ABSTRACT

Title of dissertation: THE 3' UTR OF *TURNIP CRINKLE VIRUS* INTERACTS LOCALLY AND DISTALLY TO REGULATE TRANSCRIPTION AND TRANSLATION OF THE VIRUS

Megan Y. L. Young, Doctor of Philosophy, 2012

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

Turnip crinkle virus (TCV) is a 4054 b positive-strand RNA virus of the genus Carmovirus in the Family *Tombusviridae*. Upon entry into cells, TCV is translated using host translational machinery to produce its RNA-dependent RNA polymerase (RdRp). The RNA is proposed to undergo a conformational rearrangement, mediated by recruitment of the RdRp to the 3' ends of the viral RNA, which represses translation and promotes negative-strand synthesis. A second RNA switch is proposed to occur that inhibits minus-strand synthesis and promotes recruitment of the RdRp to the 3' ends of negative-strands for the asymmetrical production of positive-strands.

Within the 3' UTR of TCV is a tRNA-shaped structure (TSS) that is capable of binding ribosomes and overlaps with structures necessary for translational enhancement. The RdRp has been shown to bind within this region and result in a widespread conformational shift. The binding of RdRp to the 3' end of the virus is very sensitive to perturbations of sequence or structure, with many mutations resulting in non-specific binding of the RdRp.

The elements within the 3' UTR have been shown to be very interactive with

alterations affecting the structure of regions hundreds of bases away. A second-site mutation study indicated that regions upstream of the 3' UTR may also be interacting with the 3' UTR. Some second-site mutations located in this upstream region were found to increase accumulation in protoplasts and additional studies are under way to explain this phenomenon. The 3' viral contribution in a luciferase reporter construct was increased to incorporate the second-site mutations. While the second-site mutations had little effect on translation, it was surprising to find that extension of the viral 3' sequence enhanced translation. Translational enhancement was mapped to just an additional twenty bases and further study revealed that a hairpin (H3) is important for viral translation and accumulation and may also be interacting with the 3' UTR.

THE 3' UTR OF *TURNIP CRINKLE VIRUS* INTERACTS LOCALLY AND DISTALLY TO REGULATE TRANSCRIPTION AND TRANSLATION OF THE VIRUS

by

Megan Y. L. Young

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
2012

Advisory Committee:

Professor Anne E. Simon, Chair
Professor Jonathan D. Dinman
Professor James N. Culver
Professor Brenda L. Fredericksen
Professor Douglas A. Julin

TABLE OF CONTENTS

Section **Page**

List of Tables

List of Figures

Chapter I

An Overview of *cis*-Acting Elements Involved in the Replication and Translation of Positive-Strand RNA Viruses

Overview of the Lifecycle of a Positive-Strand RNA Virus

Viruses that have a single positive-strand ribonucleic acid (RNA) as their genome have been known for more than 100 years (Zaitlin, 1998) when *Tobacco mosaic virus* (TMV) was first found to be a non-bacterial infectious agent. Positive-strand RNA viruses are now known to ravage plants, livestock, and humans with diseases like hepatitis C, polio, and West Nile encephalitis. Much work is being done on these viruses in the hopes of being able to combat the diseases they cause, which result in many human and animal deaths and crop damage resulting in severe economic losses.

Positive-strand RNA viruses initiate infection upon entry into host cells. Animal viruses enter a cell by receptor-mediated endocytosis or fusion to the plasma membrane, while plant viruses are inoculated into cells through human or insect activity. Upon infection, the viral RNA is uncoated from its capsid and released into the cytosol, where it becomes available for translation by the host translational machinery. pH differences, conformational changes due to binding receptors, and/or protease activity result in the uncoating of the capsid (Flint et al, 2004; Fuchs & Blaas, 2010). The viral RNA-dependent RNA polymerase (RdRp) and other viral replicase proteins are translated by the host's translational machinery, after which the viral RNA switches from a translation-competent template to a replication-competent template through a poorly understood mechanism that likely involves a change in the conformation of the RNA.

The viruses replicate through a multi-step process that utilizes a replicase complex composed of the viral RdRp, other viral proteins, and, often, host proteins (Lai, 1998; Mackenzie, 2005). The RNA is then transcribed into negative-strands that serve as templates for positive-strand synthesis. These negative-strands, along with the replication complex, are sequestered in a virus-induced, organelle-derived membrane compartments (Kopek et al, 2007; Miller & Krijnse-Locker, 2008), within which replication occurs, with release of newly synthesized positive-strands to the cytoplasm for packaging and/or further translation. Many positive-strand RNA viruses asymmetrically produce positive- and negative-strands, with up to 1000 positive-strands synthesized for every one negative-strand (Buck, 1996), which may indicate another conformational switch on nascent positive-strands that prevents them from being used for further negative-strand synthesis. Many RNA viruses produce 3' coterminal subgenomic (sg) RNAs, either through premature termination during negative-strand synthesis or internal initiation on negative strands during positive-strand synthesis, in order to enhance expression of the 3' proximal open reading frames (ORF) (White, 2002).

The viral coat protein recognizes *cis*-acting sequences and/or structures on the viral positive-strands and interacts with the viral RNA and additional coat protein to package the RNA into the viral capsid. Animal viruses either bud off from or lyse the host cell, while plant viral RNA and virions move cell-to-cell through the plasmodesmata with assistance from virally-encoded movement proteins (Flint et al, 2004).

Cis-Acting Elements Involved in Replication

Replication makes abundant use of *cis*-acting elements within the viral RNA genome, allowing the replicase complex to recognize its cognate RNA. These *cis*-acting elements are often found in untranslated regions (UTRs) and are usually specific in sequence and/or structure, allowing them to interact with proteins or other regions of the RNA (Dreher, 1999). Promoters recruit the replicase complex to positive- or negative-strand transcription start sites for accurate synthesis of genomic (g) and subgenomic (sg) RNAs (French & Ahlquist, 1988; Miller et al, 1986; Miller et al; Song & Simon, 1995; Wang & Simon, 1997). Core promoters are a specific type of *cis*-acting element defined as the minimal sequence required for replication or transcription to occur. Initially, *cis*-acting elements important for replication were found only proximal to the 3' end of the viral RNA, likely because that is where researchers looked. Recently, however, elements important for negative-strand synthesis have been identified in interior regions and near the 5' end of viral RNAs, and some are able to interact with the 3' end via RNA-RNA interactions (Alvarez et al, 2005b; Filomatori et al, 2006; Herold & Andino, 2001; Khromykh et al, 2001; You et al, 2001) or RNA-protein interactions (Frolov et al, 2001a; Herold & Andino, 2001; Isken et al, 2003; Pogany et al, 2005).

Core Promoters: Negative-Strand Synthesis

The viral RdRp must initiate transcription of negative-strands at the 3' terminus of positive-strands. RNA viruses have at their 3' termini a variety of structures that may function as core promoters for negative-strand synthesis including, but not limited to, tRNA-like structures (TLS), pseudoknots, hairpins, and short sequences with no obvious structure (Dreher, 1999). TLSes are present in eight plant virus genera and are generally

categorized by their aminoacylated 3' terminus consisting of histidine, tyrosine, or valine (Dreher 2010). Viral poly(A) tails are short, usually from twenty to 100 residues, and are found in five plant virus genera. The remaining viruses have non-TLS structures or no structure at their 3' ends and are not similar across genera (Dreher, 1999; Dreher, 2010; Zaccomer et al, 1995).

Tomato bushy stunt virus (TBSV), a member of the genus Tombusvirus, in the family *Tombusviridae*, is one of the best characterized small, positive-strand RNA plant viruses. It has a monopartite genome of 4.8 kb, encodes five proteins, and is associated with defective interfering (DI) RNAs (White & Morris, 1999). DI RNAs are of primarily helper virus-derived sequence, do not encode any proteins, and require products of the helper virus to replicate. As such, they have been extensively used to study tombusvirus replication. There are four conserved regions found in TBSV and its associated DI RNAs: RI, RII, RIII, and RIV, of which RII, RIII, and RIV are important for replication (Oster et al, 1998) with a stem loop in RII serving as a binding site for one of the replicase proteins (Pogany et al, 2005). The promoter for minus-strand synthesis is a stemloop in RIV called SL-1, located just upstream of the 3' terminal cytidylates that are required for initiation (Fabian et al, 2003). Mutagenesis of the GNRA terminal tetraloop of this stem reduced viral accumulation and disruption of the lower stem resulted in undetectable levels of accumulation that were only partially restored by compensatory mutations(Fabian et al, 2003). Upstream of SL-1 is SL-3, which has an internal asymmetric loop that interacts with the 3' terminal cytidylates to form a pseudoknot. Disrupting this pseudoknot increased *in vitro* transcription and resulted in very poor accumulation *in vivo* (Pogany et al, 2003).

The three RNAs (RNA1, RNA2, RNA3) of the tripartite *Brome mosaic virus* (BMV) have very conserved 3' terminal 200 bases, which alone are capable of directing negative-strand synthesis *in vitro* (Dreher et al, 1984). The 3' ends of the three RNAs fold into a TLS that is capable of interacting with tRNA-synthetase and nucleotidyl transferase (Bujarski et al, 1985; Dreher, 2010; Haenni et al, 1982). Mutations in four of six defined elements in the BMV TLS reduced negative-strand synthesis while the single-stranded regions of SLC were found to directly bind the RdRp *in vitro*, even in the absence of the remainder of the TLS (Bujarski et al, 1985; Chapman & Kao, 1999). RdRp binding is believed to be due to a clamped adenine motif present in the AUA triloop of SLC, as determined by NMR analysis (Kim et al, 2000).

Red clover necrotic mosaic virus has a bipartite genome consisting of RNA1 and RNA2. Two hairpins, SLDE and SLF, and their intervening sequence in the terminal seventy two bases of RNA1 were required for negative-strand synthesis *in vitro* (Iwakawa et al, 2007). In vitro, the stems of SLDE and SLF were not essential for negative-strand synthesis, but the loop sequences were critical (Iwakawa et al, 2007). The *in vitro* findings for RNA1 correspond to *in vivo* findings for elements in RNA2 that are important for accumulation (Takeda et al, 2005; Turner & Buck, 1999).

The TLS of *Turnip yellow mosaic virus* (TYMV) is the closest tRNA mimic and is similar in structure and function to $tRNA^{val}$ (de Smit et al, 2002; Dreher, 2010; Dreher & Goodwin, 1998; Rietveld et al, 1982). Only the terminal CCA bases of the TLS are required for negative-strand synthesis *in vitro* and then, only when not aminoacylated (Deiman et al, 1998; Singh & Dreher, 1998). *In vitro* negative-strand synthesis was

eliminated upon binding of eIF1A•GTP to the TLS, possibly because it interferes with the RdRp accessing the 3' end (Matsuda et al, 2004).

Tobacco mosaic virus (TMV) has a histidylatable TLS that has an acceptor stem mimic and terminates with CCA (Felden et al, 1996; van Belkum et al, 1985). It is capable of binding tRNA-synthetase, nucleotidyltransferase, eEF1A, ATP, and CTP (Dreher, 2010; Litvak et al, 1973a; Litvak et al, 1973b). The TLS is associated with an upstream pseudoknot domain, of which the 3' most proximal pseudoknot is critical for viral accumulation in plants and protoplasts and is essential for directing negative-strand synthesis *in vitro* in the presence of the TLS (Osman & Buck, 2003; Takamatsu et al, 1990).

All three RNAs of the tripartite *Alfalfa mosaic virus* (AlMV) have a ten bp hairpin (hpE) 100 bases upstream of the transcription start site. Deletion of the loop or disruption of the stem of this hairpin eliminated *in vitro* transcription, indicating that hpE is part of a core promoter (van Rossum et al, 1997b). The TLS of AlMV alone could not direct negative-strand synthesis or bind RdRp *in vitro*. Mutating the TLS resulted in aberrant transcription products, suggesting that the TLS is required for directing the RdRp to the proper initiation site at the 3' terminus (Olsthoorn & Bol, 2002). The 5' UTR of RNA1 is also required for negative-strand synthesis, though the exact mechanism remains unclear (Vlot & Bol, 2003).

The family *Flaviviridae* consists of many important human pathogens, including West Nile virus (WNV), Yellow fever virus, Kunjin virus, and Hepatitis C virus (HCV). Flaviviral genomes have complementary sequences at the 5' and 3' ends, which are suspected to circularize the genome and play a role in the virus lifecycle (Villordo &

Gamarnik, 2009). *Dengue virus* (DV) has its core promoter for negative-strand synthesis in its 5' UTR, hairpin SLA. Mutating the terminal loop of SLA or disrupting the hairpin resulted in delayed replication or undetectable viral levels, respectively. RdRp binding assays indicated that the RdRp binds the 5' UTR, but not the 3' UTR (Filomatori et al, 2006).

There are two sets of complementary sequences in DV: the 5' and 3' CS and the 5' and 3' UAR. Both sets of complementary sequences have been shown to be vital for RNA synthesis, but expendable for translation *in vitro* using a replicon system in cell culture (Alvarez et al, 2005a). A similar pair of interactions that circularize the genome is also present in WNV and is important for replication (Zhang et al, 2008). Furthermore, atomic force microscopy revealed that the RdRp is bound to a circularized DV genome, suggesting that SLA attracts the RdRp, and circularization allows transfer of the RdRp to the 3' end of the genome (Filomatori et al, 2006).

Sindbis virus (SIN) contains several conserved sequence elements (CSE) that function as promoters of transcription. CSE2, 51 bases and present near the 5' end of the genome, acts as a promoter for negative-strand synthesis from the genomic RNA (Frolov et al, 2001a; Frolov et al, 2001b; Kuhn et al, 1990; Ou et al, 1983). CSE2's location near the 5' end of the genome is thought to prevent negative-strand synthesis from subgenomic RNAs, which lack this element (Frolov et al, 2001a). CSE4, 19 bases and located just upstream of the poly(A) tail in the 3' UTR, is a copromoter for negative-strand synthesis and is proposed to interact with the 5' end of the genome (Frolov et al, 2001b; Hardy, 2006; Kuhn et al, 1990).

Core Promoters: Positive-Strand Synthesis

Positive-strand synthesis initiates at the 3' ends of viral negative-strands and has not been as extensively studied as negative-strand synthesis. *In vitro* positive-strand synthesis of DI-72 of TBSV can be directed by eleven unstructured bases at the 3' end of minus-strands. Point mutations within these 11 bases had little effect on transcription, but mutation of 9 of the 11 bases resulted in undetectable levels of transcription. Deleting most of the 11 bases led to transcription initiation at a nearby site, indicating that DI-72 has at least two 3' end sequences with *in vitro* promoter activity (Panavas et al, 2002).

In vitro positive-strand synthesis of BMV is dependent on the 26 terminal bases at the 3' ends of negative-strands. These bases are unstructured and are flanked (the -1 position) by a non-templated base that is added to the negative strand by the replicase complex (Siegel et al, 1997; Sivakumaran & Kao, 1999; Sivakumaran et al, 1999). The replicase recognizes the nucleotides at -1, +1, and +2 on negative-strands to initiate positive-strand synthesis. *In vitro* positive-strand synthesis was increased when guanylates were introduced at position +8 or +10. This increase is suggested to be due to increased stability between the template and the newly synthesized strand, resulting in the replicase being better able to remain bound during elongation (Sivakumaran et al, 1999).

Bamboo mosaic virus (BaMV) has at the 3' end of its negative-strand a large, stable hairpin; an 8 base linker; a smaller, unstable hairpin; and a 5 base tail. At least three of these elements are important for promoting positive-strand synthesis *in vitro.* Deletion of the 5 base tail nearly eliminated *in vitro* synthesis of positive-strands. Deletion of the smaller hairpin had little effect, but mutations within the larger hairpin reduced RNA synthesis efficiency *in vitro* and RNA accumulation *in vitro*. The spacing

between the two hairpins was also found to be important for *in vitro* RNA synthesis (Lin et al, 2005).

Enhancers

Replication enhancers are not essential for replication but increase the level of transcription that is directed by the core promoter. They are position-independent and can be present on either positive- or negative-strands. The untranslated DI RNAs of TBSV that are used to determine sequences important for replication contain four noncontiguous helper virus-derived regions (I, II, III, IV). Region III (RIII) has been shown to be inessential for DI RNA accumulation, but its presence results in an about 10-fold increase in accumulation. RIII was found to function in non-cognate locations and in either orientation (Ray & White, 1999). The primary activity of RIII was mapped to the 5' proximal 35 bases called segment A. Mutational analysis revealed a hairpin in segment A that was required to enhance DI RNA accumulation functions primarily in the negative-sense (Ray & White, 2003). Further study revealed the presence of two hairpins in the negative-sense RIII that are functionally redundant with regards to replication enhancement (Panavas & Nagy, 2003). It is proposed that the enhancer functions by not only binding the viral replicase complex, but also through a long-distance RNA-RNA interaction with the promoter for positive-strand synthesis. Mutations that strengthen the long-distance RNA-RNA interaction increased *in vivo* accumulation of DI RNA, suggesting that this interaction may play a role in the asymmetric accumulation of positive-strands over negative-strands (Panavas & Nagy, 2005).

Between the ORFs for the movement protein and coat protein on BMV RNA3 is an intergenic region that, when deleted, reduced negative-strand synthesis *in vitro* and *in*

vivo (French & Ahlquist, 1988; Sullivan & Ahlquist, 1999). SLC in the BMV TLS has been shown to recruit the replicase (Chapman & Kao, 1999; Dreher & Hall, 1988) and can accurately direct negative-strand synthesis from the 3' end when relocated to different locations in the genome (Ranjith-Kumar et al, 2003).

In RNA3 of AlMV is an enhancer element that is just upstream of the core promoter hpE. The enhancer is another small hairpin that, when deleted or mutated reduced *in vitro* negative-strand synthesis and *in vivo* accumulation (van Rossum et al, 1997a). The enhancer lies outside the conserved regions at the 3' ends of the three viral genomic RNAs indicating that each RNA may have a different means of enhancing negative-strand synthesis.

Upstream of RCNMV's core promoter is a series of small hairpins, two of which have been implicated in transcriptional enhancement of negative-strand synthesis. When deleted or mutated, SLDb and SLDc reduced negative-strand synthesis *in vitro* (Iwakawa et al, 2007).

CSE2, a 51 base conserved sequence near the 5' end of the SIN genome, has two structurally and phylogenetically predicted hairpins that function to enhance negative-strand synthesis *in vitro* and *in vivo* (Frolov et al, 2001a; Niesters & Strauss, 1990). Interestingly, the enhancement activity of CSE2 is more pronounced in mosquito cells than in mammalian cells (Fayzulin & Frolov, 2004; Levis et al, 1986). CSE2, while also present in DI RNAs, does not enhance transcription from these templates, indicating the virus has disparate means of regulating transcription for its genome as compared to DI RNAs(Levis et al, 1986).

Canonical Eukaryotic Translation

The majority of cellular mRNAs are translated in a cap-dependent manner and utilize various translation factors. Very briefly, eukaryotic initiation factor (eIF) 3 and eIF1A sequester the small (40S) ribosomal subunits. The ternary complex, containing eIF2, GTP, and met-tRNA$_i$, forms and then binds to the 40S subunit, forming the 43S complex. mRNA is activated in an ATP-dependent fashion by the binding of eIF4F (comprised of the eIF4E cap binding protein, the eIF4G scaffold that can also bind poly(A)-binding protein [PABP], and the helicase eIF4A) and PABP, thus circularizing the message. The 40S subunit, as part of the 43S complex, then binds the mRNA and scans, in an ATP-dependent manner, to the first AUG in a good context (Kozak, 1978), at which point eIF2•GTP hydrolyzes to eIF2•GDP and dissociates, allowing the 60S subunit to join the 40S subunit, and translation proceeds (Merrick, 2004).

When cells are stressed, cap-dependent translation is often compromised (Spriggs et al, 2008). Some cellular mRNAs utilize an internal ribosome entry site (IRES) to directly recruit the translational machinery which allows them to be translated in a cap-independent manner during these adverse conditions. IRESes vary in sequence and structure from small, unstructured sequences to large, highly structured domains (Balvay et al, 2009; Chappell et al, 2000; Kieft et al, 1999) and function by recruiting translation factors and ribosomal subunits to the 5' end of the message (Hellen & Sarnow, 2001; Jackson, 2005).

Positive-strand RNA viruses that lack either a 5' cap and/or a 3' poly(A) tail have developed alternative methods to recruit ribosomes and translation factors for translation. As such, a number of *cis*-acting sequences that serve to enhance translation have been

found in the 5' and 3' UTRs of these viruses, here grouped broadly into IRESes and translational enhancers (TEs).

Cis-acting Elements Involved in Viral Translation

Internal Ribosome Entry Sites (IRESes)

Blackcurrant reversion virus (BRV) is a bipartite nepovirus of the family *Comoviridae* that is polyadenylated but lacks a cap (Lemmetty et al, 1997). The 5' UTRs of both RNAs showed IRES activity, likely due to complementarity to the 18S rRNA (Karetnikov & Lehto, 2007; Karetnikov & Lehto, 2008). The 5' and 3' UTRs of both RNAs have at least 8 bases of complementarity between them that reduce translation when disrupted. Phylogenetic analysis of other nepoviruses did not reveal any conserved sequences, but all had the potential for a comparable 5'-3' interaction and complementarity to 18S rRNA (Karetnikov et al, 2006; Karetnikov & Lehto, 2008).

The potyvirus *Tobacco etch virus* (TEV) lacks a 5' cap and has a poly(A) tail. Within its 5' UTR are two cap-independent regulatory elements: the 5' proximal element and the 5' distal element (Niepel & Gallie, 1999). The 5' proximal element contains a pseudoknot and upstream unstructured sequence, which are required for translation (Zeenko & Gallie, 2005) in an eIF4G-dependant fashion (Gallie, 2001). The 5' UTR demonstrates IRES activity *in vivo* when placed intercistronically, thought to be due to the complementarity between the loop of the pseudoknot and 18S rRNA (Gallie, 2001).

Dicistroviruses have bicistronic genomic RNAs that have an IRES in the intergenic region (IGR) that directs translation of the downstream ORF (Wilson et al, 2000a; Wilson et al, 2000b). The IGR IRESes of dicistroviruses are particularly interesting in that they directly recruit the 40S ribosomal subunit, independent of

translation factors. The bound ribosome initiates at a non-AUG codon and immediately proceeds to elongation (Jan et al, 2003; Pestova & Hellen, 2003). The IGR IRES of *Cricket paralysis virus* (CrPV) is composed of three pseudoknots (PK1, PK2, PK3) that fold into a compact structure. PK2 and PK3 fold together and can bind 40S ribosomal subunits *in vitro* while PK1 is positioned in the P site of the ribosome (Costantino & Kieft, 2005; Costantino et al, 2008; Kanamori & Nakashima, 2001; Schüler et al, 2006; Spahn et al, 2004). The IGR IRES functions independently of translation factors both *in vitro* (Pestova & Hellen, 2003) and *in vivo* (Deniz et al, 2009), and function is enhanced when levels of ternary complex are low (Thompson et al, 2001), making it ideal for directing translation during times of cellular stress when canonical translation is downregulated (Deniz et al, 2009).

HCV-like IRESes are highly structured (Berry et al, 2011; Kieft et al, 1999; Thurner et al, 2004; Tsukiyama-Kohara et al, 1992) and bind the 40S subunit directly, in the absence of translation factors, positioning the initiator AUG in the P-site without scanning (Pestova et al, 1998; Reynolds et al, 1996; Rijnbrand et al, 1996). Unlike the dicistrovirus IRESes, translation from the IRES of *Hepatitis C virus* (HCV) requires the binding of eIF2•GTP•Met•$tRNA_i$ (Ji et al, 2004; Otto & Puglisi, 2004) to form a 48S complex and subsequent binding of eIF3 to allow the 60S subunit to join (Pestova et al, 1996a).

The IRES typified by *Encephalomyocarditis virus* (EMCV) is a large, multi-domained, highly structured region upstream of the viral initiation codon (Jang et al, 1990; Kolupaeva et al, 1998). eIF4G binds the IRES and this interaction is enhanced by eIF4A (Pestova et al, 1996b). Upon binding of the eIF4G•eIF4A complex, ATP-

dependent structural rearrangement downstream of the initiation codon occurs, presumably resulting in the formation of structures competent to bind ribosomes (Kolupaeva et al, 2003; Lomakin et al, 2000). This structural rearrangement is enhanced by *trans*-acting factors that vary from virus to virus (Yu et al, 2011). The EMCV IRES is capable of binding the 43S initiation complex in the absence of eIF1, eIF1A, and eIF4F (Pestova et al, 1996a; Pestova et al, 1998; Pestova et al, 1996b). As expected, the IRES assembles the 80S ribosome near the initiation codon thus making scanning unnecessary (Kaminski et al, 1990; Pestova et al, 1996b).

Translational Enhancers

While animal viruses tend to have highly structured 5' UTRs that contain IRESes for cap-independent translation, positive-strand plant viruses often harbor translational enhancers. Translational enhancers are *cis*-acting sequences that are thought to function through interaction elsewhere in the RNA and/or by binding factors required for translation and, surprisingly, are often found in the 3' UTRs of plant RNA viruses.

To facilitate translation, *Tomato bushy stunt virus* (TBSV) has in its 3' UTR a Y-shaped cap-independent translation enhancer (3' CITE), the core function of which maps to a region between RIII and RIV called R3.5 (Fabian & White, 2004; Na et al, 2006; Wu & White, 1999). Mutational analysis and electophoretic mobility shift assays revealed a five-base interaction with the T-shaped domain in the 5' UTR *in vitro* and *in vivo* (Fabian & White, 2004; Fabian & White, 2006). It is proposed that the 3' CITE recruits translational machinery and the long-distance RNA:RNA interaction between the 3' CITE and the 5' UTR allows transfer of the 43S complex to the 5' end where translation can commence (Fabian & White, 2004; Fabian & White, 2006). TBSV-like 3' CITEs are

conserved in all members of the Tombusvirus genus and appear possible in all genera in the family *Tombusviridae* (Fabian & White, 2004).

Satellite tobacco necrosis virus (STNV) has an extended stem loop that is a translation enhancer domain (TED) located in the 5' end of the 3' UTR (Meulewaeter et al, 1998b) that functionally replaces a 5' cap (Meulewaeter et al, 1998a) and can bind to eIF4E and eIFiso4E (Gazo et al, 2004). STNV's 5' and 3' UTRs have sequences that are complementary, but have not been shown to interact (Meulewaeter et al, 1998a). The 3' UTR also has complementarity to the 18S rRNA, which may help recruit ribosomes for cap-independent translation (Danthinne et al, 1993).

The BTE (BYDV-like translation element) is found in *Barley yellow dwarf virus* (BYDV) and all members of the Luteovirus, Necrovius, and Dianthovirus genera of the family *Tombusviridae* (Treder et al, 2008). BTEs can vary in length, but always have a 17 base conserved sequence and are functionally equivalent to the TED, although they share no sequence similarity (Wang et al, 1997). BYDV's BTE adopts a branched structure that has been shown to bind eIF4F (Treder et al, 2008) and engage in a long-distance RNA:RNA interaction with the 5' end of the genome (Guo et al, 2001; Rakotondrafara et al, 2006). Interestingly, the BTE of BYDV can be relocated within the genome and still function (Guo et al, 2000; Wang et al, 1997). When moved to the 5' UTR, the 5' interacting sequence was no longer required for cap-independent translation, indicating the purpose of the interaction is to transfer the translation factors bound at the 3' end to the 5' end (Guo et al, 2000) where translation can commence after scanning to the intiation codon (Rakotondrafara et al, 2006). The necroviruses *Tobacco necrosis virus A* and *D* (TNV-A, TNV-D) have BTEs that are capable of supporting cap-independent

translation. Sequences in the 5' and 3' UTRs support an interaction between these two regions (Shen & Miller, 2004). *Red clover necrotic mosaic virus* (RCNMV), a bipartite dianthovirus, has a translation element in the 3' UTR of RNA1 (TE-DR1) that resembles a BTE. There is the potential for base pairing between the 5' and 3' UTRs (Mizumoto et al, 2003).

The tombusvirus *Maize necrotic streak virus* (MNeSV) has in its 3' UTR a long, I-shaped structure (ISS) that enhances translation (Scheets & Redinbaugh, 2006). The ISS interacts with eIF4F through its eIF4E subunit and can also interact with the 5' UTR via a long-distance RNA:RNA interaction (Nicholson et al, 2010). This tripartite 3' CITE-5'UTR-eIF4F interaction has been shown to form *in vivo* and is required for ribosome recruitment to the 5' end for translation initiation (Nicholson et al, 2010).

RNA2 of the bipartite umbravirus *Pea enation mosaic virus* (PEMV) enhances translation through use of its PTE (PMV-like translational enhancer), found in the 3' UTR. The PTE, much like the ISS, functions by binding eIF4F via its eIF4E subunit (Wang et al, 2009). A structure adjacent to the PTE, the kissing loop tRNA-shaped structure (Kl-TSS), engages in a long-distance RNA:RNA kissing loop interaction with the 5' end, presumably for delivery of translation factors (F. Gao and A. E. Simon, unpublished). The carmovirus *Saguaro cactus virus* (SCV) also has a 3' PTE and adjacent structures required for efficient translation that engage in multiple long-distance RNA:RNA interactions with the 5' end, but have not yet been shown to bind any translation factors (Chattopadhyay et al, 2011).

tRNA-like structures (TLSes) also play a role in translational enhancement. In *Brome mosaic virus* (BMV), protein synthesis decreases when the TLS is disrupted,

though translation initiation is not dependent on tyrosylation of the TLS. Furthermore, translation can be recovered when the TLS is supplied in *trans* (Barends et al, 2004). *Tobacco mosaic virus* (TMV) has a TLS that, when coupled with the upstream pseudoknot domain, can functionally replace a poly(A) tail. This functional replacement was mapped to just the upstream pseudoknot domain, but full translational enhancement requires the 3' terminal TLS (Gallie & Walbot, 1990). Translational enhancement of TMV is also conferred by the 5' UTR, called Ω, in a sequence-dependent fashion (Gallie & Walbot, 1992). The valylated TLS of *Turnip yellow mosaic virus* (TYMV) and its upstream pseudoknot enhance translation of a reporter construct *in vivo* indicating a role in translational enhancement (Matsuda & Dreher, 2004). It was suggested that the valine on the TYMV TLS was incorporated as the first amino acid in the translated polyprotein (Barends et al, 2003), however this result was not repeatable and is currently questioned (Matsuda & Dreher, 2007).

The TCV System: A Model Positive-Strand RNA Virus Used for Studying Replication and Translation

Turnip crinkle virus (TCV) is a small, 4054 nucleotide (nt), positive-strand RNA virus in the genus Carmovirus, family *Tombusviridae*. We have been using TCV as a model to characterize *cis*-acting sequences necessary for replication and, more recently, translation. TCV has five overlapping open reading frames (ORFs) encoding two replicase proteins, two movement proteins, and one coat protein. The two replicase proteins, p28 and its readthrough product p88, are translated from the genomic RNA. p88 is the RdRp and is the only viral protein necessary for *in vitro* transcription of cognate RNAs (Rajendran et al, 2002). The two movement proteins, p8 and p9, are translated

from a 1.72 kb sgRNA and the coat protein, p38, is translated from a 1.45 kb sgRNA (Fig. 1.1). Associated with TCV is a small, non-translated satellite RNA, satC. Since satC is only replicated, it has been instrumental in identifying corresponding sequences in TCV that are necessary for replication. satC is only 356 bases and shares its 3' 166 bases with 3' sequences of TCV (Simon & Howell, 1986).

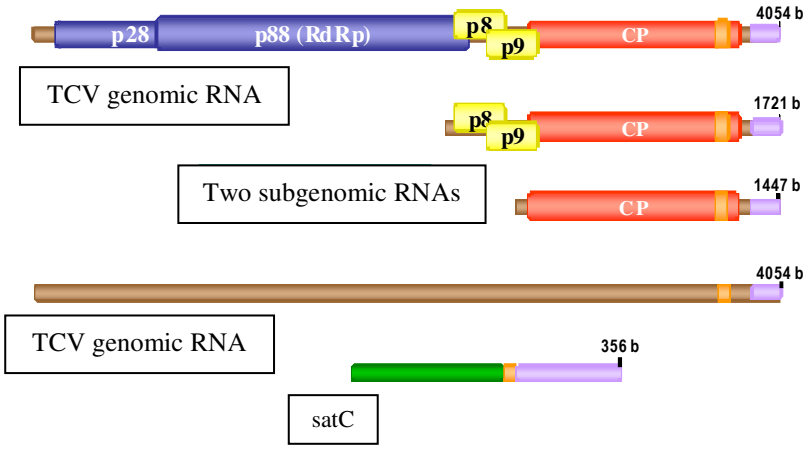


Figure 1.1 Genomic organization of TCV and satC TCV encodes five proteins: two replicase proteins (p28, p88) from the gRNA, two movement proteins (p8, p9) from sgRNA1, and coat protein (CP, p38) from sgRNA2. satC is a non-translated subviral RNA of TCV. Its 5' sequence is not similar to TCV, while its 3' sequence is derived from two regions of TCV, as indicated by color.

Core Promoter: Negative-Strand Synthesis

In satC, the core promoter for negative-strand synthesis was identified as the 29 3'-terminal bases, which form a hairpin (Pr) with a six base tail (Fig. 1.2). It was found that these bases were sufficient to direct transcription by partially purified RdRp of an attached, unrelated RNA template *in vitro* (Song & Simon, 1995). Mutational analysis of Pr indicated that the sequence and size of the loop could be varied (Song & Simon, 1995; Stupina & Simon, 1997), while disrupting the stem was detrimental to satellite accumulation *in vitro* (Song & Simon, 1995), but not *in vivo* (Stupina & Simon, 1997). *In vivo* systematic evolution of ligands by exponential enrichment (SELEX) is a useful technique for identifying important sequences, and possibly interactions, in a mutated region by allowing evolution to act on a large region of randomized bases. To further investigate the importance of the Pr hairpin, nearly the entire sequence was randomized and inoculated onto turnip plants (round 1). After 14 days, RNA was isolated from the plants and re-inoculated onto more plants (round 2). This was then done again, to result in round 3-isolated RNA. Analysis of the resultant RNA species indicated that the stability and fitness of the Pr hairpin increased with each round, which correlates with results from site-directed mutagenesis (Stupina & Simon, 1997), and the six base tail downstream of the Pr was recovered (Carpenter & Simon, 1998).

Deletion of the three terminal cytidylates of satC resulted in increased *in vitro* transcription of both full-length satC and aberrant products (Guan & Simon, 2000). Structure probing indicated that, in this deletion mutant, a complementary region upstream in the large symmetrical loop (LSL) of hairpin 5 (H5) appeared more single stranded as compared to wild type (wt) satC. This led to the supposition that these two

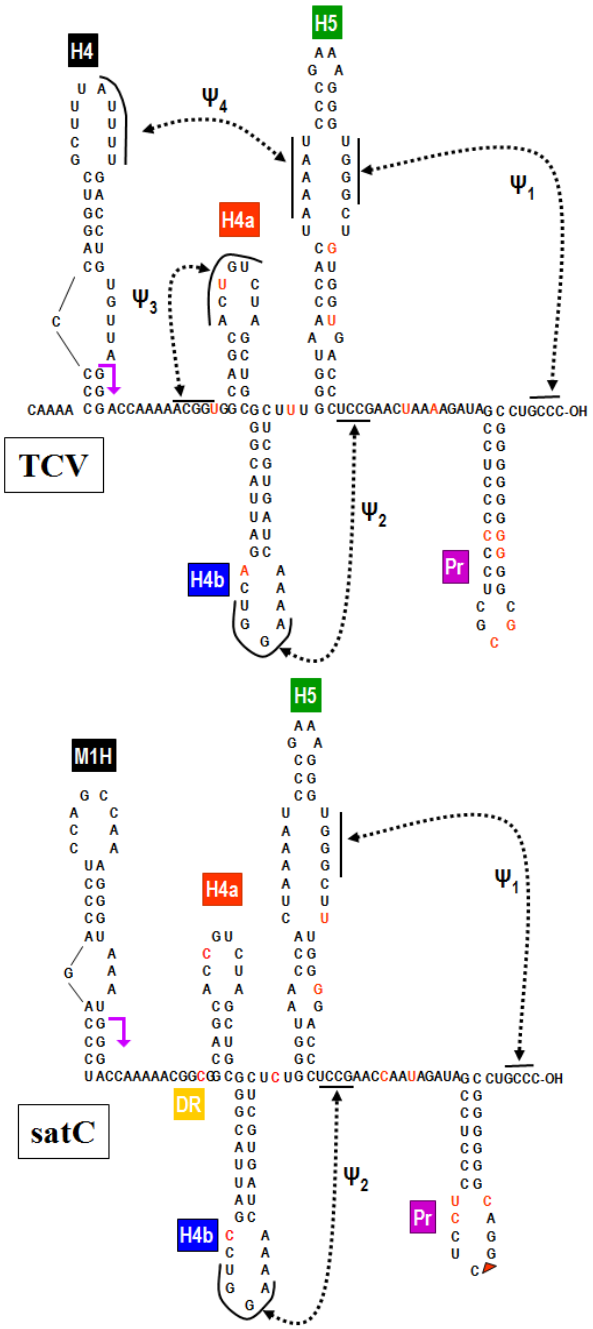


Figure 1.2 Sequence and structure of the 3' ends of TCV and satC Hairpins are labeled in colored boxes. Tertiary interactions (pseudoknots, Ψ) are depicted by double-headed arrows. Base differences between TCV and satC are shown in red, with the red triangle in the Pr loop of satC indicating a 4 base deletion. Purple arrows indicate the end of sequence similarity between TCV and satC.

regions might interact in a pseudoknot to regulate negative-strand synthesis. To test this, single and compensatory mutations were made that would disrupt or maintain the interaction, respectively. Disrupting the interaction resulted in increased *in vitro* transcription, while maintaining the interaction resulted in transcription resembling wt levels, indicating an interaction (termed pseudoknot 1, Ψ_1) between these two sequences for repression of negative-strand synthesis (Zhang et al, 2004).

satC was folded using a program called mFold (Zuker, 2003), which determines the free energies of RNA structures. The most stable structure was analyzed and it revealed a series of hairpins in the 3' terminal 140 bases; the Pr, H5, hairpin 4b (H4b), and hairpin 4a (H4a). All hairpins have been confirmed by *in vitro* structure probing (Zhang et al, 2004). The mFold structure also indicated the presence of Ψ_1 between the 3' terminal cytidylates and the 3' side of the large symmetric loop (LSL) of H5. Furthermore, phylogenetic analysis of closely related viruses identified H5s in nearly all carmoviruses. These H5s vary in the lower and upper stems, but all have stable tetra- or pentaloop terminal loops and consecutive guanylates on the 3' sides of their respective LSLs, that are predicted to basepair to the 3' terminal cytidylates (Pogany et al, 2003; Zhang et al, 2004). Indeed, a similar interaction had also been found in TBSV, which is in the Tombusvirus genus in the same family as TCV (Pogany et al, 2003).

Core Promoter: Positive-Strand Synthesis

There are also *cis*-acting elements on the negative-strand that are important for positive-strand synthesis. At the 3' end is the 3' terminal carmovirus consensus sequence (3' CCS), which is 6 bases that are conserved at the 3' ends of all carmoviral genomic, subgenomic, and subviral RNAs. The 3' CCS is a combination of three to seven

adenylates or uridylates followed by two or three cytidylates (Guan et al, 1997) and is required for positive-strand synthesis *in vivo* (Guan et al, 2000a), but not *in vitro* (Guan et al, 1997). Two additional negative-strand *cis*-acting elements that have promoter activity were identified using mutational analysis: the 3' proximal element (3'PE) located 11 bases from the 3' terminus of negative-strand satC that SELEX revealed prefers a CCS-like sequnece (Guan et al, 2000a), and the 5'PE located 41 bases from the 5' terminus of negative-strand satC which is required for positive-strand synthesis in protoplasts and *in vitro* and SELEX returned ten of the fourteen bases, indicating a primary sequence requirement (Guan et al, 2000b).

Enhancer

Motif 1 hairpin (M1H) is a 28 base hairpin in the negative-strand that functions as a replication enhancer. It is required for replication of satC *in vivo* (Nagy et al, 2001) and stimulates the synthesis of positive-strands by 10-fold and the synthesis of negative-strands by 7-fold *in vitro* (Nagy et al, 1999). Alone it is not sufficient for transcription; it needs the addition of consecutive cytidylates from which the polymerase can initiate (Nagy et al, 1999). Sequences flanking M1H may be required for full enhancer activity of M1H, as mutagenesis of these regions reduced *in vivo* accumulation (Nagy et al, 2001).

Cap-Independent Translation

Positive-strand RNA viruses must serve as template for both translation and replication, necessitating a means of regulating these two mutually exclusive processes. Ongoing work in the lab suggests that structures in TCV required for translation overlap with structures required for replication, facilitating the switch between these two roles for the RNA. H4 (see Fig. 1.2 for 3' end structures of TCV) is a perfect example of a

structure that plays important roles in both transcription and translation, although the mechanisms by which it serves these functions remain elusive. It serves as a replication enhancer in both orientations, for both positive- and negative-strand synthesis, when accompanied by the cognate core promoter (Sun & Simon, 2006); the hairpin Pr for negative-strand synthesis (Song & Simon, 1995), and the 3' carmovirus consensus sequence for positive-strand synthesis (Guan et al, 2000a; Guan et al, 1997). Disrupting the upper or lower stem reduced accumulation in protoplasts to 2% of wt levels and to no detectable accumulation, respectively, but accumulation was restored with compensatory mutations (Sun & Simon, 2006). Mutations in the asymmetric loop of H4 (H4AL) reduced translation (Stupina et al, 2008) and just one base change in H4AL (U3898G) eliminated *in vitro* transcription (Yuan et al, 2009). Deleting the five flanking adenynlates on the 3' side of H4 resulted in no accumulation in protoplasts (Sun & Simon, 2006) and mutating them to uridylates reduced translation to ~30% of wt (Stupina et al, 2008).

It has long been known that the 5' and 3' UTRs of TCV synergistically enhance translation, with the 5' UTR of sgRNA2 and the 3' UTR providing the best enhancement (Qu & Morris, 2000), but the mechanism of translational enhancement has not been elucidated. Our lab recently mapped the majority of the 3' translation enhancer activity to about 140 bases near the 3' end of the virus that includes part of the 3' UTR. Maximal translation enhancement required the entire 3' UTR as well as some sequence in the coding region of CP to be present, bases 3661-4054 (Stupina et al, 2008). This 3' translation enhancer overlaps with a 3' T-shaped structure (TSS, see Fig. 1.3) in the 3' UTR of TCV formed from three hairpins and two pseudoknots (McCormack et al, 2008).

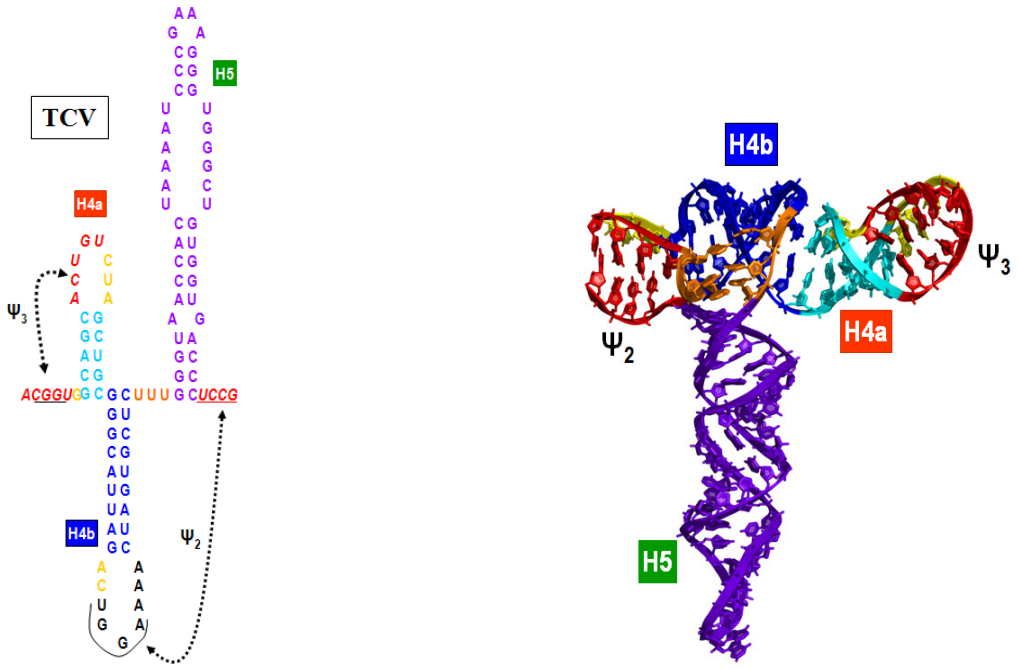


Figure 1.3 Sequence and structure of the tRNA-shaped structure (TSS) in the 3' end of TCV On the left is the sequence of TCV that folds into a tRNA-shaped structure, shown on the right. The regions are color coded the same (Stupina et al, 2008).

The TSS was found to bind both 80S ribosomes and 60S ribosomal subunits in a translation factor-independent fashion (Stupina et al, 2008). Though the sequence of nontranslated satC in this region is derived from TCV, the comparable sequence in satC does not form the TSS and is not able to bind ribosomes (Stupina et al, 2008). There is a TSS within the 3' UTR of a related virus, *Pea enation mosaic virus* (PEMV), an umbravirus. The H5 equivalent of PEMV is shorter than that of TCV, and, as such, forms an almost perfect tRNA-shaped structure. This indicates that there likely are methods of translational enhancement being utilized by positive-strand RNA viruses yet to be discovered.

Ψ_3/H4a is important for ribosome binding to the TSS (Stupina et al, 2008) and this region is also affected by binding of the RdRp (Yuan et al, 2009). Disrupting Ψ_3 resulted in structural changes not only locally, but distally in the loop of Pr (Yuan et al, 2009), which may indicate an interaction between these two regions, meaning that the Pr loop is in proximity to Ψ_3. This may indicate how this region functions: in part, the TSS forms and binds ribosomes, for which Ψ_3/H4a is particularly important, allowing translation of RdRp to occur. After a sufficient amount of translation, a threshold level of RdRp has been translated, such that it now is able to bind the 3' UTR, possibly also in the Ψ_3/H4a region, resulting in a conformational switch that prevents ribosomes from binding and, thus, prevents further translation. The Pr appears to be associated with the Ψ_3/H4a structure, which could provide a means for the RdRp to access the 3' terminus for negative-strand transcription.

Thesis Plan

Overall, the 3' end of TCV has proven to be very structured with many interactions both within the 3' end and upstream in the coat protein (CP) ORF (X. Yuan, K. Shi, A.E. Simon, unpublished; (Yuan et al, 2009; Yuan et al, 2012; Yuan et al, 2010) that are important for viral accumulation and translation. Due to the 3' end adopting a t-RNA shaped structure and binding ribosomes, I wanted to determine which of these structures were important for RdRp binding and if they overlap with structures involved in ribosome binding, possibly indicating a structural switch of the template between translation-competent and replication-competent. I discuss where the viral RdRp binds in the 3' end in Chapter II. In chapter III I address the role that upstream sequences in the CP ORF play in viral translation, with a specific emphasis on hairpin 3, which has proven to be important to viral functions. Accumulation of satC mutants identified by Systematic Evolution of Ligands by EXponential enrichment (SELEX) for functional H2 replacement is presented in Chapter IV. Finally, Chapter V presents a variety of data regarding altered RdRp binding when mutations are present in the Pr loop, 5'-3' interactions in TCV, functional domains within the 3' end of TCV, and eEF1A binding to the 3' end of TCV.

Chapter II

In Vitro RdRp Binding to a 3' End Fragment of TCV

Positive-strand RNA viral replicases, comprised of the virus-encoded RNA-dependent RNA polymerase (RdRp) together with host proteins and/or additional viral proteins (Gancarz et al, 2011; Lai, 1998; Li & Nagy, 2011; Nagy, 2011), must recognize their cognate RNAs through direct or indirect interaction with specific sequence or structural elements (*cis*-acting elements) on the template to initiate complementary strand synthesis. These RNA elements are generally present in 3' noncoding regions and assume various forms such as hairpins, pseudoknots, tRNA-like structures, poly(A) tails, or short primary sequences (Dreher, 1999). In addition to core promoters present at the 3' ends of both strands, *cis*-acting elements at the 5' end or internal elements have been found to bind RdRp, with RNA:RNA interactions or RNA-protein bridges directing the RdRp to the 3' end (Alvarez et al, 2005a; Filomatori et al, 2006; Frolov et al, 2001a; Herold & Andino, 2001; Isken et al, 2003; Klovins & van Duin, 1999; Pogany et al, 2005; Wu et al, 2001; Yoshinari & Dreher, 2000; You et al, 2001).

Since ribosomes can bind the TSS in the 3' UTR of TCV (Stupina et al, 2008) and the RdRp must start transcription at the 3' terminus, an obvious question is where (or if) in the 3' UTR the RdRp binds and whether this overlaps with structures needed for translation. Footprinting the location of RdRp binding was not a viable option, as the structure changes dramatically upon binding of the RdRp (Yuan et al, 2009; Yuan et al, 2012; Yuan et al, 2010) and there may even be transfer of the RdRp from an initial binding site to other binding sites, given that the RdRp must ultimately access the 3' terminus for transcription. Therefore, I used an *in vitro* RdRp binding assay in which

increasing amounts of radiolabled RNA was added to purified, recombinant RdRp. The reaction mixture was applied to nitrocellulose filters under vacuum and the levels of radioactivity were determined by scintillation counter. Non-linear regression was performed using GraphPad Prism4.

Materials and Methods

Preparation of Fragments

RNA fragments were *in vitro* transcribed from PCR amplified templates. The templates were made using wild type or mutation-containing oligonucleotides and wild type or mutant templates (see Table 2.1).

Table 2.1 Fragments for the filter binding assay

Fragment	Forward Oligo	Reverse Oligo	Template
Fragment 1 (F1/TSS)	T7-H4-M1H	TCV Linker	TCV66 (wt full length TCV)
F2	T7-H4-M1H	KK57	TCV66
F3	T7-M3H linker H4	TCV Linker	TCV66
F4	T7-M3H linker H4	KK57	TCV66
F4 end H4a	T7-M3H linker H4	End H4a	TCV66
F4 end H4b	T7-M3H linker H4	End H4b	TCV66
F4a	T7-F4a	KK57	TCV66
F4b	T7-M3H linker H4	F4b	TCV66
F5	H3F	KK57	TCV66
H4 + Pr	T7 TCV H4 F	KK57	TCV66
H4	T7 TCV H4 F	H4 R	TCV66
Pr	Ordered as RNA directly from IDT due to high GC stem		
MDV	Transcribed full-length from linearized plasmid containing MDV sequence		
CCFV 4a 4b in F4	T7-M3H linker H4	KK57	TCV CCFV H4a H4b
CCFV 4a 4b up in F4	T7-M3H linker H4	KK57	TCV-CCFV link3-H4b
m50 in H4	T7-M3H linker H4	H4 R	TCV-m50
m50 in F4	T7-M3H linker H4	KK57	TCV-m50
m18 in F4	T7-M3H linker H4	CCC4052AAA	TCV-m18
m51 in F4	T7-M3H linker H4	G3992C	TCV66
m52 in F4	T7-M3H linker H4	C4053G	TCV66
m51/m52 in F4	T7-M3H linker H4	C4053G/G3992C	TCV66
m44 in F1	T7-M3H linker H4	KK57	TCV-m44
m44 in F4	T7-M3H linker H4	KK57	TCV-m44
m45 in F4	T7-M3H linker H4	KK57	TCV-m45
m44/m45 in F4	T7-M3H linker H4	KK57	TCV-m44/m45

m46 in F1	T7-M3H linker H4	KK57	TCV-m46
m46 in F4	T7-M3H linker H4	KK57	TCV-m46
m47 in F1	T7-M3H linker H4	KK57	TCV-m47
m47 in F4	T7-M3H linker H4	KK57	TCV-m47
m46/47 in F1	T7-M3H linker H4	KK57	TCV-m46/m47
m46/47 in F4	T7-M3H linker H4	KK57	TCV-m46/m47
m27 in F4	T7-M3H linker H4	KK57	TCV-m27
m27 in F1	T7-H4-M1H	TCV Linker	TCV-m27
m27 in F3	T7-M3H linker H4	TCV Linker	TCV-m27
m26 in F4	T7-M3H linker H4	KK57	TCV-m26
m26 in F1	T7-H4-M1H	TCV Linker	TCV-m26
m26 in F3	T7-M3H linker H4	TCV Linker	TCV-m26
m26/m27 in F1	T7-H4-M1H	KK57	TCV-m26/m27
m26/m27 in F4	T7-M3H linker H4	KK57	TCV-m26/m27
m26/m27 in F3	T7-M3H linker H4	TCV Linker	TCV-m26/m27
m40 in F4	T7-G3912C	KK57	TCV66
m41 in F4	T7- T3923G	KK57	TCV66
m40/m41 in F4	T7-G3912C/T3923G	KK57	TCV66
m19 in F4	T7-M3H linker H4	KK57	TCV-m19
m19 in F3	T7-M3H linker H4	TCV Linker	TCV-m19
m21 in F4	T7-M3H linker H4	KK57	TCV-m21
m21 in F3	T7-M3H linker H4	TCV Linker	TCV-m21
m19/m21 in F4	T7-M3H linker H4	KK57	TCV-m19/m21
m19/m21 in F3	T7-M3H linker H4	TCV Linker	TCV-m19/m21
m38 in F4	T7-M3H linker H4	KK57	TCV-m38
m39 in F4	T7-M3H linker H4	KK57	TCV-m39
m38/m39 in F4	T7-M3H linker H4	KK57	TCV-m38/m39
m23 in F4	T7-M3H linker H4	KK57	TCV-m23
m36 in F4	T7-M3H linker H4	KK57	TCV-m36
m36 and m18 in F4	T7-M3H linker H4	CCC4052AAA	TCV-m36
m36 in F1	T7-H4-M1H	KK57	TCV-m36
m105 in F4	T7-M3H linker H4	KK57	TCV-A3975-7U
A3866U in F4a	T7-F4a A3866T	KK57	TCV66
A3867U in F4a	T7-F4a A3867T	KK57	TCV66
A3867U in F4	T7-A3867T	KK57	TCV66
A3908,9U in F4	T7-M3H linker H4	KK57	TCV-m74
A3908U in F4a	T7-F4a A3908T	KK57	TCV66
A3908U in F4	T7- A3908T	KK57	TCV66
A3909U in F4a	T7-F4a A3909T	KK57	TCV66
m36 in F4a	T7-F4a	KK57	TCV-m36
m36 in F4b	T7-F4b	KK57	TCV-m36
A3908U in F1	T7-F1 A3908T	TCV Linker	TCV66
A3909U in F1	T7-F1 A3909T	TCV Linker	TCV66
U3898G in F4	T7-M3H linker H4	KK57	TCV-Yn70
U3898G/A3975C in F4	T7-M3H linker H4	KK57	TCV-T3898G/A3975C
U3898G/A3976C in F4	T7-M3H linker H4	KK57	TCV-T3898G/A3976C
U3898G/A3977C in F4	T7-M3H linker H4	KK57	TCV-T3898G/A3977C
U3898G/A3978C in F4	T7-M3H linker H4	KK57	TCV-T3898G/A3978C

In Vitro Transcription of RNA Using T7 RNA Polymerase

To 8 µg of TCV cDNA linearized with SmaI or PCR fragments, 6 µl of 100 mM DTT, 12 µl of a ribonucleotide mix (5 mM of each base), 12 µl of 5x T7 RNA polymerase buffer (125 mM sodium chloride, 40 mM magnesium chloride, 10 mM spermidine, 200 mM Tris-HCl, pH 8.0), 0.5 µl of RNaseOut ribouclease inhibitor (40 U/µl; Invitrogen), 3 µl of T7 RNA polymerase, and water to a final volume of 60 µl were added and the reaction incubated at 37°C for 1 h. Reactions were phenol extracted using phenol/chloroform (1:1), vortexed, and then subjected to centrifugation at 13,000 rpm for 2 min at room temperature in a microcentrifuge. The aqueous layer was removed by pipet and 6 µl of 3 M sodium acetate and 120 µl of chilled 100% ethanol were added. The mixture was mixed, incubated at -80°C for at least 10 min, and centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant was decanted, the pellet washed with 70% ethanol, and then centrifuged at 13,000 rpm for 5 min at 4°C. The supernatant was discarded and the pellet dried and resuspended in water. Quality was assessed and concentration was calculated by subjecting the in-vitro transcripts to electrophoresis through a 1% agarose gel and using densitometry to compare the band intensity of the transcripts with a TCV RNA marker of a known concentration.

Expression and purification of MBP–RdRp and MBP

pMAL–c4X plasmid (New England Biolabs), containing a p88 (RdRp) insert fused to the resident maltose binding protein (MBP), was a gift from P.D. Nagy (University of Kentucky). pMAL–c4X without the p88 insert was used to prepare MBP. Plasmids were transformed into E. coli Rosetta (DE3) pLacI competent cells (Novagen) and grown on 1% agar plates containing 50 µg/ml ampicillan and 34 µg/ml chloramphenicol. A colony

was selected from the plate and grown to OD_{600} 0.5-0.6 in MB (10 g tryptone, 5 g yeast extract, 5 g NaCl per liter) containing 50 μg/ml ampicillan and 34 μg/ml chloramphenicol. IPTG (isopropylthiogalactopyranoside) was added to the culture to a concentration of 0.3 mM to induce protein expression. After 8 to 10 h of induction at 14°C, the cells were collected by centrifugation (5,000 rpm for 10 min), resuspended in column buffer (20 mM Tris, pH 8; 25 mM NaCl; 1 mM EDTA; 10 mM β-mercaptoethanol; 10 mM glycerol), and sonicated 8 times for 10 seconds each. The samples were then centrifuged (15,000 rpm for 5 min), and the supernatant was subjected to affinity-based chromatography through an amylose column. The column was washed with 100 ml column buffer, and the RdRp was eluted with 5 ml maltose-containing column buffer (10 mM maltose). All steps were carried out on ice or at 4°C. The proteins were visualized by sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis (SDS-PAGE) analysis and quantified using a Bradford assay and bovine serum albumin as a standard.

Filter Binding

Different fragments of TCV RNA were *in vitro* transcribed from PCR templates, dephosphorylated using Antarctic phosphatase (NEB), and radiolabeled with T4 Polynucleotide kinase (NEB) and γ - ^{32}P ATP. RNA was filtered through a G25 (GE) column, according to the manufacturer. Prior to the binding reaction, RNA was heated to 65°C and allowed to slowly cool to room temperature. The concentration of the RNA was determined using a spectrophotometer. Varying amounts of radiolabeled RNA and 1 μg RdRp were incubated together for 30 min at room temperature in Binding Buffer (50 mM Tris, pH 8.2; 10 mM $MgCl_2$; 10 mM DTT; 10% glycerol). The membrane filters

(Millipore) were moistened (2 x 1 ml) with Binding Buffer, the RNA:protein binding reaction applied to the filters, and the membranes again washed (2 x 3 ml) with Binding Buffer. Membrane filters were placed into scintillation vials containing 5 ml CytoScint scintillation fluid (MP Biomedical) and read on a scintillation counter (Perkin Elmer Tri-Carb 2800TR). GraphPad Prism4 (GraphPad Software) was used to perform one site binding nonlinear regression (equation: $Y=B_{max}*x/(K_d+x)$). Where Scatchard plots resulted in a nonlinear curve, which indicates a significant non-specific component to binding (see MDV control and various mutations), binding is indicated as "non-specific" and curves were fitted using a modified one site binding equation ($Y=B_{max}*x/(K_d+x) + NS*x$).

Native Polyacrylamide Gel Electrophoresis

RNA fragments were transcribed, electrophoresed through a 1% agarose gel for purification, dephosphorylated, radiolabeled, and subjected to electrophoresis through a denaturing 5% polyacrylamide gel for further purification. Purified, labeled fragments were heated to 65°C and allowed to slowly cool to room temperature. A 5% polyacrylamide gel was made with 44 mM Tris, 44 mM boric acid, and 5 mM $MgCl_2$. The fragments were electrophoresed through this native 5% polyacrylamide gel using running buffer containing 44 mM Tris, 44 mM boric acid, and 5 mM $MgCl_2$. The gel was run at 10 watts for approximately 1 hour. Following electrophoresis, the gel was dried and subjected to autoradiography.

Results

RdRp binding to 3' UTR fragments

To determine if the RdRp binds the 3' end of the virus, I first tested a series of fragments in the 3' region [F1, F2, F3, F4, F5, F4a, F4b, H4 + Pr, H4, and Pr (Fig. 2.1)]. While ribosomes bound best to F1 (Stupina et al, 2008) – the TSS alone – the RdRp bound best to F4 (K_d = 2.56 μM), which includes the TSS and upstream H4 and downstream Pr. Interestingly, deleting H4 in F4 (to produce F4a) did not change the binding of RdRp to the fragment (K_d = 2.78 μM), while deleting Pr in F4 (to make F4b) reduced RdRp binding by about half (K_d = 5.47 μM) as compared to binding to F4. An artificial fragment composed of H4 at the 5' end and Pr at the 3' end also had binding comparable to that of F4 (K_d = 2.87 μM). F4 RNA fragments truncated after either H4a (F4 end H4a) or H4b (F4 end H4b) resulted in weaker RdRp binding as compared to the intact F4 (K_d = 6.24 μM and 5.21 μM, respectively, compared to K_d = 2.56 μM).

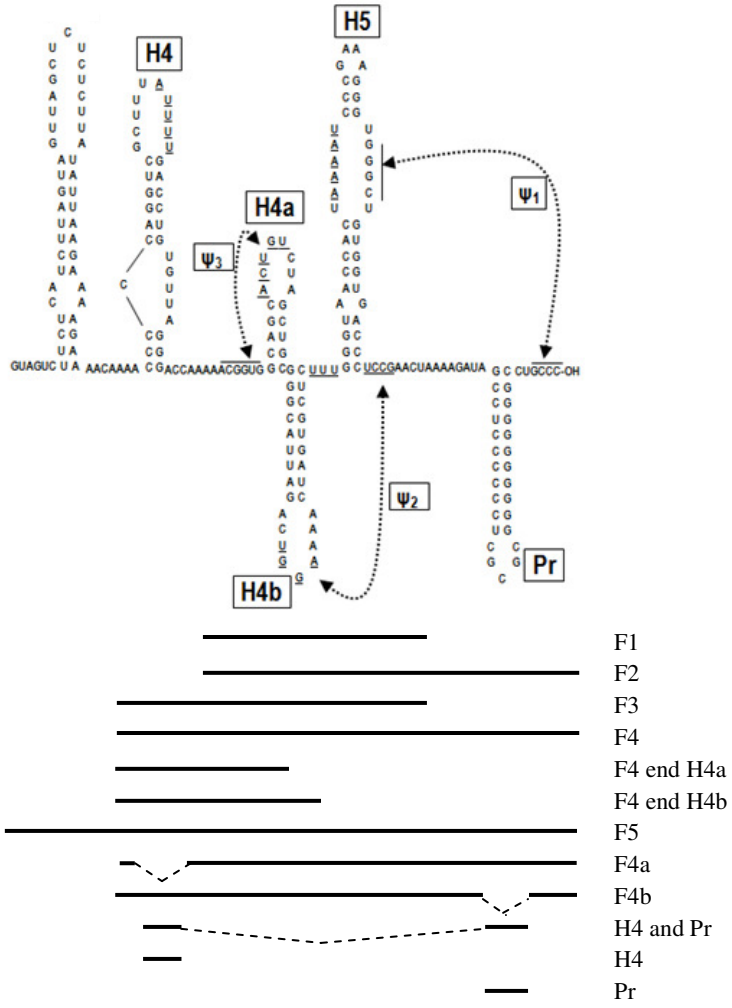
Figure 2.1 RdRp binding to 3' end fragments of TCV

A. The sequence and secondary structure of the 3' end of TCV. The lines below the structure indicate what regions are included in the various fragments. Dashed lines indicate deleted sequence.

B. K_d of RdRp binding to various fragments of the TCV 3' end and locations of the fragments.

C. GraphPad non-linear regressions of the fragments with Scatchard plot inset. Where Scatchard is non-linear, binding is denoted as non-specific.

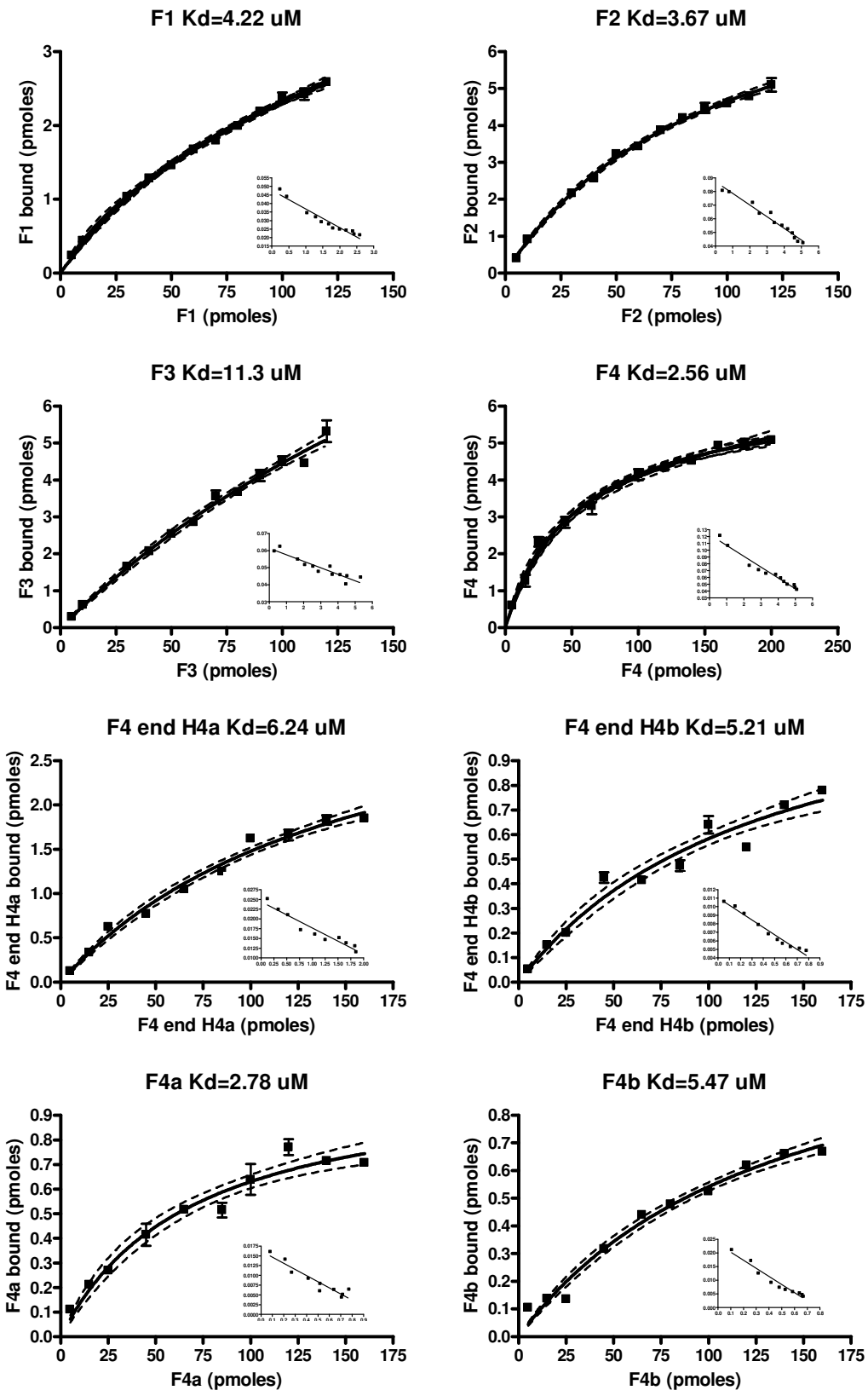
A

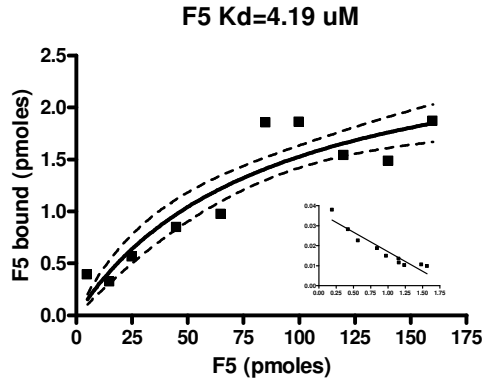


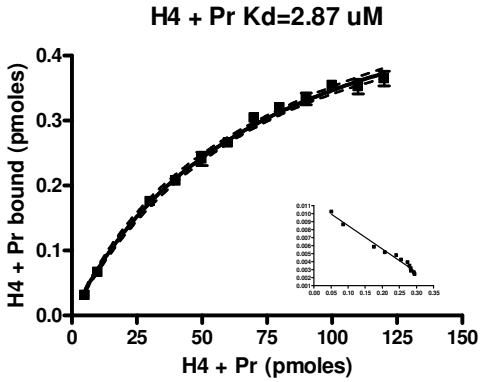
B

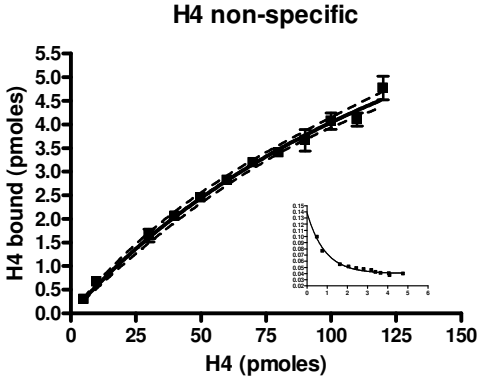
RNA	Kd, μM	Region of TCV
F1 (TSS)	4.22 ± 0.38	3901- 4017
F2	3.67 ± 0.29	3901 – 4054
F3	11.3 ± 2.47	3859 – 4017
F4	2.56 ± 0.25	3859 - 4054
F4 end H4a	6.24 ± 0.93	3859 – 3937
F4 end H4b	5.21 ± 1.15	3859 - 3966
F4a	2.78 ± 0.52	F4 Δ 3870-3903
F4b	5.47 ± 0.70	F4 Δ 4022-4048
F5	4.19 ± 0.83	3812 - 4054
H4 and Pr	2.87 ± 0.22	3870-3903 + 4022-4048
H4	non-specific	3870 - 3903
Pr	6.06 ± 1.69	4022 - 4048
MDV	non-specific	negative control

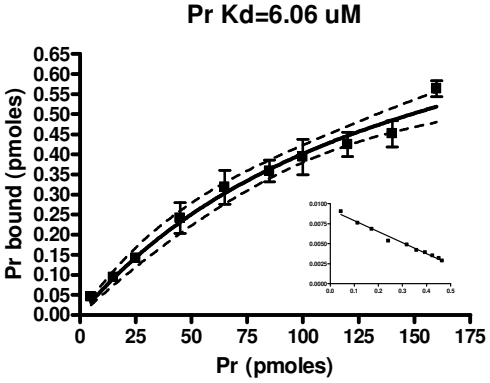
C

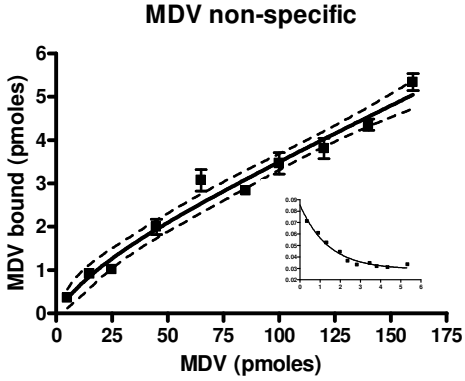

F1 Kd=4.22 uM
F1 bound (pmoles)
F1 (pmoles)
F2 Kd=3.67 uM
F2 bound (pmoles)
F2 (pmoles)
F3 Kd=11.3 uM
F3 bound (pmoles)
F3 (pmoles)
F4 Kd=2.56 uM
F4 bound (pmoles)
F4 (pmoles)
F4 end H4a Kd=6.24 uM
F4 end H4a bound (pmoles)
F4 end H4a (pmoles)
F4 end H4b Kd=5.21 uM
F4 end H4b bound (pmoles)
F4 end H4b (pmoles)
F4a Kd=2.78 uM
F4a bound (pmoles)
F4a (pmoles)
F4b Kd=5.47 uM
F4b bound (pmoles)
F4b (pmoles)


F5 Kd=4.19 uM
F5 bound (pmoles)
F5 (pmoles)


H4 + Pr Kd=2.87 uM
H4 + Pr bound (pmoles)
H4 + Pr (pmoles)


H4 non-specific
H4 bound (pmoles)
H4 (pmoles)


Pr Kd=6.06 uM
Pr bound (pmoles)
Pr (pmoles)


MDV non-specific
MDV bound (pmoles)
MDV (pmoles)

RdRp binding to F1 and F2 was reduced compared to F4 (K_d = 4.97 μM and 3.67 μM, respectively, vs. K_d = 2.56 μM), while binding to F3 was poor (K_d = 11.3 μM). In-line probing results indicate that the structure of F3 is unpredictable and unstable (X. Yuan and A. E. Simon, unpublished), possibly explaining this poor binding result. Binding was poor to single hairpins H4 or Pr (non-specific and K_d = 6.06 μM, respectively), suggesting that it is not these individual components alone that improve binding to F4 compared to F3 and the other fragments. Addition of upstream sequence (to produce F5) that is poorly structured on the positive-strand (Yuan et al, 2012) and forms a hairpin on the negative-strand (Sun & Simon, 2006) resulted in binding that is 1.6 times weaker than to F4. MDV, a small RNA associated with Qβ bacteriophage that is not a template for *in vitro* transctription by the TCV RdRp, was bound non-specifically by the RdRp.

Mutations in F4 alter RdRp binding

Since binding was best to F4, I generated mutations in this fragment that should disrupt some of the structures to determine which structures are important for RdRp binding. I present them here, element by element, for clarity.

Ψ_1

When Ψ_1, the pseudoknot connecting the 3' terminus and the 3' side of the large symmetric loop (LSL) of H5, was disrupted from either side (m18, m51, m52), RdRp binding increased (0.86, 1.36, and 1.76 μM, respectively, compared to 2.56 μM) and compensatory mutations restored binding to wt F4 levels (Fig. 2.2).

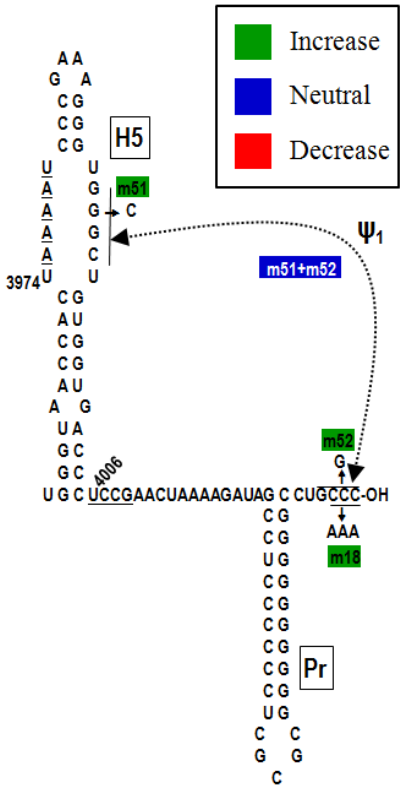
Figure 2.2 Effect of Ψ_1 mutations on RdRp binding to F4

A. The sequence and secondary structure showing Ψ_1. Mutation names are color coded to indicate an increase in binding (green), a decrease in binding (red), or no effect on binding (blue).

B. The K_d results of RdRp binding to F4 containing Ψ_1 mutations.

C. GraphPad non-linear regressions with Scatchard plot inset.

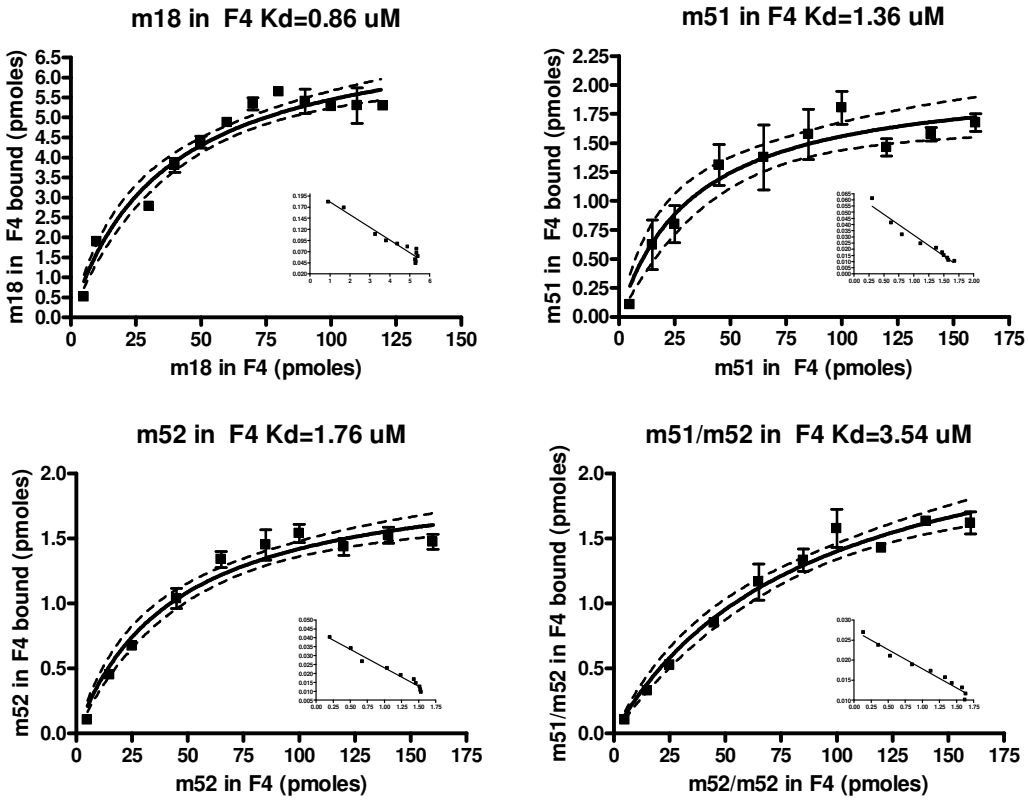
A



B

Mutation in F4	Kd, μM
F4	2.56 ± 0.25
Mutations affecting Ψ_1	
m18 in F4	0.86 ± 0.24
m51 in F4	1.36 ± 0.41
m52 in F4	1.76 ± 0.29
m51/m52 in F4	3.54 ± 0.69

C


m18 in F4 Kd=0.86 uM
m18 in F4 bound (pmoles)
m18 in F4 (pmoles)
m51 in F4 Kd=1.36 uM
m51 in F4 bound (pmoles)
m51 in F4 (pmoles)
m52 in F4 Kd=1.76 uM
m52 in F4 bound (pmoles)
m52 in F4 (pmoles)
m51/m52 in F4 Kd=3.54 uM
m51/m52 in F4 bound (pmoles)
m52/m52 in F4 (pmoles)

Ψ_2

Disrupting Ψ_2 from either side (m44, m45, m46, m47) resulted in non-specific binding (Fig. 2.3). The compensatory mutations m44/m45 did not restore specific binding, while the compensatory mutation m46/m47 did (K_d = 4.88 μM). This could indicate that there is sequence specificity in these regions or that the one set of compensatory mutations did not restore the pseudoknot.

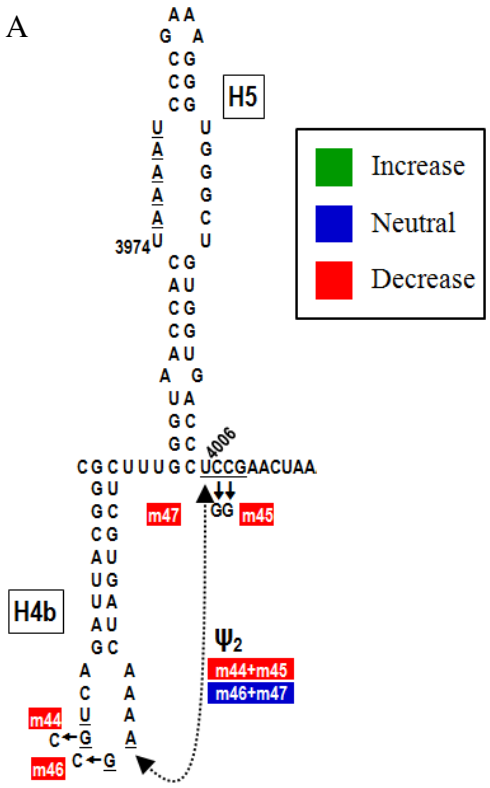
Figure 2.3 Effect of Ψ_2 mutations on RdRp binding to F4

A. The sequence and secondary structure showing Ψ_2. Mutation names are color coded to indicate an increase in binding (green), a decrease in binding (red), or no effect on binding (blue).

B. The K_d results of RdRp binding to F4 containing Ψ_2 mutations. Effect of these mutations on accumulation of full-length virus in protoplasts (McCormack et al, 2008).

C. GraphPad non-linear regressions with Scatchard plot inset. Where Scatchard is non-linear, binding is denoted as non-specific.

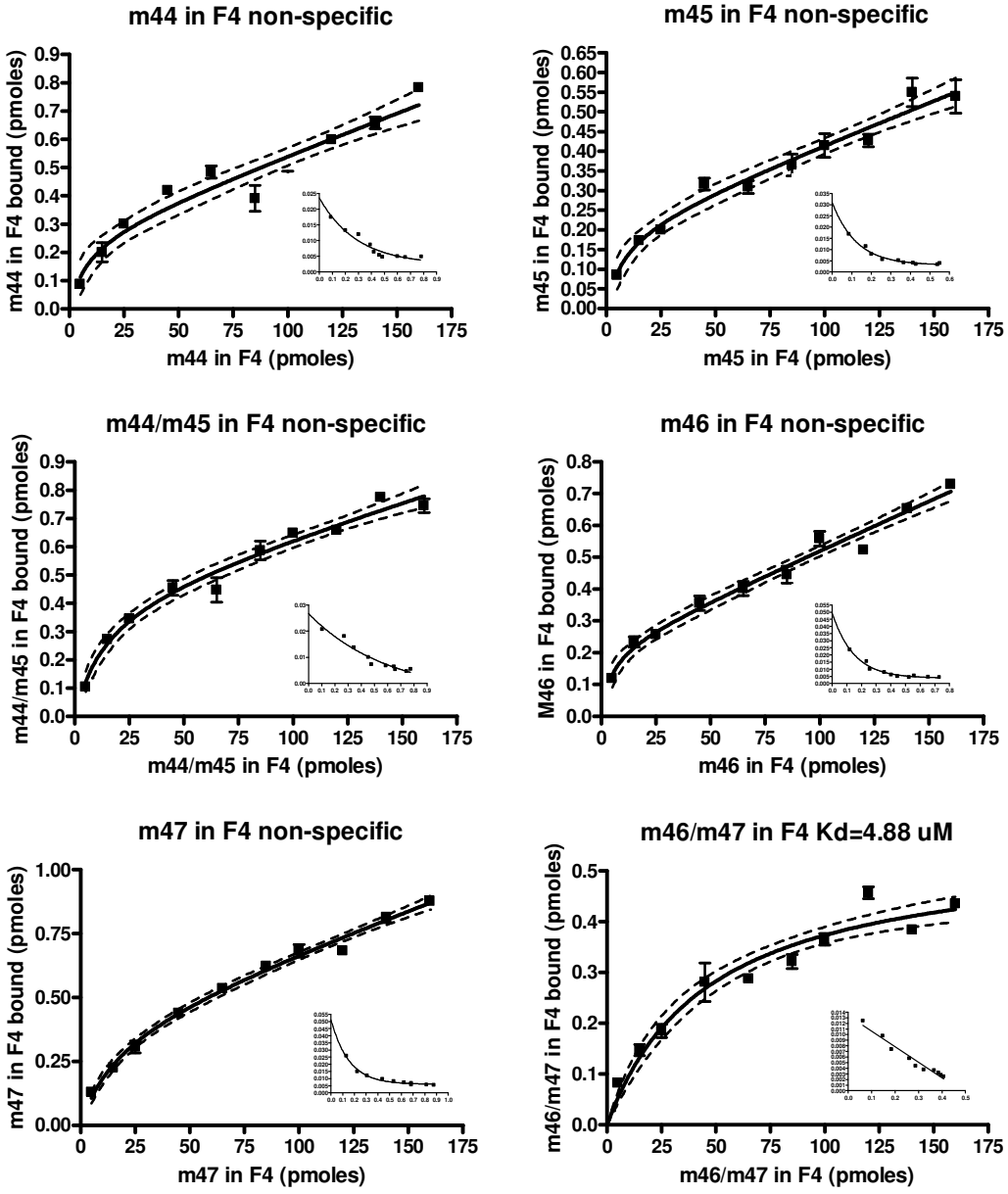
A



B

Mutation in F4	Kd, μM	Effect on accumulation
F4	2.56 ± 0.25	
Mutations affecting Ψ_2		
m44 in F4	non-specific	50%
m45 in F4	non-specific	20%
m44/m45 in F4	non-specific	58%
m46 in F4	non-specific	31%
m47 in F4	non-specific	34%
m46/47 in F4	4.88 ± 0.33	80%

C


m44 in F4 non-specific
m44 in F4 bound (pmoles)
m44 in F4 (pmoles)
m45 in F4 non-specific
m45 in F4 bound (pmoles)
m45 in F4 (pmoles)
m44/m45 in F4 non-specific
m44/m45 in F4 bound (pmoles)
m44/m45 in F4 (pmoles)
m46 in F4 non-specific
M46 in F4 bound (pmoles)
m46 in F4 (pmoles)
m47 in F4 non-specific
m47 in F4 bound (pmoles)
m47 in F4 (pmoles)
m46/m47 in F4 Kd=4.88 uM
m46/m47 in F4 bound (pmoles)
m46/m47 in F4 (pmoles)

Ψ_3

When Ψ_3 was disrupted from the 3'/H4a loop side (m27, m41), binding was enhanced (K_d = 0.94 and 0.86 μM, respectively, compared to K_d = 2.56 μM), whereas when the pseudoknot was disrupted from the 5' side (m26, m40), binding was non-specific (Fig. 2.4). The mutations combined together were not compensatory, suggesting a possible sequence requirement on the 5' side for RdRp binding.

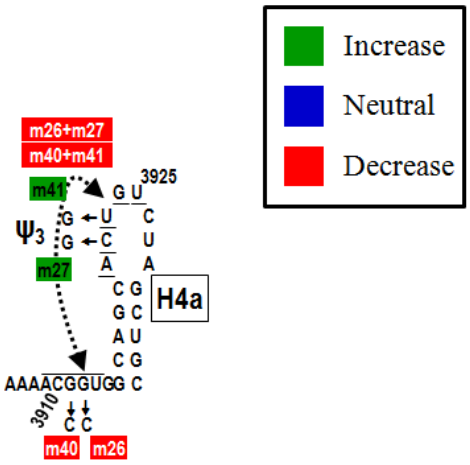
Figure 2.4 Effect of Ψ_3 mutations on RdRp binding to F4

A. The sequence and secondary structure showing Ψ_3. Mutation names are color coded to indicate an increase in binding (green), a decrease in binding (red), or no effect on binding (blue).

B. The K_d results of RdRp binding to F4 containing Ψ_3 mutations. Effect of these mutations on full-length virus accumulation in protoplasts (McCormack et al, 2008) and relative translation in protoplasts (Stupina et al, 2008). ND = not determined.

C. GraphPad non-linear regressions with Scatchard plot inset. Where Scatchard is non-linear, binding is denoted as non-specific.

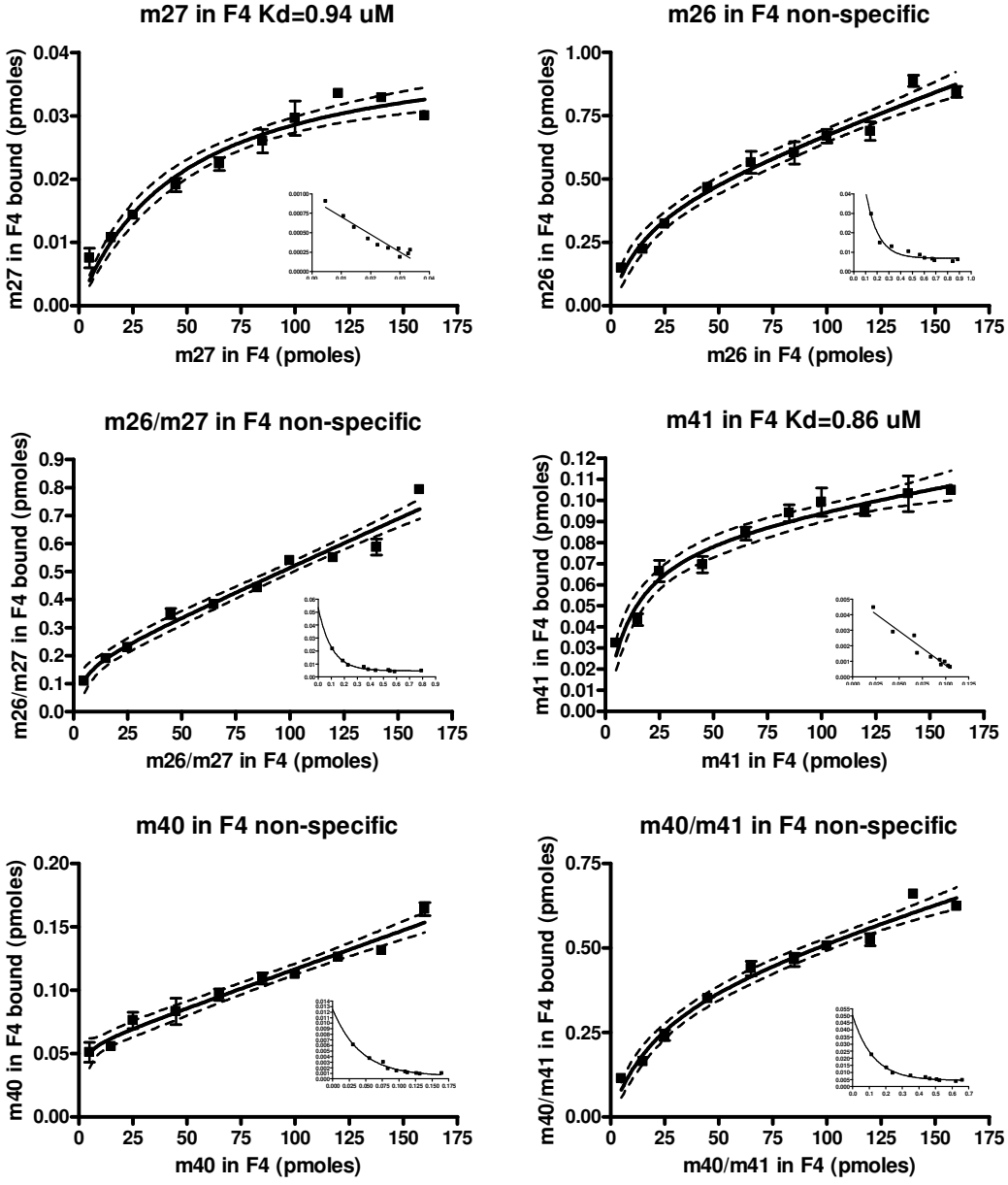
A



B

Mutation in F4	Kd, μM	Effect on accumulation	Effect on *in vivo* translation
F4	2.56 ± 0.25		
Mutations affecting Ψ_3			
m27 in F4	0.94 ± 0.17	2%	33%
m26 in F4	non-specific	1%	33%
m26/m27 in F4	non-specific	55%	77%
m41 in F4	0.86 ± 0.13	16%	ND
m40 in F4	non-specific	11%	ND
m40/m41 in F4	non-specific	22%	ND

C


m27 in F4 Kd=0.94 uM
m27 in F4 bound (pmoles)
m27 in F4 (pmoles)
m26 in F4 non-specific
m26 in F4 bound (pmoles)
m26 in F4 (pmoles)
m26/m27 in F4 non-specific
m26/m27 in F4 bound (pmoles)
m26/m27 in F4 (pmoles)
m41 in F4 Kd=0.86 uM
m41 in F4 bound (pmoles)
m41 in F4 (pmoles)
m40 in F4 non-specific
m40 in F4 bound (pmoles)
m40 in F4 (pmoles)
m40/m41 in F4 non-specific
m40/m41 in F4 bound (pmoles)
m40/m41 in F4 (pmoles)

5' side LSL of H5

Mutations in the 5' side of the large symmetric loop (LSL) in H5 were either neutral (m19, m36; K_d = 2.80 and 3.06 μM, respectively, compared to K_d = 2.56 μM) or detrimental (m23, m38, m105) for RdRp binding (Fig. 2.5). Of the detrimental mutations, only m38 had specific binding, 7.00 μM. A mutation (m39) in the terminal loop of H4 (H4TL) had little effect on RdRp binding (K_d = 2.95 μM) and combining it with m38 in the 5' side of the LSL to potentially restore Ψ_4 resulted in non-specific binding. It is possible that Ψ_4 does not form in the conditions used for *in vitro* RdRp binding or that these two mutations did not restore the pseudoknot.

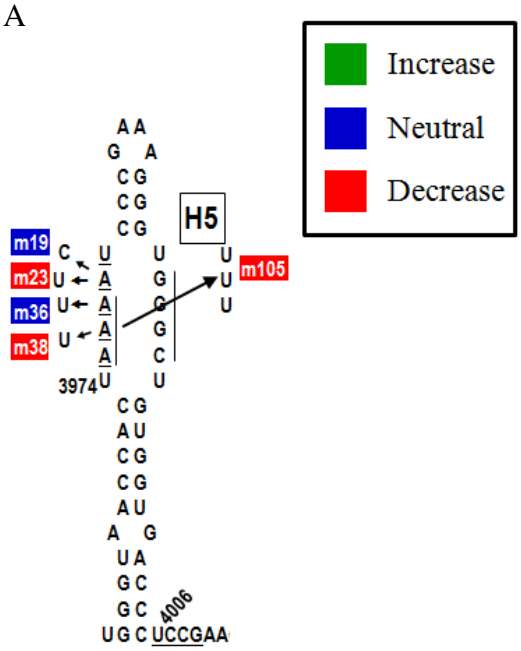
Figure 2.5 Effect of 5' side LSL mutations on RdRp binding to F4

A. The sequence and secondary structure showing the 5' side LSL mutations. Mutation names are color coded to indicate an increase in binding (green), a decrease in binding (red), or no effect on binding (blue).

B. The K_d results of RdRp binding to F4 containing 5' side LSL mutations. Effect of these mutations on full-length virus accumulation in protoplasts (m105, Yuan et al, 2009; others Stupina et al, 2008) and relative translation in protoplasts (Stupina et al, 2008). ND = not determined.

C. GraphPad non-linear regressions with Scatchard plot inset. Where Scatchard is non-linear, binding is denoted as non-specific.

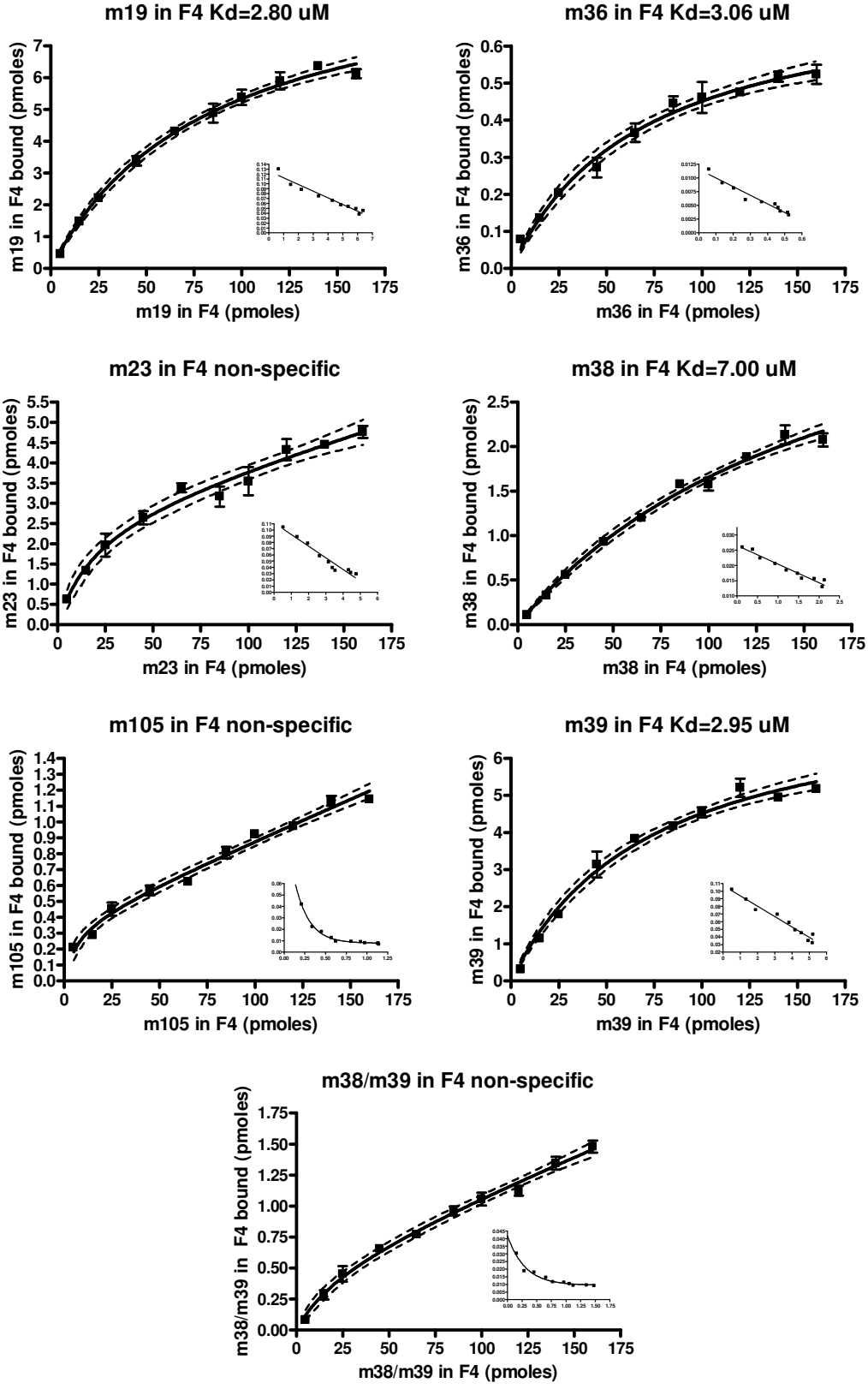
A



B

Mutation in F4	Kd, μM	Effect on accumulation	Effect on *in vivo* translation
F4	2.56 ± 0.25		
Mutations in H5 LSL			
m19 in F4	2.80 ± 0.30	74%	ND
m36 in F4	3.06 ± 0.78	ND	ND
m23 in F4	non-specific	31%	105%
m38 in F4	7.00 ± 1.06	88%	111%
m105 in F4	non-specific	6%	ND
m39 in F4	2.95 ± 0.37	55%	33%
m38/m39 in F4	non-specific	ND	ND

C


m19 in F4 Kd=2.80 uM
m19 in F4 bound (pmoles)
m19 in F4 (pmoles)
m36 in F4 Kd=3.06 uM
m36 in F4 bound (pmoles)
m36 in F4 (pmoles)
m23 in F4 non-specific
m23 in F4 bound (pmoles)
m23 in F4 (pmoles)
m38 in F4 Kd=7.00 uM
m38 in F4 bound (pmoles)
m38 in F4 (pmoles)
m105 in F4 non-specific
m105 in F4 bound (pmoles)
m105 in F4 (pmoles)
m39 in F4 Kd=2.95 uM
m39 in F4 bound (pmoles)
m39 in F4 (pmoles)
m38/m39 in F4 non-specific
m38/m39 in F4 bound (pmoles)
m38/m39 in F4 (pmoles)

H4 and downstream flanking adenylates

m21 and m50 in the asymmetric loop of H4 (H4AL) had little effect on RdRp binding (K_d = 2.89 and 3.82 μM, respectively, compared to K_d = 2.56 μM, Fig. 2.6). While changing three bases in H4AL (m21) did not substantially affect RdRp binding to F4, changing just one of the bases (U3898G, Ym70) resulted in non-specific binding. Mutating the other side of H4AL (m50) had a slight negative effect on RdRp binding (K_d = 3.82 μM compared to 2.56 μM). Mutation of adenylates flanking H4 to uridylates (m55, m74) eliminated specific RdRp binding.

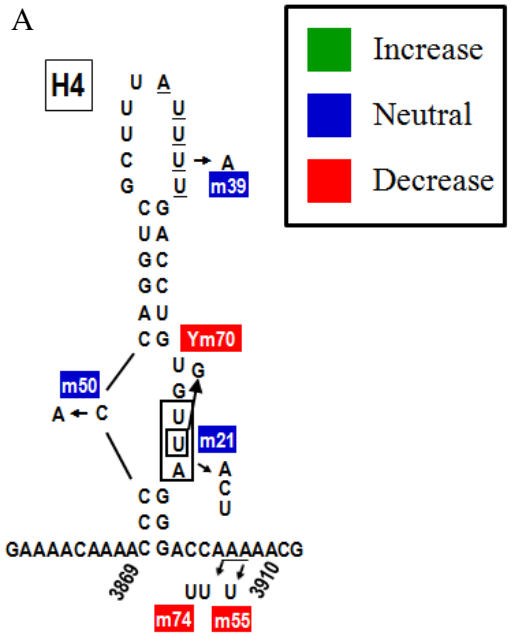
Figure 2.6 Effect of mutations in and around H4 on RdRp binding to F4

A. The sequence and secondary structure showing the mutations in and around H4. Mutation names are color coded to indicate an increase in binding (green), a decrease in binding (red), or no effect on binding (blue).

B. The K_d results of RdRp binding to F4 containing mutations in and around H4. Effect of these mutations on full-length virus accumulation in protoplasts (m74, Yuan et al, 2009; others Stupina et al, 2008) and relative translation in protoplasts (Stupina et al, 2008). In some cases, data from similar mutations were included (indicated). ND = not determined.

C. GraphPad non-linear regressions with Scatchard plot inset. Where Scatchard is non-linear, binding is denoted as non-specific.

A

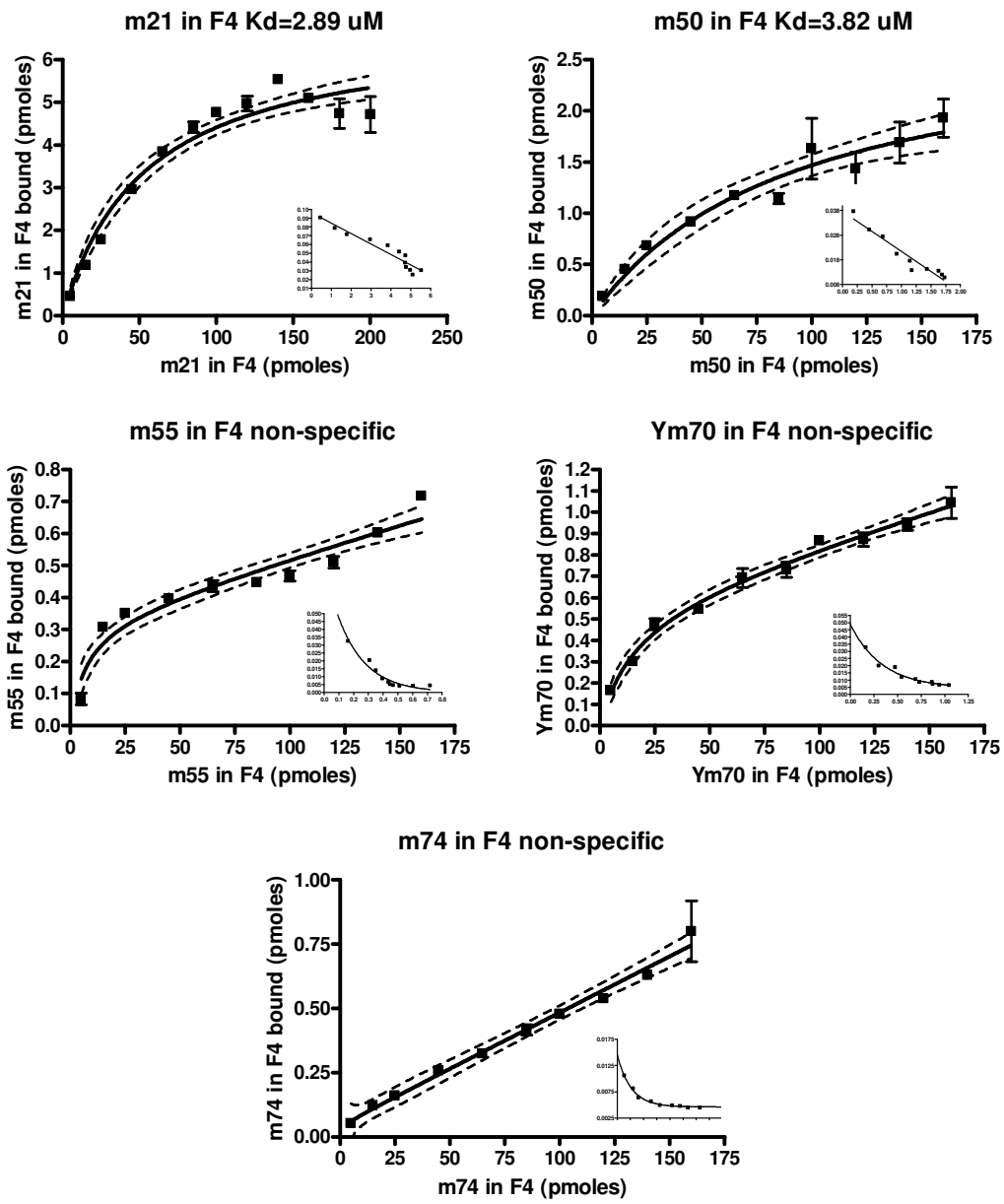


B

Mutation in F4	Kd, μM	Effect on accumulation	Effect on *in vivo* translation
F4	2.56 ± 0.25		
Mutations in/around H4			
m21 in F4	2.89 ± 0.32	34%	16%
m50 in F4	3.82 ± 1.25	9%	ND
m55 in F4	non-specific	26%	ND
Ym70 in F4	non-specific	4%	ND
m74 in F4	non-specific	3%	28% (m49)*

* m74 is AA3907UU whereas m49 is AAAAA3906UUUUU, so they are similar mutations, but not the same

C


m21 in F4 Kd=2.89 uM
m21 in F4 bound (pmoles)
m21 in F4 (pmoles)
m50 in F4 Kd=3.82 uM
m50 in F4 bound (pmoles)
m50 in F4 (pmoles)
m55 in F4 non-specific
m55 in F4 bound (pmoles)
m55 in F4 (pmoles)
Ym70 in F4 non-specific
Ym70 in F4 bound (pmoles)
Ym70 in F4 (pmoles)
m74 in F4 non-specific
m74 in F4 bound (pmoles)
m74 in F4 (pmoles)

Addition of monovalent cations does not eliminate non-specific binding

The conditions utilized in the RdRp binding assay came from previous studies in the lab (Sun & Simon, 2006). The buffer is 50 mM Tris, pH 8.2; 10 mM $MgCl_2$; 10 mM DTT; and 10% glycerol. However, in the *in vitro* transcription RdRp assay, in which purified TCV RdRp is used to transcribe complementary RNA strands, the reaction contains potassium glutamate (100 mM) as well as $MgCl_2$. To determine whether the presence of potassium glutamate affects the filter binding results, I performed filter binding with some mutations in F4 in the presence of 20 or 100 mM potassium glutamate. As an RNA engages in tertiary structure, the charge density increases as the phosphates in the backbone of the RNA molecule come into closer proximity. More structured RNAs have higher charge densities, necessitating cations for neutralization. K^+ can help neutralize this increase in charge density that occurs (Draper, 2004). Additionally, since K^+ is a somewhat large cation, I also performed filter binding with just a few mutations in F4 in the presence of NaCl, to determine if the smaller Na^+ cation affects binding differently.

Potassium glutamate (100 mM) did not have a significant effect on the filter binding results, as compared to conditions without any monovalent cations (Fig. 2.7). Some specific fragments or mutations had different results (F4a, m74, m18), but the results were similar between the conditions of 100 mM potassium glutamate and no cations. The binding of RdRp to F4 in all four conditions was comparable, bolstering the hypothesis that F4 is a functional domain that folds into a very stable structure. Of particular interest is that mutations in Ψ_3 performed similarly to no monovalent cation in both K^+ conditions, with disruption in the flanking region resulting in non-specific

binding (m26), disruption from the H4a side maintaining specific binding (m27), and the two together increasing non-specific binding (m26/27). Different results were obtained for these mutations in the presence of Na^+, with m26 and m26/27 maintaining specific binding and m27 resulting in non-specific binding. Altogether, these results did not warrant determining K_d values for all the mutations in the presence of 100 mM potassium glutamate.

Figure 2.7 RdRp binding to the 3' end of TCV in the presence of a monovalent cation

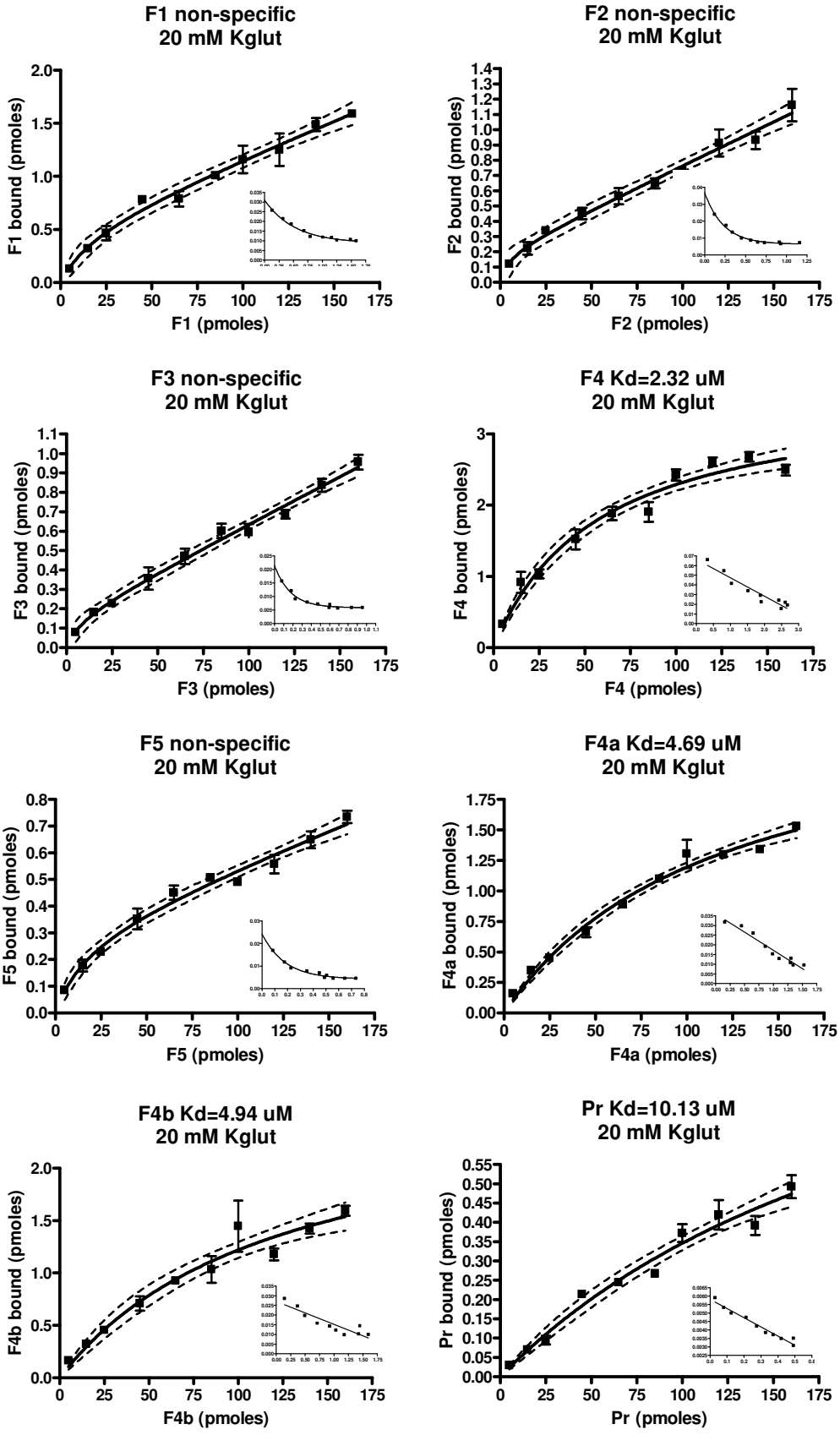
A. Binding of RdRp to different fragments and different mutations in F4 is shown in μM. Binding assays were conducted in the presence of 20 mM or 100 mM potassium glutamate or 20 mM sodium chloride. The original binding values with no monovalent cation in the binding buffer are shown for reference ("none"). ND = not determined.

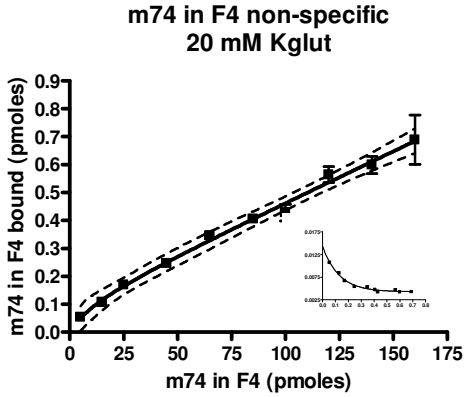
B. GraphPad non-linear regressions with Scatchard plot inset. Where Scatchard is non-linear, binding is denoted as non-specific.

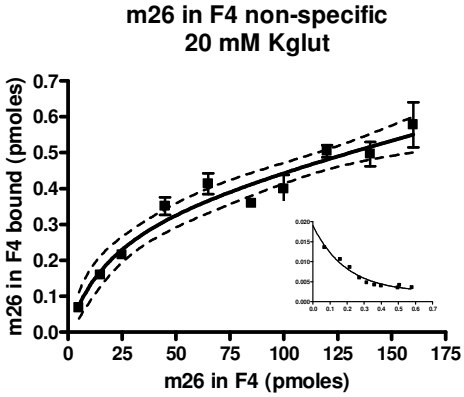
A

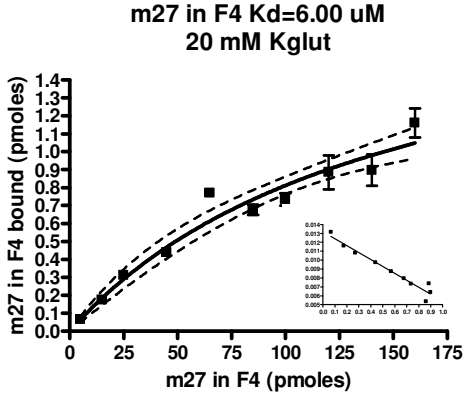
RNA	none	20 mM Kglut	100 mM Kglut	20 mM NaCl
F1	4.97 ± 0.85	non-specific	2.93 ± 0.65	ND
F2	3.67 ± 0.29	non-specific	3.40 ± 0.88	ND
F3	11.3 ± 2.47	non-specific	non-specific	ND
F4	2.56 ± 0.25	2.32 ± 0.37	3.01 ± 0.89	1.49 ± 0.39
F5	4.19 ± 0.83	non-specific	5.86 ± 1.60	ND
F4a	2.78 ± 0.52	4.69 ± 0.71	non-specific	ND
F4b	5.47 ± 0.70	4.94 ± 1.51	5.94 ± 1.04	ND
Pr	6.06 ± 1.69	10.13 ± 3.35	9.74 ± 2.43	ND
m74 in F4	44.64 ± 49	non-specific	5.60 ± 0.91	ND
m26 in F4	non-specific	non-specific	non-specific	3.07 ± 0.61
m27 in F4	0.94 ± 0.17	6.00 ± 1.86	5.98 ± 1.98	non-specific
m26/27 in F4	non-specific	non-specific	non-specific	3.71 ± 1.09
m105 in F4	non-specific	non-specific	non-specific	ND
m18	0.86 ± 0.24	5.43 ± 0.84	6.96 ± 1.81	ND

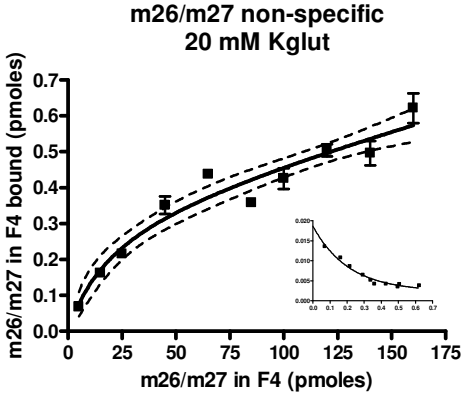
B

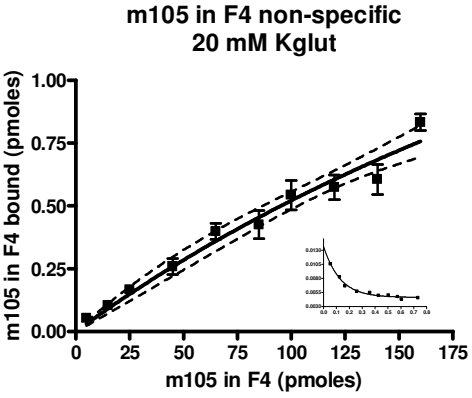

F1 non-specific
20 mM Kglut
F1 bound (pmoles)
F1 (pmoles)
F2 non-specific
20 mM Kglut
F2 bound (pmoles)
F2 (pmoles)
F3 non-specific
20 mM Kglut
F3 bound (pmoles)
F3 (pmoles)
F4 Kd=2.32 uM
20 mM Kglut
F4 bound (pmoles)
F4 (pmoles)
F5 non-specific
20 mM Kglut
F5 bound (pmoles)
F5 (pmoles)
F4a Kd=4.69 uM
20 mM Kglut
F4a bound (pmoles)
F4a (pmoles)
F4b Kd=4.94 uM
20 mM Kglut
F4b bound (pmoles)
F4b (pmoles)
Pr Kd=10.13 uM
20 mM Kglut
Pr bound (pmoles)
Pr (pmoles)

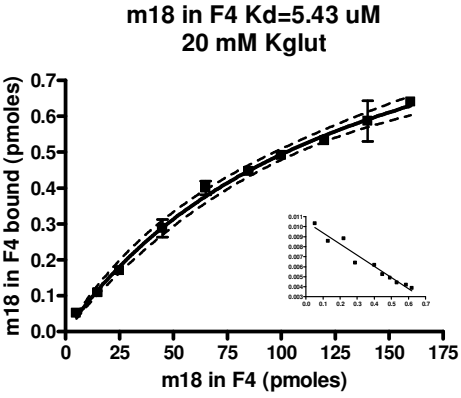

m74 in F4 non-specific
20 mM Kglut
m74 in F4 bound (pmoles)
m74 in F4 (pmoles)

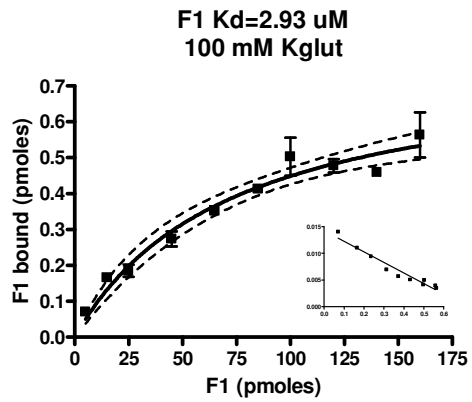

m26 in F4 non-specific
20 mM Kglut
m26 in F4 bound (pmoles)
m26 in F4 (pmoles)

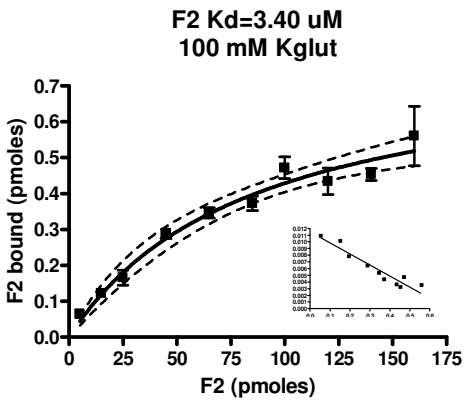

m27 in F4 Kd=6.00 uM
20 mM Kglut
m27 in F4 bound (pmoles)
m27 in F4 (pmoles)

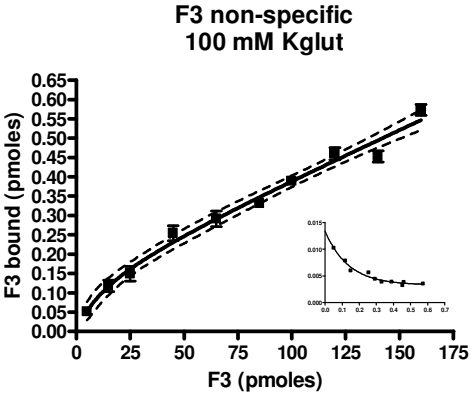

m26/m27 non-specific
20 mM Kglut
m26/m27 in F4 bound (pmoles)
m26/m27 in F4 (pmoles)

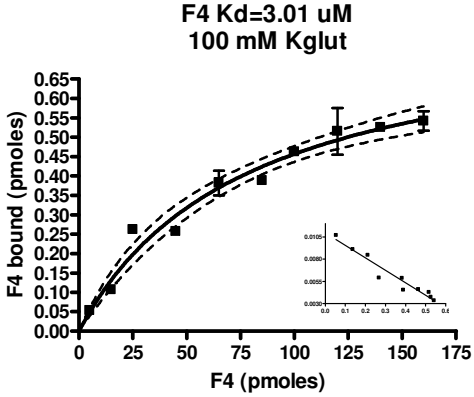

m105 in F4 non-specific
20 mM Kglut
m105 in F4 bound (pmoles)
m105 in F4 (pmoles)

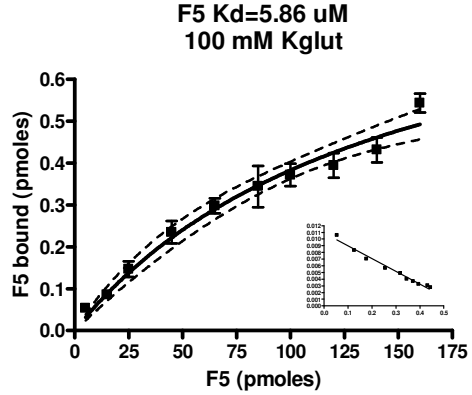

m18 in F4 Kd=5.43 uM
20 mM Kglut
m18 in F4 bound (pmoles)
m18 in F4 (pmoles)

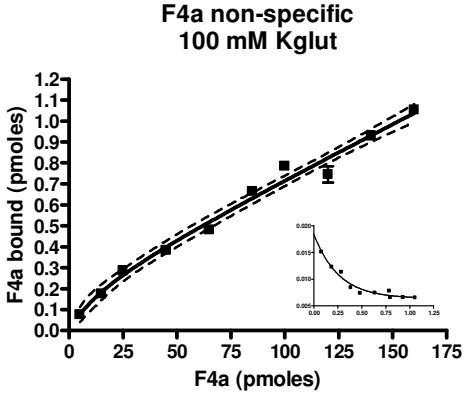

F1 Kd=2.93 uM
100 mM Kglut
F1 bound (pmoles)
F1 (pmoles)

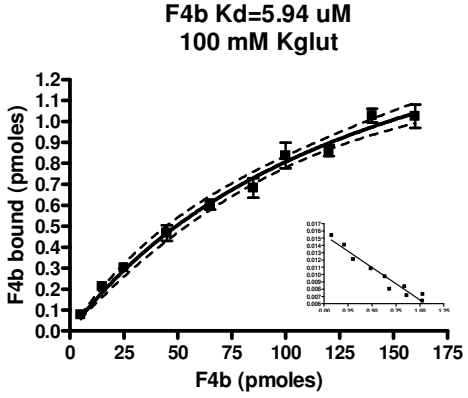

F2 Kd=3.40 uM
100 mM Kglut
F2 bound (pmoles)
F2 (pmoles)

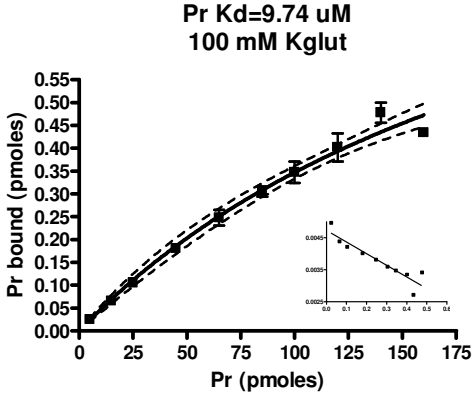

F3 non-specific
100 mM Kglut
F3 bound (pmoles)
F3 (pmoles)

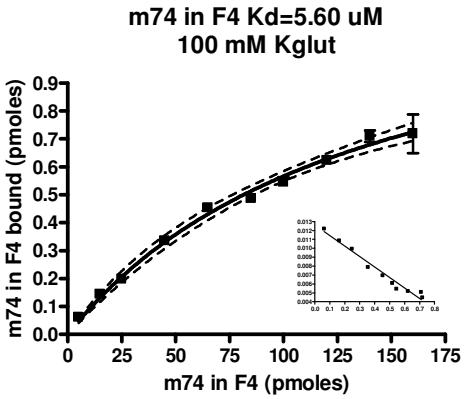

F4 Kd=3.01 uM
100 mM Kglut
F4 bound (pmoles)
F4 (pmoles)

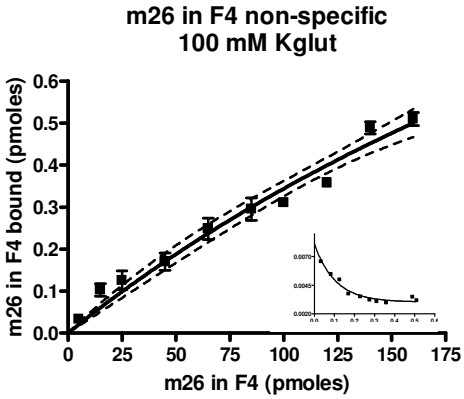

F5 Kd=5.86 uM
100 mM Kglut
F5 bound (pmoles)
F5 (pmoles)

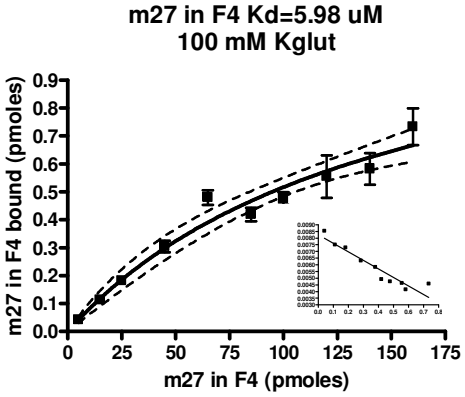

F4a non-specific
100 mM Kglut
F4a bound (pmoles)
F4a (pmoles)

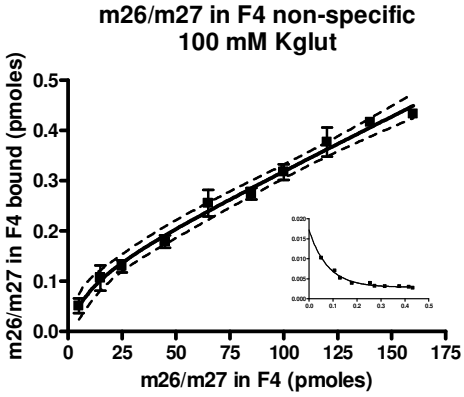

F4b Kd=5.94 uM
100 mM Kglut
F4b bound (pmoles)
F4b (pmoles)

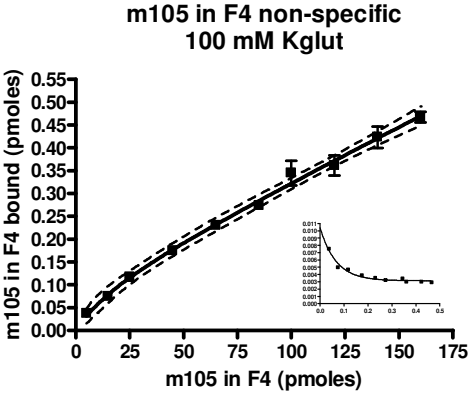

Pr Kd=9.74 uM
100 mM Kglut
Pr bound (pmoles)
Pr (pmoles)

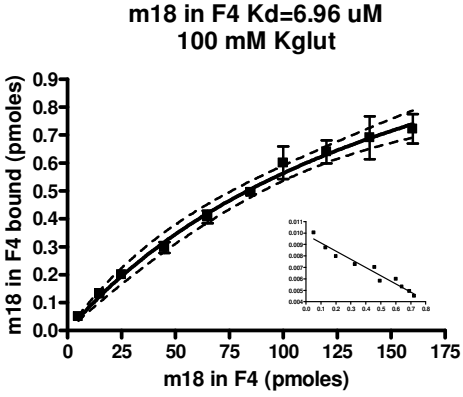

m74 in F4 Kd=5.60 uM
100 mM Kglut
m74 in F4 bound (pmoles)
m74 in F4 (pmoles)

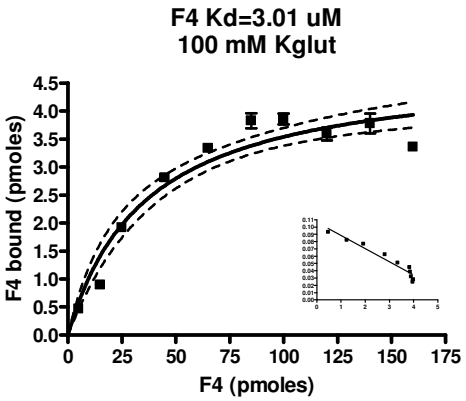

m26 in F4 non-specific
100 mM Kglut
m26 in F4 bound (pmoles)
m26 in F4 (pmoles)

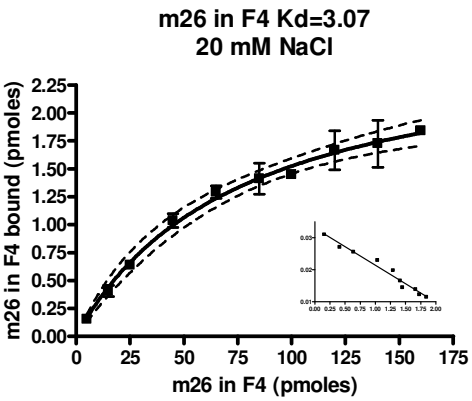

m27 in F4 Kd=5.98 uM
100 mM Kglut
m27 in F4 bound (pmoles)
m27 in F4 (pmoles)

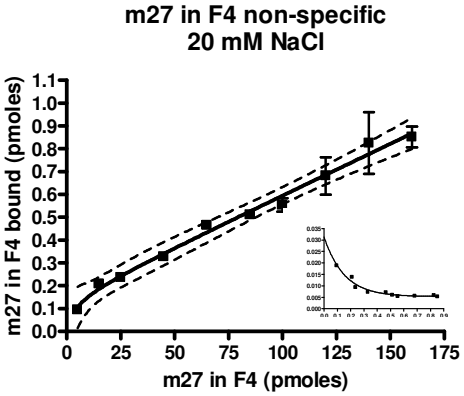

m26/m27 in F4 non-specific
100 mM Kglut
m26/m27 in F4 bound (pmoles)
m26/m27 in F4 (pmoles)

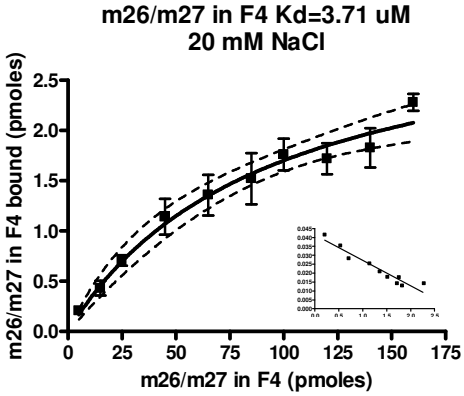

m105 in F4 non-specific
100 mM Kglut
m105 in F4 bound (pmoles)
m105 in F4 (pmoles)


m18 in F4 Kd=6.96 uM
100 mM Kglut
m18 in F4 bound (pmoles)
m18 in F4 (pmoles)


F4 Kd=3.01 uM
100 mM Kglut
F4 bound (pmoles)
F4 (pmoles)


m26 in F4 Kd=3.07
20 mM NaCl
m26 in F4 bound (pmoles)
m26 in F4 (pmoles)


m27 in F4 non-specific
20 mM NaCl
m27 in F4 bound (pmoles)
m27 in F4 (pmoles)


m26/m27 in F4 Kd=3.71 uM
20 mM NaCl
m26/m27 in F4 bound (pmoles)
m26/m27 in F4 (pmoles)

Mutations that result in non-specific binding do not adopt a different equilibrium of structural species, as assayed by native gel electrophoresis

A possible explanation for some mutations resulting in non-specific binding is that they may significantly alter the global folding of the fragments and, therefore, the binding of RdRp. If these changes are sufficiently pronounced, they might be visualized as a change in the mobility of species in a native polyacrylamide gel. I tested wt F4 and two F4 mutants to see if I could visualize any global difference among migrating species. m51 disrupts Ψ_1 with a mutation in the LSL of H5. It increases RdRp binding and results in a massive increase in *in vitro* transcribed products (Yuan et al, 2009). Ym70 is a mutation present in the asymmetric loop of H4 that results in non-specific RdRp binding and undetectable levels of *in vitro* transcription (Yuan et al, 2010). Radiolabeled samples were gel purified through a 5% denaturing polyacrylamide gel to ensure the presence of only one species prior to subjecting the samples to electrophoresis on a nondenaturing polyacrylamide gel. The purified samples were slow cooled to allow them to fold prior to electrophoresis through a 5% polyacrylamide gel prepared without urea or EDTA but with 5 mM $MgCl_2$. The fragments containing mutations did not result in a different pattern than the wt F4 (Fig. 2.8), indicating that any differences in folding due to these mutations are not detectable by this assay. Overexposure (Fig. 2.8, right panel) also did not reveal any differences among these fragments.

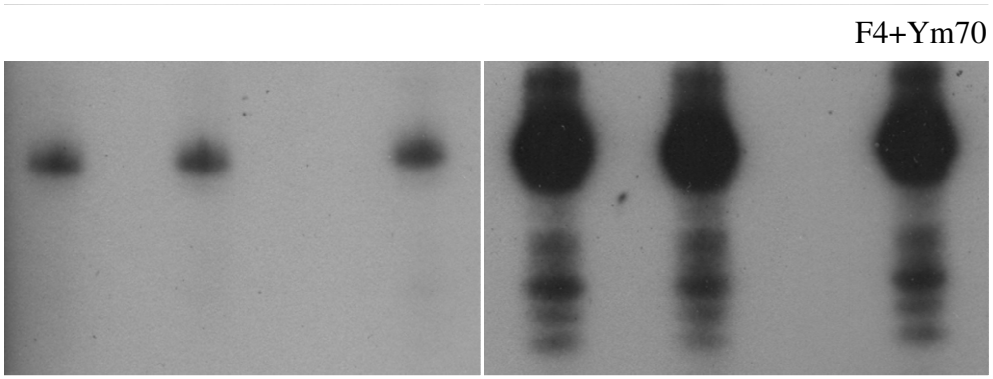


Figure 2.8 Native gel of F4 and F4 containing mutations

Labeled F4 and F4 containing mutations were run on a native 5% polyacrylamide gel in the presence of 5 mM $MgCl_2$. Autoradiography indicated no discernible differences among these TCV fragments, even upon overexposure (right panel).

Discussion

It has been shown that a tRNA-shaped structure within the 3' UTR of TCV (McCormack et al, 2008; Zuo et al, 2010) overlaps with structures required for ribosome binding and translational enhancement (Stupina et al, 2008) and that binding of the RdRp to the 3' end of TCV results in an extensive conformational shift that would likely preclude ribosome binding (Yuan et al, 2009). For this chapter, a filter binding assay was used to determine what sequences are important for RdRp binding within the 3' end region. This involved the use of mutations that disrupted various structures to better understand which structures are important for RdRp binding.

RdRp binds best to F4, which includes the region between H4 and the 3' end. Binding to F3, which does not include the Pr hairpin and flanking sequences, was quite poor. The Pr, which is a strong promoter of transcription in satC but not TCV (Zhang et al, 2006b) and is phylogenetically conserved in all carmoviruses except *Galinsoga mosaic virus* (Chattopadhyay et al, 2011), did not bind RdRp as well as F4. The Pr of

TCV has additional functions including participating in the switch between translation and replication, and thus it is not surprising that binding to the element alone is poor. The Pr has further been shown to be involved in a long-distance RNA:RNA interaction to facilitate ribosomal readthrough of the p28 stop codon in order to generate p88 (the RdRp) (Cimino et al, 2011), indicating its involvement in a multitude of steps in the viral life cycle. Poor binding to F3 is therefore not a consequence of the absence of Pr but likely due to an altered conformation of the remaining fragment (X. Yuan and A. E. Simon, unpublished). Binding to H4 alone was non-specific, however binding to H4 and Pr together, which artificially juxtaposes these two distal hairpins, was comparable to binding to F4. It has previously been shown that a fragment containing H4 and Pr joined by H4 flanking sequence and Pr flanking sequence was an effective competitor for binding of RdRp in a gel shift assay (Sun & Simon, 2006). This may suggest that, when brought into artificial proximity to each other, these hairpins can interact or fold in such a way as to bind RdRp well, whereas, shown here, the individual components bind RdRp poorly or non-specficially. F5 included a (-)-strand hairpin (M3H) upstream of H4. The presence of this additional sequence reduced binding to the RdRp by 1.6 fold.

While fragment F3 ends in the linker region between H5 and Pr, Pr was specifically deleted from F4 to make F4b. F4b, unlike F3, contains the Pr flanking sequences. When tested, deletion of the Pr in F4b resulted in binding about 2-fold worse than binding to wt F4. The H4a/Ψ_3 region has been shown to be important for ribosome binding, translation (Stupina et 2008), accumulation in protoplasts (McCormack et al, 2008), and RdRp binding, and the Pr is known to interact with this region (Yuan et al,

2009). It is therefore possible that deletion of the Pr alters the structure of the fragment and worsens RdRp binding to F4b.

Disruption of Ψ_1

Ψ_1 is phylogenetically conserved in all carmoviruses (Chattopadhyay et al, 2011) and many tombusviruses (Pogany et al, 2003). When Ψ_1 was disrupted in a defective interfering RNA of *Tomato bushy stunt tombusvirus*, *in vitro* (-)-strand synthesis increased, however the interaction proved vital for accumulation in protoplasts (Pogany et al, 2003). In TCV, mutations that disrupt Ψ_1 do not affect ribosome binding to F2 (Stupina et al, 2008) and are known to increase *in vitro* transcription (Yuan et al, 2009; Zhang et al, 2004), though many of these products are aberrantly initiated. Mutations in Ψ_1 behaved as expected in the filter binding assays: improved binding when Ψ_1 is disrupted, with compensatory mutations restoring binding to wt levels. This indicates that the formation of Ψ_1 restricts access of the polymerase to the 3' terminal nucleotides, impedes RdRp binding to the F4 fragment.

Disruption of Ψ_2

Ψ_2 has the potential to form in five carmoviruses, including TCV. Mutations disrupting Ψ_2 resulted in non-specific binding in F4 and F1, suggesting there is sequence or structure specificity required for RdRp binding. When combined, only m46/m47 restored specific binding, possibly indicating that the other set of mutations did not restore the pseudoknot in these conditions. These results reflect similar results seen when these mutations were introduced in the full-length virus, decreased accumulation in protoplasts with the individual mutations and only one set of compensatory mutations (m46/m47) restoring high levels (80%) of accumulation (McCormack et al, 2008).

Disruption of Ψ_3

Within Carmoviruses, Ψ_3 is present only in TCV and CCFV. When a mutation disrupting Ψ_3 was in the loop of H4a, binding was enhanced. However, when the mutation disrupting the pseudoknot was in the 5' flanking region of H4a, binding was non-specific. Compensatory mutations did not restore specific binding. Similar results were seen when the identical mutations were incorporated into F1. This indicates that the presence of Ψ_3 has a negative effect on binding, since binding improves when it is disrupted. However, it also indicates that the 5' flanking region of H4a sequence is important for binding, because specific binding is eliminated when it is mutated, even when it is coupled with likely compensatory mutations in the loop of H4a. Mutations that disrupt Ψ_3 in F1 were detrimental to ribosome binding and compensatory mutations restored ribosome binding and a similar reduction in, and restoration of, translation of a reporter construct was seen with mutations that disrupt and restore, respectively, Ψ_3 (Stupina et 2008). Reduced and recovered accumulation in protoplasts was also seen with mutations that disrupt or restore Ψ_3, respectively (McCormack et al, 2008). Taken together, these data implicate Ψ_3 as a region where the RdRp is specifically interacting, leading to the RdRp-mediated conformational shift.

Disruption of Ψ_4, H4 and flanking adenylates

m21 alters three bases in the 3' side of the asymmetric loop of H4 and reduces accumulation, translation, and ribosome binding (Stupina et al, 2008). The three-base m21 mutation had a minimal effect on RdRp binding, but a one-base change (Ym70) in the same area resulted in non-specific RdRp binding and reduced accumulation of the full-length virus in protoplasts to 8% of wt levels (Yuan et al, 2009). Interestingly,

binding was enhanced when m21 was coupled with m19 in the LSL of H5, both of which had a minimal effect on RdRp biding alone. When the bulged C in the asymmetric loop of H4 was mutated to an A in F4, binding was similar to wt F4. In addition, when this C to A change was tested in H4 alone, binding was similar to wt H4, however this same mutation reduced accumulation in protoplasts to 9% of wt levels (Yuan et al, 2009). Given that this mutation has a minimal effect on RdRp binding but is detrimental to viral accumulation, it is likely this mutation causes a defect in translation, which affects the overall accumulation levels. m39, in the terminal loop of H4 and involved in Ψ_4, also had a minimal effect on binding, but when it was coupled with its partner mutation, m38, in the LSL of H5, binding was non-specific. m38 alone resulted in poor binding of the RdRp to F4 and might explain why the two mutations together did not compensate.

Of the mutations that disrupt Ψ_4, m19 is the only one that was not detrimental, but it also was not beneficial; RdRp bound to it comparably as to wt F4. m23, m36, and m38 in F4 all were detrimental for RdRp binding, with m38 being the only one with specific binding, albeit at 2.7-fold worse than wt F4. These mutations are of interest because they are the opposite of ribosome binding: ribosome binding was enhanced when mutations were introduced in this region. Interestingly, binding of RdRp to m36 in F1 was about twice as good as to wt F1 and about comparable to binding to wt F4. This is one of the few instances where the data for F1 and F4 differ.

When the lower stem and flanking regions of H4 were mutated, it was found that mutating the flanking regions had deleterious effects on the virus' ability to accumulate in protoplasts (Yuan et al, 2009). Coupling this with the filter binding results for mutations in H4 (no severe negative effect) led to the idea that the flanking adenylates on

either side of H4 may be sequence important for RdRp binding. It was hypothesized that when these adenylates were mutated, it might have made the RdRp less efficient at binding, resulting in lowered accumulation. Furthermore, when SELEX of the flanking sequences of M1H of satC (placed similarly to H4 of TCV) was performed, a run of adenylates on the 5' side and some adenylates on the 3' side were recovered, again suggesting their importance (Sun & Simon, 2003). To test this, a fragment (F4a) in which H4 is deleted was made. Single A to U mutations in the flanking regions of H4 were made in this new fragment. Only one (m53) of the four mutations in F4a in the 3' flanking adenylates had specific binding and it was worse than binding to F4 and F4a. This supports the idea that the adenylates on either side of H4 may be important for RdRp binding and that H4 itself (which, when mutated or deleted had no severe negative effects on RdRp binding) serves merely to bridge the two flanking sequences and bring them into appropriate proximity to each other for binding. Mutation of the adenylates on the 3' side of H4 in F1 resulted in non-specific binding, as in F4a. Truncation of F4 after H4a made binding of RdRp about 2.5 fold worse than to F4. Truncating F4 after H4b was also worse for RdRp binding, but not as bad as truncation after H4a, suggesting that H4a and H4b act as a unit, as has been previously indicated (Guo et al, 2009; McCormack et al, 2008).

Global folding of fragments

I hypothesized that the presence of the monovalent cations K^+ or Na^+ may help the fragments fold properly and may reduce how many mutations resulted in non-specific binding. Addition of monovalent cations to the binding buffer had varied results. Several fragments and mutations that demonstrated specific binding when potassium was not

present resulted in non-specific binding in the presence of 20 mM potassium glutamate. Similar results were seen in the presence of 100 mM potassium glutamate, where more fragments and mutations were non-specific than without the potassium present. Fewer fragments and mutations were tested in the presence of 20 mM NaCl, but of those that were, one was still non-specific. Interestingly, the presence of Na^+ resulted in specific binding to m26, non-specific binding to m27, and specific binding to m26/m27, exactly the opposite of results obtained when no monovalent cation was present.

It is possible that various mutations could result in a change in the global folding of the fragments (Liu et al, 2012) that might affect binding. I looked at F4 containing either a mutation beneficial or detrimental for RdRp binding using a non-denaturing polyacrylamide gel. If these mutations alter the global folding of the fragment, it was not detectable by this assay. While this does not rule out altered structure as part of the reason we obtain non-specific binding, it suggests that any structural alterations that may be occurring are subtle.

Overall, both Ψ_1 and Ψ_3 proved detrimental to RdRp binding, as binding was enhanced when they were disrupted. Ψ_3 was also important for ribosome binding (Stupina et 2008) and may be one of the key areas that play a role in both transcription and translation. The interactivity of the 3' end of TCV (Yuan et al, 2009; Yuan et al, 2010; Yuan et al, 2012) makes it possible that the RdRp first binds a region and then is transferred to the 3' terminus for transcription.

Table 2.2 RdRp binding to 3' fragments of TCV and mutations in 3' fragments of TCV, complete results

RNA	Kd, μM	Purpose of mutation
* F1 (TSS)	4.22 ± 0.38	3901- 4017
* F2	3.67 ± 0.29	3901 – 4054
* F3	11.3 ± 2.47	3859 – 4017
* F4	2.56 ± 0.25	3859 - 4054
* F4 end H4a	6.24 ± 0.93	3859 – 3937
* F4 end H4b	5.21 ± 1.15	3859 - 3966
* F4a	2.78 ± 0.52	Δ3870-3903
* F4b	5.47 ± 0.70	Δ4022-4048
* F5	4.19 ± 0.83	3812 - 4054
* H4 and Pr	2.87 ± 0.22	3870-3903 + 4022-4048
* H4	non-specific	3869 - 3903
* Pr	6.06 ± 1.69	4022 - 4048
* MDV	non-specific	negative control
CCFV 4a 4b in F4	non-specific	CCFV H4a/H4b
CCFV 4a 4b up in F4	1.78 ± 0.30	CCFV Ψ3/H4a/H4b
m50 in H4	7.28 ± 0.58	5' side H4AL
* m50 in F4	3.82 ± 1.24	5' side H4AL
* m18 in F4	0.86 ± 0.24	disrupts Ψ1
* m51 in F4	1.36 ± 0.41	disrupts Ψ1
* m52 in F4	1.76 ± 0.29	disrupts Ψ1
* m51/m52 in F4	3.54 ± 0.69	restores Ψ1
m44 in F1	non-specific	disrupts Ψ2
* m44 in F4	non-specific	disrupts Ψ2
* m45 in F4	non-specific	disrupts Ψ2
* m44/m45 in F4	non-specific	restores Ψ2
m46 in F1	non-specific	disrupts Ψ2
* m46 in F4	non-specific	disrupts Ψ2
m47 in F1	non-specific	disrupts Ψ2
* m47 in F4	non-specific	disrupts Ψ2
m46/47 in F1	non-specific	restores Ψ2
* m46/47 in F4	4.88 ± 0.33	restores Ψ2
* m27 in F4	0.94 ± 0.17	disrupts Ψ3
m27 in F1	1.96 ± 0.34	disrupts Ψ3
m27 in F3	non-specific	disrupts Ψ3
* m26 in F4	non-specific	disrupts Ψ3
m26 in F1	non-specific	disrupts Ψ3
m26 in F3	3.73 ± 0.45	disrupts Ψ3
m26/m27 in F1	non-specific	restores Ψ3
* m26/m27 in F4	non-specific	restores Ψ3
* m40 in F4	non-specific	disrupts Ψ3
m41in F4	0.86 ± 0.13	disrupts Ψ3
* m40/m41 in F4	non-specific	restores Ψ3

* m19 in F4	2.80 ± 0.30	disrupts Ψ4
m19 in F3	1.96 ± 0.38	disrupts Ψ4
* m21 in F4	2.89 ± 0.32	forces Ψ4
m21 in F3	2.62 ± 0.34	forces Ψ4
m19/m21 in F4	1.03 ± 0.25	Ψ4 and H4AL
m19/m21 in F3	2.95 ± 0.61	Ψ4 and H4AL
* m38 in F4	7.00 ± 1.06	disrupts Ψ4
* m39 in F4	2.95 ± 0.37	disrupts Ψ4
* m38/m39 in F4	non-specific	restores Ψ4
* m23 in F4	6.46 ± 0.44	disrupts Ψ4
* m36 in F4	3.06 ± 0.78	disrupts Ψ4
m36 and m18 in F4	non-specific	H5 LSL and Ψ1
m36 in F1	2.79 ± 0.41	disrupts Ψ4
* m105 in F4	non-specific	H5 LSL
m53 in F4a	3.97 ± 0.80	As 5' of H4
m54 in F4a	non-specific	As 5' of H4
m54 in F4a	non-specific	As 5' of H4
* m74 in F4	non-specific	As 3' of H4
m55 in F4a	non-specific	As 3' of H4
* m55 in F4	non-specific	As 3' of H4
m56 in F4a	non-specific	As 3' of H4
m36 in F4a	non-specific	disrupts Ψ4
m36 in F4b	non-specific	disrupts Ψ4
m55 in F1	non-specific	As 3' of H4
m56 in F1	non-specific	As 3' of H4
* Ym70 in F4	non-specific	H4AL
Ym70/A3975C in F4	non-specific	H4AL and Ψ4
Ym70/A3976C in F4	non-specific	H4AL and Ψ4
Ym70/A3977C in F4	non-specific	H4AL and Ψ4
Ym70/A3978C in F4	non-specific	H4AL and Ψ4

*=presented in this chapter

Chapter III

Interaction Between the 3' end of TCV and the CP ORF Reveals a Role for H3 and its Upstream Sequences in Translation

Positive-strand RNA viruses must regulate the competing processes of translation and transcription (Barton et al, 1999; Gamarnik & Andino, 1998) and do so, in part, through long-distance interactions and conformational switches (Edgil & Harris, 2006; Miller & White, 2006; Simon & Gehrke, 2009). Recent studies have demonstrated overlapping RNA structures for translation and replication at the ends of viral genomes that bind translation or replication factors and presumably prevent the opposing process (Stupina et al, 2008; Villordo & Gamarnik, 2009; Wu & White, 1999). Turnip crinkle virus (TCV) is a small, positive-strand virus that only encodes 5 proteins. Its small size and limited coding capacity make it ideal for studying RNA structure/function relationships as they pertain to translation and replication of the virus.

It has become increasingly clear that elements in the 3' UTR of TCV have many tertiary interactions within this region (Yuan et al, 2009). The asymmetric loop of H4 (H4AL) is known to be important for both translation and replication (Stupina et al, 2008, Yuan et al, 2009) and may function through short- or long-distance interactions. To determine specific interactions in which H4AL might engage, a mutation (m21) in the asymmetric loop was used to generate second-site mutations (Yuan et al, 2010). m21 reduced accumulation in protoplasts to 34% of wt, reduced translation of a reporter construct in protoplasts to 23% (Stupina et al, 2008), and eliminated *in vitro* transcription of a 3' end fragment of TCV (Yuan et al, 2009). The primary mutation site had partially reverted in all the clones recovered and second-site mutations were present in the 3' UTR

as well as upstream in the coat protein (CP) ORF. While mutations in the 3' UTR were not unexpected, given the many interactions therein, it was surprising to find mutations in the CP ORF, which seems to indicate that the 3' UTR may be interacting upstream.

In this chapter, I explore whether the second-site mutations have an effect on viral accumulation when CP effects are eliminated by using a construct that does not make CP. I also use a luciferase reporter construct to determine whether the second-site mutations have an effect on translation. Expanding the viral contribution at the 3' end of the luciferase reporter construct resulted in increased translation and I performed a deletion series to map the additional translation. Ultimately, these studies indicated that an upstream hairpin, H3, is playing a role in translation and its structure and stability are important for the increased translation seen.

Materials and Methods

In Vitro Transcription of RNA Using T7 RNA Polymerase

As described in Chapter II.

Culturing of Arabidopsis Callus

Approximately 100 *Arabidopsis thaliana*, ecotype Col-0, seeds were aliquotted into a 1.5 ml microcentrifuge tubes and vapor sterilized for 3-4 hours by subjecting them to the gasses produced by adding 3 ml HCl to 100 ml bleach. To generate new callus, surface-sterilized seeds were placed onto MS plates and each plate individually sealed with parafilm. MS plates are made by combining 30 g sucrose, 4.4 g of MS Salts (Sigma-Aldrich or Phytotechnology Laboratories), 10 ml of 100x vitamins/glycine stock, and water to a total volume of 1 liter. The pH was adjusted to 5.8 using sodium hydroxide.

Bacteriological agar was added to a concentration of 1% and the mixture autoclaved. When the temperature of the mixture was less than 60°C, 250 µl of 2,4-D (2 mg/ml) and 250 µl of kinetin (2 mg/ml) were added to 400 ml of medium and poured into sterile petri dishes. Every 3 weeks, callus was passaged in a sterile environment using sterilized forceps to mechanically break the callus and transferring it to new MS plates. Plates were incubated in a Percival Scientific I-36LL incubator at 20°C using a photoperiod of 16 hr light and 8 hr dark, and an illuminance of 35 µmol/m^2S.

Preparation and Inoculation of Callus Culture Protoplasts with Infectious Viral RNA or Luciferase Constructs Using Polyethylene Glycol

In a sterile environment, 0.6 M mannitol was added to plates of callus and the callus was broken into smaller pieces using a 14.6 cm Pasteur pipet melted into an L-shape. Callus and mannitol were then poured into a 50 ml centrifuge tube and agitated at room temperature using a rotating shaker for 20 min at 100 rpm. For every three plates of callus being processed, 50 ml of PIM was dispensed into a sterile glass bottle of appropriate size along with 0.5 g cellulase (10 KU/g dry weight; Calbiochem) and 0.1 g pectinase (3 KU/g dry weight; Calbiochem). PIM was made by combining 1 ml of 1000x vitamin stock (0.02 g thiamine HCl, 0.01 g pyridoxine HCl, 0.01 g nicotinic acid, and 2 g myo-inositol added to water, volume adjusted to 20 ml), 0.5 ml of 2000x hormone stock (0.004 g 2,4-D, 0.004 g kinetin, and 0.5 ml 1 N potassium hydroxide added to water, volume adjusted to 10 ml), 4.4 g of MS salts, 34.2 g of sucrose, 0.585 g of MES, 91 g of mannitol, and 0.555 g calcium chloride to 1 liter of water. The pH was adjusted to 5.8 with potassium hydroxide. The callus/mannitol was centrifuged at 3,000 rpm for 5 min at room temperature and the supernatant poured off. All centrifugations in this procedure

were performed in a Beckman GPR centrifuge. The callus was transferred to the bottle containing the PIM/enzymes mixture. The bottle was then wrapped completely in aluminum foil and incubated at room temperature for 4 hr on a rotating shaker at 100 rpm. The solution containing the protoplasts was filtered through a sterilized 53 μm nylon mesh (Small Parts) into a 50 ml centrifuge tube using a sterile funnel.

Protoplasts were centrifuged at 1,000 rpm for 5 min at 4°C. The supernatant was decanted and the pellet resuspended by gentle shaking of the tube. Twenty milliliters of cold (4°C) 0.6 M mannitol was added and the tube inverted several times to wash the protoplasts. The cells were centrifuged at 1,000 rpm for 5 min at 4°C. The washes were repeated for a total of three times. After the final wash and centrifugation, protoplasts were resuspended in cold 0.6 M mannitol and kept on ice. Protoplasts were then quanitifed using a hemocytometer and 5 x 10^6 cells (for full-length virus) or 7 x 10^6 cells (for luciferase constructs) were aliquotted into 50 ml centrifuge tubes, one for each inoculation. Each tube was centrifuged at 1,000 rpm for 5 min at 4°C and the supernatant poured off leaving about 100 μl of solution. Protoplasts were kept on ice until inoculation. For each inoculation, 20 μg of *in vitro* transcribed TCV RNA or 30 μg luciferase construct transcript plus 10 μg renilla luciferase transcript, was combined with 8 μl of 1 M calcium chloride and the volume brought up to 430 μl with water. Each inoculation mix was kept on ice until needed. The inoculation mix was added to the protoplasts and the mixture swirled to combine. Next, 2.2 ml of 50% (w/v) PEG (25 g of PEG 1540 in 50 mM Tris-HCl, pH 7.5; final volume 50 ml) was added, mixed well by swirling for 20 sec, and incubated at room temperature for 20 sec. Cold 0.6 M mannitol containing 1 mM calcium chloride was added to a volume of 40 ml and the tube

incubated on ice for 15 min. Protoplasts were then centrifuged at 1,000 rpm for 5 min at 4°C and the supernatant decanted. Twenty milliliters of cold 0.6 M mannitol containing 1 mM calcium chloride was added, centrifuged at 1,000 rpm for 5 min at 4°C, and the supernatant decanted. The wash and centrifugation were repeated two more times. After the final centrifugation, the supernatant was completely decanted and each tube of protoplasts was resuspended in protoplast culture medium (PCM: 1 ml 1000x vitamin stock, 0.5 ml 2000x hormone stock, 4.4 g MS salts, 34.2 g sucrose, 0.585 g MES, 72.8 g mannitol, and water to 1 liter; pH adjusted to 5.8 using 1N potassium hydroxide). The protoplasts were then poured into a 60 x 15 mm petri dish. For full-length virus, the protoplasts were kept in the dark and incubated at room temperature for 40 hr. For luciferase constructs, protoplasts were kept in the light and incubated at room temperature for 18 hr.

Extraction of Total RNA from Arabidopsis Protoplasts

The condition of the protoplasts at 40 hpi was checked by examining the plates under a light microscope. Plates were tilted to evenly distribute the cells and collected by transferring to two 1.5 ml microcentrifuge tubes. Protoplasts were then subjected to centrifugation at 13,000 rpm for 30 sec at room temperature. All centrifugations were performed for this procedure using a microcentrifuge. The supernatant was removed by pipet and one of the tubes was stored as a backup sample by freezing at -80°C. Two hundred µl of 1:1 phenol/chloroform containing 0.1% (w/v) 8-hydroxyquinoline and 200 µl of RNA special extraction buffer (50 mM Tris-HCl, pH 7.5; 5 mM EDTA, pH 8.0; 100 mM sodium chloride; 1% (w/v) SDS) were added to each tube and the mixture vortexed. The samples were centrifuged at 13,000 rpm for 2 min at 4°C and the aqueous layer

transferred to a new tube. Twenty five µl of 3 M sodium acetate, pH 5.2, and 575 µl of 100% ethanol were added and the mixture, mixed, and incubated at -80°C for 5 min. The tubes were then centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant was discarded and 500 µl of 70% ethanol was added, mixed, and the sample centrifuged at 13,000 rpm for 5 min at 4°C. The supernatant was removed and the pellet dried and resuspended in 20 µl of water. RNA concentration was determined using a spectrophotometer at an optical density (OD) of 260.

Northern Blotting

One and a half micrograms of total RNA and water up to 5 µl were combined with 5 µl of 2x formamide loading buffer [800 µl of formamide added to 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 incubated at 65°C for 5 min, quenched on ice, and electrophoresed through a 1% agarose gel. The gel was rinsed briefly with water and soaked in a 6% formaldehyde solution for 1 hr with shaking. The formaldehyde was decanted and the gel soaked in 10x SSC for 10 min followed by another 15 min in 10x SSC with a 45 µm pure nitrocellulose membrane, with shaking. The 10x SSC solution contains 350.5 g of sodium chloride and 176.4 g of sodium citrate, trisodium salt, dehydrate, and water to a final volume of 2 liters. The pH was adjusted to 7.0 with HCl. RNA was transferred to the membrane using the capillary transfer method and the membrane rinsed with 10x SSC. The membrane was then placed, face-down, on an ultraviolet light box for 2 min and dried at 80°C for 5 min. Both the membrane and gel were analyzed under ultraviolet light to verify the transfer. To probe for positive-sense TCV, the membrane was prehybridized for at least 8 hrs at 42°C using a 30% (v/v)

formamide prehybrization buffer. Prehybridization buffer was prepared by combining 3 ml of formamide, 2 ml of 50x Denhardt's reagent (contains 5g of Ficoll, Type 400; 5 g of polyvinylpyrrollidone; 5 g of bovine serum albumin, Fraction V; and water to 500 ml), 2.5 ml of 20x SSPE (combined 175.3 g sodium chloride; 27.6 g of sodium phosphate, anhydrate, monobasic; 40 ml of 0.5 M EDTA, pH 8; and adjusted to pH 7.4 with sodium hydroxide and to a volume of 1 liter with water), 2.5 ml of water, 200 µl of 10 mg/ml denatured single-stranded DNA, and 200 µl of 10% (w/v) SDS. During the prehybridization incubation, oligonucleotides were radiolabeled by combining 9.5 µl of water, 6 µl of 10 pmol/µl oligonucleotide, 2 µl of 10x T4 polynucleotide kinase buffer (700 mM Tris-HCl, pH 7.6; 100 mM magnesium chloride; 50 mM DTT), 1.5 µl of γ-^{32}P ATP (10 mCi/ml), and 1 µl of T4 polynucleotide kinase (10,000 U/ml; New England Biolabs), and the mixture incubated for 1 hour at 37°C. The radiolabeled oligonucleotide was added to the prehybridization buffer and incubated for at least 8 hrs at 42°C. The prehybridization buffer was decanted and the membrane washed for 10 min with Northern wash I (made by combining 690 ml of water, 300 ml of 20x SSPE, and 10 ml of 10% (w/v) SDS) at 42°C. Next, the membrane was washed for 10 min with Northern wash II (made by combining 985 ml of water, 5 ml of 20x SSPE, and 10 ml of 10% (w/v) SDS) at 42°C. The membrane was allowed to dry at room temperature, covered in plastic wrap, and audioradiography performed at -80°C. QuantityOne was used to determine relative levels.

In vivo Translation Assay

A single-luciferase reporter construct, T7-63-Fluc-3661, was used to assay for translation *in vivo*. T7-63-Fluc-3661 contains the exact 63 base 5' UTR of TCV. The 3' end of the

construct contains the 3' region of TCV (position 3661 to 4054). Modifications were made to this construct to introduce mutations into either or both of the 5' and 3' ends or to lengthen the viral sequence present at the 3' end. These constructs were linearized with SspI and used as templates for T7 polymerase-driven RNA synthesis *in vitro*. Thirty μg of uncapped *in vitro*-transcribed single-luciferase constructs was inoculated into protoplasts along with 10 μg of uncapped transcripts containing internal control Rluc. Protoplasts were harvested by centrifugation for 30 s at 13,000 rpm at 18 hr postinfection (hpi). Cells were lysed in 1X Passive lysis buffer (Promega) and luciferase activity assayed using a TD 20/20 luminometer (Turner Designs) and a dual reporter assay system (Promega).

In-Line Probing of RNA

RNA transcripts were purified from agarose gels, dephosphorylated with Antarctic phosphatase (NEB), 5' end labeled using T4 polynucleotide kinase (NEB) and [γ–^{32}P]ATP, and then purified by polyacrylamide gel electrophoresis. 5' end labeled fragments were denatured at 75°C and slowly cooled to 25°C. Approximately five picomoles of end-labeled RNA was incubated at 25°C in 50 mM Tris-HCl (pH 8.5) and 20 mM $MgCl_2$ for 14 hr. RNA cleavage ladders were made by incubating 5 pmol of end-labeled RNA in 1 μg yeast tRNA, 50 mM $NaHCO_3$-Na_2CO_3 (pH 9.2), and 1 mM EDTA for 5 min at 95°C. RNase T_1 digests were produced by incubating 10 pmol of denatured end-labeled RNA in 1 μg yeast tRNA, 20 mM sodium citrate (pH 5.0), 1 mM EDTA, 7 M urea, and 1 U RNase T_1 (Ambion) for 3 min at room temperature. All reaction mixtures were ethanol precipitated, resuspended with gel loading buffer II (Ambion), heated at 95°C for 2 min, and subjected to electrophoresis through an 8% denaturing

polyacrylamide gel followed by drying and autoradiography. At least three independent in-line probing assays were produced for each fragment. Profile differences were noted only if found in all replicate gels.

Results

Second-site mutations in the CP ORF can enhance viral accumulation

Our previous studies demonstrated that the 3' UTR of TCV is involved in a complex network of interactions (Yuan et al, 2009; Yuan et al, 2010; Yuan et al, 2012). Hairpin H4 was shown to be important for transcription and translation, with mutations in H4 disrupting intra-hairpin interactions and those between H4 and other proximal structures in the 3' UTR (Yuan et al, 2009; Yuan et al, 2010). As described in Chapter II, RdRp binds to a 3' end fragment of TCV (F4) containing H4 with a K_d of 2.56 μM. Binding of RdRp is minimally affected by the deletion of H4 from this fragment (F4a, 2.78 μM) and H4 alone binds RdRp non-specifically. To better understand how H4 functions in transcription and translation a 3-base mutation (m21) was made in the H4 asymmetric loop (H4AL, Fig. 3.1). m21 reduced translation to 23% and reduced accumulation to 34% (Stupina et al, 2008). Second-site mutations were generated through serial passaging of the m21-containing virus in plants (Yuan et al, 2010). Several second-site mutations were located in the 3' UTR and were found to reduce RdRp transcription, enhance translation and RdRp binding, and alter the structures of both the tRNA shaped structure (TSS) and H4 (Yuan et al, 2010).

Second-site mutations arising from the primary mutation in H4AL were also found upstream in the coat protein (CP) open reading frame (ORF) (Fig. 3.1). These upstream second-site mutations were associated with a partial reversion of the primary

mutation site [UUA (wt) → ACU (m21) → UUU (rev1), Fig 3.1B]. The purpose of generating second-site mutations is to reveal possible sequences or regions that interact

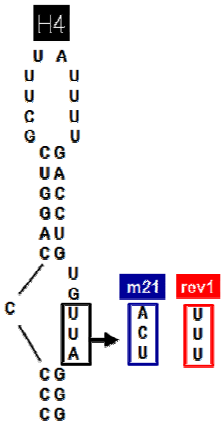
Figure 3.1 Location of primary and second site alterations recovered in the CP ORF of TCV accumulating after three passages through host plants.

A. Primary site mutation and partial reversion. The rev1 sequence was found in all recovered clones. Second-site mutations generated and recovered by X. Yuan and K. Shi, unpublished.

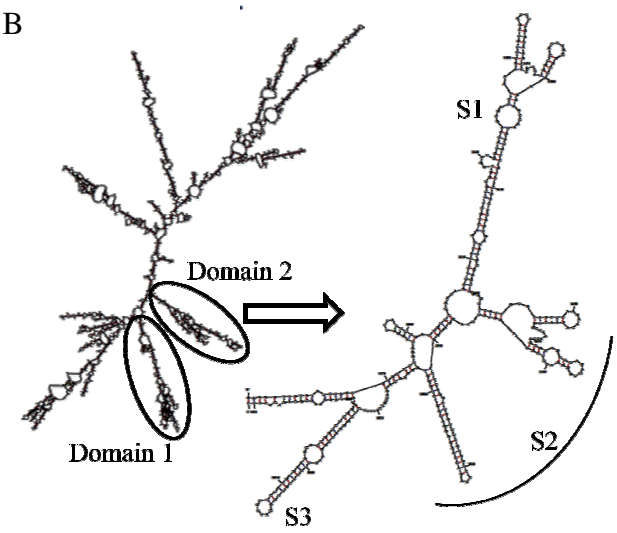
B. mFold prediction of full-length virus with constraints from SHAPE of bases 800-900. On the right, Domain 2, which contains most of the second-site mutations, is shown larger.

C. Sequence and location of specific second-site mutations in S1 and S2 in the CP ORF. Mutations are color-coded where red and green denote an increase or decrease, respectively, in accumulation in protoplasts. Black denotes a mutation that was not tested. Inverted sequence repeat shown in blue. Repeated "UGUUA" of primary mutation site is boxed. S1 and S2 are labeled based on an mFold structure predicted prior to upstream SHAPE work.

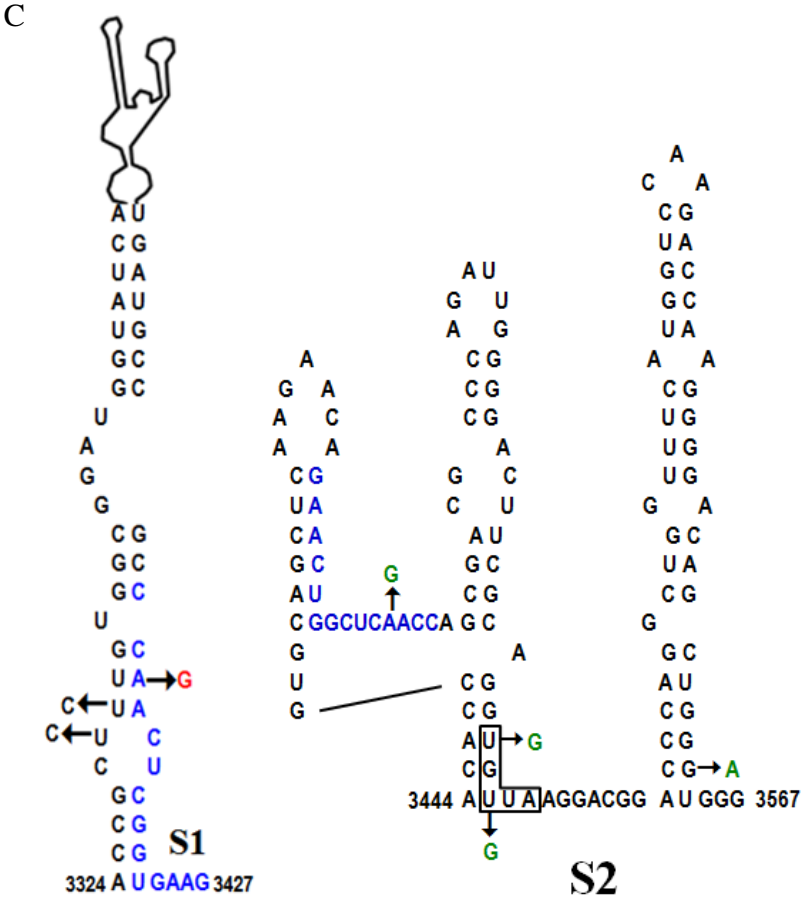
A


H4
m21
rev1

B


S1
Domain 2
Domain 1
S2
S3

C


S1
3324 A U GAAG 3427
C GGCUCAACCA G C
3444 A U U A AGGACGGG A U GGG 3567
S2

with the primary mutation site that mutate to compensate for the deleterious primary mutation. Since the second-site mutations were found with partial reversion of the primary mutation site, it was unclear whether the second-site mutations arose before or after the partial reversion. To help determine this, the second-site mutations were introduced into three different viral backbones and assayed for accumulation in protoplasts (Fig. 3.2A). The backbones were: wt, wild-type TCV; m21, the original primary mutation in H4AL; and rev1, the partial reversion in H4AL. In this way, each mutation or set of mutations could be evaluated in the context of the primary mutation (m21) or the partial reversion (rev1) to determine the sequence of events and discover in which context(s), if any, the second-site mutations are compensatory for virus accumulation.

One of the second-site mutations, G3561A, is a transition that also changes a glycine residue in the CP to glutamic acid (Fig. 3.2A). When introduced into wt, this mutation was surprisingly beneficial, accumulating in *Arabidopsis* Col-0 protoplasts to 133% of wt (Fig. 3.2A). The primary mutation, m21, accumulated in protoplasts to only 23% of wt and this mutation in conjunction with m21 (m21-G3561A) did not appreciably improve accumulation (27%). As found previously, the partial revertant (rev1) accumulated to higher levels than m21, 62% of wt. Unexpectedly, in rev1, G3561A accumulated to levels nearly five times greater than rev1 alone (rev1-G3561A compared to rev1) and three times greater than wt (304%). Thus, while rev1 alone is modestly detrimental and G3561A alone is one-third better than wt, together they vastly enhance viral accumulation in protoplasts, despite G3561A causing a mutation in the CP, the silencing suppressor for TCV (Qu et al, 2003).

Figure 3. 2 Second-site mutations in the UGA background result in increased viral accumulation in protoplasts

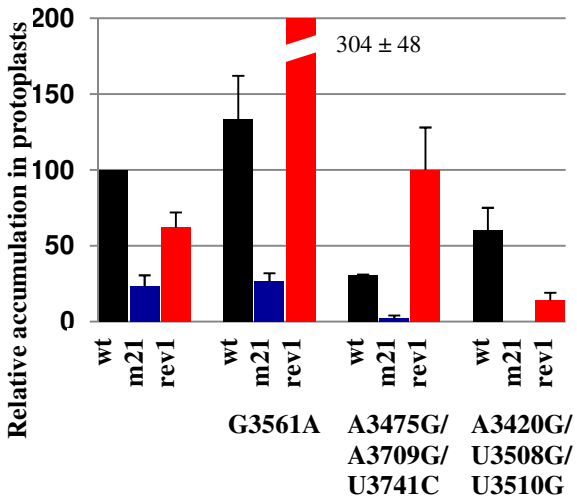
A. Second-site mutations tested in wt, primary mutation (m21), or partial reversion (rev1) background. Right panel shows amino acid changes in the CP as a result of these second-site mutations. Left panel is accumulation in protoplasts as determined by northern blot. Data in A is from X. Yuan, K. Shi, and A. E. Simon, unpublished.

B. Second-site mutations tested in the CP null (UGA) background.

C. Accumulation in protoplasts, as determined by northern blot, of other alterations at or around G3561 performed by X. Yuan, unpublished. For protoplasts, relative values are from at least three independent experiments and standard deviation is shown.

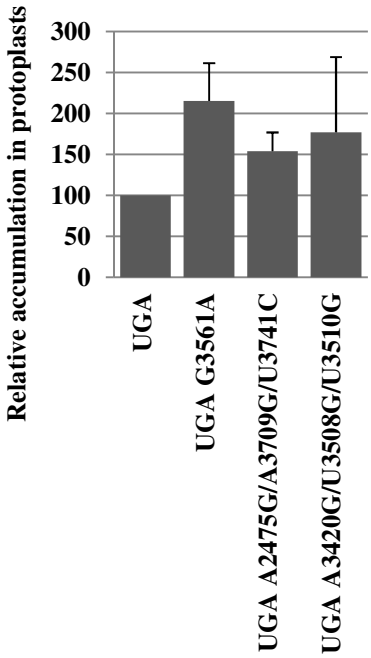
A

Second-site mutation	Amino acid change in CP
G3561A	Gly→Glu
A3475G	No change
A3709G	No change
U3741C	Val→Ala
A3420G	Gln→Arg
U3508G	No change
U3510G	Val→Gly

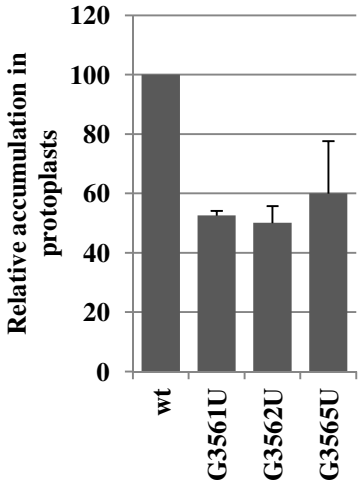


Above work performed by X. Yuan and K. Shi, unpublished

B



C



Perfomed by X. Yuan, unpublished

Because the G3561A mutation had such a surprisingly enhancing effect on viral accumulation in protoplasts, other mutations at or near G3561 designed to be silent mutations or to modify the same residue as G3561A were assayed for their effect on accumulation (Fig. 3.2C). G3561C, which alters a glycine in the CP to a valine, accumulated to only 52% of wt. G3562U, a silent mutation, accumulated to only 50% of wt. G3565U, also a silent mutation, accumulated to 60% of wt. These results indicate that there is something specific to the guanine to adenine transition at this location that results in increased viral accumulation in protoplasts. Due to these unanticipated findings, G3561A is undergoing further study.

Mutations A3475G, A3709G, and U3741C were found in the same clone and were also associated with the rev1 partial reversion mentioned above. Of these three mutations, only the U3741C transition altered the CP, changing a valine residue to an alanine. These three mutations together in the wt background accumulated in protoplasts to only 30% of wt. They were even more detrimental in conjunction with the primary m21 mutation, accumulating to only 2% of wt and 10% of m21 alone. When added to rev1, these mutations compensated for the deleterious partial reversion, accumulating to 100% of wt and 161% of rev1 alone.

Mutations A3420G, U3508G, and U3510G were also found together in the same clone along with rev1. Of these three mutations, the A3420G transition and the U3510G transversion altered the CP, changing a glutamine to arginine and a valine to glycine, respectively. Interestingly, U3508G and U3510G are located in an upstream copy of the sequence that was altered by the primary mutation (UGUUA, see Fig. 3.1A and C). These three mutations together in the wt background accumulated in protoplasts to 60% of wt.

Combining these mutations with the original m21 mutation resulted in no detectable accumulation. Virus containing rev1 and these three mutations accumulated in protoplasts to only 14% of wt and 23% of rev1 alone. Given the lack of accumulation when present with the primary m21 mutation, it is likely that for this set of second-site mutations the partial reversion occurred prior to, or concurrent with, the generation of the second-site mutations, though even that combination failed to regenerate a robust virus.

These results demonstrate that some second-site mutations complement rev1 and one enhanced accumulation to better than wt (G3561A). To determine if the alteration of the CP or the RNA sequence was important for the phenotype, mutations were introduced into TCV unable to produce CP, due to a nonsense mutation replacing the sixth amino acid of the CP (UGA, (Manfre & Simon, 2008)). G3561A, which increased viral accumulation in wt and rev1 backgrounds, was still beneficial in the UGA background, accumulating in protoplasts to 215% relative to TCV with UGA alone. Similarly, A3475G/A3709G/U3741C, which increased accumulation in the rev1 background, accumulated to 154% of UGA when in the UGA background. Somewhat surprisingly, A3420G/U3508G/U3510G, which was detrimental in all other backgrounds, also accumulated to above UGA levels (177%) when in the UGA background. This indicates that the enhancement of viral accumulation in protoplasts seen with G3561A and A3475G/A3709G/U3741C cannot be attributed to CP alterations. However, the negative impact of A3420G/U3508G/U3510G is likely due to the two alterations in the CP, as these three mutations together in the UGA background are no longer detrimental.

Mutations in upstream putative structures S1 and S2 do not affect accumulation of the virus in the absence of the CP

Several second-site mutations that enhance accumulation in protoplasts in the wt background (Fig. 3.2A) were clustered in an mFold (Zuker, 2003) predicted structure that contains two stem loops, S1 and S2. mFold predicts RNA secondary structure based on thermodynamics. Tertiary interactions can stabilize RNA structure, but are not taken into account in mFold predictions. SHAPE (Selective 2'-Hydroxyl Acylation analyzed by Primer Extension) was subsequently performed on an upstream region of TCV, bases 800-900 (M. Kuhlmann and A. E. Simon, unpublished). Using the constraints from this SHAPE, full-length TCV was refolded and the 3' 800 bases in the majority of structures obtained folded into two domains, Domain1 and Domain2. Domain2 contains the region where second-site mutations were found. This new structural prediction still forms S1, but folds the region of S2 differently (Fig. 3.1C compared to Fig. 3.3B), however neither structural prediction of this region has been experimentally confirmed. Deletions in this area were made without the SHAPE information, so deletion endpoints are somewhat arbitrary, but are based on the earlier structure, which was predicted by mFold.

Since these mutations are clustered in a region that, when mutated, enhances accumulation in protoplasts, this suggests that the normal function of this region might be to act as a repressor. U3508G/U3510G are present in a copy of the original m21 mutation (UGUUA in H4AL, boxed in Fig. 3.1C). These two U to G mutations in the H4AL repeat in S2 in the wt background accumulated in protoplasts better than wt (126%). A3420G in S1 in the wt background resulted in accumulation in protoplasts of only ~18% of wt. When these three mutations were combined in the wt background (A3420G/ U3508G/

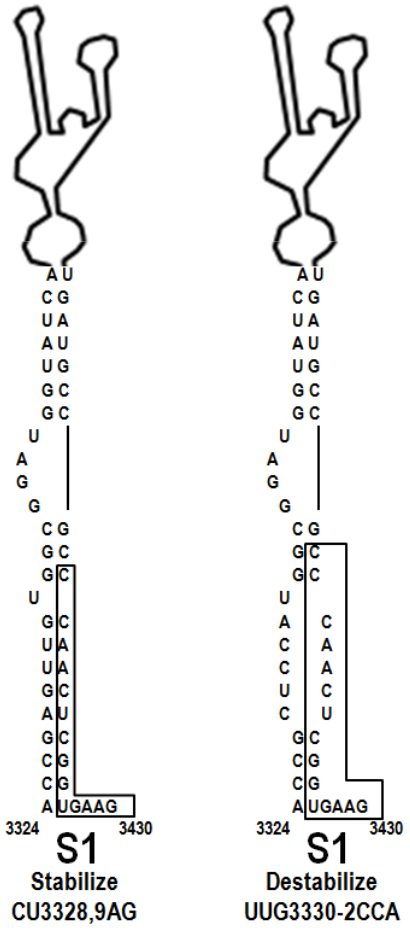
Figure 3.3 Effect of S1 and S2 mutations on accumulation in protoplasts

A. Stabilization or destabilization of S1. Inverted sequence repeat is boxed.

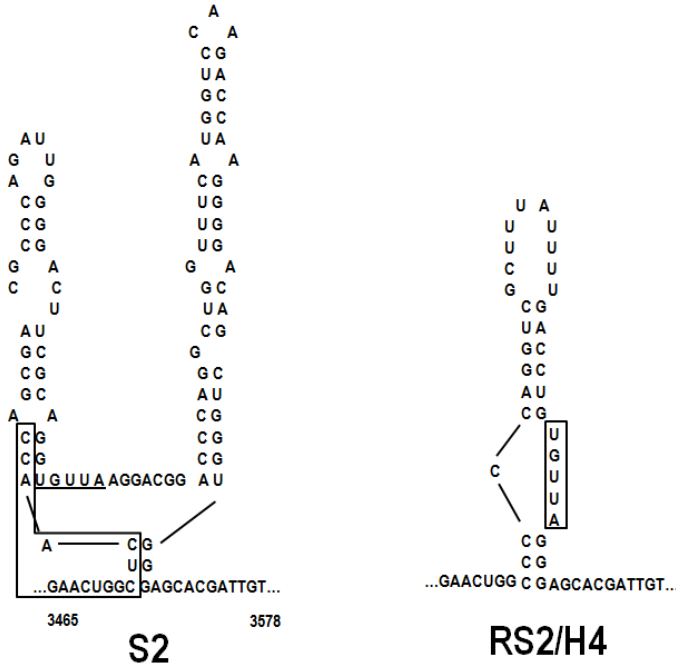
B. On left, S2, with inverted sequence repeat boxed and repeated UGUUA underlined. On right, H4 replacing S2, with UGUUA of primary mutation boxed. Structure of S2 is from the mFold prediction done prior to upstream SHAPE.

C. On left, accumulation in protoplasts, as determined by northern blot, of full-length TCV genomic RNA containing S1 stabilizing/destabilizing mutations. On right, accumulation in protoplasts, as determined by northern blot, of virus lacking S2 or containing H4 in place of S2. For protoplasts, relative values are from at least three independent experiments and standard deviation is shown.

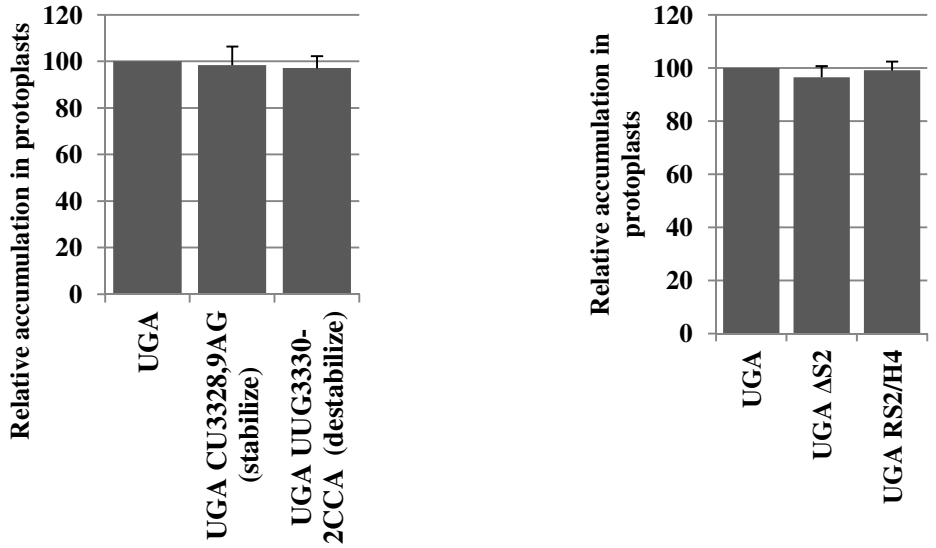
A


3324
3430
S1
Stabilize
CU3328,9AG
3324
3430
S1
Destabilize
UUG3330-2CCA

B


...GAACUGGC GAGCACGATTGT...
3465
3578
S2
...GAACUGG C G AGCACGATTGT..
RS2/H4

C


Relative accumulation in protoplasts
120
100
80
60
40
20
0
UGA
UGA CU3328,9AG (stabilize)
UGA UUG3330-2CCA (destabilize)
Relative accumulation in protoplasts
120
100
80
60
40
20
0
UGA
UGA ΔS2
UGA RS2/H4

U3510G), accumulation in protoplasts was reduced to 60% of wt, indicating that the two downstream mutations in the H4AL sequence repeat (U3508G/U3510G) can compensate for the negative mutation in S1. This suggests a possible interaction between S1 and S2. We noticed an interesting inverted sequence repeat (GAAc/gUGGCUCAACC, boxed in Fig. 3.3A, B) that could be indicative of how a repressor might function: the sequence in the upstream structure (S1), when S1 is destabilized, might be able to replace the sequence in the downstream structure (S2), compromising the ability of the region to repress as normal, thus increasing accumulation in protoplasts. Similarly, stabilizing S1 should result in a decrease in accumulation in protoplasts. Mutations were generated in the UGA background in S1, but not in the inverted sequence repeat, to either stabilize or destabilize the stem and assayed for accumulation in protoplasts. Neither stabilizing nor destabilizing S1 in the UGA background had an effect viral on accumulation, indicating that the potential strand invasion is not occurring (Fig 3.3C).

Because S2 has a copy of the UGUUA sequence that was mutated in H4AL as the primary mutation (see Fig 3.1), it is possible that S2 could be functioning as an additional H4 or in a manner similar to H4. To see if S2 could be functionally replaced by H4, S2 was deleted from the UGA construct and replaced with H4 (ΔS2 and RS2/H4, respectively, Fig. 3.3B). Neither deleting S2 nor replacing it with H4 affected viral accumulation in the UGA background in protoplasts (Fig. 3.3C), even though deleting S2 deletes a region that contained four second-site mutations (see Fig. 3.1C). These results indicate that S1/S2 are not acting as a repressor, as they can be mutated or deleted without detriment to accumulation, and that S2, though bearing a sequence repeat of the primary mutation, is not merely acting as an additional H4.

Second-site mutations and mutations in S1 and S2 do not affect translation of a reporter construct

The second-site mutations were not detrimental to viral accumulation when CP effects were eliminated and actually increased viral accumulation in the UGA background. Because viral accumulation in protoplasts is a product of both transcription and translation of the virus, these mutations were introduced into a firefly luciferase (Fluc) reporter construct to determine what effect they have on translation of the virus. The parental TCV Fluc construct is 63-Fluc-3661, where 63 indicates the 63-base 5' UTR of TCV and 3661 indicates TCV sequence starting at base 3661 of TCV (includes the 3' 140 bases of the CP ORF and the 3' UTR of TCV). Since these mutations are upstream of TCV base 3661, the luciferase construct was extended to start at TCV base 3281 (380 additional bases: 63-Fluc-3281, see Fig. 3.4A).

Figure 3.4D shows the parental 63-Fluc-3661 construct as the 100% level and the second-site mutations introduced into the longer 63-Fluc-3281 construct. As a control, the additional sequence (3281-3660) was inverted to ensure that any differences seen were not due to potentially increased stability of a longer construct. 63-Fluc-3660-3281, the inverted sequence control, reduced translation to 75% of 3661 levels, while the extended construct, 3281, increased translation to 153%. G3561A increased translation to 153%, A3475G/A3709G/U3741C increased translation to 147%, and A3420G/U3508G/ U3510G increased translation to 172% of wt. These levels are comparable to translation of 63-Fluc-3281, the extended construct into which the second-site mutations were introduced, indicating that, while some of the second-site mutations enhanced viral accumulation, they did not improve translation of this reporter construct.

Figure 3.4 Effect of second-site mutations on relative translation of a reporter construct in protoplasts

A. Diagram of luciferase construct, showing viral sequence and the 5' and 3' ends.

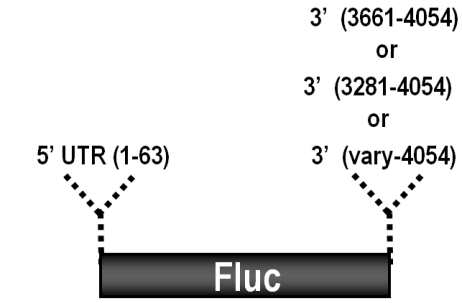
B. Schematic representation of deletion constructs. Dashed line denotes inversion of sequence from 3281-3660 (construct 3660-3281). Approximate location of mutation in m3609-3614 is shown. Locations of TSS and 3' hairpins are shown. Relative values are from at least three independent experiments and standard deviation is shown.

C. mFold predicted structure of TCV bases 3851-4054 showing putative hairpin and interaction.

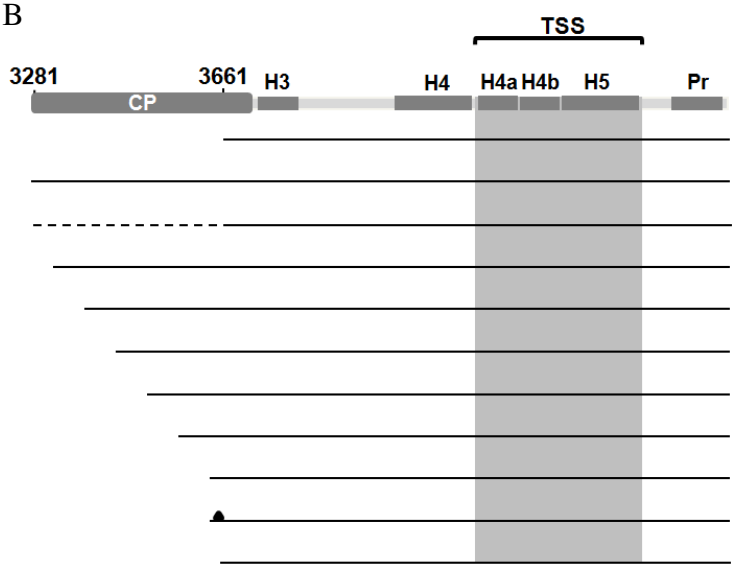
D. Relative translation in protoplasts of second-site mutations and S1/S2 mutations. *=$p<0.05$, **=$p<0.01$, Student's *t*-test.

E. Relative translation of decreasing amounts of viral RNA at the 3' end of the reporter construct. ***=$p<0.001$, Student's *t*-test.

A

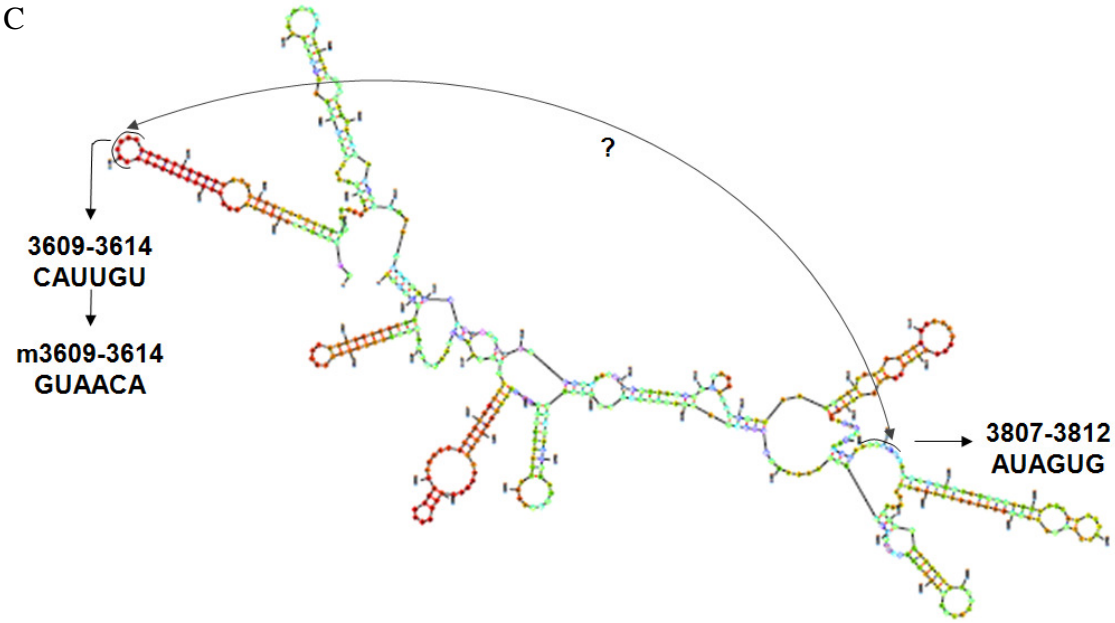


B

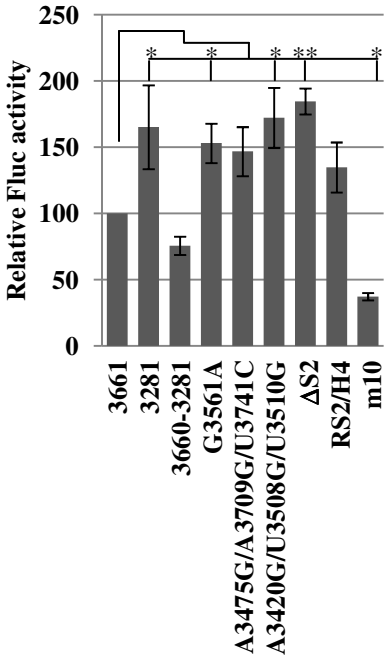


% Luciferase activity	Construct
100	3661
153.5 ± 5.8	3281
75.7 ± 6.9	3660-3281
167.6 ± 9.3	3331
170.7 ± 1.4	3381
218.7 ± 23.0	3431
238.3 ± 9.9	3481
219.2 ± 15.5	3531
291.6 ± 25.2	3581
266.6 ± 30.0	3581 m3609-3614
266.2 ± 32.8	3641

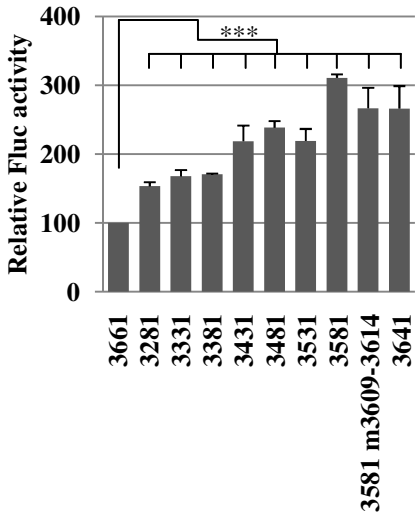
C


?
3609-3614
CAUUGU
m3609-3614
GUAACA
3807-3812
AUAGUG

D


Relative Fluc activity
250
200
150
100
50
0
*
*
*
**
*
3661
3281
3660-3281
G3561A
A3475G/A3709G/U3741C
A3420G/U3508G/U3510G
ΔS2
RS2/H4
m10

E


Relative Fluc activity
400
300
200
100
0
***
3661
3281
3331
3381
3431
3481
3531
3581
3581 m3609-3614
3641

Deleting S2 and replacing it with H4 did not affect viral accumulation, but these mutations could still be having an effect on translation. To determine this, S2 was deleted from the extended 3281 luciferase construct and replaced with H4 (Fig. 3.4D). Deleting S2 increased translation to 185% and replacing it with H4 increased translation to 134%. These levels are comparable to translation of 63-Fluc-3281, indicating that these mutations do not substantially affect translation of the reporter construct just as they also do not affect viral accumulation.

m10 changes the two consecutive uridylates in H4AL to adenylates and translates at about 16% relative to 63-Fluc-3661 (Stupina et al, 2008). As a negative control m10 was also introduced into the extended construct. 63-Fluc-3281 m10 reduces translation to 37% of 3661, which is about 24% of 3281, demonstrating the expected detrimental effect.

Taken together these results indicate that second-site mutations generated by a primary mutation in H4 in the 3' UTR are not detrimental to viral accumulation when CP effects are accounted for and do not affect translation of a reporter construct. These mutations were generated to compensate for the deleterious m21 mutation and many were clustered in an upstream region, but determining exactly how these second-site mutations may have helped improve viral function remains elusive. What is interesting is that extending the 3' viral contribution enhanced translation about 1.6 fold (3281 as compared to 3661). Inverting the additional bases did not enhance translation (3660-3281 as compared to 3281), so the increased translation is not due to having additional sequence.

Extending the viral 3' region increases translation of a reporter construct

Extending the 3' viral region in the reporter construct increased translation in a sequence-specific manner. To determine what in this additional sequence was contributing to increased translation, a deletion series was generated to narrow down what sequence was necessary for the increased translational enhancement (Fig. 3.4B and E). The additional 5' positioned 380 bases (compared with the original 3661) were deleted in 50-base increments from the 5' end and and reporter constructs assayed for luciferase activity. Surprisingly, the smallest fragment (63-Fluc-3581), which was only 80 bases longer than 63-Fluc-3661, provided the most translational enhancement (Fig. 3.4E).

mFold was used to examine possible structures in a fragment of TCV starting at base 3581, corresponding to the shortest of the deletion series (Fig. 3.4B), which could provide information relevant to the increased translation. All of the proposed structures contained a hairpin, the loop of which is capable of possibly pairing with downstream sequences (see Fig. 3.4C). To test if this is the method by which this fragment has increased translation, the loop of the proposed hairpin was mutated to disrupt the potential interaction (3581m3609-3614) and the fragment was shortened to just past the proposed hairpin, to entirely eliminate the proposed hairpin (3641). Mutating the loop to disrupt the potential interaction had no effect on translation and deletion of the proposed hairpin also had no effect (Fig. 3.4E). 3641, which is only 20 bases larger than the original 3661 construct, maintained the increased translation.

Increasing the 3' viral region in the reporter construct increases translation through increased stability of H3

mFold was used to ascertain if there were any obvious structural differences between the TCV fragments starting at 3641 or 3661 that could account for the increased translation that occured from inclusion of the 20 additional bases. Just upstream of the UAG stop codon for the CP ORF is a hairpin (H3) that mFold indicates reliably folds in the fragment starting with 3641, but sometimes does not fold in the fragment starting with 3661. This suggested that H3 may have increased stability in the longer fragment (3641) which may account for the increased translation. In-line probing (McCormack et al, 2008; Winkler et al, 2002) was performed to determine structural differences in fragments starting at either 3661 or 3641 (Fig 3.5A, B). No differences were present in the structure of H3 in these two fragments, but there were flexibility differences upstream of H3.

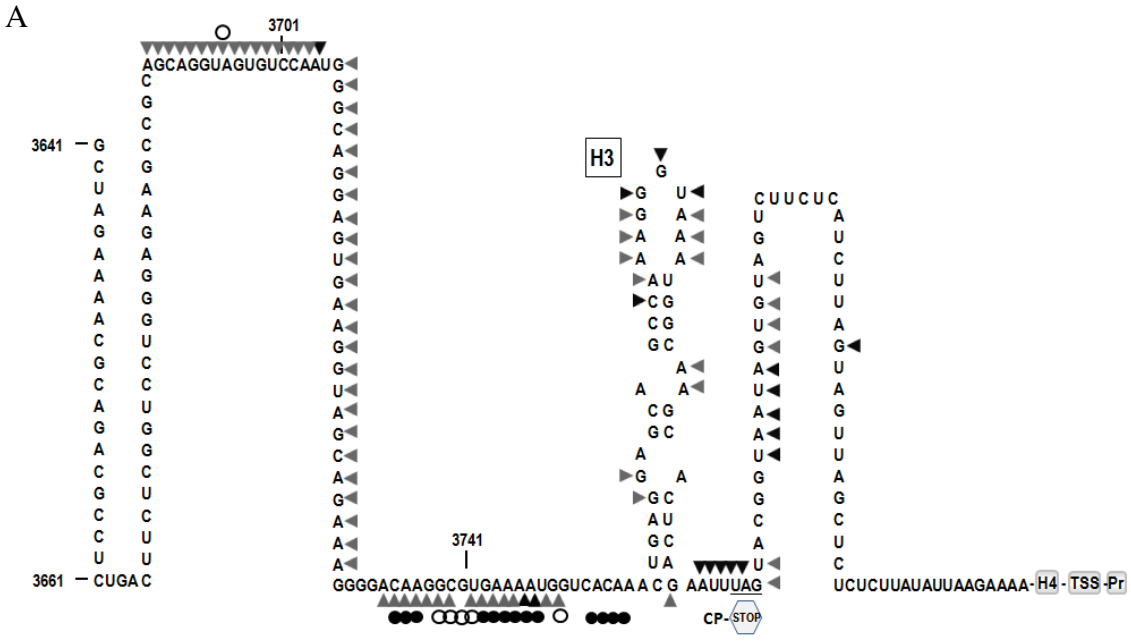
The region 3735-3750 was flexible in 3661 (triangles in Fig. 3.5) and this flexibility was altered in the longer 3641 fragment. The region 3735-3755 was generally more flexible in 3641 than in 3661, with the exception of four bases (GCGU 3739-42, open circles in Fig. 3.5) upstream of H3, which were significantly less flexible in 3641 than in 3661. This suggests these four bases may be participating in an interaction that facilitates H3 stability or translation.

.

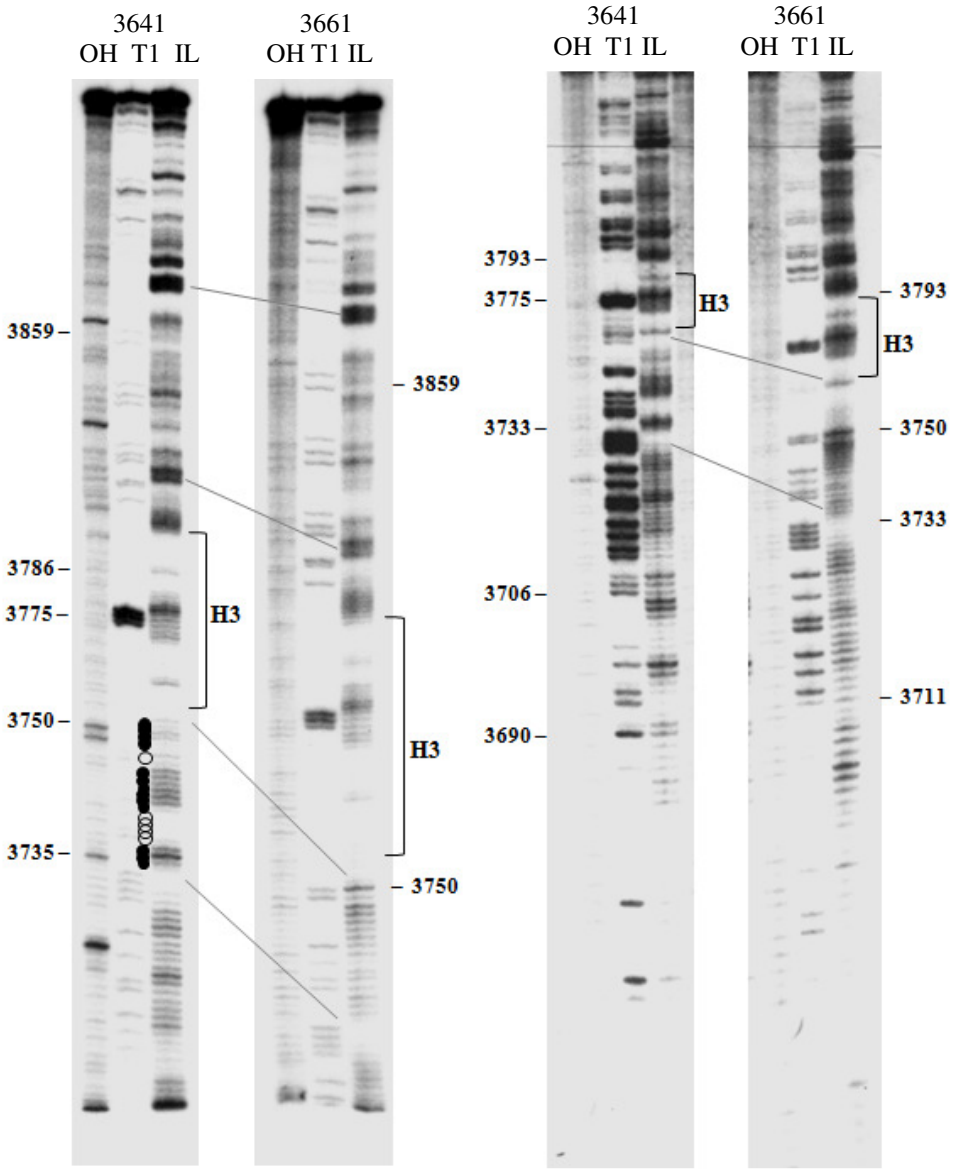
Figure 3.5 Effect of fragment length on structure of H3 and surrounding regions

A. Flexibility of residues in the 3661 fragment as determined by in-line probing. Darker triangles denote higher flexibility. Open circles indicate residues that have less flexibility in in-line probing of the 3641 fragment and closed circles indicate residues that have increased flexibility in the 3641 fragment.

B. In-line cleavage of fragments starting at either TCV base 3641 or 3661. Open or solid circles indicate decreased or increased, respectively, cleavage of 3641 relative to 3661.



B


3641
OH T1 IL
3661
OH T1 IL
3641
OH T1 IL
3661
OH T1 IL
3859
3786
3775
3750
3735
H3
3859
H3
3750
3793
3775
3733
3706
3690
H3
3793
H3
3750
3733
3711

To test whether the structure and/or stability of H3 is important for translation, mutations were made in the luciferase construct (Fig. 3.6A) that alter the loops of H3 (mLL H3, mML H3, mTL H3); eliminate H3 (ΔH3); replace H3 with the comparable sequence from *Cardamine chlorotic fleck virus* (CCFV), the virus most closely related to TCV, (CCFV H3); or disrupt the upper stem of H3 on either side and together, for possible compensation, (C3768G, G3782C, C3768G/G3782C). Additionally, just upstream of H3 are several adenylates that a separate second-site mutation study indicated might be important (Yuan et al, 2012). Therefore, constructs were generated where the 3' viral sequence begins just before these As (3743) or where the 3' viral sequence begins at the base of H3 (3756) were made. Finally, a construct in which the 4 less flexible bases seen by in-line probing (see Fig. 3.5) in fragment 3641 were mutated (GCGU 3739-42 CACA) was also generated.

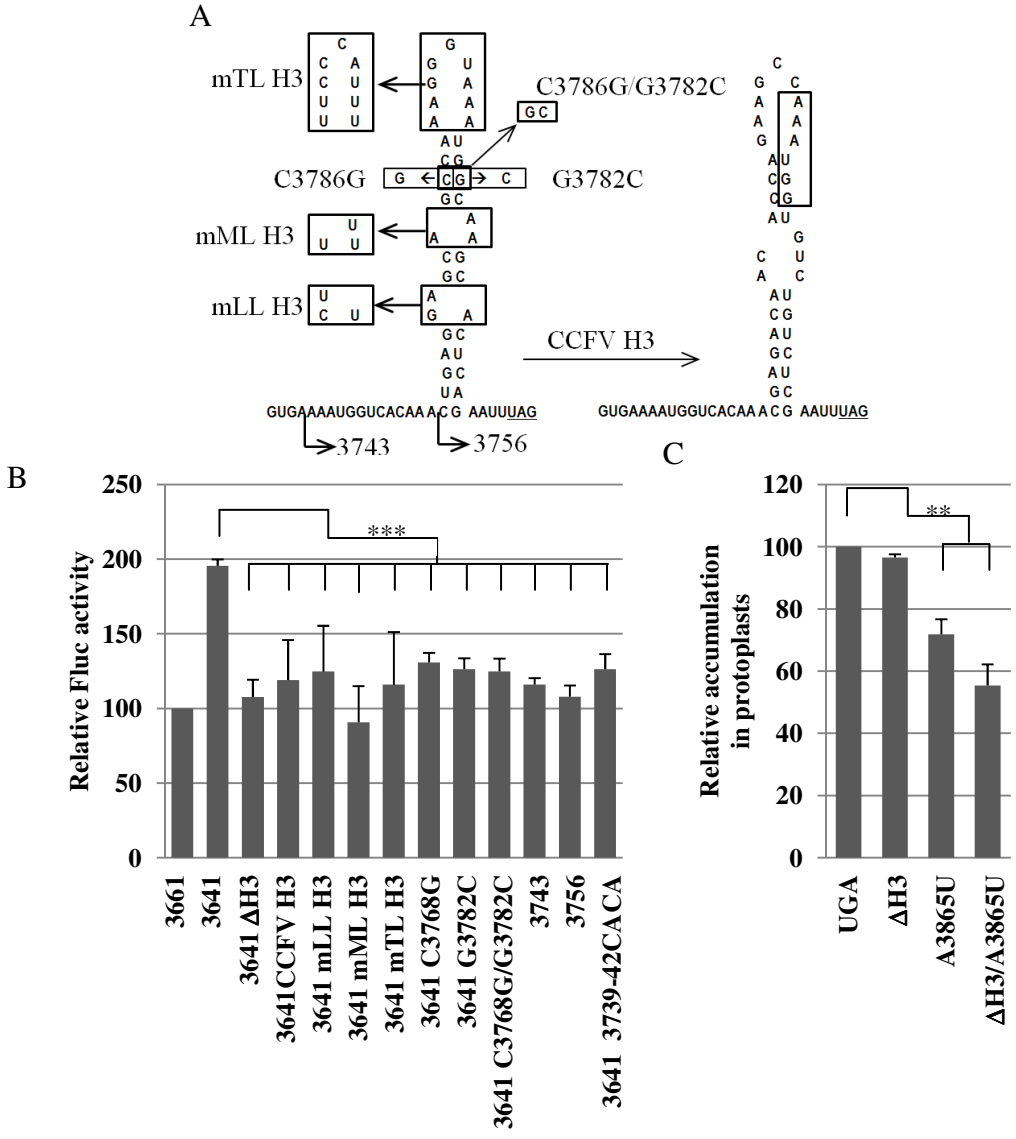
Deleting H3 (ΔH3) in 3641 reduced translation to 55% of 3641(the construct with the 3' sequence beginning at 3641) and replacing H3 with the comparable CCFV sequence (CCFV H3) in 3641 reduced translation to 61% of 3641. Mutations that altered any of the loops of H3 reduced translation to near 3661 levels, with the altered lower loop of H3 (mLL) reducing translation to 64%, the altered middle loop of H3 (mML) reducing translation to 46%, and the altered terminal loop of H3 (mTL) reducing translation to 59% of 3641 levels. This indicates that deleting or disrupting H3 reduces or eliminates the increased translation seen with the extended viral sequence (Fig. 3.6B).

Figure 3.6 Effect of mutations in and around H3 on translation and accumulation in protplasts.

A. Location and identities of mutations. Boxed sequence in CCFV H3 is conserved between TCV and CCFV.

B. Relative translation of mutations in A. ***=p<0.001, Student's *t*-test.

C. Relative accumulation in protoplasts following deletion of H3, A3865U, and coupling ΔH3 with A3865U in full-length TCV containing the UGA mutation eliminating CP translation. For protoplasts, relative values are from at least three independent experiments and standard deviation is shown. **=p<0.01, Student's *t*-test.



Disrupting either side of the upper stem of H3 (C3768G and G3782C) in 3641 reduced translation to 67% and 65% of 3641 levels. The two mutations together (C3768G/G3782C) in 3641 did not restore the enhanced translation, with translation levels at 64% of 3641. Starting the viral fragment at the 5' flanking As (3743) or at the base of H3 (3756) also eliminated the increased translation, likely due to H3 not folding properly in these constructs. Mutating the four less flexible bases (3739-42CACA) seen in in-line probing also eliminated the increased translation, translating at only 65% of 3641. Together, these results indicate that the formation and stability of H3 is important for enhanced translation.

Deletion of H3 reduces accumulation of the full-length virus in protoplasts

To determine if H3 plays a similarly important role in the full-length virus, ΔH3, A3865U (a primary mutation that generated a second-site mutation in H3, Yuan et 2012), and both alterations together were incorporated into the CP null background (UGA) to see how they affect accumulation of the virus in protoplasts (Fig. 3.6C). Surprisingly, deleting H3 had a minimal effect on accumulation (96%). A3865U reduced accumulation to 71% and the two together were worse, accumulating to 55% of the UGA wt. It is possible that deleting H3 would compensate for the deleterious effect of the A3865U mutation, as the second-site mutation generated in H3 was compensatory (Yuan et al, 2012), however this did not occur. A3865U affects the structure of Pr loop (Yuan et al, 2012) and Pr loop is important for proper translational readthrough to generate RdRp (Cimino et al, 2011), so it is likely that having no sequence/structure at all in the place of H3 was unable to correct these defects.

Discussion

The known importance of H4 in both translation and replication (Stupina et al, 2008, Yuan et al, 2009) suggested that H4 might interact locally or distally. Allowing a virus to self-evolve and generate second-site mutations that can compensate for a negative primary mutation is a well established method for detecting interactions in RNA viruses (Yuan et al, 2010; Zhang et al, 2006). In this study, second-site mutations that arose in response to primary mutations in H4AL (m21) were associated with a partial reversion of the primary mutations, resulting in the H4AL sequence being more similar to wt (UUA mutated to ACU partially reverting to UUU). The second-site mutations in the 3' UTR were not unexpected given the highly interactive nature of the 3' end of TCV (Yuan et al, 2009; Yuan et al, 2012), however second-site mutations generated upstream in the CP ORF suggested a heretofore unsuspected interaction between the 3' UTR and upstream regions.

Of the upstream second-site mutations, G3561A is of particular interest because it greatly increased viral accumulation in protoplasts when in the wt and rev1 backgrounds. This increased accumulation was maintained when G3561A was placed in the UGA background, which eliminates CP silencing suppressor and/or stabilizing effects as a consideration. Other mutations at or near G3561 did not display this accumulation increase, indicating an effect specific to the G3561A transition. To help determine why G3561A has this effect, additional studies are being performed, such as infection of plants (where viral movement is a consideration) and SHAPE of the region with and without the mutation.

mFold predicted that the region containing many of the upstream second-site mutations folds into two large stem loops. However, in the UGA background, deletion of one of these stem loops, which contains four of the second-site mutations, had no effect on viral accumulation. As the second-site mutations were generated from a primary mutation in the 3' UTR, it was assumed that there was an interaction between the upstream region and the region of the primary mutation. The deletion of S2, and, therefore, the elimination of any such interaction, had no effect on accumulation, suggesting that there is something other than a direct interaction occurring between these two regions.

Introduction of the upstream second-site mutations into the luciferase reporter construct revealed that additional viral sequence enhances translation. A deletion series revealed that enhanced translation was achieved with just an additional twenty viral bases. mFold suggested that these additional bases may help stabilize H3 and mutations in and around H3 indicated its importance for the increased translation seen. While H3 plays a role in translational enhancement of a reporter construct, deleting H3 in the UGA background had a minimal effect on viral accumulation. Furthermore, SHAPE of an upstream region from bases 800-900 revealed no differences when H3 had been deleted (M. Kuhlmann and A. E. Simon, unpublished). This region is where the Pr (the 3'-most hairpin in the 3' UTR) interacts to facilitate readthrough to generate the RNA-dependent RNA polymerase (Cimino et al, 2011). Because it is known that the Pr interacts locally with H4 (Yuan et al, 2010), and another second-site study with a primary mutation in an H4-flanking region generated second-site mutations in and around H3 (X. Yuan and A. E.

Simon, unpublished), it was possible that deletion of H3 would alter the structure or interactions of the 3'UTR thereby affecting this upstream region.

The elements within the 3' UTR have been shown to be very interactive with alterations affecting the structure of regions hundreds of bases away. Here, a second-site mutation study indicated that regions upstream of the 3' UTR may also be interacting with the 3' UTR. Some second-site mutations located in this upstream region were found to increase accumulation in protoplasts and additional studies are under way to explain this phenomenon. The 3' viral region in a luciferase reporter construct was increased to incorporate the second-site mutations. While the second-site mutations had little effect on translation of the luciferase construct, it was surprising to find that extension of the viral 3' sequence enhanced translation. Translational enhancement was mapped to just an additional twenty bases and further study revealed that a hairpin (H3) is important for viral translation and accumulation. Another second-site study revealed a possible interaction between H3 and the 3' UTR (Yuan et al, 2012), but the exact role of H3 in viral accumulation and translation is not yet known.

Chapter IV

Sequence and Length Requirements for a Hairpin

in the 5' end of a Satellite RNA Associated with TCV

Viruses can be associated with subviral RNAs [defective interfering (DI) RNAs or satellite (sat) RNAs] that are dependent on the helper virus for replication, movement, and encapsidation (Simon et al, 2004). Unlike DI RNAs, which have sequence primarily derived from the helper virus, most satRNAs have little or no sequence similarity to contiguous portions of the viral genome. Subviral RNAs are dependent on their helper viruses for *trans*-acting factors necessary for replication and often accumulate to the detriment of the helper virus, leading to marked attenuation of viral symptoms. This is especially true for DI RNAs, since they have extensive helper virus-derived sequence. satRNAs, in contranst, generally have little sequence similarity with the helper virus, and therefore have varied effects on helper virus accumulation and symptoms. Some satRNAs, like most of those found associated with *Cucumber mosaic virus* (CMV), reduce the accumulation of CMV RNA and attenuate symptoms. However some satRNAs associated with CMV are pathogenic, resulting in necrosis or programmed cell death, depending on the host (Simon et al, 2004; Xu & Roossinck, 2000). Symptoms can also be intensified by satRNAs, like satC of TCV and the satRNA of *Panicum mosaic virus* (PMV). While the satRNA of PMV enhances viral accumulation in infected plants (Scholthof et al, 1999), satC interferes with TCV accumulation (Li et al, 1989; Simon & Howell, 1986).

Due to their dispensability for and interference with helper virus accumulation and/or symptoms, subviral RNAs were considered parasites. However, evolutionarily, it

is not likely that viruses have maintained inessential, and often detrimental, subviral RNAs that do not confer a selective advantage to the virus (Simon et al, 2004). The satRNA of the umbravirus *Groundnut rosette virus* is required for viral encapsidation, to the obvious benefit of the virus (Robinson et al, 1999). There is a mutualistic relationship between satC and TCV, with both gaining fitness from the association. satC indirectly enhances movement of TCV by interfering with virion formation (Zhang & Simon, 2003) which results in an increase in available coat protein which, as the silencing suppressor, is needed to overcome post-transcriptional gene silencing (Qu et al, 2003; Thomas et al, 2003).

Subviral RNAs do not encode their own replicase but still must have sequences and/or structures that the viral RdRp can recognize for replication. DI RNAs and other subviral RNAs that share sequence with their helper viruses can be used to help determine what *cis*-acting elements are important for replication of the virus by also being important for replication of the subviral RNA. The DI RNAs of TBSV are well characterized (Chernysheva & White, 2005; Fabian et al, 2003), as is satC of TCV, which is similar to TCV at its 3'end (Simon & Howell, 1986; Zhang & Simon, 2005). The 5' part of satC is derived from another subviral RNA associated with TCV, satD. The TCV-like 3' end of satC has been extensively studied to gain insight into the functioning of TCV, but the satD-like 5' region has been largely unexplored. After SELEX (Systematic Evolution of Ligands by EXponential enrichment) of a putative hairpin near the 5' end of satC (bases 48-123, D. Kushner and A. E. Simon, unpublished), I inoculated the winners into protoplasts and performed northern blots to assay how well the SELEX winners were able to accumulate, relative to wild type satC.

Materials and Methods

In Vitro Transcription of RNA Using T7 RNA Polymerase

As described in Chapter II.

Culturing of Arabidopsis Callus

As described in Chapter III.

Preparation and Inoculation of Callus Culture Protoplasts with Infectious Viral RNA Using Polyethylene Glycol

As described in Chapter III.

Preparation and Inoculation of Callus Culture Protoplasts with Infectious Viral RNA or Luciferase Constructs Using Polyethylene Glycol

As described in Chapter III, except for each inoculation, 20 µg of *in vitro* transcribed wt TCV RNA plus 2 µg of wt or mutant satC was combined with 8 µl of 1 M calcium chloride and the volume brought up to 430 µl with water.

Extraction of Total RNA from Arabidopsis Protoplasts

As described in Chapter III.

Northern Blotting

As described in Chapter III, except using a probe that detects satC.

Results

Previous studies indicated a role for the satD-like 5' end of satC in satellite replication and regulation of the ratio of positive-strand monomers to dimers (Carpenter et al, 1991; Simon et al, 1988). There are two putative hairpins in the 5' 200 bases of satC: H2 and

H6 (Fig. 4.1). To determine what role H2 plays in the life cycle of satC, SELEX was performed on replacements of part or all of H2. Bases 67-103 (the top part of H2) were randomized, or the entire H2 haripin (bases 48-123) was replaced with either 76, 38, or 19 random bases (the full length of H2 is 76 bases, 38 bases is half of that, and 19 bases is one quarter) to assay for sequence and length requirements in H2 (D. Kushner and A. E. Simon, unpublished results). As a control, H2 was deleted and not replaced with any sequence and was allowed to self-evolve *in planta*. The winners of each of these constructs after five rounds of SELEX in turnip plants, or three rounds of self-evolution for the deletion construct, (Table 4.1) were inoculated into *Arabidopsis* protoplasts and a northern blot performed to determine accumulation relative to wild type satC.

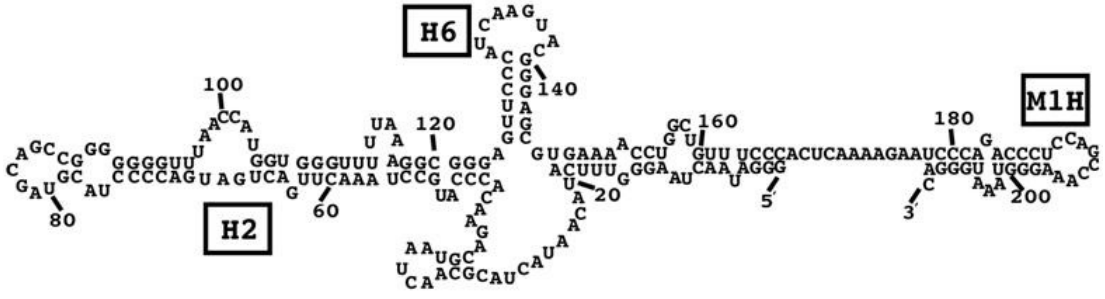


Figure 4.2 mFold prediction of 5' 210 bases of satC

H2 and H6 are mFold-predicted hairpins present in the 5' end of satC. H2 spans bases 48-123 and was the target of a SELEX assay.

SELEX	# of random bases	Sample name	Recovered sequences[a]	Short name
Δ 48-123[b] 76 bases	0	Δ H2	No random sequences inserted to replace H2	ΔH2
		Δ H2 + 38[c]	Flanking sequence duplicated, 124-140, 27-47	+38
nt 67-103 37 bases	37	R control 1	Pre-SELEX, random sequence	R
		R51a.3	(1) UAAAAGCCAGGGAUGUACCCCCAUCGUGU AUACUGGUG	R1
		R52b.1	(13) AGGCCACGUGCCCAGAAGUUCGCUAGAC UUUUGAUUU	R2
nt 48-123 76 bases	76	Y control 13	Pre-SELEX, random sequence	Y
		Y52a.2	(8) AACGUGUCUAUUCGGUCUGAAAACUGAG CGAUCGGUGC	Y1
		Y2	(3) AACGUGCCUAUUCGGUCUGAAGACUGAG CGAUCGGUGC[d]	Y2
nt 48-123 76 bases	38	B control 3	Pre-SELEX, random sequence	B
		B51c.1	(2) AGCUCCAACGGCCAAAGUCAGCCUGUCGU CCCAGGAGC	B1
		B51c.3	(9) GACAUUCAGAGGUCUCCCCUGACGGUGAC CGGGUCAUU	B2
		B52c.1	(12) GGACUGAAAGCAACGGACGGCAAAGCUC GUACUAACCU	B3
nt 48-123 76 bases	19	O control 1	Pre-SELEX, random sequence	O
		O51a.1	(12) GAGAUCCGGCGUGAAUCAU	O1
		O51ba.5	UCUGGAAAGUCUUUUUUUGGUAGUUCCCAG AAAGUCUUUUUUGGUaguuccca	O2
		O51b.2	(3) GAGAUCCGGCGUGAAUCAUAGAGAUCCG GCGUGAAUCAU[e]	O3
		O52b.2	(5) UGCGGUAUCAUUCAACUGC	O4
		O52b.5	wt	O5

Table 4.1 Recovered sequences after five rounds of each of the four SELEX experiments

a. The number in parentheses indicates how many times each sequence was recovered.

b. Not a SELEX, but self-evolution (no randomized bases included).

c. Duplicated flanking sequence to fill deletion *in vivo* after self-evolution. Final sequence: 1-47, [124-140, 27-47,] 124-end. Region in brackets is the location of the deletion that got filled by 38 bases of flanking sequence.

d. Sequence is identical to the prior sequence in the Table, with the exception of the 2 red, underlined nts.

e. A duplication of the sequence of O1, with an adenylate between the repeats.

(Data: D. Kushner and A. E. Simon, unpublished results.)

All of the randomized input sequences (R, Y, B, O) accumulated poorly (17%, 9%, 7%, and <1% of wt satC, respectively). The red SELEX winners, R1 and R2, were able to evolve greater fitness and accumulated to 88% and 81% of wt satC, respectively. Of the two yellow winners, Y1 and Y2, Y1 evolved moderate improvement to accumulate to 30% of wt satC, but Y2 still accumulated poorly to only 9% of wt satC. The three blue winners all accumulated similarly to about 30% of wt satC levels, despite their different recovered sequences. Of the five orange winners, O1 and O4 only accumulate to 2% and 8% of wt satC levels, respectively. O2 and O3 accumulated better at 57% and 50% of wt satC levels, respectively. O5 accumulated to 82% of wt levels, but it was discovered that O5 is essentially wt (O5 has not been completely sequenced, but of the portion that was sequenced, it was wt sequence). Overall, the SELEX winners were less fit than wild type satC, as they did not accumulate to wild type satC levels (Fig. 4.2). Interestingly, many of the winners altered the dimer:monomer ratio (Table 4.2), and this phenomenon is still being investigated.

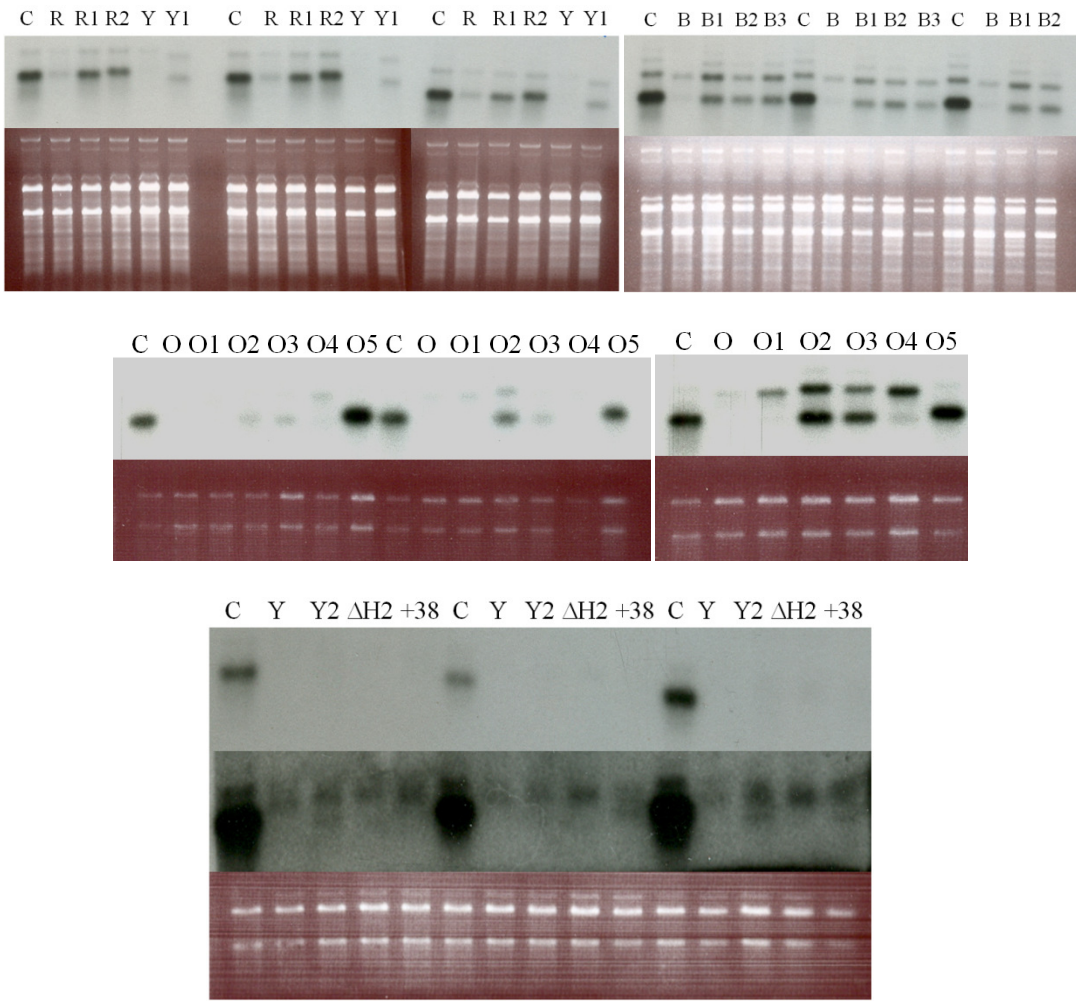


Figure 4.2 Accumulation of satC and satC SELEX winners

Relative accumulation of satC (C) and the satC SELEX winners (see table 4.1) in protoplasts at 40 hours post inoculation. Upper panels are the Northern blots. Lower bands in the blots are the monomeric form of satC, while upper bands are the dimeric form. Lower panels are the ethidium-stained agarose gels. The final blot has a middle panel showing the overexposed film to demonstrate that the mutants are present at very low levels. Monomers and dimers are quantified in Table 4.2.

	monomer		dimer		d/m
C	100	2.16	100	1.44	0.13
R	17.18	0.84	41.5	1.82	0.28
R1	88.11	1.58	101.86	5.79	0.17
R2	81.33	5.46	101.93	2.55	0.21
Y	8.54	0.75	34.23	2.12	0.76
Y1	29.91	1.66	102.87	3.84	0.45
Y2	8.99	1.31	ND	ND	ND

	monomer		dimer		d/m
C	100	1.32	100	0.97	0.25
B	6.65	1	18.19	1.21	1.75
B1	30.33	1.85	111.44	2.36	1.04
B2	26.17	0.81	74.58	1.13	0.79
B3	27.25	0.55	92.64	1.14	0.92

	monomer		dimer		d/m
C	100	1.09	100	1.58	0.04
O	0.27	0.04	88.77	2.61	8.49
O1	1.57	0.35	355.23	8.28	8.41
O2	56.6	1.55	837.18	8.31	0.66
O3	49.94	3.55	587.81	9.02	0.5
O4	8.08	0.15	731.28	6.73	4.75
O5	82.26	4.51	72.95	5.08	0.05

C	100	1.14
ΔH2	7.33	0.81
ΔH2+38	7.45	1.23

Table 4.2 Accumulation of satC and satC mutants

Both monomer and dimer levels at 40 hours post inoculation were determined from three independent experiments. The values and error are shown. The dimer:monomer ratio was determined from the average dimer and monomer densitometry values. ND = not determined.

Deletion and randomization of H2 revealed that the satellite requires sequence in this region, but not specific sequence. Generally, the recovered sequences resulted in 38 or 39 bases that have the potential to fold into a hairpin, suggesting H2 may function as a spacer, bringing 5' and 3' flanking sequences into appropriate spatial configurations for proper biological function. It is interesting that, when allowed to evolve its own sequence, satC prefers sequence one-half the amount of wild type sequence (38 bases as compared to full-length H2, 76 bases). Accumulation of the construct with H2 deleted was negligible, as was the self-evolved ΔH2+38, which used flanking sequence to fill in an additional 39 bases. satC with H2 deleted or replaced with 38 random nt accumulated very poorly, at less than 10% of wt. The northern blot was overexposed to ensure the presence of these mutants (Fig. 4.2). The dimer:monomer ratio was not determined for these constructs, or Y2, as there was barely any detectable satellite present.

Chapter V

Other Experiments

Materials and Methods

In Vitro Transcription of RNA Using T7 RNA Polymerase

As described in Chapter II.

Filter Binding

As described in Chapter II, except RNA fragments were incubated with varying amounts of eEF1a for binding.

Culturing of Arabidopsis Callus

As described in Chapter III.

Preparation and Inoculation of Callus Culture Protoplasts with Infectious Viral RNA Using Polyethylene Glycol

As described in Chapter III.

Preparation and Inoculation of Callus Culture Protoplasts with Infectious Viral RNA or Luciferase Constructs Using Polyethylene Glycol

As described in Chapter III.

Extraction of Total RNA from Arabidopsis Protoplasts

As described in Chapter III.

Northern Blotting

As described in Chapter III.

In vivo Translation Assay

As described in Chapter III.

Electrophoretic Mobility Shift Assay

All RNAs were transcribed *in vitro* using T7 RNA polymerase. RdRp was expressed in *E.* coli and purified as a recombinant protein with Maltose Binding Protein (MBP), as previously described ((Rajendran et al, 2002) see Chapter II). Each reaction contained approximately 10 ng of satD-, labeled with ^{32}P at its 5' end; 1 μg of RdRp; and competitor RNAs in excess, as indicated. The RNAs and RdRp were incubated together for 30 minutes at room temperature in a buffer containing 50 mM Tris pH 8.2, 10 mM $MgCl_2$, 10 mM dithiothreitol, and 10% glycerol. 1.5 μg of MBP was used as a protein control and yeast tRNA was used as a control competitor. Reactions were electrophoresed on a 1% agarose gel in an ice bath at 25 mA for 35 minutes. The gel was then dried and subjected to autoradiography followed by densitometry scanning for quantification (QuantityOne).

Results

Pr loop mutants enhance RdRp binding

The second-site mutation study discussed in Chapter III also generated second-site mutations in the 3'UTR. Some of these mutations were present in the Pr loop. The Pr loop mutations were found to affect *in vitro* transcription, *in vivo* translation and accumulation, and the structure of upstream regions (Yuan et al, 2010). I wanted to determine if these mutations resulted in altered RNA-dependent RNA polymerase (RdRp) binding to the 3' end. I used an electrophoretic mobility shift assay (EMSA) to

assay binding of RdRp to either wild type F4 or F4 containing mutations in the Pr loop (Fig. 5.1). Figure 5.1 shows that some of the Pr loop mutants have an enhanced ability to bind the RdRp. UCG4032AA in F4 and wt F4 compete to similar levels against satD- for RdRp binding. However, U4032C in F4 and G4034A in F4 are better competitors than wt F4 for binding to TCV RdRp, with U4032C being the stronger competitor. Interestingly, these results indicate that RdRp binding is inversely correlated to template activity in the *in vitro* transcription assay (Yuan et al, 2010).

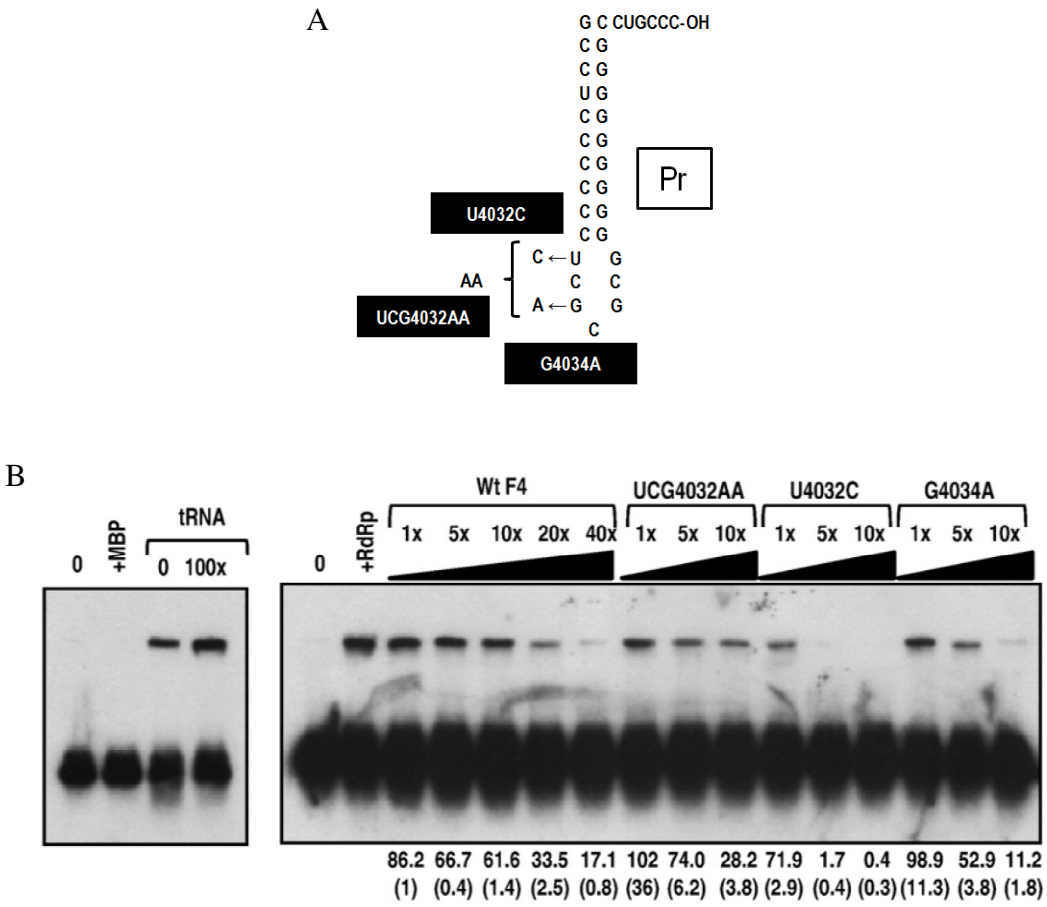


Figure 5.1 RdRp binding to Pr loop mutants. Electrophoretic mobility shift competition assays were conducted using radiolabeled satD minus strands (satD-) and increasing amounts of unlabeled competitor RNAs.

A. Sequence and location of Pr loop mutants.

B. Left panel, controls with no RdRp (0), maltose binding protein (+MBP) or TCV RdRp (brackets) with 0 or 100x yeast tRNA as competitor. Right panel, no RdRp (0); plus RdRp alone (+RdRp); or plus RdRp and increasing amounts of wt and mutant competitor RNAs (brackets). The assays were conducted four times. Average values for percent shifted RNA are given and standard deviation is in parentheses.

There is no direct 5'-3' RNA interaction in TCV, as assayed by a translation reporter construct

Small RNA plant viruses, such as *Blackcurrant reversion virus*, the viruses with BTEs, and *Tomato bushy stunt virus*, are known to circularize via RNA:RNA interactions (Karetnikov & Lehto, 2008; Treder et al, 2008; Fabian & White, 2006). While it is known that there is synergistic translational enhancement between the 5' and 3' UTRs of TCV (Qu & Morris, 2000; Stupina et al, 2008), it is not known how this enhancement occurs. Given that canonical eukaryotic translation proceeds through a circularized message and that some other small RNA plant viruses are known to circularize, it is possible that circularization of the genome plays at least some role in translational enhancement in TCV. There are sequences in the 5' and 3' UTRs that have complementarity and could circularize the genome.

Ψ_4 exists at least some of the time *in vivo* (Stupina et al, 2008, Yuan et al, 2009). Since the 5' side sequence of the large symmetric loop (LSL: UAAAAU) is repeated in the 5' UTR (bases 29-34), it is possible that the partner sequence, the terminal loop of H4 (H4TL), interacts with this sequence at the 5' end when not interacting with the same sequence in the 5' side of the LSL to form Ψ_4. A comparable potential interaction is present in the most closely related virus to TCV, *Cardamine chlorotic fleck virus* (CCFV), with its 5' LSL sequence and Ψ_4 partner (ACGAAAUC) present in its 5' UTR.

To test this potential interaction, I made mutations in a luciferase construct and assayed for translation in protoplasts. These constructs have mutations in the 5' potential interacting sequence that would disrupt the putative interaction. I also tested each of these three 5' end changes with three different 3' ends: wt TCV 3' 400, m10 (UGUUA in H4

asymmetric loop to UGAAA) in the TCV 3' 400, and H4TL (terminal loop of H4 changed from GCUUUAUUUU to GCAAAUAAAA) in the TCV 3' 400 (see Fig. 5.2). I tested the 5' end mutations in conjunction with m10 in the 3' UTR because it has the potential to base pair with the sequences of interest in the 5' UTR and because it is known that H4AL interacts with H4TL and destabilizes Ψ_4 (Yuan et al, 2010). Additionally, mutations in the 5' side of the LSL of H5, which is the 3' side of Ψ_4 and contains the sequence that is repeated in the 5' UTR, show structural effects in H4AL in addition to changes in H4TL. Mutations in H4TL also affect the structure in H4AL (Yuan et al, 2009).

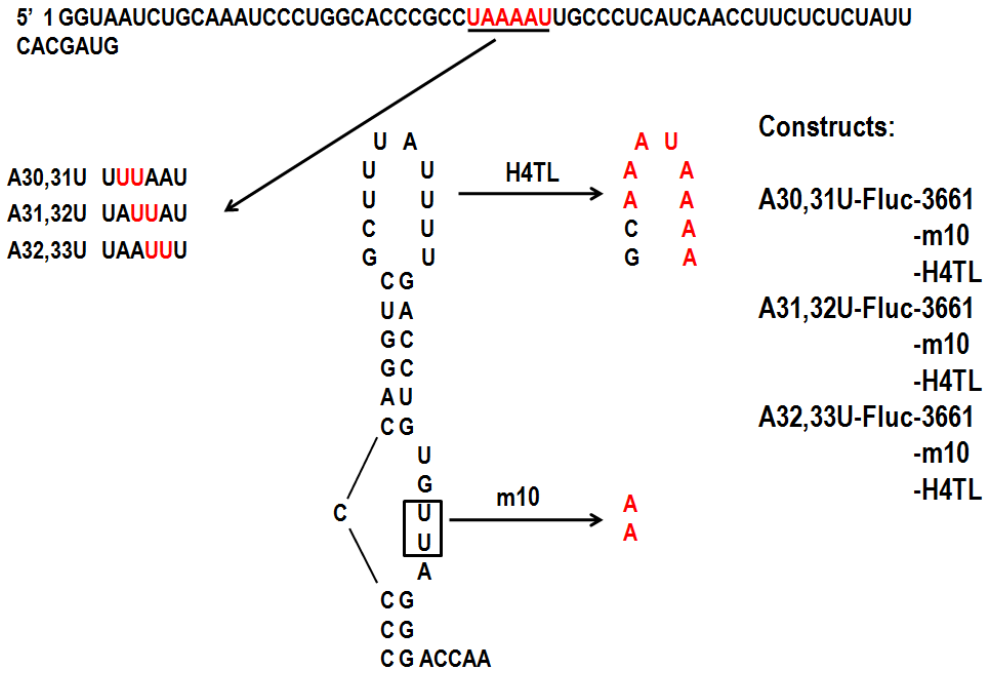


Figure 5.2 Potential 5' – 3' interacting sequence in TCV

The red, underlined "UAAAAU" in the 5' UTR is the same as the sequence found on the 5' side of the LSL of H5 in the 3' UTR. In the 3' UTR, this sequence interacts with the H4 terminal (H4TL) loop to form Ψ_4. The 5' sequence was mutated and tested in conjunction with three different 3' ends, resulting in the constructs listed at right.

The 5' mutations in conjunction with the wild type 3' end had minimal effect on translation, with the most severe (A30,31U-Fluc-3661) still translating to 75% of wild type levels (Fig. 5.3). That same 5' end mutation actually improved translation when coupled with m10 or H4TL mutations in the 3' end, as compared to those 3' mutations coupled with a wild type 5' end (A30,31U-Fluc-m10, 60% ; A30,31U-Fluc-H4TL, 26%; compared to 63-Fluc-m10, 12%; 63-Fluc-H4TL, 15%). The other 5' mutations coupled with the 3' mutations translated to levels comparable to the 3' end mutations alone.

Since the asymmetric loop and terminal loop of H4 are known to interact (Yuan et 2009), I wanted to see if the H4TL mutation in conjunction with m10 in the asymmetric loop was viable for translation. These two mutations together did not increase translation over their individual levels (63-Fluc-H4TL+m10, 9%; 63-Fluc-m10, 12%; 63-Fluc-H4TL, 15%). Indeed, since these two mutations together aren't substantially more detrimental than either alone, it suggests that they affect the same function. m38 is a mutation in the Ψ_4 interacting sequence in the LSL of H5. Again, since the asymmetric loop of H4 (H4AL) is known to negatively affect the formation of Ψ_4, I thought perhaps m38, which disrupts Ψ_4, coupled with m10 in the H4AL, might be viable for translation. While m38 alone translates to wild type levels, m38 coupled with m10 translates to only 8% of wild type levels, indicating that the detrimental effect of m10 alone (12%) is not overcome by the presence of m38, which disrupts Ψ_4 (Fig. 5.3).

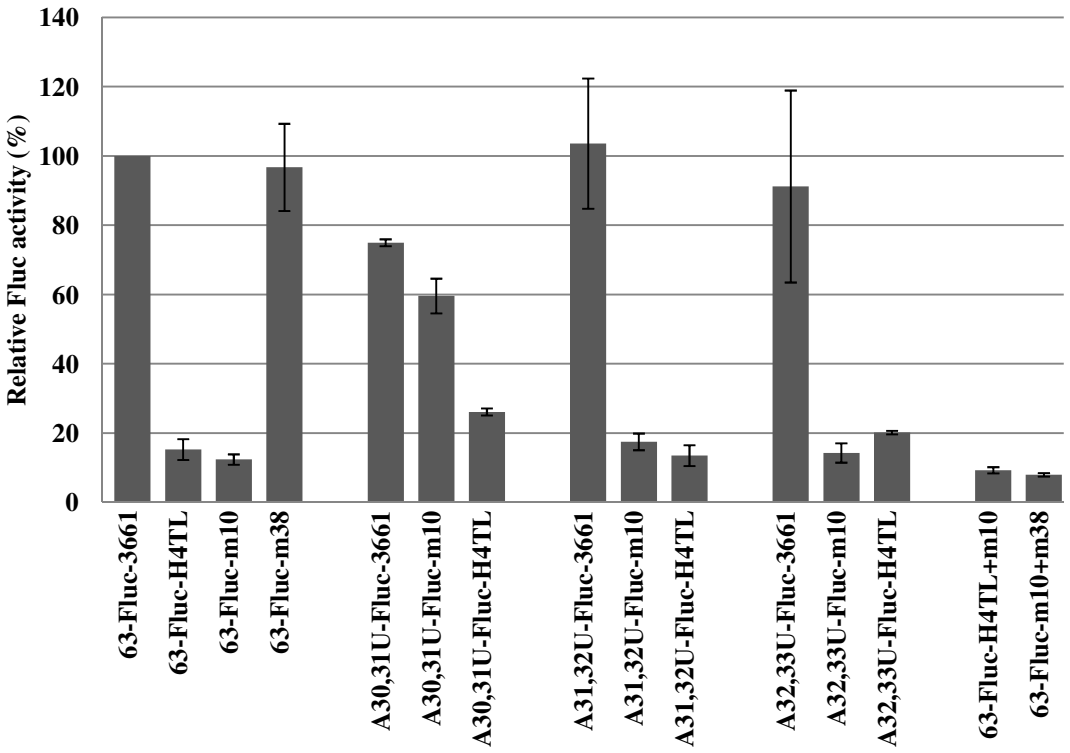


Figure 5.3 Relative translation of mutants that disrupt the possible 5'-3' interaction

Relative translation of mutant TCV luciferase constructs in protoplasts at 18 hours post inoculation. All values are averages from at least three experiments. Standard deviation bars are indicated.

While testing these mutations did not indicate an interaction between the two regions, any interaction would likely be complicated by the many interactions in the 3' end (Yuan et al, 2009; Yuan et al, 2010; Yuan et al, 2012). Interestingly, it has been shown that the 3' UTR of TCV works synergistically with the 5' UTR of *Saguaro cactus virus* (SCV; K. Shi and A. E. Simon, unpublished) and with the 5' UTR of PEMV (X. Yuan and A. E. Simon, unpublished) in the translational luciferase assay. This is

interesting because the 63 base 5' UTR of TCV, the 39 base 5' UTR of SCV, and the 20 base 5' UTR of PEMV have very little sequence similarity, so it is unclear why the 5' UTRs of SCV and PEMV in conjunction with the 3' UTR of TCV are able to direct efficient translation from the luciferase construct. Furthermore, unlike TCV, SCV and PEMV are known to circularize their genomes via a long-distance RNA:RNA interaction (Chattopadhyay et al, 2011, F. Gao and A. E. Simon, umpublished), further obfuscating what may be occurring with the disparate sequences in the luciferase construct.

Functional Domains in the 3' UTR of TCV

Previous work in the lab investigated the functional domains present in the 3' UTR of TCV through mutational analysis and replacement of TCV elements with the analogous elements of *Cardamine chlorotic fleck virus* (CCFV, see Fig 3 of McCormack et al, 2008), the virus most closely related to TCV (65% identity in the 3'UTR). It was found that the 3' 142 bases of CCFV (Fig. 5.4) could functionally replace the equivalent region in TCV, with that chimeric construct accumulating to levels comparable to wild type. Additional chimeric constructs were made that replaced the various elements found in the 3' end of TCV indicating that some elements are replaceable while others are less so (Table 5.1, left, from McCormack et al, 2008). Since the construct containing CCFV Ψ_3→H4b resulted in increased accumulation in protoplasts (McCormack et al, 2008), I wanted to make and test additional constructs to narrow down exactly which elements are required for increased accumulation. I made 5 constructs: one containing just the DR of CCFV (the upstream flanking sequence of H4a, which forms Ψ_3), one containing Ψ_3 and H4a of CCFV, one containing just the DR and H4b of CCFV, one contiaing Ψ_3 through H5 of CCFV, and one containing H4a through Ψ_2 of CCFV (refer to Fig. 5.4 for

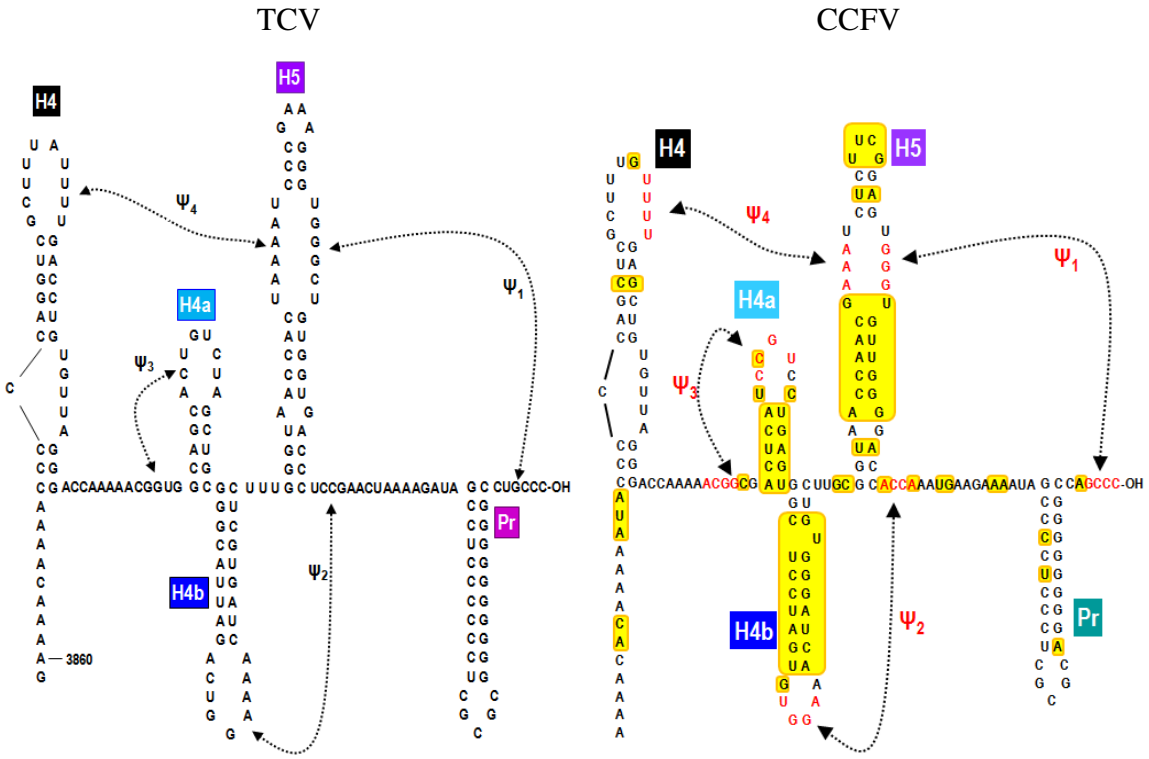


Figure 5.4 Sequence and structure of the 3' ends of TCV and CCFV

Hairpins are labeled in colored boxes. Tertiary interactions (pseudoknots, Ψ) are depicted by double-headed arrows with paired sequence colored red. Base differences between TCV and CCFV are shown boxed in yellow in the CCFV sequence. The CCFV structures are predicted based on similarity to TCV.

structures). The right side of Table 5.1 shows the results. Replacing just the DR did not affect accumulation, while replacing $\Psi_3 \rightarrow$H4a or H4a$\rightarrow \Psi_2$ resulted in accumulation of 41% of wild type. Replacing the DR and H4b as well as $\Psi_3 \rightarrow$H5 resulted in no detectable accumulation in protoplasts. Interestingly, replacing $\Psi_3 \rightarrow$H5 resulted in no detectable accumulation, but replacing $\Psi_3 \rightarrow$H4b increased accumulation to 128%. The difference between these two substitutions is just H5 and the small linker between H4b and H5. This suggests that H5, in the context of upstream regions, is detrimental to accumulation even though replacing H5 alone accumulated to 14% of wild type (McCormack et al, 2008).

Construct*	Accumulation (%)
pTSNL5 (wt)	100
CCFV H4a	21 ± 9
CCFV H4b	62 ± 12
CCFV H4a,H4b	46 ± 15
CCFV $\Psi_3 \rightarrow$H4b	128 ± 46
CCFV H5	14 ± 4

Construct	Accumulation (%)
pTSNL5 (wt)	100
CCFV DR	98 ± 6
CCFV Ψ_3, H4a	41 ± 17
CCFV DR, H4b	0
CCFV $\Psi_3 \rightarrow$H5	0
CCFV H4a$\rightarrow \Psi_2$	41 ± 19

*made and tested by J.C.McCormack (McCormack et al, 2008)

Table 5.1 Accumulation in protoplasts of chimeric TCV containing 3' elements from CCFV On the left is data from McCormack et 2008. On the right are new constructs. Mutations were made in pTSNL5, which is wt TCV with a SnaBI site added in the linker between H5 and Pr and serves here as the wt 100% control.

This work could be continued by constructing additional chimeras. The construct $\Psi_3 \rightarrow$Link2 (Link2 is the small linker between H4b and H5) would either implicate or rule out this small linker in reducing accumulation in protoplasts to undetectable levels. If Link2 proved to not be detrimental to accumulation, as is expected given the very limited differences in this region between TCV and CCFV (refer to Fig. 5.4), the different regions of H5 could be stepwise replaced to determine which differences result in the

reduced accumulation. H5 of TCV and CCFV differ in many areas, including the terminal loop, the upper stem, the large symmetric loop, the middle stem, the lower loop, and the lower stem. In-line probing of the 3' ends of the two different chimeras that accumulate so differently (Ψ_3→H4b, 128%; Ψ_3→H5, undetectable) may reveal structural differences that would elucidate what is occurring to result in such different accumulation levels.

Binding of eEF1A to 3' end fragments of TCV

The eukaryotic translation elongation factor 1a (eEF1a) is known to interact with plant viral genomes and play a role in viral replication. During *Tobacco mosaic virus* (TMV) infection, eEF1a has been shown to interact with the RNA-dependent RNA polymerase (RdRp) of TMV as well as the viral 3' end (Yamaji et al, 2006; Zeenko et al, 2002). Furthermore, when eEF1a is downregulated in host cells, TMV accumulation is correspondingly reduced (Yamaji et al, 2010). In *Tomato bushy stunt virus* (TBSV), eEF1a binds a 3' hairpin, RSE, that is a *cis*-acting element involved in replicase assembly and eEF1a is also found in purified TBSV replication complexes (Li et al, 2009). eEF1a is also found in the replication vesicles induced by infection with *Turnip mosaic virus* (TuMV). The RdRp and VPg-protease proteins of TuMV also interact with eEF1a (Thivierge et al, 2008). eEF1a•GTP interacts with the aminoacylated viral genomes of *Turnip yellow mosaic virus* and downregulates negative-strand synthesis, presumably early in infection to allow the competing process of translation to occur (Matsuda et al, 2004).

Since TCV has in its 3' end a t-RNA shaped structure (TSS) that is known to bind ribosomes (Stupina et al, 2008) and that, with additional flanking sequence, binds the

viral RdRp (see Chapter II), I investigated whether it also binds eEF1a. I started with of series of fragments in the 3' region [F1, F2, F3, F4, H4 + Pr, and H4 (Fig. 5.5)]. Binding was best to F1, the TSS. However, because most of the RdRp binding (Chapter II) and ribosome binding (Stupina et al, 2008) was done in F3 or F4, I looked at selected mutations in these fragments to help determine the role the various structures play.

The mutations in F3 either had no effect on binding or resulted in non-specific binding (Fig.5.6, left table). Mutations in the asymmetric loop of H4 (H4AL, m21) or that disrupt Ψ_4 had non-specific binding, while mutations that disrupt (m26, m27) or restore (m26/m27) Ψ_3 had no effect on eEF1a binding to F3. Surprisingly, most mutations in F4 increased eEF1a binding to the fragment (Fig. 5.6, right table). m18, which disrupts Ψ_1, increased binding by more than threefold, while mutations that disrupt Ψ_4 (m19) or Ψ_3 (m26, m27) had more moderate increases to binding, ranging from 2.5 to 1.5, respectively. m21 in H4AL reduced eEF1a binding to F4 by almost half. Combining m21 in H4AL with m19, which disrupts Ψ_4, only very slightly improved eEF1a binding. While the individual mutations that disrupt Ψ_3 improved binding, combining them to restore the pseudoknot reduced binding to near F4 levels.

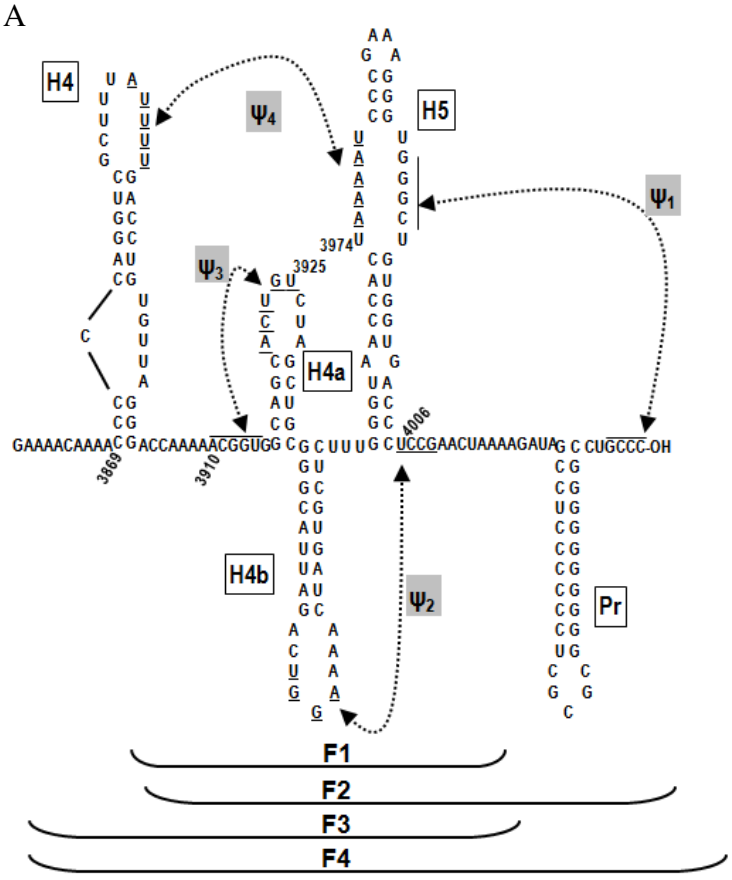
Figure 5.5 eEF1a binding to 3' end fragments of TCV

A. Diagram of part of the 3' UTR of TCV that indicates what sequence is contained in each fragment.

B. The binding of eEF1a to the various fragments.

C. GraphPad non-linear regressions with Scatchard plot inset. Where Scatchard is non-linear, binding is indicated as non-specific.

A



B

RNA fragment	Kd, μM	Region of TCV
F 1, TSS	0.11 ± 0.01	3901- 4017
F 2	1.02 ± 0.12	3901 – 4054
F 3	0.33 ± 0.01	3859 – 4017
F 4	0.59 ± 0.01	3859 - 4054
H4 + Pr	non-specific	3870-3903 + 4022-4048
H 4	0.66 ± 0.09	3870 - 3903

C

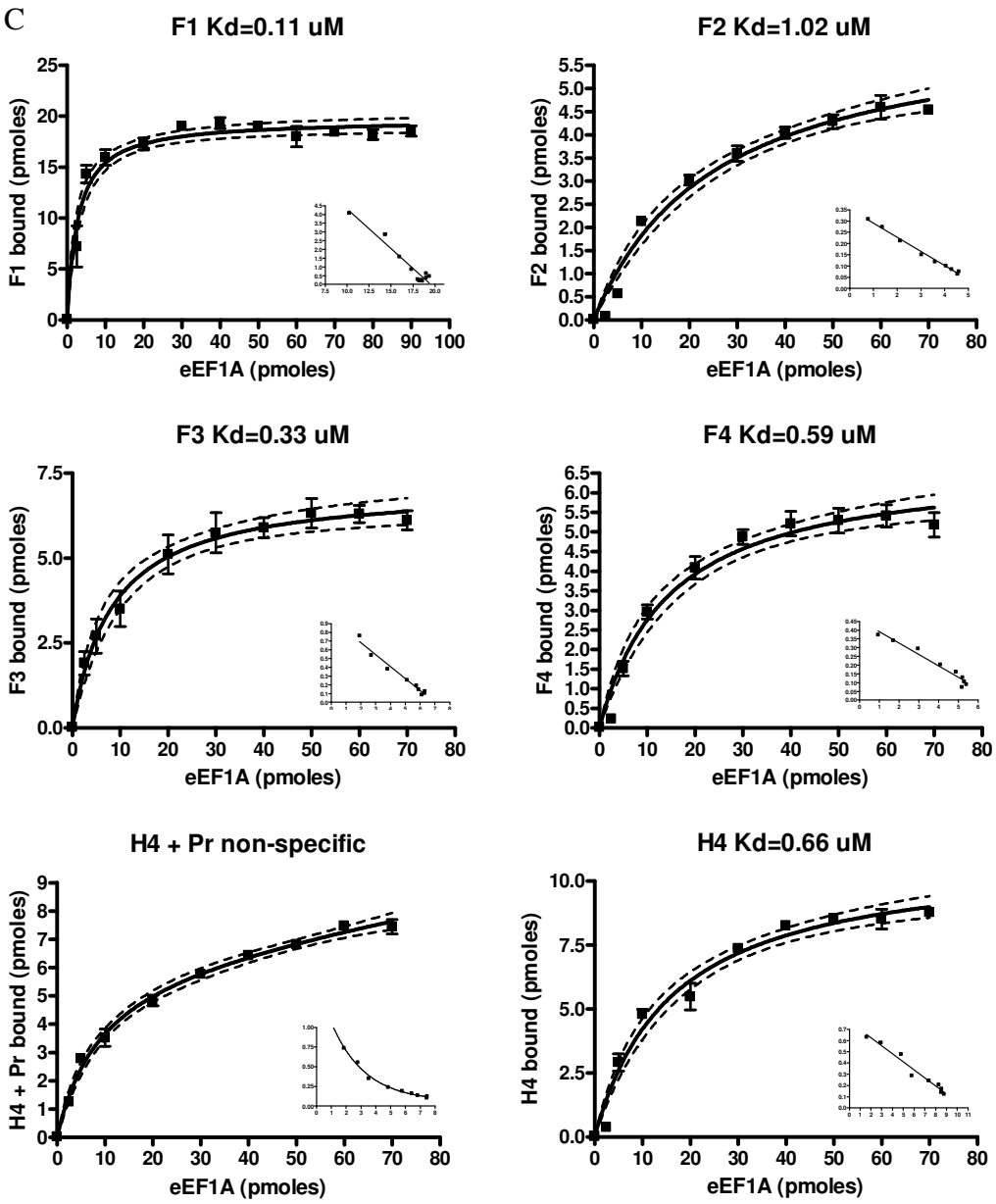

F1 Kd=0.11 uM
F1 bound (pmoles)
eEF1A (pmoles)
F2 Kd=1.02 uM
F2 bound (pmoles)
eEF1A (pmoles)
F3 Kd=0.33 uM
F3 bound (pmoles)
eEF1A (pmoles)
F4 Kd=0.59 uM
F4 bound (pmoles)
eEF1A (pmoles)
H4 + Pr non-specific
H4 + Pr bound (pmoles)
eEF1A (pmoles)
H4 Kd=0.66 uM
H4 bound (pmoles)
eEF1A (pmoles)

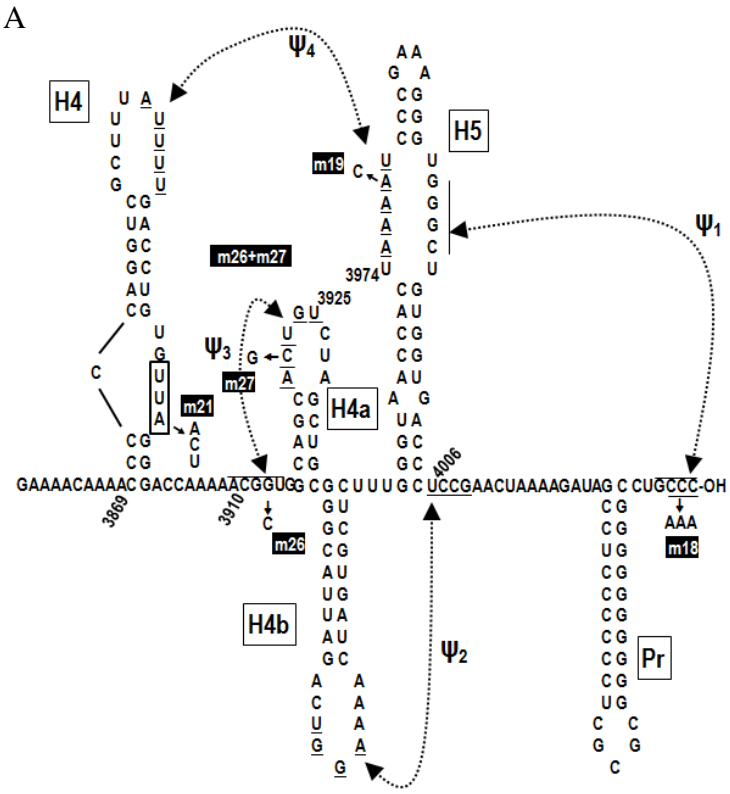
Figure 5.6 eEF1a binding to 3' end fragments of TCV containing mutations

A. Diagram of part of the 3' UTR of TCV that indicates the sequence and location of the mutations in F4.

B. The binding of eEF1a to the various fragments.

C. GraphPad non-linear regressions with Scatchard plot inset. Where Scatchard is non-linear, binding is indicated as non-specific.

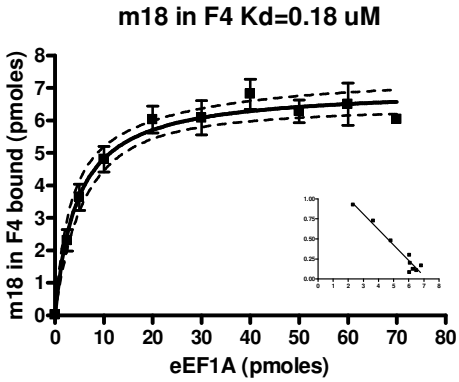
A

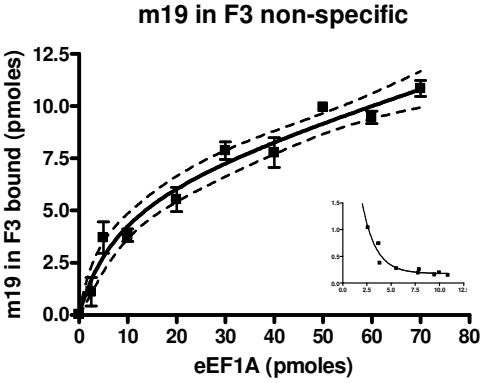


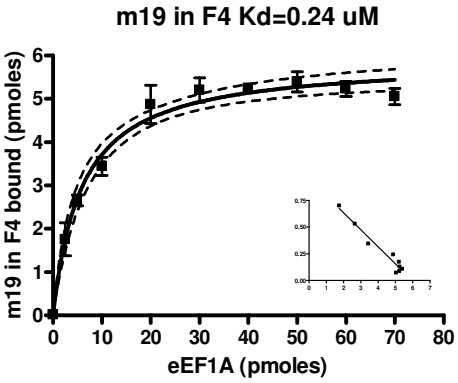
B

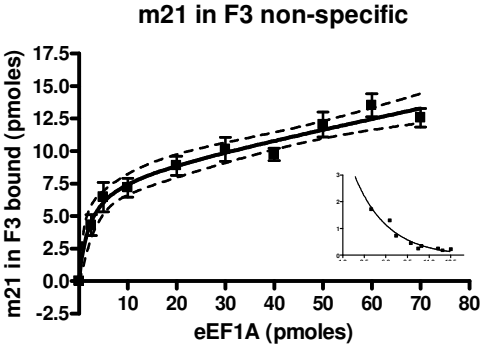
RNA fragment	Kd, μM
F3	0.33 ± 0.01
m19 in F3	non-specific
m21 in F3	non-specific
m19/m21 in F3	non-specific
m27 in F3	0.22 ± 0.04
m26 in F3	0.30 ± 0.05
m26/m27 in F3	0.42 ± 0.04

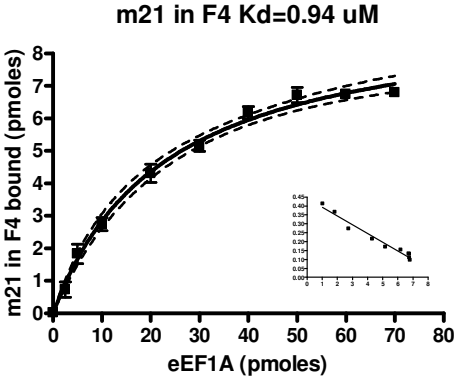
RNA fragment	Kd, μM
F4	0.59 ± 0.01
m18 in F4	0.18 ± 0.03
m19 in F4	0.24 ± 0.03
m21 in F4	0.94 ± 0.01
m19/m21 in F4	0.45 ± 0.05
m27 in F4	0.23 ± 0.02
m26 in F4	0.40 ± 0.06
m26/m27 in F4	0.65 ± 0.01

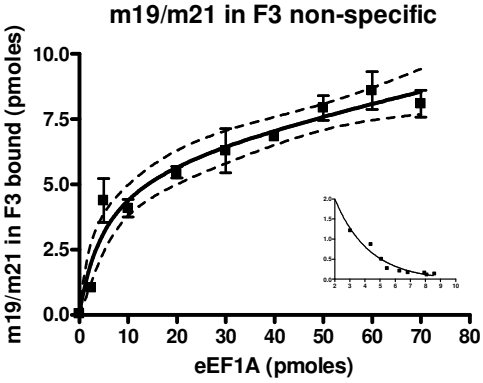

m18 in F4 Kd=0.18 uM
m18 in F4 bound (pmoles)
eEF1A (pmoles)

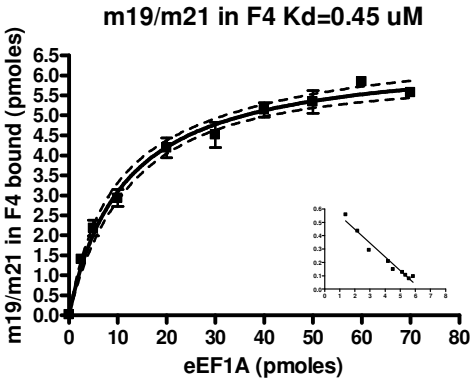

m19 in F3 non-specific
m19 in F3 bound (pmoles)
eEF1A (pmoles)

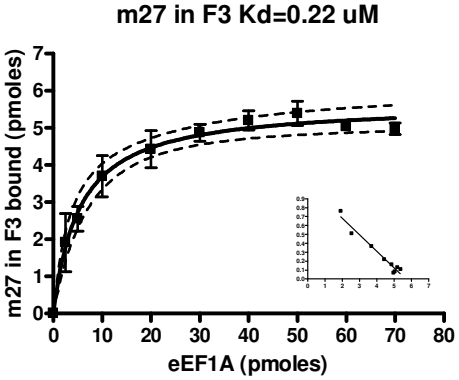

m19 in F4 Kd=0.24 uM
m19 in F4 bound (pmoles)
eEF1A (pmoles)

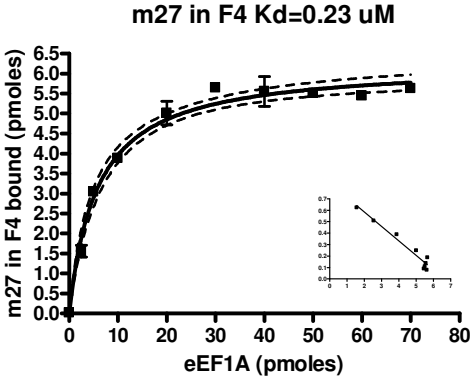

m21 in F3 non-specific
m21 in F3 bound (pmoles)
eEF1A (pmoles)

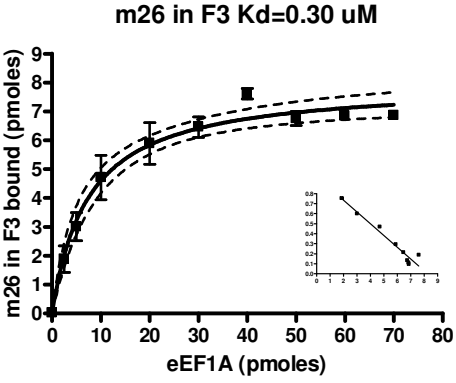

m21 in F4 Kd=0.94 uM
m21 in F4 bound (pmoles)
eEF1A (pmoles)

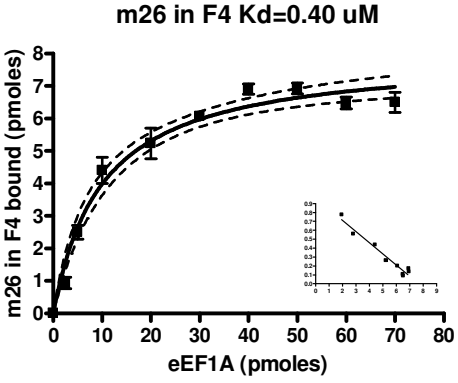

m19/m21 in F3 non-specific
m19/m21 in F3 bound (pmoles)
eEF1A (pmoles)

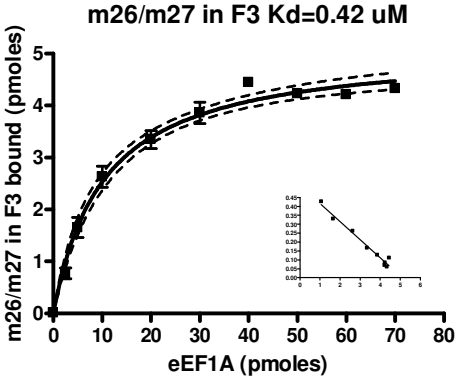

m19/m21 in F4 Kd=0.45 uM
m19/m21 in F4 bound (pmoles)
eEF1A (pmoles)

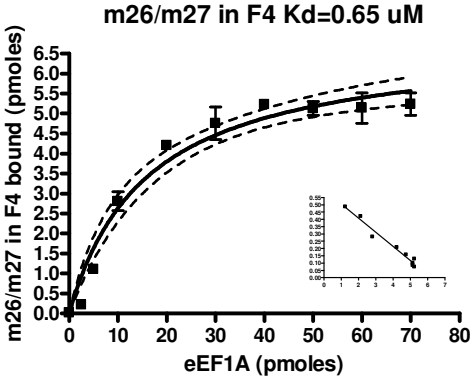

m27 in F3 Kd=0.22 uM
m27 in F3 bound (pmoles)
eEF1A (pmoles)


m27 in F4 Kd=0.23 uM
m27 in F4 bound (pmoles)
eEF1A (pmoles)


m26 in F3 Kd=0.30 uM
m26 in F3 bound (pmoles)
eEF1A (pmoles)


m26 in F4 Kd=0.40 uM
m26 in F4 bound (pmoles)
eEF1A (pmoles)


m26/m27 in F3 Kd=0.42 uM
m26/m27 in F3 bound (pmoles)
eEF1A (pmoles)


m26/m27 in F4 Kd=0.65 uM
m26/m27 in F4 bound (pmoles)
eEF1A (pmoles)

Conclusions

These studies revealed that RdRp binding to the 3' UTR of TCV is very sensitive to perturbations of sequence or structure. Furthermore, poor or increased binding does not necessarily correlate to reduced or improved *in vitro* transcription or accumulation. Many mutations resulted in the RdRp binding non-specifically to the RNA fragment, as it does to a non-template RNA (MDV). This non-specificity was not rectified by the addition of monovalent cations to the buffer and cannot be attributed to different global folding of the RNAs. Given how interconnected and cooperative the structure is (Yuan et al, 2009; Yuan et al, 2010; Yuan et al, 2012), it is not surprising that altering the structure through mutation or deletion reduces the specificity of the RdRp•RNA interaction.

While specific RdRp binding to 3' end fragments was often eliminated with mutations, mutations were generally favorable for eEF1a binding. This is not unexpected as RdRp must have specificity for its cognate RNA and not host mRNAs, while eEF1a is a host protein that is likely better able to bind a variety of structures, given its presence in replicase complexes in other plant viruses (Li et al, 2009; Thivierge et al, 2008). Since eEF1a is known in some instances to downregulate minus-strand synthesis (Matsuda et al, 2004), it would be interesting to determine whether eEF1a and RdRp binding to the 3' end of TCV are mutually exclusive functions or whether they are compatible. Mutations that abrogated specific RdRp binding were not detrimental to eEF1a and it would be interesting to determine if binding of eEF1a to these mutant fragments would then allow specific binding of the RdRp.

Second-site mutations located in the CP ORF revealed that the interactivity of the 3' UTR extends even farther upstream than we had considered. Of particular interest is

G3561A, which enhances accumulation in protoplasts. Additional work is being done to ascertain why this is so. While selected second-site mutations did not reduce translation of our reporter construct, it did reveal that translation could be further enhanced by the addition of increased viral sequence. The increased translational enhancement was mapped to just twenty additional bases. Structural analysis suggested the enhanced translation was due to increased stability of H3 in the CP ORF, and mutational analysis confirmed the importance of H3 for both translation and accumulation.

A primary site mutation located in the 5' flanking adenylates of H4 results in structural differences in the Pr loop. Second-site mutations located upstream in the CP ORF rectify this Pr loop structural change (Yuan et al, 2012). How these regions interact and the precise role that H3 plays is still unclear, but this clearly demonstrates how the genomic RNA of TCV is highly structured for proper function.

Given the interactivity displayed by the TCV genome, one interaction that has not been found is a direct RNA:RNA interaction between the 5' and 3' ends. I tested several possible interacting sequences between the 5' and 3' ends in the luciferase construct but did not find that any of them appeared to be interacting. Whether the genome of TCV circularizes and, if so, what mediates it is still an open question.

Figure 5.7 outlines the lifecycle of TCV. It begins with the virus being inoculated into a plant cell (1). In the lab, this is achieved through manual or mechanical means, while in nature the vector is the flea beetle. The RNA itself is infectious, though virions can also initiate infection. If a virion enters a cell, it is then ucoated (2), likely due to environmental differences between the interior and exterior of a cell. The genomic RNA serves as a template for translation (3) and generates viral proteins p28 (replicase-

associated protein) and p88 (RdRp). When sufficient levels of RdRp have been translated, it then can bind the genomic RNA resulting in conformational alterations (4) to the RNA that prevent it from being accessed by ribosomes for further translation. The RdRp is then able to transcribe (-)-strands (5) from the genomic strand, which serve as templates for (+)-strand synthesis and sgRNA (+)-strand synthesis (6). It is believed that nascent (+)-strands adopt a conformation that does not make them templates for further (-)-strand synthesis, as there is asymmetric production of (+)-strands relative to (-)-strands (7). The movement proteins and coat protein are translated from the subgenomic RNAs (8), which help the virus to move intercellularly and encapsidate new viral genomes (9, 10), respectively.

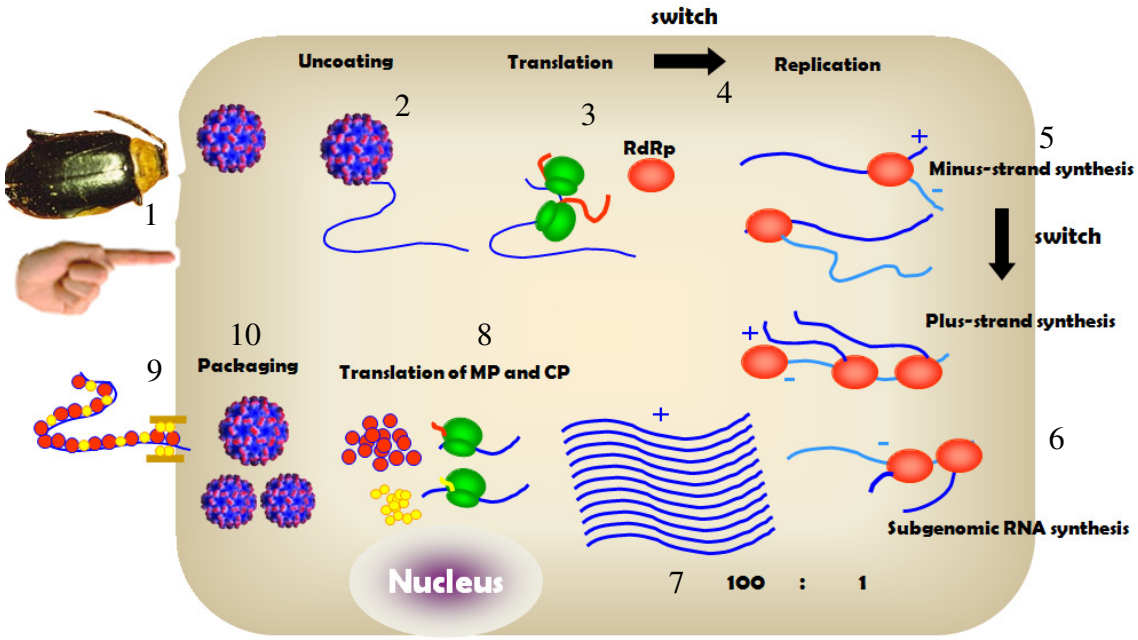


Figure 5.7 Overview of the lifecycle of TCV

See text for details.

The RdRp binding work I performed (Chapter II) revealed the sensitivity of RdRp binding, given that many mutations resulted in poorer or non-specific binding. This pertains to (4) in Figure 5.7, the switch between translation and replication. Binding of RdRp to the 3' end of TCV is known to substantially alter the structure of the region (Yuan et al, 2009; Yuan et al, 2010; Yuan et al, 2012) and my work reveals that the binding is often altered in the presence of mutations, which could affect the ability of the RNA to undergo the conformational switch required to transition to transcription from translation.

My work in Chapter III details the importance of H3 in translation (3 in Fig. 5.7). H3 is also known to be the packaging signal (Qu and Morris, 1997). Packaging concerns were eliminated from my studies as I either used reporter constructs, (luciferase constructs for translation) or the UGA construct that does not produce CP. The exact role H3 plays in translation is still unknown, as well as its overall role in the virus lifecycle. When performing accumulation assays in protoplasts, we do not routinely assay for accumulation of (-)-strands, so mutations could result in a defect in (-)-strand synthesis that goes unobserved. Additionally, protoplasts are individual plant cells, so a defect in intercellular mobility would not be observed.

REFERENCES

Alvarez D, De Lella Ezcurra A, Fucito S, Gamarnik A (2005a) Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. *Virology* **339:** 200-212

Alvarez D, Lodeiro M, Ludueña S, Pietrasanta L, Gamarnik A (2005b) Long-range RNA-RNA interactions circularize the dengue virus genome. *J Virol* **79:** 6631-6643

Balvay L, Soto Rifo R, Ricci EP, Decimo D, Ohlmann T (2009) Structural and functional diversity of viral IRESes. *Biochim Biophys Acta* **1789:** 542-557

Barends S, Bink H, van den Worm S, Pleij C, Kraal B (2003) Entrapping ribosomes for viral translation: tRNA mimicry as a molecular Trojan horse. *Cell* **112:** 123-129

Barends S, Rudinger-Thirion J, Florentz C, Giegé R, Pleij C, Kraal B (2004) tRNA-like structure regulates translation of Brome mosaic virus RNA. *J Virol* **78:** 4003-4010

Barton D, Morasco B, Flanegan J (1999) Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. *J Virol* **73:** 10104-10112

Berry KE, Waghray S, Mortimer SA, Bai Y, Doudna JA (2011) Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning. *Structure* **19:** 1456-1466

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

Bujarski JJ, Dreher TW, Hall TC (1985) Deletions in the 3'-terminal tRNA-like structure of brome mosaic virus RNA differentially affect aminoacylation and replication in vitro. *Proc Natl Acad Sci U S A* **82:** 5636-5640

Carpenter C, Cascone P, Simon A (1991) Formation of multimers of linear satellite RNAs. *Virology* **183:** 586-594

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

Chapman MR, Kao CC (1999) A minimal RNA promoter for minus-strand RNA synthesis by the brome mosaic virus polymerase complex. *J Mol Biol* **286:** 709-720

Chappell SA, Edelman GM, Mauro VP (2000) A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proc Natl Acad Sci U S A* **97:** 1536-1541

Chattopadhyay M, Shi K, Yuan X, Simon AE (2011) Long-distance kissing loop interactions between a 3' proximal Y-shaped structure and apical loops of 5' hairpins enhance translation of Saguaro cactus virus. *Virology* **417:** 113-125

Chernysheva OA, White KA (2005) Modular arrangement of viral cis-acting RNA domains in a tombusvirus satellite RNA. *Virology* **332:** 640-649

Cimino PA, Nicholson BL, Wu B, Xu W, White KA (2011) Multifaceted regulation of translational readthrough by RNA replication elements in a tombusvirus. *PLoS Pathog* **7:** e1002423

Costantino D, Kieft JS (2005) A preformed compact ribosome-binding domain in the cricket paralysis-like virus IRES RNAs. *RNA* **11:** 332-343

Costantino DA, Pfingsten JS, Rambo RP, Kieft JS (2008) tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat Struct Mol Biol* **15:** 57-64

Danthinne X, Seurinck J, Meulewaeter F, Van Montagu M, Cornelissen M (1993) The 3' untranslated region of satellite tobacco necrosis virus RNA stimulates translation in vitro. *Mol Cell Biol* **13:** 3340-3349

de Smit MH, Gultyaev AP, Hilge M, Bink HH, Barends S, Kraal B, Pleij CW (2002) Structural variation and functional importance of a D-loop-T-loop interaction in valine-accepting tRNA-like structures of plant viral RNAs. *Nucleic Acids Res* **30:** 4232-4240

Deiman BA, Koenen AK, Verlaan PW, Pleij CW (1998) Minimal template requirements for initiation of minus-strand synthesis in vitro by the RNA-dependent RNA polymerase of turnip yellow mosaic virus. *J Virol* **72:** 3965-3972

Deniz N, Lenarcic EM, Landry DM, Thompson SR (2009) Translation initiation factors are not required for Dicistroviridae IRES function in vivo. *RNA* **15:** 932-946

Draper DE (2004) A guide to ions and RNA structure. *RNA* **10:** 335-343

Dreher T (1999) FUNCTIONS OF THE 3'-UNTRANSLATED REGIONS OF POSITIVE STRAND RNA VIRAL GENOMES. *Annu Rev Phytopathol* **37:** 151-174

Dreher TW (2010) Viral tRNAs and tRNA-like structures. *Wiley Interdiscip Rev RNA* **1:** 402-414

Dreher TW, Bujarski JJ, Hall TC (1984) Mutant viral RNAs synthesized in vitro show altered aminoacylation and replicase template activities. *Nature* **311:** 171-175

Dreher TW, Goodwin JB (1998) Transfer RNA mimicry among tymoviral genomic RNAs ranges from highly efficient to vestigial. *Nucleic Acids Res* **26:** 4356-4364

Dreher TW, Hall TC (1988) Mutational analysis of the sequence and structural requirements in brome mosaic virus RNA for minus strand promoter activity. *J Mol Biol* **201:** 31-40

Edgil D, Harris E (2006) End-to-end communication in the modulation of translation by mammalian RNA viruses. *Virus Res* **119:** 43-51

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

Fabian M, 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:** 28862-28872

Fabian MR, White KA (2006) Analysis of a 3'-translation enhancer in a tombusvirus: a dynamic model for RNA-RNA interactions of mRNA termini. *RNA* **12:** 1304-1314

Fayzulin R, Frolov I (2004) Changes of the secondary structure of the 5' end of the Sindbis virus genome inhibit virus growth in mosquito cells and lead to accumulation of adaptive mutations. *J Virol* **78:** 4953-4964

Felden B, Florentz C, Giegé R, Westhof E (1996) A central pseudoknotted three-way junction imposes tRNA-like mimicry and the orientation of three 5' upstream pseudoknots in the 3' terminus of tobacco mosaic virus RNA. *RNA* **2:** 201-212

Filomatori C, Lodeiro M, Alvarez D, Samsa M, Pietrasanta L, Gamarnik A (2006) A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes Dev* **20:** 2238-2249

Flint M, Logvinoff C, Rice C, McKeating J (2004) Characterization of infectious retroviral pseudotype particles bearing hepatitis C virus glycoproteins. *J Virol* **78:** 6875-6882

French R, Ahlquist P (1988) Characterization and engineering of sequences controlling in vivo synthesis of brome mosaic virus subgenomic RNA. *J Virol* **62:** 2411-2420

Frolov I, Hardy R, Rice C (2001a) Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *RNA* **7:** 1638-1651

Frolov I, Hardy R, Rice CM (2001b) Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *RNA* **7:** 1638-1651

Fuchs R, Blaas D (2010) Uncoating of human rhinoviruses. *Rev Med Virol* **20:** 281-297

Gallie D (2001) Cap-independent translation conferred by the 5' leader of tobacco etch virus is eukaryotic initiation factor 4G dependent. *J Virol* **75:** 12141-12152

Gallie D, Walbot V (1990) RNA pseudoknot domain of tobacco mosaic virus can functionally substitute for a poly(A) tail in plant and animal cells. *Genes Dev* **4:** 1149-1157

Gallie D, Walbot V (1992) Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. *Nucleic Acids Res* **20:** 4631-4638

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

Gancarz BL, Hao L, He Q, Newton MA, Ahlquist P (2011) Systematic identification of novel, essential host genes affecting bromovirus RNA replication. *PLoS One* **6:** e23988

Gazo B, Murphy P, Gatchel J, 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:** 13584-13592

Guan H, Carpenter C, Simon A (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:** 345-354

Guan H, Carpenter C, Simon A (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:** 355-363

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

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

Guo L, Allen E, Miller W (2000) Structure and function of a cap-independent translation element that functions in either the 3' or the 5' untranslated region. *RNA* **6:** 1808-1820

Guo L, Allen E, Miller W (2001) Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. *Mol Cell* **7:** 1103-1109

Guo R, Lin W, Zhang J, Simon A, Kushner D (2009) Structural plasticity and rapid evolution in a viral RNA revealed by in vivo genetic selection. *J Virol* **83:** 927-939

Haenni AL, Joshi S, Chapeville F (1982) tRNA-like structures in the genomes of RNA viruses. *Prog Nucleic Acid Res Mol Biol* **27:** 85-104

Hardy RW (2006) The role of the 3' terminus of the Sindbis virus genome in minus-strand initiation site selection. *Virology* **345:** 520-531

Hellen CU, Sarnow P (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15:** 1593-1612

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

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

Iwakawa HO, Kaido M, Mise K, Okuno T (2007) cis-Acting core RNA elements required for negative-strand RNA synthesis and cap-independent translation are separated in the 3'-untranslated region of Red clover necrotic mosaic virus RNA1. *Virology* **369:** 168-181

Jackson R (2005) Alternative mechanisms of initiating translation of mammalian mRNAs. *Biochem Soc Trans* **33:** 1231-1241

Jan E, Kinzy TG, Sarnow P (2003) Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. *Proc Natl Acad Sci U S A* **100:** 15410-15415

Jang SK, Pestova TV, Hellen CU, Witherell GW, Wimmer E (1990) Cap-independent translation of picornavirus RNAs: structure and function of the internal ribosomal entry site. *Enzyme* **44:** 292-309

Ji H, Fraser CS, Yu Y, Leary J, Doudna JA (2004) Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc Natl Acad Sci U S A* **101:** 16990-16995

Kaminski A, Howell MT, Jackson RJ (1990) Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism. *EMBO J* **9:** 3753-3759

Kanamori Y, Nakashima N (2001) A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA* **7:** 266-274

Karetnikov A, Keränen M, Lehto K (2006) Role of the RNA2 3' non-translated region of Blackcurrant reversion nepovirus in translational regulation. *Virology* **354:** 178-191

Karetnikov A, Lehto K (2007) The RNA2 5' leader of Blackcurrant reversion virus mediates efficient in vivo translation through an internal ribosomal entry site mechanism. *J Gen Virol* **88:** 286-297

Karetnikov A, Lehto K (2008) Translation mechanisms involving long-distance base pairing interactions between the 5' and 3' non-translated regions and internal ribosomal entry are conserved for both genomic RNAs of Blackcurrant reversion nepovirus. *Virology* **371:** 292-308

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

Kieft JS, Zhou K, Jubin R, Murray MG, Lau JY, Doudna JA (1999) The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J Mol Biol* **292:** 513-529

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

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

Kolupaeva VG, Lomakin IB, Pestova TV, Hellen CU (2003) Eukaryotic initiation factors 4G and 4A mediate conformational changes downstream of the initiation codon of the encephalomyocarditis virus internal ribosomal entry site. *Mol Cell Biol* **23:** 687-698

Kolupaeva VG, Pestova TV, Hellen CU, Shatsky IN (1998) Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *J Biol Chem* **273:** 18599-18604

Kopek B, Perkins G, Miller D, Ellisman M, Ahlquist P (2007) Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biol* **5:** e220

Kozak M (1978) How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* **15:** 1109-1123

Kuhn RJ, Hong Z, Strauss JH (1990) Mutagenesis of the 3' nontranslated region of Sindbis virus RNA. *J Virol* **64:** 1465-1476

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

Lemmetty A, Latvala S, Jones A, Susi P, McGavin W, Lehto K (1997) Purification and properties of a new virus from black currant, its affinities with nepoviruses, and its close association with black currant reversion disease. *Phytopathology* **87:** 404-413

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

Li X, Heaton L, Morris T, Simon A (1989) Turnip crinkle virus defective interfering RNAs intensify viral symptoms and are generated de novo. *Proc Natl Acad Sci U S A* **86:** 9173-9177

Li Z, Nagy PD (2011) Diverse roles of host RNA binding proteins in RNA virus replication. *RNA Biol* **8:** 305-315

Li Z, Pogany J, Panavas T, Xu K, Esposito AM, Kinzy TG, Nagy PD (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:** 245-260

Lin JW, Chiu HN, Chen IH, Chen TC, Hsu YH, Tsai CH (2005) Structural and functional analysis of the cis-acting elements required for plus-strand RNA synthesis of Bamboo mosaic virus. *J Virol* **79:** 9046-9053

Litvak S, Tarrago-Litvak L, Chapeville F (1973a) Turnip Yellow Mosaic Virus RNA as a Substrate of the Transfer RNA Nucleotidyltransferase II. Incorporation of Cytidine 5'-Monophosphate and Determination of a Short Nucleotide Sequence at the 3' End of the RNA. *J Virol* **11:** 238-242

Litvak S, Tarragó A, Tarragó-Litvak L, Allende JE (1973b) Elongation factor-viral genome interaction dependent on the aminoacylation of TYMV and TMV RNAs. *Nat New Biol* **241:** 88-90

Liu F, Kim Y, Cruickshank C, Theimer CA (2012) Thermodynamic characterization of the Saccharomyces cerevisiae telomerase RNA pseudoknot domain in vitro. *RNA* **18:** 973-991

Lomakin IB, Hellen CU, Pestova TV (2000) Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. *Mol Cell Biol* **20:** 6019-6029

Mackenzie J (2005) Wrapping things up about virus RNA replication. *Traffic* **6:** 967-977

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

Matsuda D, Dreher T (2004) The tRNA-like structure of Turnip yellow mosaic virus RNA is a 3'-translational enhancer. *Virology* **321:** 36-46

Matsuda D, Dreher TW (2007) Cap- and initiator tRNA-dependent initiation of TYMV polyprotein synthesis by ribosomes: evaluation of the Trojan horse model for TYMV RNA translation. *RNA* **13:** 129-137

Matsuda D, Yoshinari S, Dreher TW (2004) eEF1A binding to aminoacylated viral RNA represses minus strand synthesis by TYMV RNA-dependent RNA polymerase. *Virology* **321:** 47-56

McCormack J, Yuan X, Yingling Y, Kasprzak W, Zamora R, Shapiro B, Simon A (2008) Structural domains within the 3' untranslated region of Turnip crinkle virus. *J Virol* **82:** 8706-8720

Merrick W (2004) Cap-dependent and cap-independent translation in eukaryotic systems. *Gene* **332:** 1-11

Meulewaeter F, Danthinne X, Van Montagu M, Cornelissen M (1998a) 5'- and 3'-sequences of satellite tobacco necrosis virus RNA promoting translation in tobacco. *Plant J* **14:** 169-176

Meulewaeter F, Van Montagu M, Cornelissen M (1998b) Features of the autonomous function of the translational enhancer domain of satellite tobacco necrosis virus. *RNA* **4:** 1347-1356

Miller S, Krijnse-Locker J (2008) Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol* **6:** 363-374

Miller W, Bujarski J, Dreher T, 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:** 537-546

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

Miller W, White K (2006) Long-distance RNA-RNA interactions in plant virus gene expression and replication. *Annu Rev Phytopathol* **44:** 447-467

Mizumoto H, Tatsuta M, Kaido M, Mise K, Okuno T (2003) Cap-independent translational enhancement by the 3' untranslated region of red clover necrotic mosaic virus RNA1. *J Virol* **77:** 12113-12121

Na H, Fabian M, White K (2006) Conformational organization of the 3' untranslated region in the tomato bushy stunt virus genome. *RNA* **12:** 2199-2210

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

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

Nagy PD (2011) The roles of host factors in tombusvirus RNA recombination. *Adv Virus Res* **81:** 63-84

Nicholson BL, Wu B, Chevtchenko I, White KA (2010) Tombusvirus recruitment of host translational machinery via the 3' UTR. *RNA* **16:** 1402-1419

Niepel M, Gallie D (1999) Identification and characterization of the functional elements within the tobacco etch virus 5' leader required for cap-independent translation. *J Virol* **73:** 9080-9088

Niesters HG, Strauss JH (1990) Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. *J Virol* **64:** 1639-1647

Olsthoorn RC, Bol JF (2002) Role of an essential triloop hairpin and flanking structures in the 3' untranslated region of Alfalfa mosaic virus RNA in in vitro transcription. *J Virol* **76:** 8747-8756

Osman TA, Buck KW (2003) Identification of a region of the tobacco mosaic virus 126- and 183-kilodalton replication proteins which binds specifically to the viral 3'-terminal tRNA-like structure. *J Virol* **77:** 8669-8675

Oster S, Wu B, White K (1998) Uncoupled expression of p33 and p92 permits amplification of tomato bushy stunt virus RNAs. *J Virol* **72:** 5845-5851

Otto GA, Puglisi JD (2004) The pathway of HCV IRES-mediated translation initiation. *Cell* **119:** 369-380

Ou JH, Strauss EG, Strauss JH (1983) The 5'-terminal sequences of the genomic RNAs of several alphaviruses. *J Mol Biol* **168:** 1-15

Panavas T, Nagy PD (2003) The RNA replication enhancer element of tombusviruses contains two interchangeable hairpins that are functional during plus-strand synthesis. *J Virol* **77:** 258-269

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

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

Pestova TV, Hellen CU (2003) Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev* **17:** 181-186

Pestova TV, Hellen CU, Shatsky IN (1996a) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol Cell Biol* **16:** 6859-6869

Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU (1998) A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* **12:** 67-83

Pestova TV, Shatsky IN, Hellen CU (1996b) Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol Cell Biol* **16:** 6870-6878

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

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

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

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

Rajendran K, Pogany J, 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:** 1707-1717

Rakotondrafara A, Polacek C, Harris E, Miller W (2006) Oscillating kissing stem-loop interactions mediate 5' scanning-dependent translation by a viral 3'-cap-independent translation element. *RNA* **12:** 1893-1906

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

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

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

Reynolds JE, Kaminski A, Carroll AR, Clarke BE, Rowlands DJ, Jackson RJ (1996) Internal initiation of translation of hepatitis C virus RNA: the ribosome entry site is at the authentic initiation codon. *RNA* **2:** 867-878

Rietveld K, Van Poelgeest R, Pleij CW, Van Boom JH, Bosch L (1982) The tRNA-like structure at the 3' terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA. *Nucleic Acids Res* **10:** 1929-1946

Rijnbrand RC, Abbink TE, Haasnoot PC, Spaan WJ, Bredenbeek PJ (1996) The influence of AUG codons in the hepatitis C virus 5' nontranslated region on translation and mapping of the translation initiation window. *Virology* **226:** 47-56

Robinson DJ, Ryabov EV, Raj SK, Roberts IM, Taliansky ME (1999) Satellite RNA is essential for encapsidation of groundnut rosette umbravirus RNA by groundnut rosette assistor luteovirus coat protein. *Virology* **254:** 105-114

Scheets K, Redinbaugh MG (2006) Infectious cDNA transcripts of Maize necrotic streak virus: infectivity and translational characteristics. *Virology* **350:** 171-183

Scholthof KB, Jones RW, Jackson AO (1999) Biology and structure of plant satellite viruses activated by icosahedral helper viruses. *Curr Top Microbiol Immunol* **239:** 123-143

Schüler M, Connell SR, Lescoute A, Giesebrecht J, Dabrowski M, Schroeer B, Mielke T, Penczek PA, Westhof E, Spahn CM (2006) Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nat Struct Mol Biol* **13:** 1092-1096

Shen R, Miller W (2004) The 3' untranslated region of tobacco necrosis virus RNA contains a barley yellow dwarf virus-like cap-independent translation element. *J Virol* **78:** 4655-4664

Siegel RW, Adkins S, Kao CC (1997) Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. *Proc Natl Acad Sci U S A* **94:** 11238-11243

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

Simon A, 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:** 3423-3428

Simon A, Roossinck M, Havelda Z (2004) Plant virus satellite and defective interfering RNAs: new paradigms for a new century. *Annu Rev Phytopathol* **42:** 415-437

Simon AE, Gehrke L (2009) RNA conformational changes in the life cycles of RNA viruses, viroids, and virus-associated RNAs. *Biochim Biophys Acta* **1789:** 571-583

Singh RN, Dreher TW (1998) Specific site selection in RNA resulting from a combination of nonspecific secondary structure and -CCR- boxes: initiation of minus strand synthesis by turnip yellow mosaic virus RNA-dependent RNA polymerase. *RNA* **4:** 1083-1095

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

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

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

Spahn CM, Jan E, Mulder A, Grassucci RA, Sarnow P, Frank J (2004) Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* **118:** 465-475

Spriggs KA, Stoneley M, Bushell M, Willis AE (2008) Re-programming of translation following cell stress allows IRES-mediated translation to predominate. *Biol Cell* **100:** 27-38

Stupina V, Meskauskas A, McCormack J, Yingling Y, Shapiro B, Dinman J, Simon A (2008) The 3' proximal translational enhancer of Turnip crinkle virus binds to 60S ribosomal subunits. *RNA* **14:** 2379-2393

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

Sullivan ML, Ahlquist P (1999) A brome mosaic virus intergenic RNA3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA in vivo. *J Virol* **73:** 2622-2632

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

Takamatsu N, Watanabe Y, Meshi T, Okada Y (1990) Mutational analysis of the pseudoknot region in the 3' noncoding region of tobacco mosaic virus RNA. *J Virol* **64:** 3686-3693

Takeda A, Tsukuda M, Mizumoto H, Okamoto K, Kaido M, Mise K, Okuno T (2005) A plant RNA virus suppresses RNA silencing through viral RNA replication. *EMBO J* **24:** 3147-3157

Thivierge K, Cotton S, Dufresne PJ, Mathieu I, Beauchemin C, Ide C, Fortin MG, Laliberté JF (2008) Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. *Virology* **377:** 216-225

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

Thompson SR, Gulyas KD, Sarnow P (2001) Internal initiation in Saccharomyces cerevisiae mediated by an initiator tRNA/eIF2-independent internal ribosome entry site element. *Proc Natl Acad Sci U S A* **98:** 12972-12977

Thurner C, Witwer C, Hofacker IL, Stadler PF (2004) Conserved RNA secondary structures in Flaviviridae genomes. *J Gen Virol* **85:** 1113-1124

Treder K, Kneller E, Allen E, Wang Z, Browning K, Miller W (2008) The 3' cap-independent translation element of Barley yellow dwarf virus binds eIF4F via the eIF4G subunit to initiate translation. *RNA* **14:** 134-147

Tsukiyama-Kohara K, Iizuka N, Kohara M, Nomoto A (1992) Internal ribosome entry site within hepatitis C virus RNA. *J Virol* **66:** 1476-1483

Turner RL, Buck KW (1999) Mutational analysis of cis-acting sequences in the 3'- and 5'-untranslated regions of RNA2 of red clover necrotic mosaic virus. *Virology* **253:** 115-124

van Belkum A, Abrahams JP, Pleij CW, Bosch L (1985) Five pseudoknots are present at the 204 nucleotides long 3' noncoding region of tobacco mosaic virus RNA. *Nucleic Acids Res* **13:** 7673-7686

van Rossum CM, Brederode FT, Neeleman L, Bol JF (1997a) Functional equivalence of common and unique sequences in the 3' untranslated regions of alfalfa mosaic virus RNAs 1, 2, and 3. *J Virol* **71:** 3811-3816

van Rossum CM, Reusken CB, Brederode FT, Bol JF (1997b) The 3' untranslated region of alfalfa mosaic virus RNA3 contains a core promoter for minus-strand RNA synthesis and an enhancer element. *J Gen Virol* **78 (Pt 11):** 3045-3049

Villordo S, Gamarnik A (2009) Genome cyclization as strategy for flavivirus RNA replication. *Virus Res* **139:** 230-239

Vlot AC, Bol JF (2003) The 5' untranslated region of alfalfa mosaic virus RNA 1 is involved in negative-strand RNA synthesis. *J Virol* **77:** 11284-11289

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

Wang S, Browning K, Miller W (1997) A viral sequence in the 3'-untranslated region mimics a 5' cap in facilitating translation of uncapped mRNA. *EMBO J* **16:** 4107-4116

Wang Z, Treder K, Miller WA (2009) Structure of a viral cap-independent translation element that functions via high affinity binding to the eIF4E subunit of eIF4F. *J Biol Chem* **284:** 14189-14202

White K, Morris T (1999) Defective and defective interfering RNAs of monopartite plus-strand RNA plant viruses. *Curr Top Microbiol Immunol* **239:** 1-17

White KA (2002) The premature termination model: a possible third mechanism for subgenomic mRNA transcription in (+)-strand RNA viruses. *Virology* **304:** 147-154

Wilson JE, Pestova TV, Hellen CU, Sarnow P (2000a) Initiation of protein synthesis from the A site of the ribosome. *Cell* **102:** 511-520

Wilson JE, Powell MJ, Hoover SE, Sarnow P (2000b) Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol Cell Biol* **20:** 4990-4999

Winkler WC, Cohen-Chalamish S, Breaker RR (2002) An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A* **99:** 15908-15913

Wu B, Vanti W, White K (2001) An RNA domain within the 5' untranslated region of the tomato bushy stunt virus genome modulates viral RNA replication. *J Mol Biol* **305:** 741-756

Wu B, White K (1999) A primary determinant of cap-independent translation is located in the 3'-proximal region of the tomato bushy stunt virus genome. *J Virol* **73:** 8982-8988

Xu P, Roossinck MJ (2000) Cucumber mosaic virus D satellite RNA-induced programmed cell death in tomato. *Plant Cell* **12:** 1079-1092

Yamaji Y, Kobayashi T, Hamada K, Sakurai K, Yoshii A, Suzuki M, Namba S, Hibi T (2006) In vivo interaction between Tobacco mosaic virus RNA-dependent RNA polymerase and host translation elongation factor 1A. *Virology* **347:** 100-108

Yamaji Y, Sakurai K, Hamada K, Komatsu K, Ozeki J, Yoshida A, Yoshii A, Shimizu T, Namba S, Hibi T (2010) Significance of eukaryotic translation elongation factor 1A in tobacco mosaic virus infection. *Arch Virol* **155:** 263-268

Yoshinari S, Dreher T (2000) Internal and 3' RNA initiation by Qbeta replicase directed by CCA boxes. *Virology* **271:** 363-370

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

Yu Y, Abaeva IS, Marintchev A, Pestova TV, Hellen CU (2011) Common conformational changes induced in type 2 picornavirus IRESs by cognate trans-acting factors. *Nucleic Acids Res* **39:** 4851-4865

Yuan X, Shi K, Meskauskas A, Simon A (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*

Yuan X, Shi K, Simon AE (2012) A local, interactive network of 3' RNA elements supports translation and replication of Turnip crinkle virus. *J Virol* **86:** 4065-4081

Yuan X, Shi K, Young MY, Simon AE (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* **402:** 271-280

Zaccomer B, Haenni AL, Macaya G (1995) The remarkable variety of plant RNA virus genomes. *J Gen Virol* **76 (Pt 2):** 231-247

Zaitlin M (1998) *The Discovery of the Causal Agent of the Tobacco Mosaic Disease*.

Zeenko V, Gallie D (2005) Cap-independent translation of tobacco etch virus is conferred by an RNA pseudoknot in the 5'-leader. *J Biol Chem* **280:** 26813-26824

Zeenko VV, Ryabova LA, Spirin AS, Rothnie HM, Hess D, Browning KS, Hohn T (2002) Eukaryotic elongation factor 1A interacts with the upstream pseudoknot domain in the 3' untranslated region of tobacco mosaic virus RNA. *J Virol* **76:** 5678-5691

Zhang B, Dong H, Stein D, Iversen P, Shi P (2008) West Nile virus genome cyclization and RNA replication require two pairs of long-distance RNA interactions. *Virology* **373:** 1-13

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

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

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

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

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

Zuo X, Wang J, Yu P, Eyler D, Xu H, Starich MR, Tiede DM, Simon AE, Kasprzak W, Schwieters CD, Shapiro BA, Wang YX (2010) Solution structure of the cap-independent translational enhancer and ribosome-binding element in the 3' UTR of turnip crinkle virus. *Proc Natl Acad Sci U S A* **107:** 1385-1390