol* **87**(Pt 3)**,** 705-14.

Nicholson, B., Wu, B., Chevtchenko, I., and White, K. (2010). Tombusvirus recruitment of host translational machinery via the 3' UTR. *RNA* **16**(7)**,** 1402-19.

Nicholson, B. L., and White, K. A. (2008). Context-influenced cap-independent translation of Tombusvirus mRNAs in vitro. *Virology* **380**(2)**,** 203-12.

Nishikiori, M., Dohi, K., Mori, M., Meshi, T., Naito, S., and Ishikawa, M. (2006). Membrane-bound tomato mosaic virus replication proteins participate in RNA synthesis and are associated with host proteins in a pattern distinct from those that are not membrane bound. *J Virol* **80**(17)**,** 8459-68.

O'Reilly, E., and Kao, 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**(2)**,** 287-303.

Olsthoorn, R., Mertens, S., Brederode, F., and Bol, J. (1999). A conformational switch at the 3' end of a plant virus RNA regulates viral replication. *EMBO J* **18**(17)**,** 4856-64.

Panavas, T., Hawkins, C., Panaviene, Z., and Nagy, P. (2005). The role of the p33:p33/p92 interaction domain in RNA replication and intracellular localization of p33 and p92 proteins of Cucumber necrosis tombusvirus. *Virology* **338**(1)**,** 81-95.

Panavas, T., and Nagy, P. (2003). The RNA replication enhancer element of tombusviruses contains two interchangeable hairpins that are functional during plus-strand synthesis. *J Virol* **77**(1)**,** 258-69.

Panavas, T., Pogany, J., and Nagy, P. (2002). Analysis of minimal promoter sequences for plus-strand synthesis by the Cucumber necrosis virus RNA-dependent RNA polymerase. *Virology* **296**(2)**,** 263-74.

Pogany, J., Fabian, M., White, K., and Nagy, P. (2003). A replication silencer element in a plus-strand RNA virus. *EMBO J* **22**(20)**,** 5602-11.

Pogany, J., White, K. A., and Nagy, P. D. (2005). Specific binding of tombusvirus replication protein p33 to an internal replication element in the viral RNA is essential for replication. *Journal of Virology* **79**(8)**,** 4859-4869.

Prod'homme, D., Jakubiec, A., Tournier, V., Drugeon, G., and Jupin, I. (2003). Targeting of the turnip yellow mosaic virus 66K replication protein to the chloroplast envelope is mediated by the 140K protein. *J Virol* **77**(17)**,** 9124-35.

Prod'homme, D., Le Panse, S., Drugeon, G., and Jupin, I. (2001). Detection and subcellular localization of the turnip yellow mosaic virus 66K replication protein in infected cells. *Virology* **281**(1)**,** 88-101.

Prody, G., Bakos, J., Buzayan, J., Schneider, I., and Bruening, G. (1986). Autolytic processing of dimeric plant virus satellite RNA. *Science* **231**(4745)**,** 1577-80.

Qu, F., and Morris, T. (1997). Encapsidation of turnip crinkle virus is defined by a specific packaging signal and RNA size. *J Virol* **71**(2)**,** 1428-35.

Qu, F., Ren, T., and Morris, T. (2003). The coat protein of turnip crinkle virus suppresses posttranscriptional gene silencing at an early initiation step. *J Virol* **77**(1)**,** 511-22.

Quadt, R., Ishikawa, M., Janda, M., and 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 U S A* **92**(11)**,** 4892-6.

Quadt, R., Kao, C., Browning, K., Hershberger, R., and Ahlquist, P. (1993). Characterization of a host protein associated with brome mosaic virus RNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A* **90**(4)**,** 1498-502.

Rajendran, K., Pogany, J., and Nagy, P. (2002). Comparison of turnip crinkle virus RNA-dependent RNA polymerase preparations expressed in Escherichia coli or derived from infected plants. *J Virol* **76**(4)**,** 1707-17.

Rao, A. L. N. (2006). Genome packaging by spherical plant RNA viruses. *Annual Review of Phytopathology* **44,** 61-87.

Ray, D., and White, K. (1999). Enhancer-like properties of an RNA element that modulates Tombusvirus RNA accumulation. *Virology* **256**(1)**,** 162-71.

Ray, D., and White, K. A. (2003). An internally located RNA hairpin enhances replication of Tomato bushy stunt virus RNAs. *Journal of Virology* **77**(1)**,** 245-257.

Reichel, C., and Beachy, R. (1998). Tobacco mosaic virus infection induces severe morphological changes of the endoplasmic reticulum. *Proc Natl Acad Sci U S A* **95**(19)**,** 11169-74.

Reid, C., and Lazinski, D. (2000). A host-specific function is required for ligation of a wide variety of ribozyme-processed RNAs. *Proc Natl Acad Sci U S A* **97**(1)**,** 424-9.

Resende, R. O., de Haan, P., de Avila, A., Kitajima, E., Kormelink, R., Goldbach, R., and Peters, D. (1991). Generation of envelope and defective interfering RNA mutants of tomato spotted wilt virus by mechanical passage. *J Gen Virol* **72 (Pt 10),** 2375-83.

Restrepo-Hartwig, M., and Ahlquist, P. (1996). Brome mosaic virus helicase- and polymerase-like proteins colocalize on the endoplasmic reticulum at sites of viral RNA synthesis. *J Virol* **70**(12)**,** 8908-16.

Restrepo-Hartwig, M., and Ahlquist, P. (1999). Brome mosaic virus RNA replication proteins 1a and 2a colocalize and 1a independently localizes on the yeast endoplasmic reticulum. *J Virol* **73**(12)**,** 10303-9.

Ritzenthaler, C., Laporte, C., Gaire, F., Dunoyer, P., Schmitt, C., Duval, S., Piéquet, A., Loudes, A., Rohfritsch, O., Stussi-Garaud, C., and Pfeiffer, P. (2002). Grapevine fanleaf virus replication occurs on endoplasmic reticulum-derived membranes. *J Virol* **76**(17)**,** 8808-19.

Robinson, D., Ryabov, E., Raj, S., Roberts, I., and Taliansky, M. (1999). Satellite RNA is essential for encapsidation of groundnut rosette umbravirus RNA by groundnut rosette assistor luteovirus coat protein. *Virology* **254**(1)**,** 105-14.

Roossinck, M., Sleat, D., and Palukaitis, P. (1992). Satellite RNAs of plant viruses: structures and biological effects. *Microbiol Rev* **56**(2)**,** 265-79.

Russo, M., Di Franco, A., and Martelli, G. (1983). The fine structure of Cymbidium ringspot virus infections in host tissues. III. Role of peroxisomes in the genesis of multivesicular bodies. *J Ultrastruct Res* **82**(1)**,** 52-63.

Schaad, M., Jensen, P., and Carrington, J. (1997). Formation of plant RNA virus replication complexes on membranes: role of an endoplasmic reticulum-targeted viral protein. *EMBO J* **16**(13)**,** 4049-59.

Scholthof, K., Jones, R., and Jackson, A. (1999). Biology and structure of plant satellite viruses activated by icosahedral helper viruses. *Curr Top Microbiol Immunol* **239,** 123-43.

Scholthof, K., Scholthof, H., and Jackson, A. (1995). The tomato bushy stunt virus replicase proteins are coordinately expressed and membrane associated. *Virology* **208**(1)**,** 365-9.

Schwartz, M., Chen, J., Janda, M., Sullivan, M., den Boon, J., and Ahlquist, P. (2002). A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. *Mol Cell* **9**(3)**,** 505-14.

Shapiro, B., Kasprzak, W., Grunewald, C., and Aman, J. (2006). Graphical exploratory data analysis of RNA secondary structure dynamics predicted by the massively parallel genetic algorithm. *J Mol Graph Model* **25**(4)**,** 514-31.

Shapiro, B., and Wu, J. (1997). Predicting RNA H-type pseudoknots with the massively parallel genetic algorithm. *Comput Appl Biosci* **13**(4)**,** 459-71.

Siegel, R., Adkins, S., and Kao, C. (1997). Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. *Proc Natl Acad Sci U S A* **94**(21)**,** 11238-43.

Siegel, R., Bellon, L., Beigelman, L., and Kao, C. (1999). Use of DNA, RNA, and chimeric templates by a viral RNA-dependent RNA polymerase: evolutionary implications for the transition from the RNA to the DNA world. *J Virol* **73**(8)**,** 6424-9.

Silhavy, D., and Burgyán, J. (2004). Effects and side-effects of viral RNA silencing suppressors on short RNAs. Trends Plant Sci 9(2), 76-83.

Simon, A., and Howell, S. (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**(13)**,** 3423-3428.

Simon, A. E., and Bujarski, J. J. (1994). RNA-RNA Recombination and evolution in virus-infected plants. *Annual Review of Phytopathology* **32,** 337-362.

Simon, A. E., and Gehrke, L. (2009). RNA conformational changes in the life cycles of RNA viruses, viroids, and virus-associated RNAs. *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms* **1789**(9-10)**,** 571-583.

Simon, A. E., Roossinck, M. J., and Havelda, Z. (2004). Plant virus satellite and defective interfering RNAs: New paradigms for a new century. *Annual Review of Phytopathology* **42,** 415-437.

Sit, T., Vaewhongs, A., and Lommel, S. (1998). RNA-mediated trans-activation of transcription from a viral RNA. *Science* **281**(5378)**,** 829-32.

Sivakumaran, K., and Kao, C. (1999). Initiation of genomic plus-strand RNA synthesis from DNA and RNA templates by a viral RNA-dependent RNA polymerase. *J Virol* **73**(8)**,** 6415-23.

Sivakumaran, K., Kim, C., Tayon, R. J., and 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**(3)**,** 667-82.

Song, C., and Simon, A. (1995). Requirement of a 3'-terminal stem-loop in in vitro transcription by an RNA-dependent RNA polymerase. *J Mol Biol* **254**(1)**,** 6-14.

Song, C. Z., and Simon, A. E. (1994). RNA-dependent RNA-polymerase from plants infected with Turnip crinkle virus can transcribe (+)-strands and (-)-strands of virus-associated RNAs. *Proceedings of the National Academy of Sciences of the United States of America* **91**(19)**,** 8792-8796.

Stupina, V., and Simon, A. E. (1997). Analysis in vivo of turnip crinkle virus satellite RNA C variants with mutations in the 3'-terminal minus-strand promoter. *Virology* **238**(2)**,** 470-477.

Stupina, V. A., Meskauskas, A., McCormack, J. C., Yingling, Y. G., Shapiro, B. A., Dinman, J. D., and Simon, A. E. (2008). The 3 ' proximal translational enhancer of Turnip crinkle virus binds to 60S ribosomal subunits. *Rna-a Publication of the Rna Society* **14**(11)**,** 2379-2393.

Sullivan, M., and 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**(4)**,** 2622-32.

Sullivan, M. L., and Ahlquist, P. (1997). cis-acting signals in bromovirus RNA replication and gene expression: Networking with viral proteins and host factors. *Seminars in Virology* **8**(3)**,** 221-230.

Sun, X. P., and Simon, A. E. (2003). Fitness of a turnip crinkle virus satellite RNA correlates with a sequence-nonspecific hairpin and flanking sequences that enhance replication and repress the accumulation of virions. *Journal of Virology* **77**(14)**,** 7880-7889.

Sun, X. P., and Simon, A. E. (2006). A cis-replication element functions in both orientations to enhance replication of Turnip crinkle virus. *Virology* **352**(1)**,** 39-51.

Sun, X. P., Zhang, G. H., and Simon, A. E. (2005). Short internal sequences involved in replication and virion accumulation in a subviral RNA of Turnip crinkle virus. *Journal of Virology* **79**(1)**,** 512-524.

Taliansky, M., and Robinson, D. (1997). Down-regulation of groundnut rosette virus replication by a variant satellite RNA. *Virology* **230**(2)**,** 228-35.

Tatsuta, M., Mizumoto, H., Kaido, M., Mise, K., and Okuno, T. (2005). The red clover necrotic mosaic virus RNA2 trans-activator is also a cis-acting RNA2 replication element. *J Virol* **79**(2)**,** 978-86.

Thivierge, K., Cotton, S., Dufresne, P. J., Mathieu, I., Beauchemin, C., Ide, C., Fortin, M. G., and Laliberte, J. F. (2008). Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. *Virology* **377**(1)**,** 216-225.

Thomas, C., Leh, V., Lederer, C., and Maule, A. (2003). Turnip crinkle virus coat protein mediates suppression of RNA silencing in Nicotiana benthamiana. *Virology* **306**(1)**,** 33-41.

Tijerina, P., Mohr, S., and Russell, R. (2007). DMS footprinting of structured RNAs and RNA-protein complexes. *Nat Protoc* **2**(10)**,** 2608-23.

Tomita, Y., Mizuno, T., Díez, J., Naito, S., Ahlquist, P., and Ishikawa, M. (2003). Mutation of host DnaJ homolog inhibits brome mosaic virus negative-strand RNA synthesis. *J Virol* **77**(5)**,** 2990-7.

Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**(4968)**,** 505-10.

Turner, K., Sit, T., Callaway, A., Allen, N., and Lommel, S. (2004). Red clover necrotic mosaic virus replication proteins accumulate at the endoplasmic reticulum. *Virology* **320**(2)**,** 276-90.

van Regenmortel, M., Mayo, M., Fauquet, C., and Maniloff, J. (2000). Virus nomenclature: consensus versus chaos. *Arch Virol* **145**(10)**,** 2227-32.

Villarreal, L. P. (2005). Viruses and the Evolution of Life, ASM Press, Washington DC pp. 395.

Voinnet, O. (2008). Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends Plant Sci* **13**(7)**,** 317-28.

Vázquez, A., Alonso, J., and Parra, F. (2000). Mutation analysis of the GDD sequence motif of a calicivirus RNA-dependent RNA polymerase. *J Virol* **74**(8)**,** 3888-91.

Wang, J. L., and Simon, A. E. (2000). 3 '-end stem-loops of the subviral RNAs associated with turnip crinkle virus are involved in symptom modulation and coat protein binding. *Journal of Virology* **74**(14)**,** 6528-6537.

Wang, R., Stork, J., and Nagy, P. (2009). A key role for heat shock protein 70 in the localization and insertion of tombusvirus replication proteins to intracellular membranes. *J Virol* **83**(7)**,** 3276-87.

Wang, R. Y. L., and Nagy, P. D. (2008). Tomato bushy stunt virus co-opts the RNA-binding function of a host metabolic enzyme for viral genomic RNA synthesis. *Cell Host & Microbe* **3**(3)**,** 178-187.

Wang, X., and Gillam, S. (2001). Mutations in the GDD motif of rubella virus putative RNA-dependent RNA polymerase affect virus replication. *Virology* **285**(2)**,** 322-31.

Wang, Y., Xiao, M., Chen, J., Zhang, W., Luo, J., Bao, K., Nie, M., and Li, B. (2007). Mutational analysis of the GDD sequence motif of classical swine fever virus RNA-dependent RNA polymerases. *Virus Genes* **34**(1)**,** 63-5.

Weber-Lotfi, F., Dietrich, A., Russo, M., and Rubino, L. (2002). Mitochondrial targeting and membrane anchoring of a viral replicase in plant and yeast cells. *J Virol* **76**(20)**,** 10485-96.

Wei, T., Huang, T., McNeil, J., Laliberté, J., Hong, J., Nelson, R., and Wang, A. (2010). Sequential recruitment of the endoplasmic reticulum and chloroplasts for plant potyvirus replication. *J Virol* **84**(2)**,** 799-809.

White, K. A., and Nagy, P. D. (2004). Advances in the molecular biology of tombusviruses: Gene expression, genome replication, and recombination. *In* "Progress in Nucleic Acid Research and Molecular Biology, Vol 78", Vol. 78, pp. 187-226.

Wu, B., Oliveri, S., Mandic, J., and White, K. (2010). Evidence for a premature termination mechanism of subgenomic mRNA transcription in a carmovirus. *J Virol* **84**(15)**,** 7904-7.

Wu, B. D., Vanti, W. B., and White, K. A. (2001). An RNA domain within the 5 ' untranslated region of the tomato bushy stunt virus genome modulates viral RNA replication. *Journal of Molecular Biology* **305**(4)**,** 741-756.

Wu, B. D., and White, K. A. (1998). Formation and amplification of a novel tombusvirus defective RNA which lacks the 5 ' nontranslated region of the viral genome. *Journal of Virology* **72**(12)**,** 9897-9905.

Wu, G., and Kaper, J. (1995). Competition of viral and satellite RNAs of cucumber mosaic virus for replication in vitro by viral RNA-dependent RNA polymerase. *Res Virol* **146**(1)**,** 61-7.

Xu, P., and Roossinck, M. (2000). Cucumber mosaic virus D satellite RNA-induced programmed cell death in tomato. *Plant Cell* **12**(7)**,** 1079-92.

Xu, W., and White, K. A. (2009). RNA-Based Regulation of Transcription and Translation of Aureusvirus Subgenomic mRNA1. *Journal of Virology* **83**(19)**,** 10096-10105.

Yoshinari, S., and Dreher, T. (2000). Internal and 3' RNA initiation by Qbeta replicase directed by CCA boxes. *Virology* **271**(2)**,** 363-70.

Yuan, X., Shi, K., Young, M., and Simon, A. (2010). The terminal loop of a 3' proximal hairpin plays a critical role in replication and the structure of the 3' region of Turnip crinkle virus. *Virology*.

Yuan, X. F., Shi, K. R., Meskauskas, A., and Simon, A. E. (2009). The 3 ' end of Turnip crinkle virus contains a highly interactive structure including a translational enhancer that is disrupted by binding to the RNA-dependent RNA polymerase. *Rna-a Publication of the Rna Society* **15**(10)**,** 1849-1864.

Zechmann, B., Müller, M., and Zellnig, G. (2003). Cytological modifications in zucchini yellow mosaic virus (ZYMV)-infected Styrian pumpkin plants. *Arch Virol* **148**(6)**,** 1119-33.

Zhang, F. L., and 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**(1)**,** 8-13.

Zhang, G. H., and Simon, A. E. (2003b). A multifunctional turnip crinkle virus replication enhancer revealed by in vivo functional SELEX. *Journal of Molecular Biology* **326**(1)**,** 35-48.

Zhang, G. H., Zhang, J. C., George, A. T., Baumstark, T., and Simon, A. E. (2006a). Conformational changes involved in initiation of minus-strand synthesis of a virus-associated RNA. *Rna-a Publication of the Rna Society* **12**(1)**,** 147-162.

Zhang, G. H., Zhang, J. C., and Simon, A. E. (2004). Repression and derepression of minus-strand synthesis in a plus-strand RNA virus replicon. *Journal of Virology* **78**(14)**,** 7619-7633.

Zhang, J. C., and Simon, A. E. (2005). Importance of sequence and structural elements within a viral replication repressor. *Virology* **333**(2)**,** 301-315.

Zhang, J. C., Stuntz, R. M., and Simon, A. E. (2004). Analysis of a viral replication repressor: sequence requirements for a large symmetrical internal loop. *Virology* **326**(1)**,** 90-102.

Zhang, J. C., Zhang, G. H., Guo, R., Shapiro, B. A., and Simon, A. E. (2006b). A pseudoknot in a preactive form of a viral RNA is part of a structural switch activating minus-strand synthesis. *Journal of Virology* **80**(18)**,** 9181-9191.

Zhang, J. C., Zhang, G. H., McCormack, J. C., and Simon, A. E. (2006c). Evolution of virus-derived sequences for high-level replication of a subviral RNA. *Virology* **351**(2)**,** 476-488.

Zuker, M. (1989). Computer prediction of RNA structure. *Methods Enzymol* **180,** 262-88.

Zuker, M. (1994). Prediction of RNA secondary structure by energy minimization. *Methods Mol Biol* **25,** 267-94.

Zuo, X. B., Wang, J. B., Yu, P., Eyler, D., Xu, H., Starich, M. R., Tiede, D. M., Simon, A. E., Kasprzak, W., Schwieters, C. D., Shapiro, B. A., and Wang, Y. X. (2010). Solution structure of the cap-independent translational enhancer and ribosome-binding element in the 3 ' UTR of turnip crinkle virus. *Proceedings of the National Academy of Sciences of the United States of America* **107**(4)**,** 1385-1390.