3' UTRs of Carmoviruses

Anne E. Simon

Department of Cell Biology and Molecular Genetics

University of Maryland College Park

College Park, MD 20742

Phone: 301-405-8975

Fax: 301-314-7930

Email: simona@umd.edu

Running Title: 3' UTRs of Carmoviruses

Keywords: Carmovirus; Turnip crinkle virus; RNA structure/function; cap-independent translation; satellite RNAs

Abstract

Carmovirus is a genus of small, single-stranded, positive-strand RNA viruses in the *Tombusviridae*. One member of the carmoviruses, *Turnip crinkle virus* (TCV), has been used extensively as a model for examining the structure and function of RNA elements in 3' UTR as well as in other regions of the virus. Using a variety of genetic, biochemical and computational methods, a structure for the TCV 3' UTR has emerged where secondary structures and tertiary interactions combine to adopt higher order 3-D structures including an internal, ribosome-binding tRNA-shaped configuration that functions as a 3' cap-independent translation enhancer (3'CITE). The TCV 3'CITE also serves as a scaffold for non-canonical interactions throughout the 3' UTR and extending into the upstream open reading frame, interactions that are significantly disrupted upon binding by the RNA-dependent RNA polymerase. Long-distance interactions that connect elements in the 3' UTR with both the 5' end and the internal ribosome recoding site suggest that 3' UTR of carmoviruses are intimately involved in multiple functions in the virus life cycle. Although carmoviruses share very similar genome organizations, lengths of 5' and 3' UTRs, and structural features at the 3' end, the similarity rapidly breaks down the further removed from the 3' terminus revealing different 3'CITEs and unique virus-specific structural features. This review summarizes 20 years of work dissecting the structure and function of the 3' UTR of TCV and other carmoviruses. The astonishing structural complexity of the 3' UTRs of these simple carmoviruses provides lessons that are likely applicable to many other plant and animal RNA viruses.

Introduction

The carmovirus genus contains some of the simplest single-stranded, positive-sense RNA viruses (Simon, 2010). Carmoviruses have short 5' UTRs (24 nt to 134 nt), five open reading frames (ORFs), and 3' UTRs of 224 nt to 406 nt (Table 1). The five ORFs code for (5' to 3'): a replicase-associated protein; the RNA-dependent RNA polymerase (RdRp), which is synthesized when ribosomes read through the 5' ORF amber termination codon; two small movement-associated proteins, and a single capsid protein (CP) that also serves as the RNA silencing suppressor (Carrington et al., 1989; Hacker et al., 1992; Li et al., 1998; Qu et al., 2003; Riviere and Rochon, 1990; Thomas et al., 2003) (Fig. 1A). The genomic (g)RNA is the messenger RNA for the replication proteins, and two subgenomic (sg)RNAs encode the movement proteins and the CP (Jiwan and White, 2011).

Like many small plant RNA viruses, the gRNA and sgRNAs of carmoviruses are not capped and lack a poly(A) tail. Thus, their translation is dependent on non-canonical mechanisms to attract ribosomes to the 5' end of the gRNA for translation initiation (Firth and Brierley, 2012; Kneller et al., 2006). Translational control of many small RNA plant viruses, including the carmoviruses, rests with structured elements located within the 3' UTRs, and in some cases extending into the CP ORF (Simon and Miller, 2013) known as 3' cap-independent translation enhancers (3'CITEs). 3'CITEs are binding sites for translation initiation factors or ribosomes/ribosomal subunits, which are usually transferred to the 5'end via long-distance RNA:RNA kissing-loop interactions with 5' UTR or coding-region hairpins for use in translation initiation. In addition to 3'CITEs, the 3' UTRs must contain elements that attract the RdRp to the 3' terminus for initiation of replication. Since replication is incompatible with translation, the 3' UTR should also contain elements that provide a means for switching between these two

fundamental viral activities.

Much of our knowledge on the structure and function of carmovirus 3' UTRs comes from work with *Turnip crinkle virus* (TCV) due to the pioneering efforts of Stephen H. Howell and T. Jack Morris who first established TCV as a model system (Altenbach and Howell, 1981; Carrington et al., 1989). Further development of TCV as a model for RNA virus translation and replication was assisted by its fortuitous association with two small, non-coding satellite (sat)RNAs (Altenbach and Howell, 1981; Simon and Howell, 1986). Although the original definition of a satRNA posited that their RNA sequence be substantially or completely different from that of the required helper virus (Murant and Mayo, 1982), satC was unusual in being a chimeric RNA formed from two recombination events between a second TCV non-coding satRNA (satD) and two portions from the 3' region of the TCV gRNA, including the 3' terminal 180 nt (Simon and Howell, 1986) (Fig. 1A). By analyzing sequences and structures present in the region shared by either satC or TCV that are critical for satC accumulation, TCV accumulation, or both, elements were distinguished that either contribute specifically to replication or to translation or were multifunctional. However, it has become clear in the course of these studies that some elements originating from TCV have evolved both structurally and functionally to support the unique accumulation requirements of satC (Guo et al., 2009).

1. General organization of the 3' UTR of carmoviruses

Figure 1B illustrates the secondary and tertiary structures in the 3' UTR of TCV (McCormack et al., 2008). All carmoviruses share the two 3' terminal hairpins (Pr and H5), which are separated by a similar length linker sequence (L1) of unknown function (Fig. 2). Nine of 15 carmoviruses contain a hairpin just upstream of H5 known as H4b. In four of these

viruses, a 4 nt sequence in the terminal loop of H4b is known or predicted to pair with sequence just downstream of H5, forming a pseudoknot known as Ψ_2 (McCormack et al., 2008). Three viruses (TCV, CCFV, and JINRV) have a hairpin directly upstream of H4b termed H4a, and all three viral gRNA are known or predicted to form an H-type pseudoknot (Ψ_3) between the H4a loop and upstream adjacent sequence. As will be discussed below, H4a, H4b, Ψ_2, and Ψ_3 form a higher order 3-D T-shaped structure (TSS) that (for at least TCV) serves as the 3'CITE (McCormack et al., 2008; Stupina et al., 2008) (Fig. 2B). Six or seven carmoviruses are known or predicted to have a Panicum mosaic virus-like translation enhancer known as the PTE (Batten et al., 2006; Wang et al., 2011; Chattopadhyay et al., 2011), whereas one virus (CbMV) has a predicted translation enhancer domain (TED)-like 3'CITE (Gazo et al., 2004; A. E. Simon, unpublished), and MNSV has an I-shaped structure (ISS)-type 3'CITE (Truniger et al., 2008). TCV and CCFV have a common hairpin upstream of the TSS called H4 (Sun and Simon, 2006), which is just downstream from a loosely structured hairpin, M3H (Carpenter et al., 1995). Each of these elements will be discussed in more detail below.

2. Structure and function of carmoviruses 3' UTR elements

The 3'terminal tail. All carmoviruses contain a short 3' terminal sequence just downstream of the Pr hairpin known as the 3' terminal tail. The tail sequence, which ranges from 4 to 9 nt, nearly always terminates with three cytidylates (Table 1). The four 3' terminal residues of the tail are known or predicted to pair with sequence on the 3' side of a large symmetrical loop (LSL) within upstream hairpin H5, and this pseudoknot (Ψ_1) is critical for accumulation of both satC and TCV (Zhang et al., 2004a; Zhang et al., 2006c) (Fig. 3). Similar interactions are also found in most umbraviruses and all tombusviruses. For umbravirus *Pea enation mosaic virus*

(PEMV RNA2), the interaction is between the 3' terminal residues and the terminal loop of the upstream hairpin (these umbraviruses also have a Pr hairpin). In tombusviruses, which do not have an equivalent Pr hairpin, the interaction is between the 3' terminal residues and a large asymmetrical loop located in a hairpin that is a replication repressor (Pogany et al., 2003). Disruption of Ψ_1 in satC causes a substantial enhancement in the transcription of complementary strands, with many transcripts having aberrant sizes (Zhang et al., 2004a). This suggests that Ψ_1 is critical for both controlling the levels of minus-strand synthesis and for correct 3' end recognition by the RdRp. The importance of the 3' tail was also demonstrated by the existence of three different mechanisms to repair deletions of the tail sequences in satC (Carpenter and Simon, 1996; Guan and Simon, 2000; Nagy et al., 1997). One mechanism made use of short primers complementary to the 3' tail synthesized by the RdRp abortive cycling during transcription initiation of the TCV gRNA. The structural impediment that led to the RdRp abortively terminating transcription was not the stable Pr hairpin just upstream of the tail as the RdRp did not terminate transcription prematurely on a template consisting of satC with the 3' terminal 100 nt of TCV (Nagy et al., 1997).

Pr Hairpin. The sequences and proposed structures of carmovirus Pr hairpins are presented in Fig. 4A. The hairpin was originally described for satC, where it functions as a core promoter (Song and Simon, 1995; Stupina and Simon, 1997; Sun and Simon, 2006). Requirements for specific sequences and structures in the satC Pr were determined using a novel in vivo genetic selection method (in vivo SELEX), whereby the entire hairpin was randomized allowing for helper virus selection of satC with a functional promoter (Carpenter and Simon, 1998). Selected satC had wild-type sequence at the base of the hairpin whereas the loop sequence showed little

sequence specificity. The satC Pr was able to independently support transcription in vitro using purified recombinant TCV RdRp (Song and Simon, 1995). The upper stem may also have some sequence specificity as compensatory mutations did not produce viable satRNA. However, it was not determined if the mutations designed to be compensatory restored the stem structure of the Pr.

The more stable Pr of TCV contains properties that differ significantly from that of its satRNA. TCV Pr cannot promote transcription independent of upstream elements in vitro (Sun and Simon, 2006) and is not by itself a good substrate for the RdRp (Zhang et al., 2006c). In addition, satC cannot use the Pr of TCV as an efficient promoter, and TCV does not accumulate to detectable levels with the Pr of satC (Wang and Simon, 2000; Zhang et al., 2006c). Furthermore, the TCV CP binds to the TCV Pr in vitro but does not bind specifically to the satC Pr (Wang and Simon, 2000). The terminal loop of the TCV Pr was recently found to engage in a phylogenetically conserved long-distance interaction with a bulge loop in a hairpin required for stop-codon readthrough necessary for synthesis of the RdRp (Cimino et al., 2011) (Fig. 4B). This long-distance interaction is conserved throughout the *Tombusviridae* but can differ in the position of the 3' partner sequence. Eleven of 15 carmoviruses connect with the readthrough element using the terminal loop of their Pr hairpins (Fig. 4A, asterisks). In contrast, three carmoviruses use the terminal loop of upstream hairpin H5 for this long-distance interaction (Fig. 5) and one has complementary sequence in the linker region between H5 and Pr (Cimino et al., 2011). The Pr of satC has several positional changes compared with the Pr of TCV and it has been suggested that these changes evolved to allow the hairpin to engage in functions specifically required for satC accumulation (e.g., serving as a more efficient core promoter) since the long-distance interaction is no longer possible or required (Zhang et al., 2006c). Mutations

in the TCV Pr loop reduced transcription of a 3' truncated fragment in vitro and affected RdRp binding to the fragment, suggesting that the TCV Pr is also important for TCV replication. These mutations also enhanced translation of a luciferase reporter construct containing the 3' UTR and 3' 400 nt of TCV (Yuan et al., 2010). Interestingly, even though the TCV Pr and CCFV Pr have similar stem sequences and terminal loops (see Fig. 4A), TCV is unable to accumulate efficiently with the Pr of CCFV (Zhang et al., 2006c).

Hairpin H5. Upstream of the Pr is hairpin H5, which is separated from the Pr by the L1 linker sequence that is unpaired within 3' fragments (Fig. 5). The structure of all carmovirus H5 can be described as: an upper stem (US) of 2 to 9 bp capped mainly by a tetraloop; a large symmetrical internal bulge loop (LSL) with 5 to 9 nt on both sides of the loop; and a lower stem (LS) containing (in most viruses) a small internal bulge loop. Six carmoviruses have identical sequences in their H5 LSL, with four adenylates flanked by single uridylates on the 5' side of the loop and the Ψ_1 pairing sequence (5'GGGC) flanked by uridylates on the 3' side. The H5 LSL in the remaining carmoviruses are variations on this theme with 3' side adenylates/uridylates and 5' side Ψ_1 pairing sequences.

The structure and function of H5 has mainly been studied using satC. SatC was unable to accumulate with the similar H5 of CCFV and JINRV, but accumulated more efficiently with the slightly more stable H5 of TCV. Substituting individual H5 regions with those of CCFV H5 combined with in vivo genetic selection of the various H5 regions revealed that satC did not have strong sequence requirements in the US, whereas the sequence in the LSL and LS were critical (Zhang and Simon, 2005; Zhang et al., 2004b). For TCV, site specific mutations revealed similar requirements for the LSL and LS including a requirement for the symmetrical lower

interior bulge loop (McCormack and Simon, 2004). Mutations in the LSL unexpectedly resulted in the appearance of (apparently) randomly scattered mutations throughout the 3' UTR. This led to the suggestion that H5 might be important for a replicase maturation step, as had been previously proposed for the similar hairpin in the tombusviruses (Panaviene et al., 2005). The H5-equivalent hairpin in the tombusviruses also has the partner sequence of the 3' terminal bases (Fig. 3). However, the hairpin differs from carmovirus H5 by having a large asymmetrical bulge loop in the position of the LSL with the absence of the 5' positioned consecutive adenylates (Pogany et al., 2003).

H4a, H4b, and Ψ_3. H5, Pr and Ψ_1 are the only elements that are ubiquitous in the 15 carmoviruses. In 9 carmoviruses, a hairpin (H4b) has either been shown or predicted to be located just upstream of H5 (McCormack et al., 2008) (Fig. 2; Table 1). The loop of TCV H4b forms a pseudoknot (Ψ_2) with sequence just downstream from H5 and within the L1 linker (Fig. 1B), and a pseudoknot in this position is also conserved in three additional carmoviruses (McCormack et al., 2008). This pseudoknot is also required for satC accumulation, suggesting (at least) a role in replication (Guo et al., 2009; Zhang et al., 2006a; Zhang et al., 2006b). In satC, H4b and adjacent upstream hairpin H4a function as a single unit. This was demonstrated by the efficient accumulation of satC containing a single hairpin replacing H4a and H4b that evolved following in vivo selection of satC containing random sequences in place of H4a and H4b (Guo et al., 2009). In addition, satC accumulated more effectively when H4a and H4b were together replaced with the hairpins from CCFV, but not when individual hairpins were replaced (Zhang et al., 2006b). In both satC and TCV, the loop of H4a engages in an H-type pseudoknot (Ψ_3) with adjacent upstream sequence (Guo et al., 2009; McCormack et al., 2008). In satC, Ψ_3 is

not required for accumulation in protoplasts, but satC with sequence capable of forming the pseudoknot are more fit in direct competition experiments in whole plants. Although Ψ_3 is not required for replication of satC, the sequence upstream of H4a that participates in Ψ_3 is critical (Sun et al., 2005). This short GC-rich sequence (known as "DR") has been proposed to mediate a conformational shift in satC from a pre-active, basal structure, to an actively replicated structure (Zhang et al., 2006a; Zhang et al., 2006b). In TCV and CCFV (and possibly JINRV), Ψ_3 is a required component of the T-shaped, 3'CITE formed from Ψ_3, Ψ_2, H4a, H4b, and H5 (McCormack et al., 2008; Stupina et al., 2008) (Fig. 2B; see below). Similar to satC, substitution of hairpins and pseudoknot sequences with those of CCFV demonstrated that efficient TCV accumulation was dependent on a combination of H4b, H4a and Ψ_3, or all components of the TSS, originating from CCFV (McCormack et al., 2008). In contrast, TCV was unable to use any single or combination of JINRV hairpins.

Linker sequence L3. The A-rich L3 sequence just upstream of Ψ_3 in TCV (and CCFV) is critical for virus accumulation. Mutations throughout the five adenylate residues in TCV strongly reduced both transcription in vitro and translation of luciferase reporter constructs in vivo (Yuan et al., 2009). The identical L3 sequence is also present in satC and in vivo SELEX of the sequence revealed that substitution with other sequences could still produce viable satC, suggesting that the consecutive adenylates were not critical for satRNA replication (Sun et al., 2005). Mutations in TCV L3 caused structural changes in Ψ_3, and vice-versa, suggesting that the consecutive adenylates might be physically associated with Ψ_3 (Yuan et al., 2009). Disruption of Ψ_3 enhanced transcription in vitro by the RdRp (Yuan et al., 2009) and reduced translation of a reporter construct in vivo (Stupina et al., 2008), suggesting that in TCV, the Ψ_3 element within

226 the TSS is important for translation.

227

228 *Hairpin H4.* Just upstream of L3 is hairpin H4, which is nearly identical in sequence and
229 structure in TCV and CCFV. SatC also has a hairpin with a similar secondary structure in the
230 same location that evolved from three sequences brought together by two RNA recombination
231 events (Cascone et al., 1990; Simon and Howell, 1986). Much of the early analysis of H4 in
232 TCV involved addition of the hairpin to poorly replicating templates, and assaying for
233 transcription and RdRp binding in vitro (Sun and Simon, 2006). Unfortunately, these studies
234 were conducted before the knowledge that the flanking adenylates (included with the hairpin
235 additions) were also critical for TCV accumulation. Thus, many of the early results cannot
236 distinguish between effects of adding the hairpin and adding the flanking sequences. Assaying
237 TCV gRNA containing mutations in H4 revealed critical individual residues in the terminal and
238 internal loops that reduced viral RNA to near background levels (Yuan et al., 2012). Curiously,
239 some of these mutations reduced in vitro transcription to background levels, and translation of
240 reporter constructs assayed in cell culture were also reduced below the level of detection,
241 whereas other mutations in the same element had a positive effect on transcription while
242 moderately reducing translation Disruption of the upper and lower stems of H4 strongly
243 reduced TCV viral RNA accumulation in protoplasts (Sun and Simon, 2006) but had little effect
244 on in vitro transcription assays. H4 also played a peripheral role in ribosome binding to the TSS.
245 Disruption of a possible pseudoknot between the four consecutive uridylates in the H4 terminal
246 loop and the four consecutive adenylates in the H5 LSL enhanced ribosome binding to the TSS.
247 The in vivo validation of this potential pseudoknot was not possible because of the critical
248 association of the H5 adenylates with replication (Stupina et al., 2008).

M3H/USR. The region upstream of H4 has been termed the "unstructured" region (USR) (Yuan et al., 2012), however a weakly structured hairpin can form within this region in TCV and CCFV (X. Yuan and A. E. Simon, unpublished). Deletion of the weak hairpin eliminates translational enhancement by the 3' UTR (Stupina et al., 2008), and strengthening the stability of the hairpin was also detrimental to virus accumulation (X. Yuan and A. E. Simon, unpublished). M3H also exists in the minus-sense orientation and is a key hot-spot for recombination between TCV and satD (Carpenter et al., 1995).

3. Higher-order structures: Carmovirus 3'CITEs

The TCV TSS. Three hairpins (H4a, H4b and H5) and two pseudoknots (Ψ_2 and Ψ_3) present in TCV and CCFV (and possibly JINRV) were predicted by RNA2D3D and molecular modeling to fold into a T-shaped structure (TSS) (McCormack et al., 2008), which was confirmed by SAXS-NMR (Zuo et al., 2010). The TSS of TCV structurally resembles a canonical tRNA, with H5 mimicking the anticodon stem and H4a/Ψ_3 composing the amino-acceptor stem. The TSS binds to 80S ribosomes and 60 ribosomal subunits with high nanomolar dissociation constants, and competes for binding with an acylated tRNA that specifically binds to the P-site (Stupina et al., 2008). Mutations that reduce ribosome binding had a similar negative effect on translation of a reporter construct in vivo, suggesting that ribosome binding is important for the TSS functioning as a 3'CITE.

Although satC contains the TCV region that harbors the TSS, six positional differences between satC and TCV in this region disrupt the TSS structure (Stupina et al., 2008), and satC is unable to bind to ribosomes (Guo et al., 2011; Stupina et al., 2008). Step-wise conversion of

satC-specific residues to those of TCV resulted in a template that was equally capable of ribosome binding. The most critical difference that affected ribosome binding to satC was a residue in the terminal loop of H4b, upstream of the residues that participate in Ψ_2. Interestingly, satC accumulated to normal levels when ribosome binding was "restored", suggesting that these alterations from the original TCV sequence were for reasons other than eliminating potentially disruptive ribosome binding to a non-translated RNA. Wild-type satC was more fit in direct competition assays with satC restored to the original TCV sequence in the TSS region, but how these changes enhanced fitness is not known.

A key question is how ribosomes bound to the 3' UTR of TCV affect translation initiating at the 5' end of the template. The TCV 5' UTR is 62 nt with no obvious structural features (Stupina et al., 2011). Unlike many other 3'CITEs, no RNA:RNA interaction is detectable between the 3' and 5' ends of TCV. Lack of such an interaction is also suggested by the ability of both full-length TCV gRNA, and reporter constructs with the 3' UTR of TCV, to use the 5' UTR of JINRV as efficiently as the TCV 5' UTR. The two 5' UTRs share little sequence similarity outside of the initial 5' residues that comprise the carmovirus consensus sequence (CCS; Table 1), which in satC is necessary for complementary strand synthesis (Guan et al., 2000). A polyUC sequence just upstream of the TCV p28 start codon was required for high levels of 3' translational enhancement and was also a site of 40S ribosomal subunit binding as revealed by toeprinting and reduction in flexibility of bound residues (Stupina et al., 2011). The flexibility of residues in the TCV gRNA 5' UTR sequence (especially in the polyUC region) when bound by 80S ribosomes was significantly altered when the 3' UTR was added *in trans*. This led to a ribosome bridging model whereby circularization of the template is mediated by 60S bound to the TSS joining with 40S bound to the polyUC sequence. The absence of a

polyUC sequence in the JINRV 5' UTR suggests that more than one sequence may be capable of attracting 40S subunits. However, the CCFV 5' UTR contains a similar polyUC sequence as the 5' UTR of TCV yet is inefficiently used by TCV gRNA. The TCV gRNA was also not viable with the 5' UTR of *Saguaro cactus virus* (SCV) (Stupina et al., 2011), nor was the TCV 3' region compatible with the SCV 5' UTR in reporter constructs. As noted below, SCV uses a different 3'CITE (a PTE) and also engages in a long-distance RNA-RNA interaction that connects the 3' and 5' ends. Addition of out-of-frame initiation codons to the TCV 5' UTR reduced translation significantly, suggesting that a scanning step is likely needed (Stupina et al., 2011). However, whether the scanning requirement is for the initial, pioneer round of translation initiation or initiation that occurs during subsequent steady-state translation is not known.

The SCV PTE. The first characterized PTE was from the 3' UTR of *Panicum mosaic virus* (PMV), where it served as a translation enhancer in vitro and contains a putative long-distance RNA:RNA interaction with a hairpin in the 5' UTR (Batten et al., 2006). A PTE was also found in the umbravirus PEMV RNA2 where its properties have been well characterized (Wang et al., 2009; Wang et al., 2011). All PTE contain an upper three-way branched structure atop a stem with a large guanylate-rich asymmetric loop that engages in a pseudoknot with sequences between the upper two hairpins (Fig. 6). The PEMV RNA2 PTE binds to eIF4E (the cap-binding protein) and, unlike the PMV PTE, it does not engage in any interaction with the 5' end. Rather, a long-distance interaction is performed by an adjacent, upstream 3'CITE, the ribosome-binding T-shaped element known as the kl-TSS (Gao et al., 2012).

As shown in Fig. 2C, seven carmoviruses have PTE-type 3'CITEs. With the exception of JINRV, all PTE are similar to the PMV PTE in having the 5' side hairpin loop engage in a

known or putative long-distance kissing-loop interaction with a hairpin located either at the 5' terminus or approximately 100 nt from the 5' end, which is usually within the coding sequence (Chattopadhyay et al., 2011). The partner sequences involved in all long-distance interactions between 3'CITEs and the 5' end in carmoviruses are virtually identical (UGCCA/UGGCA) with either sequence located in the PTE (Simon and Miller, 2013). Unlike the PEMV RNA 2 and PMV PTEs, all carmoviruses have either been shown or are predicted to have the PTE resting atop an extended stem just upstream of H4b (or H5 for HCRSV) (Chattopadhyay et al., 2014) (Fig. 2C). Currently, it is not known if the lower stem and the conserved hairpin that extends from the 5' side of the lower stem are required for translation.

The SCV PTE, similar to the PTEs of HnRSV and PFBV, spans the junction between the CP ORF and the 3' UTR (Chattopadhyay et al., 2011) (Fig. 6). The SCV PTE pairing partner is located just within the p26 ORF, as demonstrated by compensatory mutagenesis. Moving the interacting sequence to the terminal loop of the 5' terminal hairpin (the location of the pairing sequence for two of the PTE-containing carmoviruses) reduced translation to 25% of wild-type (Chattopadhyay et al., 2014). Since deletion of the 5'terminal hairpin decreased translation by a similar amount, this suggested that the 5' terminal hairpin may be playing an unknown role in translation that was affected when the pairing sequence was transferred. The location of the pairing partner in the 5' UTR of the smaller of the two sgRNA is strikingly similar for all PTE, between 62 and 65 nt from the 5' end (Chattopadhyay et al., 2011) (Fig. 6). Relocating the sgRNA pairing partner 27 nt closer to the 5' end did not significantly decrease translation, suggesting that the precise location of the pairing partner in the sgRNA was for reasons other than simple translation efficiency (Chattopadhyay et al., 2014). As mentioned above, the 5' UTR of SCV was unable to substitute for the 5' UTR of TCV (Stupina et al., 2011), suggesting

that the SCV 5' UTR sequence/structure evolved to function with a 3'CITE that engages in a long-distance interactions such as the PTE and is missing components needed to attract 40S subunits, a putative requirement for TSS-type 3'CITEs.

The ISS of MNSV. *Melon necrotic spot virus* (MNSV) contains an I-shaped (ISS)-type 3'CITE in its 3' UTR (Truniger et al., 2008). Structural predictions using mFold currently place the ISS atop an extended lower stem just upstream from hairpin H5 (Fig. 2C), similar to the placement of PTEs. The ISS is known to interact with eIF4E and plants resistant to MNSV encode an altered eIF4E. One virulent strain of MNSV acquired an insertion of a new, eIF4E-independent 3'CITE by recombination with a polerovirus (Miras et al., 2014). The 55 nt element adopts a two hairpin structure, and both the ISS and the 55 nt element require the MNSV 5' UTR for efficient translation. There are two possible 5' interacting sites for the ISS terminal loop sequence (C<u>UGGC</u>UA; the conserved sequence also is also found in carmovirus PTEs is underlined): one is located in the 5' UTR at position 10 (UA<u>GCC</u>GG; conserved sequence is underlined), and the second is located just within the 5' ORF at position 125 (UA<u>GCCA</u>), similar to the location of sequences complementary to the terminal loops of 3'CITEs in SCV (position 104), PFBV (position 100), HnRSV (position 104), PSNV (position 107) and CbMV (position 118). The former MNSV sequence would require a G:U pair, not found in any other carmovirus 3'CITE RNA:RNA interaction. Since only constructs containing exact 5' and 3' UTRs were used in analysis of the MNSV ISS, it is currently not known which interacting sequence is used in MNSV gRNA. Function of the 55-nt element as a 3'CITE also requires the MNSV 5' UTR, although no obvious interaction is discernable. This suggests that an alternative mechanism allowing for transfer of translation initiation components from the 55 nt element may be present.

4. Higher order 3' UTR conformations in TCV.

A valid, rarely addressed question is whether discrete RNA elements are interconnected through non-canonical interactions within a region such as a UTR. Binding of the RdRp to various sized TCV 3' gRNA fragments led to a widespread conformational shift that was nearly identical in the overlapping regions of different sized fragments and was reversible when the RdRp levels were reduced (Yuan et al., 2009; Yuan et al., 2012) (Fig. 7C). Major structural changes in the RdRp-bound conformation included a hairpin upstream of the 3' UTR (hairpin H3), the terminal and internal loops of H4, the TSS, all linkers, and the sequence adjacent to the terminal loop of Pr. This widespread conformational shift was suggested to disrupt ribosome binding to the TSS, possibly leading to a transition between a translated template and a transcribed template.

Mutation of different elements followed by in-line RNA structure probing also revealed numerous connections between elements, with H4 and the TSS serving as central scaffold "hubs" (Fig. 7B). Specific alterations in the H4 loops and the A-rich sequence upstream of H4 resulted in identical structural changes in the Pr within the RdRp-bound conformation and significant reductions in TCV accumulation. Second-site mutations that compensated for both the structural changes and reduced TCV accumulation were scattered throughout the 3' UTR and an upstream hairpin, supporting the proposition that this region contains substantial interconnections between elements. Interestingly, compensatory second-site mutations were also located throughout the CP ORF, and the CP (the virus's silencing suppressor) and an active silencing system in the cells is required for the compensatory effects (Chattopadhyay, Stupina, Yuan, and Simon, manuscript in preparation).

5. Conclusions

Carmoviruses, arguably among the simplest of the RNA viruses, have 3' UTRs with a dazzling complexity of elements needed for their cap-independent life-cycle. Disruption of even single nucleotides within hairpins and linker sequences can result in dramatic reductions in virus viability, with structural changes occurring throughout the region and extending into coding sequence. The dynamic, interactive nature of the carmovirus 3' UTR is clearly illustrated by the extensive, reversible conformational changes that occur upon RdRp binding, which likely require a highly interactive higher order structure combining all of the discrete 3' elements into a functional whole. An emerging theme is that virtually all of the secondary structures and linker regions of TCV gRNA have evolved to support replication and translation functions of the virus. These local interactions are just one layer in what is likely a highly compact gRNA structure stabilized by long-distance RNA:RNA interactions (Nicholson and White, 2014; Wu et al., 2013). These interactions must bring the Pr loop to the vicinity of the recoding structure element for pairing that leads to RdRp synthesis and bring the TSS proximal to locations where translation is terminating for ribosome/ribosomal subunit "pick-up", which then must be delivered to the 5' end. Recent work on umbravirus PEMV RNA2, with a 3' UTR over three times the size of carmovirus 3' UTR, has revealed the presence of two ribosome-binding TSS-like structures, one PTE, and numerous additional hairpins that function in replication and/or translation. Thus the lessons learned from analyses of the 3' UTRs of carmoviruses may be applicable to other, possibly many other, positive-strand RNA viruses.

Acknowledgements

I gratefully acknowledge the many graduate students and postdoctoral associates over the years who have contributed to the analysis of TCV and PEMV. I also gratefully thank the National Science Foundation for funding this work from the beginning, when the virus was being described as "an insignificant laboratory oddity". Current work is funded by NSF (MCB 1157906).

References

Altenbach, S.B., Howell, S.H., 1981. Identification of a satellite RNA associated with Turnip crinkle virus. Virology 112(1), 25-33.

Batten, J.S., Desvoyes, B., Yamamura, Y., Scholthof, K.B.G., 2006. A translational enhancer element on the 3'-proximal end of the Panicum mosaic virus genome. FEBS Lett 580, 2591-2597.

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

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

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

433 Carrington, J.C., Heaton, L.A., Zuidema, D., Hillman, B.I., Morris, T.J., 1989. The genome
434 structure of Turnip crinkle virus. Virology 170, 219-226.
435 Cascone, P.J., Carpenter, C.D., Li, X.H., Simon, A.E., 1990. Recombination between satellite
436 RNAs of Turnip crinkle virus. EMBO J 9, 1709-1715.
437 Chattopadhyay, M., Kuhlmann, M.M., Kumar, K., Simon, A.E., 2014. Position of the kissing-
438 loop interaction associated with PTE-type 3'CITEs can affect enhancement of cap-
439 independent translation. Virology 458, 43-52.
440 Chattopadhyay, M., Shi, K., Yuan, X., Simon, A.E., 2011. Long-distance kissing loop
441 interactions between a 3 ' proximal Y-shaped structure and apical loops of 5 ' hairpins
442 enhance translation of Saguaro cactus virus. Virology 417, 113-125.
443 Cimino, P.A., Nicholson, B.L., Wu, B., Xu, W., White, K.A., 2011. Multifaceted regulation of
444 translational readthrough by RNA replication elements in a Tombusvirus. PLoS Path 7),
445 e1002423.
446 Firth, A.E., Brierley, I., 2012. Non-canonical translation in RNA viruses. J Gen Virol 93, 1385–
447 1409.
448 Gao, F., Gulay, S.P., Kasprzak, W., Dinman, J.D., Shapiro, B.A., Simon, A.E., 2013. The
449 Kissing-Loop T-Shaped Structure Translational Enhancer of Pea Enation Mosaic Virus
450 Can Bind Simultaneously to Ribosomes and a 5 ' Proximal Hairpin. J Virol 87, 11987-
451 12002.
452 Gao, F., Kasprzak, W., Stupina, V.A., Shapiro, B.A., Simon, A.E., 2012. A ribosome-binding, 3 '
453 translational enhancer has a T-shaped structure and engages in a long-distance RNA-
454 RNA interaction. J Virol 86, 9828-9842.

Gao, F., Kasprzak, W.K., Szarko, C., Shapiro, B.A., Simon, A.E., 2014. The 3 ' untranslated region of Pea enation mosaic virus contains two T-shaped, ribosome-binding, cap-independent translation enhancers. J Virol 89, 11696-11712.

Gazo, B.M., Murphy, P., Gatchel, J.R., Browning, K.S., 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.C., Carpenter, C.D., Simon, A.E., 2000. 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.C., Simon, A.E., 2000. Polymerization of nontemplate bases before transcription initiation at the 3' ends of templates by an RNA-dependent RNA polymerase: An activity involved in 3' end repair of viral RNAs. Proc Natl Acad Sci USA 97, 12451-12456.

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

Guo, R., Meskauskas, A., Dinman, J.D., Simon, A.E., 2011. Evolution of a helper virus-derived, ribosome binding translational enhancer in an untranslated satellite RNA of Turnip crinkle virus. Virology 419, 10-16.

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

Jiwan, S.D., White, K.A., 2011. Subgenomic mRNA transcription in Tombusviridae. RNA Biol 8, 287-294.

Kneller, E.L.P., Rakotondrafara, A.M., Miller, W.A., 2006. Cap-independent translation of plant viral RNAs. Virus Res 119, 63-75.

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

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

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

Miras, M., Sempere, R.N., Kraft, J.J., Miller, W.A., Aranda, M.A., Truniger, V., 2014. Interfamilial recombination between viruses led to acquisition of a novel translation-enhancing RNA element that allows resistance breaking. New Phytol 202, 233-246.

Murant, A.F., Mayo, M.A., 1982. Satellites of plant viruses. Ann Rev Phytopathol 20, 49-70.

Nagy, P.D., Carpenter, C.D., Simon, A.E., 1997. A novel 3'-end repair mechanism in an RNA virus. Proc Natl Acad Sci USA 94(4), 1113-1118.

Nicholson, B.L., White, K.A., 2014. Functional long-range RNA-RNA interactions in positive-strand RNA viruses. Nat Rev Microbiol 12, 493-504.

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

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

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

500 Riviere, C.J., Rochon, D.M., 1990. Nucleotide sequence and genomic organization of Melon
501 necrotic spot virus. J Gen Virol 71, 1887-1896.
502 Simon, A.E., 2010. Genus Carmovirus (Tombusviridae). The Springer Index of Viruses, edited
503 by Tidona, C.A., C. Darai, Buchen-Osmond, C. Springer, Berlin Heidelberg.
504 Simon, A.E., Howell, S.H., 1986. The virulent satellite RNA of Turnip crinkle virus has a major
505 domain homologous to the 3' end of the helper virus genome. EMBO J 5, 3423-3428.
506 Simon, A.E., Miller, W.A., 2013. 3 ' Cap-independent translation enhancers of plant viruses. In:
507 Gottesman, S. (Ed.), Annual Review of Microbiology, Vol 67. Vol. 67, pp. 21-42.
508 Song, C.Z., Simon, A.E., 1995. Requirement of a 3'-terminal stem-loop in in vitro transcription
509 by an RNA-dependent RNA-polymerase. J Mol Biol 254, 6-14.
510 Stupina, V., Simon, A.E., 1997. Analysis in vivo of turnip crinkle virus satellite RNA C variants
511 with mutations in the 3'-terminal minus-strand promoter. Virology 238, 470-477.
512 Stupina, V.A., Meskauskas, A., McCormack, J.C., Yingling, Y.G., Shapiro, B.A., Dinman, J.D.,
513 Simon, A.E., 2008. The 3' proximal translational enhancer of Turnip crinkle virus binds
514 to 60S ribosomal subunits. RNA 14, 2379-2393.
515 Stupina, V.A., Yuan, X., Meskauskas, A., Dinman, J.D., Simon, A.E., 2011. Ribosome binding
516 to a 5 ' translational enhancer is altered in the presence of the 3 ' untranslated region in
517 cap-independent translation of Turnip crinkle virus. J Viol 85, 4638-4653.
518 Sun, X.P., Simon, A.E., 2006. A cis-replication element functions in both orientations to enhance
519 replication of Turnip crinkle virus. Virology 352, 39-51.
520 Sun, X.P., Zhang, G.H., Simon, A.E., 2005. Short internal sequences involved in replication and
521 virion accumulation in a subviral RNA of Turnip crinkle virus. J Virol 79, 512-524.

522 Thomas, C.L., Leh, V., Lederer, C., Maule, A.J., 2003. Turnip crinkle virus coat protein
523 mediates suppression of RNA silencing in Nicotiana benthamiana. Virology 306, 33-41.
524 Truniger, V., Nieto, C., Gonzalez-Ibeas, D., Aranda, M., 2008. Mechanism of plant eIF4E-
525 mediated resistance against a Carmovirus (Tombusviridae): cap-independent translation
526 of a viral RNA controlled in cis by an (a)virulence determinant. Plant J 56, 716-727.
527 Wang, J.L., Simon, A.E., 2000. 3 '-end stem-loops of the subviral RNAs associated with turnip
528 crinkle virus are involved in symptom modulation and coat protein binding. J Virol 74,
529 6528-6537.
530 Wang, Z., Treder, K., Miller, W.A., 2009. Structure of a viral cap-independent translation
531 element that functions via high affinity binding to the eIF4E subunit of eIF4F. J Biol
532 Chem 284, 14189-14202.
533 Wang, Z., Parisien, M., Scheets, K., Miller, W.A., 2011. The cap-binding translation initiation
534 factor, elF4E, binds a pseudoknot in a viral cap-independent translation element.
535 Structure 19, 868-880.
536 Wu, B., Grigull, J., Ore, M.O., Morin, S., White, K.A., 2013. Global Organization of a Positive-
537 strand RNA Virus Genome. PLoS Path 9(5).
538 Yuan, X., Shi, K., Young, M.Y.L., Simon, A.E., 2010. The terminal loop of a 3 ' proximal
539 hairpin plays a critical role in replication and the structure of the 3 ' region of Turnip
540 crinkle virus. Virology 402, 271-280.
541 Yuan, X.F., Shi, K.R., Meskauskas, A., Simon, A.E., 2009. The 3' end of Turnip crinkle virus
542 contains a highly interactive structure including a translational enhancer that is disrupted
543 by binding to the RNA-dependent RNA polymerase. RNA 15, 1849-1864.

544 Yuan, X.F., Shi, K.R., Simon, A.E., 2012. A Local, Interactive Network of 3 ' RNA Elements
545 Supports Translation and Replication of Turnip Crinkle Virus. J Virol 86(8), 4065-4081.
546 Zhang, G.H., Zhang, J.C., George, A.T., Baumstark, T., Simon, A.E., 2006a. Conformational
547 changes involved in initiation of minus-strand synthesis of a virus-associated RNA. RNA
548 12, 147-162.
549 Zhang, G.H., Zhang, J.C., Simon, A.E., 2004a. Repression and derepression of minus-strand
550 synthesis in a plus-strand RNA virus replicon. J Virol 78, 7619-7633.
551 Zhang, J.C., Simon, A.E., 2005. Importance of sequence and structural elements within a viral
552 replication repressor. Virology 333, 301-315.
553 Zhang, J.C., Stuntz, R.M., Simon, A.E., 2004b. Analysis of a viral replication repressor:
554 sequence requirements for a large symmetrical internal loop. Virology 326, 90-102.
555 Zhang, J.C., Zhang, G.H., Guo, R., Shapiro, B.A., Simon, A.E., 2006b. A pseudoknot in a
556 preactive form of a viral RNA is part of a structural switch activating minus-strand
557 synthesis. J Virol 80, 9181-9191.
558 Zhang, J.C., Zhang, G.H., McCormack, J.C., Simon, A.E., 2006c. Evolution of virus-derived
559 sequences for high-level replication of a subviral RNA. Virology 351, 476-488.
560 Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic
561 Acids Res 31, 3406-3415.
562 Zuo, X., Wang, J., Yu, P., Eyler, D., Xu, H., Starich, M.R., Tiede, D.M., Simon, A.E., Kasprzak,
563 W., Schwieters, C.D., Shapiro, B.A., Wang, Y.-X., 2010. Solution structure of the cap-
564 independent translational enhancer and ribosome-binding element in the 3 ' UTR of
565 turnip crinkle virus. Proc Natl Acad Sci USA 107, 1385-1390.

566

Figure Legends

Fig. 1. Genome organization and 3' UTR structure of carmovirus TCV. (A) TCV gRNA. p8 and p9 are translated from gRNA1 and CP is translated from sgRNA2. The readthrough stop codon is shown at the termination of p28. SatC and satD are two TCV satRNAs. SatC is composed of satD at the 5' end and two regions of TCV at the 3' end including a large portion of the TCV 3' UTR (in black). (B) Secondary structure and tertiary interactions in the 3' UTR of TCV. The 3 pseudoknots are represented by hatched lines. Arrow in hairpin H4 denotes the 5' end point of the region shared with satC. M3H is a weakly structured hairpin and this region is frequently referred to as the "unstructured" region (USR). The TSS is a T-shaped, ribosome-binding structure that functions as a 3'CITE. L1, linker 1; L3, linker 3.

Fig. 2. Comparison of secondary and tertiary structures in the 3' UTRs of carmoviruses. (A) Color-coded representation of the 3' UTR of TCV (see Fig. 1B). Triangle denotes the start of the 3' UTR. (B) Color-coded 3-D structure model of the TCV TSS 3'CITE. (C) Color coded models of the 3' UTRs of remaining carmoviruses using TCV structure as a reference. See legend to Table 1 for the full virus names. With the exception of SCV, structures are putative and are based on phylogenetic comparisons and mFold computational modeling (Zuker, 2003). Pink structures are PTE-like 3'CITEs, green structure is a TED-like 3'CITE and brown structure is an ISS-type 3'CITE. Red terminal loops are conserved sequences either known or predicted to engage in kissing-loop RNA:RNA interactions with 5' proximal hairpins. Although JINRV appears to have elements that might form a TSS and a PTE, it is not known if either putative 3'CITE is functional. The 3'CITEs for AnFBV, SYMMV, CPMoV and NLVCV have not been

identified.

Fig. 3. Comparison of the pseudoknot between the 3' terminal tail and nearby upstream hairpins in carmovirus TCV, tombusvirus *Tomato bushy stunt virus* (TBSV) (Pogany et al., 2003), and PEMV RNA2 (Gao et al., 2014).

Fig. 4. The Pr 3' terminal hairpin. (A) Comparison of carmovirus Pr hairpins. Asterisks denote hairpins whose terminal loops are engaged in a long-distance interaction with the recoding structure element (RSE) located just downstream of the readthrough termination codon, as shown in (B). (B) Interaction between the RSE and the Pr in TCV. Interacting residues are in shaded boxes. The p28 termination codon is boxed.

Fig. 5. The H5 hairpin. (A) Comparison of carmovirus H5 hairpins. US, upper stem; LSL, large symmetrical loop; LS, lower stem. Note the strong sequence similarities in the LSL. Sequences in or adjacent to these hairpins that are complementary to the RSE bulge loop are in shaded boxes.

Fig. 6. The PTE of SCV. Residues involved in the long-distance kissing-loop interaction between the PTE and coding region hairpin near the 5' end of the gRNA, and between the PTE and a hairpin in the sgRNA 5' UTR, are in shaded boxes. Initiation codons are shown.

Fig. 7. Interactions link elements in the 3' UTR of TCV. (A) Relative positions of elements in the 3' region of TCV as shown in Fig. 1B. (B) Network of putative tertiary interactions that

connect elements in the TCV 3' region. Elements are colored as in (A). Circular bands indicate how mutations in the elements affect translation (red) or transcription (blue). Thick and thin bands denote that disruption of elements enhances or reduces translation/transcription, respectively. Broken bands indicate that mutations of different residues within the element can have opposite effects on translation/transcription. Absence of a band indicates that mutations had no significant effect on transcription/translation. Putative tertiary interactions are denoted by arrows. Thick double-headed arrows denote pseudoknots. Black thin lines denote that disruption of one element affects the structure of the element as assayed by in-line structure probing. Bidirectional arrows reflect reciprocal effects (McCormack et al., 2008; Yuan et al., 2010; Yuan et al., 2012). L1, linker 1; L3, linker 3. (C) In-line probing of a 3' UTR fragment of TCV in the absence (In-L) or presence (In-L+RdRp) of purified TCV RdRp. In-line probing can inform on the structure of short RNA fragments by monitoring if residues are flexible (i.e., single-stranded) and can thus assume an orientation (among many) where the 2'OH, the adjacent phosphorus center, and adjacent 5´ oxygen are positioned in a straight line. This configuration results in non-enzymatic nucleophilic attack of the 2'OH on the phosphorus center, followed by backbone cleavage of (in this case) a 5' radiolabeled transcript of the TCV 3'UTR. Absence of a band denotes that the residue is held in position (by base-pairing or base-stacking) and cannot assume the orientation necessary for cleavage. OH is a hydroxide generated ladder and T1 is a partial RNase T1 digest denoting the position of guanylates. This substantial conformational shift in the presence of the RdRp extends into the coding region (Yuan et al., 2012).

Table(s)

Table 1. Properties of Carmoviruses

Name	gRNA length	5'UTR length	5' terminal carmovirus consensus sequence	3'UTR length	3'UTR Components							3'CITE long distance interaction?
					3' terminal tail	H5/Pr/ Ψ_1	H4b	H4a	Ψ_2	Ψ_3	3'CITE	
TCV	4054 nt	63 nt	GGUAAU	253 nt	CUGCCC	Y	Y	Y	Y	Y	TSS	no
CCFV	4041 nt	37 nt	GGUUUU	274 nt	CAGCCC	Y	Y	Y	Y	Y	TSS	no
JINRV	4014 nt	31 nt	GGGUAAA	271 nt	GGCCC	Y	Y	Y	Y	Y	TSS/PTE?	no
SCV	3879 nt	39 nt	GGGUAA	224 nt	CCGCCC	Y	Y	N	N	N	PTE	5'ORF
HnRSV	3956 nt	33 nt	GGGGUUUU	246 nt	GCCC	Y	N	N	N	N	PTE	5'ORF
PFBV	3912 nt	32 nt	GGGAUA	236 nt	AACCCGCCC	Y	Y	N	N	N	PTE	5'ORF
CarMV	4003 nt	69 nt	GGGUAA	288 nt	GCCC	Y	Y	N	N	N	PTE	5'hairpin
HCRSV	3911 nt	30 nt	GGGAAA	284 nt	CAGCCC	Y	N	N	N	N	PTE	5'hairpin
PSNV	4048 nt	134 nt	GGGGAUAAUUAUA	263 nt	GCCCC	Y	N	N	N	N	TED	5'UTR
CbMV	3919 nt	36 nt	CGAUAAA	234 nt	CCGCCCG	Y	Y	N	Y	N	ISS	5'ORF
MNSV	4266 nt	88 nt	GGAUUA	279 nt	CCGCCC	Y	N	N	N	N	ISS	5'ORF
AnFBV	3964 nt	24 nt	GGGUAA	257 nt	GCCAA	Y	N	N	N	N	?	?
SYMMV	4009 nt	49 nt	GGGUAA	281 nt	CCCCCC	Y	N	N	N	N	?	?
CPMoV	4029 nt	34 nt	NGGUAU	252 nt	CCCCCC	Y	N	N	N	N	?	?
NLVCV	4172 nt	59 nt	none	406 nt	GCCC	Y	Y	N	Y	N	?	?

* TCV (*Turnip crinkle virus*); CCFV (*Cardamine chlorotic fleck virus*); JINRV (*Japanese iris necrotic ring virus*); SCV (*Saguaro cactus virus*); HnRSV (*Honeysuckle ringspot virus*); PFBV (*Pelargonium flower break virus*); CarMV (*Carnation mottle virus*); HCRSV (*Hibiscus chlorotic ringspot virus*); PSNV (*Pea stem necrosis virus*); CbMV (*Calibrachoa mottle virus*); MNSV (*Melon necrotic spot virus*); AnFBV (*Angelonia flower break virus*); SYMMV (*Soybean yellow mottle mosaic virus*); CPMoV (*Cowpea mottle virus*); NLVCV (*Nootka lupine vein clearing virus*).

Figure(s)

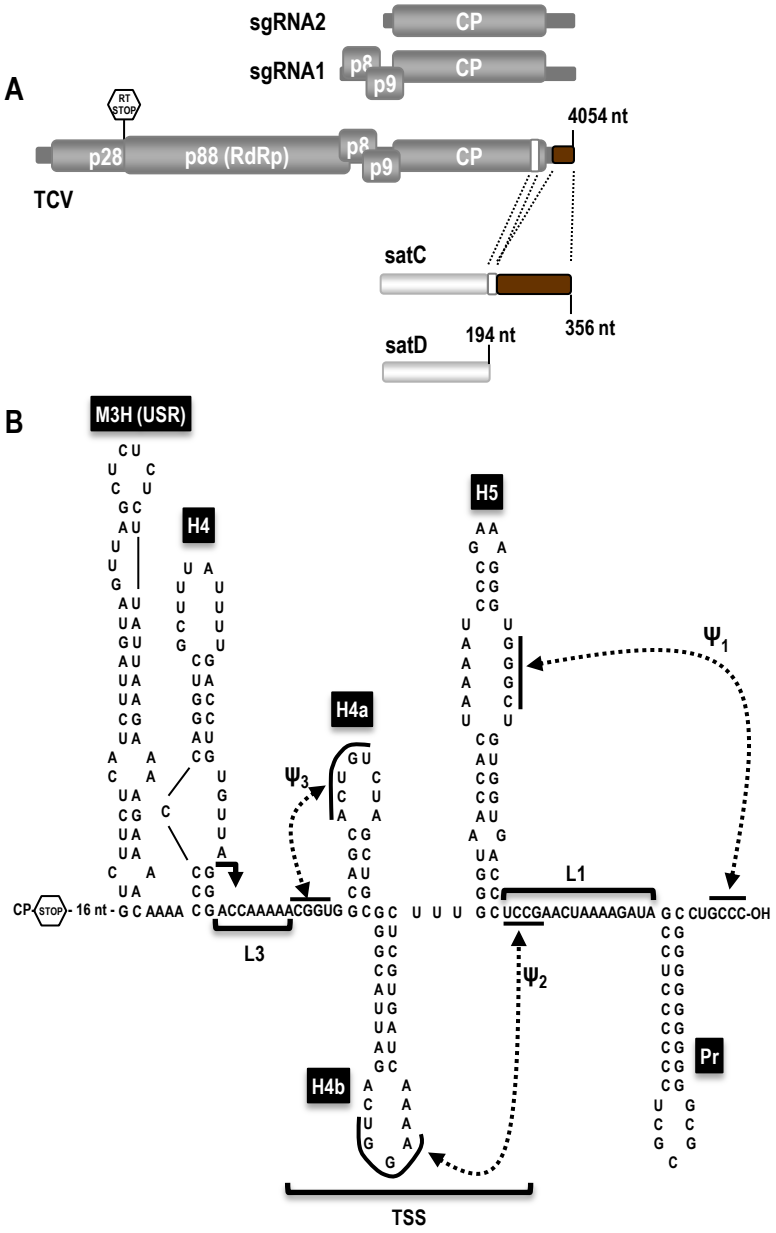

Figure 1

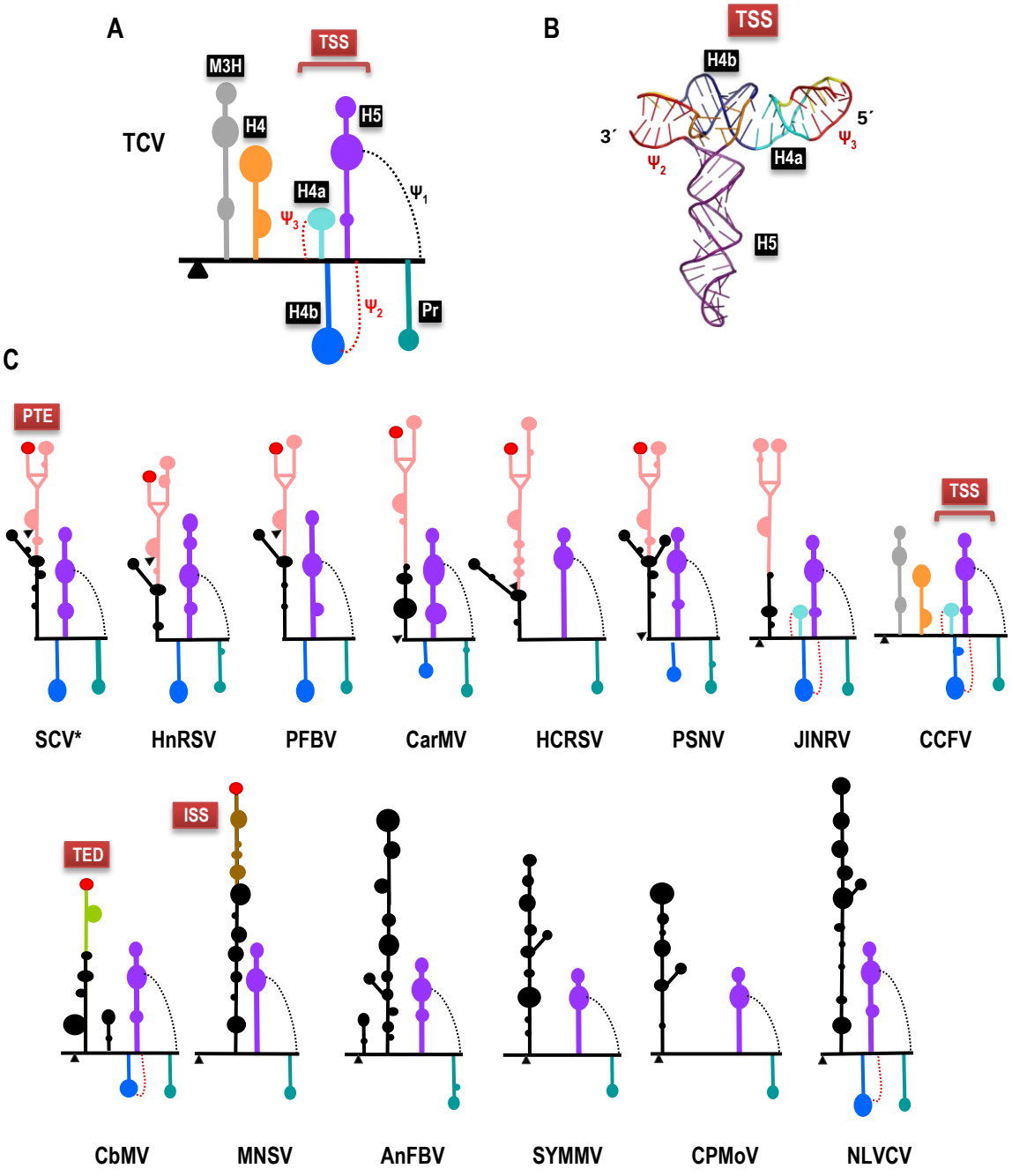
A
TSS
M3H
TCV
H4
H5
H4a
Ψ1
Ψ3
H4b
Ψ2
Pr
B
TSS
H4b
3´
5´
Ψ3
Ψ2
H4a
H5
C
PTE
TSS
SCV*
HnRSV
PFBV
CarMV
HCRSV
PSNV
JINRV
CCFV
TED
ISS
CbMV
MNSV
AnFBV
SYMMV
CPMoV
NLVCV

Figure 2

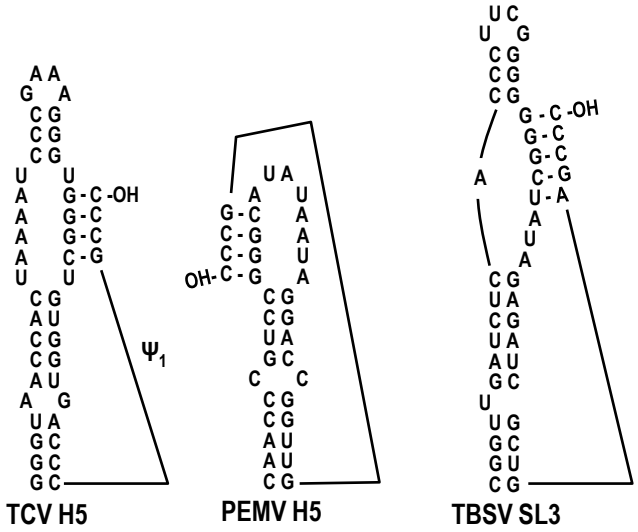

Figure 3

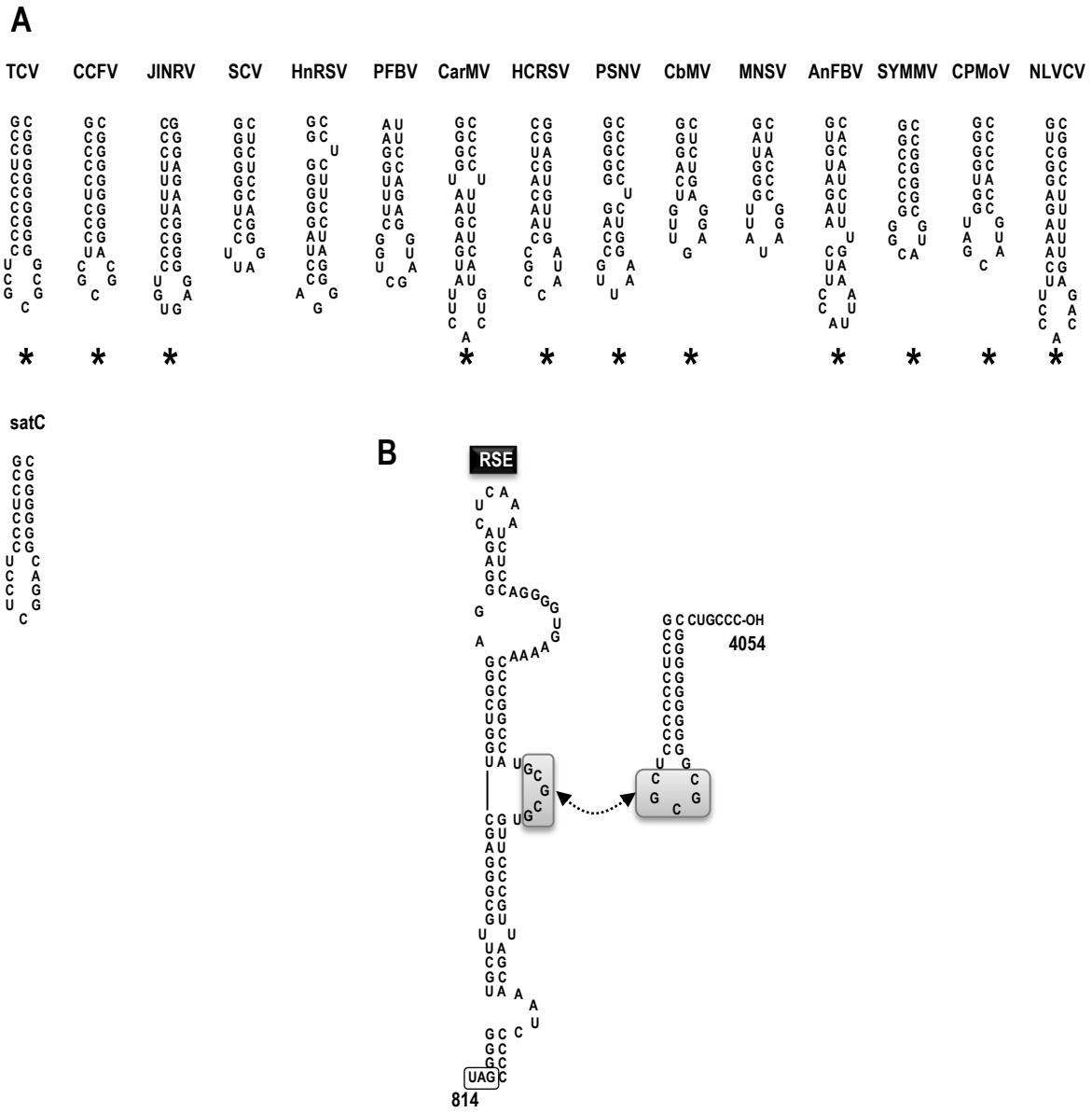
A
TCV
CCFV
JINRV
SCV
HnRSV
PFBV
CarMV
HCRSV
PSNV
CbMV
MNSV
AnFBV
SYMMV
CPMoV
NLVCV
satC
B
RSE
GC CUGCCC-OH
4054
UAG
814

Figure 4

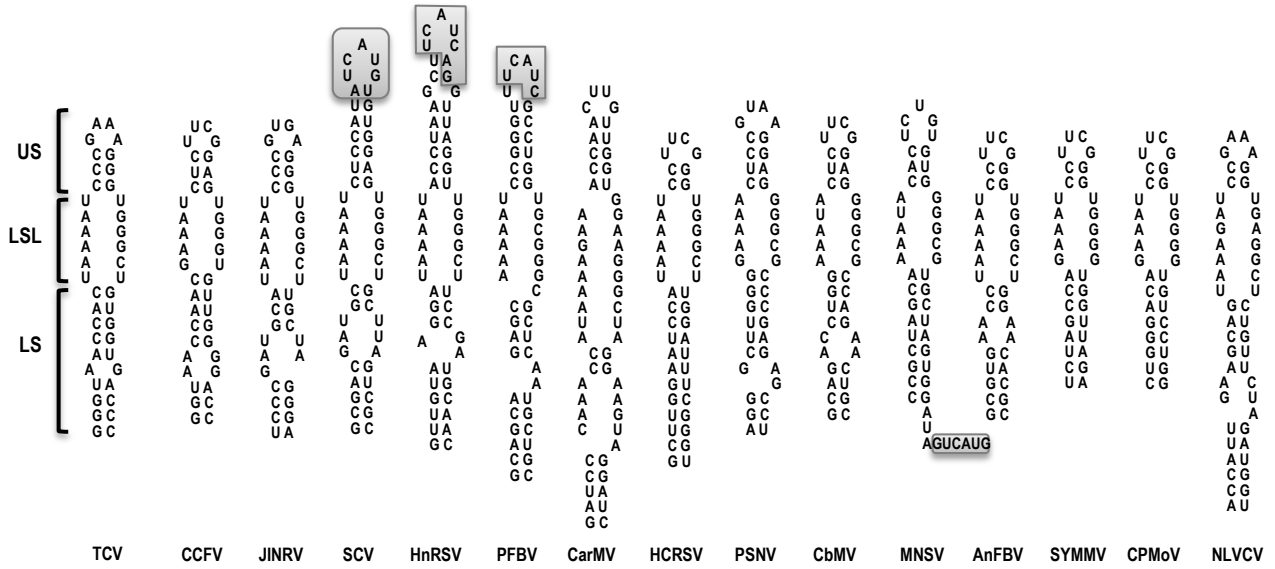
US
LSL
LS
AGUCAUG
TCV
CCFV
JINRV
SCV
HnRSV
PFBV
CarMV
HCRSV
PSNV
CbMV
MNSV
AnFBV
SYMMV
CPMoV
NLVCV

Figure 5

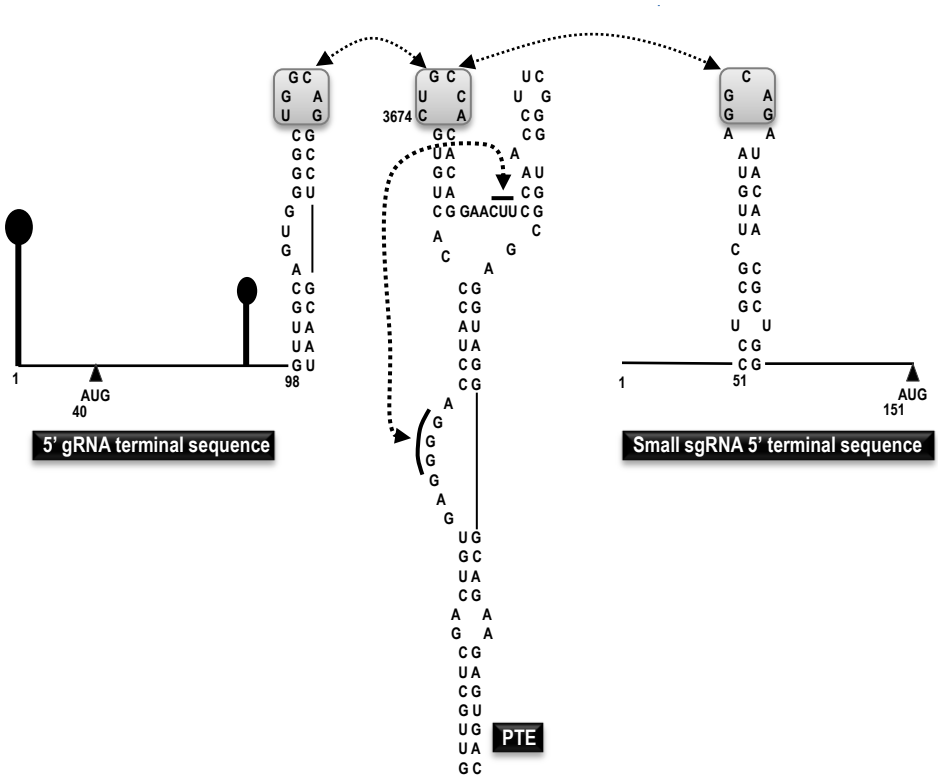
3674
1
AUG
40
98
5' gRNA terminal sequence
PTE
1
51
AUG
151
Small sgRNA 5' terminal sequence

Figure 6

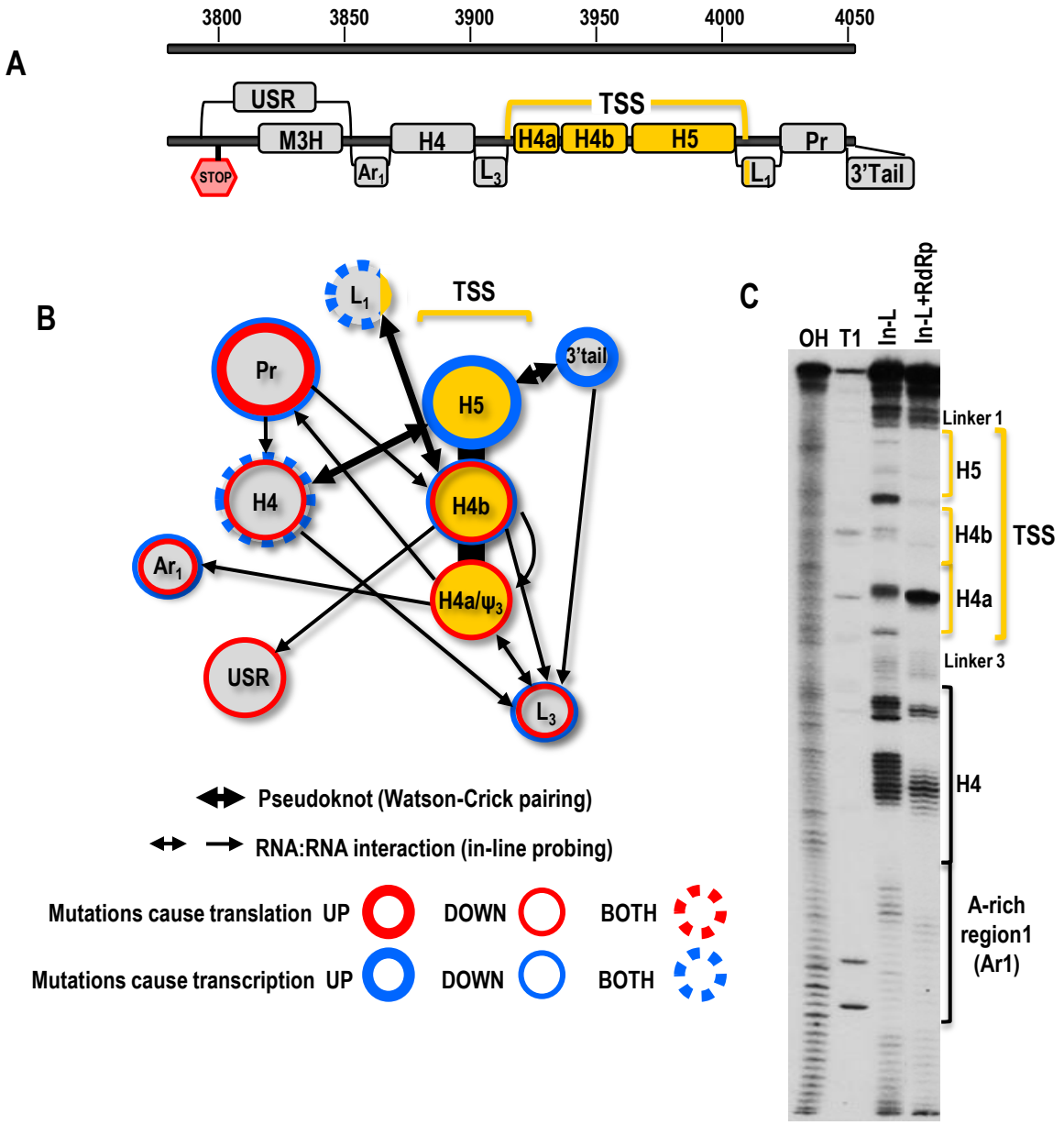
A
3800
3850
3900
3950
4000
4050
USR
M3H
H4
TSS
H4a
H4b
H5
Pr
STOP
Ar1
L3
L1
3'Tail
B
L1
TSS
Pr
3'tail
H5
H4
H4b
Ar1
H4a/ψ3
USR
L3
Pseudoknot (Watson-Crick pairing)
RNA:RNA interaction (in-line probing)
Mutations cause translation UP
DOWN
BOTH
Mutations cause transcription UP
DOWN
BOTH
C
OH
T1
In-L
In-L+RdRp
Linker 1
H5
H4b
TSS
H4a
Linker 3
H4
A-rich region1 (Ar1)

Figure 7