

# Tospovirus ambisense genomic RNA segments use almost complete repertoire of stable tetraloops in the intergenic region

Max T. B. Clabbers<sup>1</sup>, Rene C. L. Olsthoom<sup>1</sup> and Alexander P. Gultyaev<sup>1,2,\*</sup>

<sup>1</sup>Institute of Chemistry, Leiden University, 2300 RA Leiden and <sup>2</sup>Department of Viroscience, Erasmus Medical Center, 3000 CA Rotterdam, The Netherlands

Associate Editor: Ivo Hofacker

## ABSTRACT

**Summary:** The intergenic regions of the ambisense RNA segments of viruses from the *Tospovirus* genus form large extended RNA structures that regulate virus replication. Using comparative structure analysis, we show the presence of conserved alternative conformations at the apical parts of these structures. In one conformation, a branched Y-shape, the 5'-proximal hairpin arms are mostly capped by exceptionally stable tetraloop motifs. The tetraloop hairpins are folded in both virus and virus-complementary sense RNAs, and different tetraloops can functionally replace each other. Folding simulations show that the branched Y-shape structures can undergo a conformational transition to alternative extended rod-like conformations. Functional importance of both alternatives is supported by nucleotide covariations. The balanced equilibrium between alternative structures is evidenced by native gel electrophoresis of mutant RNA transcripts with shifted equilibria. The tetraloops play a role in the stability and dynamics of structures but may also be recognized by proteins involved in translation and/or replication.

**Contact:** gultyaevap2@chem.leidenuniv.nl

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

Received on November 22, 2013; revised on February 5, 2014; accepted on February 26, 2014

## 1 INTRODUCTION

Viruses of the genus *Tospovirus* are plant-infecting pathogens that possess ambisense genomic RNA segments (Nguyen and Haenni, 2003). These ambisense segments, M (medium) and S (small), yield viral RNA transcripts of opposite polarities with non-overlapping protein-coding regions. The 3' ends of vRNA sense (the polarity of the negative-sense RNA packaged in the virion) and vcRNA (virus-complementary) transcripts overlap in the intergenic regions (IGR) (Fig. 1A).

The tospovirus M and S RNA transcripts are not polyadenylated like the majority of eukaryotic mRNAs. Instead, they contain A-rich sequences in the 3'UTR of mRNAs, encoded in the IGRs that contain long stretches of A-rich and U-rich sequences. This results in the folding of extended stem-loop structures that have been suggested to regulate the translation and/or transcription termination of mRNAs of both polarities (de Haan *et al.*, 1990; Geerts-Dimitriadou *et al.*, 2012; van Knippenberg *et al.*, 2005).

The determinants of the regulatory functions of tospovirus IGR structured domains are not yet known. The structures (frequently called hairpins in the tospovirus literature) may fold into various shapes with branched or unbranched apical parts at the top of the A-U-rich stem-loops (e.g. de Haan *et al.*, 1990; Geerts-Dimitriadou *et al.*, 2012; Knierim *et al.*, 2006; van Knippenberg *et al.*, 2005). Here, we investigate the hairpins formed by the tospovirus M and S RNA IGR domains, searching for conserved structural features.

## 2 METHODS

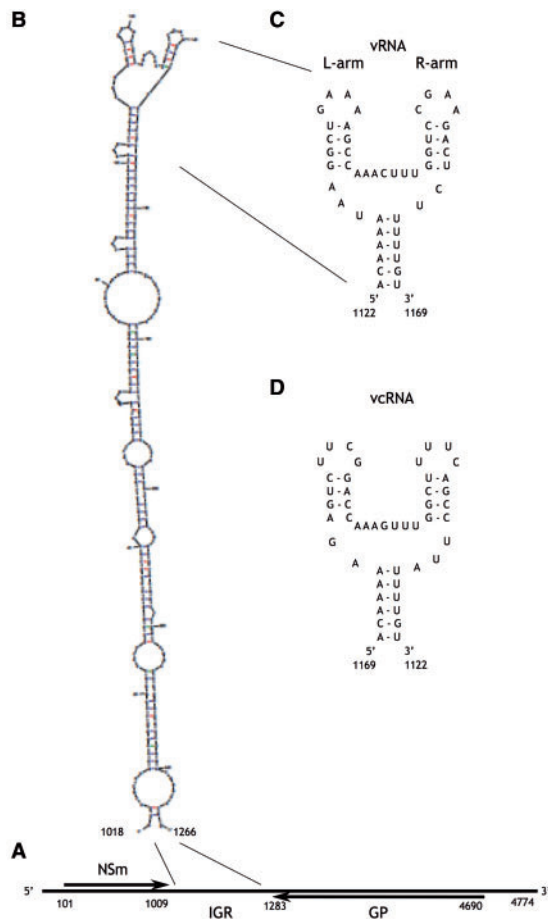
The sequences of tospovirus RNA segments were retrieved from the GenBank database using the Taxonomy Browser (NCBI Resource Coordinators, 2013). The complete list of viruses studied and abbreviations is given in the Supplementary Data. The structures in the IGR regions were predicted using free energy minimization algorithms of the Mfold server (Zuker, 2003) and RNAalifold program (Bernhart *et al.*, 2008). On the localization of the structure, formed by the pairing between A- and U-rich regions in a particular IGR region, the alternative apical structures were explored by inspection of the suboptimal structures yielded by Mfold program for the whole structural domain. To ensure the exhaustive analysis of homologous structures in all known (partial) sequences of the M and S segments, the sequences of various isolates of the same virus were also searched by BLAST program using sequences of previously found domains as queries. The kinetic folding pathways were studied by simulations of cotranscriptional folding using the genetic algorithm of STAR package (Gultyaev *et al.*, 1995) and kinefold server (Xayaphoummine *et al.*, 2005). Native gel electrophoresis detection of alternative conformations is described in Supplementary Data.

## 3 RESULTS

### 3.1 Conserved Y-shape tetraloop-containing structure and an alternative stem-loop folding in tospovirus M RNA

The A- and U-rich regions in the tospovirus M segment IGRs form extended stem-loop structures in both vRNA and vcRNA sense (Geerts-Dimitriadou *et al.*, 2012; van Knippenberg *et al.*, 2005). Our comparative RNA structure analysis, combining thermodynamic and kinetic folding calculations with sequence alignments of all available sequences, shows the presence of characteristic RNA motifs in the apical regions of large A-U-rich stem-loop domains (Supplementary Fig. S1). For instance, a Y-shaped motif with two hairpins can be folded at the top of the TSWV vRNA stem-loop structure (Fig. 1B and C). The two hairpins, denoted here as L-arm and R-arm, contain loops of four nucleotides. The L-arm loop sequence, GAAA, belongs to

\*To whom correspondence should be addressed.



**Fig. 1.** The structure of the IGR region in the TSWV M segment vRNA. (A) The ambisense gene structure with NSm and GP protein-coding regions (arrows show ORF orientations) is shown for the TSWV-T strain (accession AY870389) as an example. (B) The vRNA IGR stem-loop structure predicted by free energy minimization. (C) The apical Y-shape conformation in the vRNA structure. (D) The vcRNA apical Y-shape conformation

the GNRA type of extra stable tetraloops (Heus and Pardi, 1991; Woese *et al.*, 1990) (here N is an arbitrary nucleotide, and R is a purine). Remarkably, the 5'-proximal arm in the predicted vcRNA structure, which is the mirror image of the vRNA R-arm (Fig. 1D), also contains a tetraloop, UUCG, of another extra stable UNCG type (Cheong *et al.*, 1990; Tuerk *et al.*, 1988; Woese *et al.*, 1990).

The Y-shaped configuration is conserved in M segments of other tospoviruses, except BeNMV and SVNaV that have an extra hairpin inserted between two tetraloop hairpins, converting the Y-shape into a trident configuration (Supplementary Fig. S2). All apical hairpins are supported by covariations and base-pair indels (Supplementary Fig. S3). Furthermore, the 5'-proximal hairpin loops in both vRNA and vcRNA structures exhibit a preference for one of the stable tetraloop types (Table 1). The vRNA L-arm is always a tetraloop, and diverse vRNA and vcRNA loop sequences satisfy one of the known tetraloop motifs: GNRA (Heus and Pardi, 1991; Woese *et al.*, 1990), UNCG (Cheong *et al.*, 1990; Tuerk *et al.*, 1988; Woese

*et al.*, 1990) or YNMG (Proctor *et al.*, 2002), where Y is C or U, and M is C or A. The tospovirus tetraloop motifs can functionally substitute each other: the 5'-proximal tetraloops are GNRA and UNCG types in vRNA and vcRNA, respectively, or the reverse combination: GNRA in vcRNA and UNCG or YNMG in vRNA (Table 1).

Insertion of an extra adenine in the GAAAA loop of GTV (Table 1) is known to preserve the GNRA geometry in the GNRNA or GNR(X)nA motifs (Legault *et al.*, 1998; Lemieux and Major, 2006). A cUUA loop, predicted in the INSV vcRNA, is one of the most stable tri-loops (Shu and Bevilacqua, 1999). However, the TNRV and IYSV uAAAA loops do not belong to the known exceptionally stable or frequent tri-loops (Gardner *et al.*, 2011; Shu and Bevilacqua, 1999; Thulasi *et al.*, 2010).

Structure predictions and multiple alignments (Supplementary Fig. S1) show that all viruses of the Eurasian and the BeNMV/SVNaV clades (Chen *et al.*, 2013; de Oliveira *et al.*, 2012) can also form an alternative extended rod-like structure competing with branched structures with two or three hairpins. Apparently, both alternatives and/or equilibrium between them are functional because they are stable thermodynamically, with rather close free energy values, and are supported by covariations (Supplementary Fig. S1). The RNA folding pathway simulations (see Section 2) demonstrate that the tospovirus M RNA branched structures and extended stem-loops can form sequentially in the cotranscriptional folding pathway (Supplementary Fig. S4). In these simulations, the L-arm vRNA hairpin (or R-arm complement in vcRNA folding) is always formed first. On RNA elongation, the simulations suggest either the addition of the second arm followed by the disruption of both hairpin arms in favor of the more stable extended stem-loop or direct initiation of the stem-loop without the second arm formation. Some simulations predicted the Y-shape or trident conformations to be the final fold. However, calculations of free energies of alternative structures using Mfold server (Zuker, 2003) show that such a simulation result is determined by deviations in used thermodynamic parameters rather than by kinetic trapping in the metastable structure. Nevertheless, the native gel electrophoresis experiments with wild-type M segment fragments show that such trapping is possible *in vitro* because the alternative structures can be detected (Supplementary Fig. S5). Moreover, introduction of mutations stabilizing one of the alternatives leads to the expected equilibrium shifts.

### 3.2 Tetraloop hairpins and alternative structure in tospovirus S RNA

Both tetraloop-containing structures and rod-like alternatives were also predicted to fold in S segment vRNAs and vcRNAs (Supplementary Fig. S6). The S segment structural motifs seem to be less constrained than those of M segment. The two conserved hairpin arms are mostly separated by spacers of various lengths that may form their own structure. In four viruses (MeSMV, INSV, MYSV and PSMV), it was not possible to identify an R-arm homolog, and the unusually short (183 nucleotides) IGR of the PolRSV S segment does not form an A-U-rich structure at all because of the absence of the vRNA U-rich stretch (Ciuffo *et al.*, 2008). Nevertheless, the majority of

Table 1. Tospovirus M segment L-arm tetraloops

vRNA	vcRNA	Tetraloop type	Viruses
uGAAAa	cUUCGg	GNRA/UNCG	American clade
uGAAAa	cUUCGg	GNRA/UNCG	TSWV,CSNV,ZLCV,TCSV,GRSV
uGAAAa	cUUCAg	GNRA/...	TSWV-IT-TO
cUCCGg	cUUAg	UNCG/...	TSWV-YN
			INSV
			Eurasian clade
cUUAGg	uGAAAa	YNMG/GNRA	CaCV,GBNV,TZSV,CCSV
cUUAGg	cGAAAg	YNMG/GNRA	MYSV
cCUAGg	uGAAAa	YNMG/GNRA	WBNV,GBNV
cCAAGg	uGAAAa	YNMG/GNRA	WSMoV
cUUAGg	uGAAAa	YNMG/GNRNA	GTV
cUACGg	uGAAAa	UNCG/GNRA	WSMoV
cUUCGg	uAAAAa	UNCG/...	TNRV
cCAAGg	uAAAAa	YNMG/...	IYSV
cCAAGg	uAAAAa	YNMG/...	SLNSV
cCAAGg	cCCAAAa	YNMG/.....	PolRSV
cCAAGg	uUAAACa	YNMG/.....	TYRV
			SVNaV-BeNMV clade
cUUCGg	cGAAAg	UNCG/GNRA	SVNaV,BeNMV

Note: The loop-closing pairs are shown in lowercase. The nucleotides in the loops other than stable tetraloops are shown by dots in the tetraloop type column.

S segment vRNAs and about half of vcRNAs have a tetraloop in the 5'-proximal arm of IGR structure (Supplementary Table S2). A tetraloop of the UMAC type, which resembles GNRA fold in conformation and stability (Zhao *et al.*, 2012), not seen in M segments, is folded in the MeSMV S segment. The rod-like structure is not conserved in S segments of American clade viruses. In spite of that, nucleotide covariations in Eurasian clade structures suggest the importance of both Y-shape and its alternative (Supplementary Fig. S6). Base-pair indels and covariations occur simultaneously in both arms, preserving not only the double-hairpin structure but also the alternative stem-loop conformation (Fig. 2).

4 DISCUSSION

Comparison of predicted tospovirus IGR structures shows that their conserved hairpin arms exhibit a striking bias in the loop sequences to form tetraloops in the 5'-proximal arms of both vRNA and vcRNA structures, using a considerable part of the known repertoire of extra stable tetraloop motifs. A probability of such a bias to be determined by chance alone is negligible. Given a structure with two tetranucleotide hairpin loops, what is a constrained motif by itself, two stable tetraloops such as GNRA and YNMG would randomly occur with probability 1/512 in an uniform nucleotide distribution (for GNRA/UNCG: 1/1024). Conservation of these motifs in different viruses is not determined just by shared ancestry because various tetraloop types are found in homologous hairpins, showing that the tetraloops were evolved several times. A probability of such independent events in different viruses and segments to be determined by chance, being the product of sequence motif probabilities, is negligible. This shows that the observed bias is determined by strong evolutionary constraints.

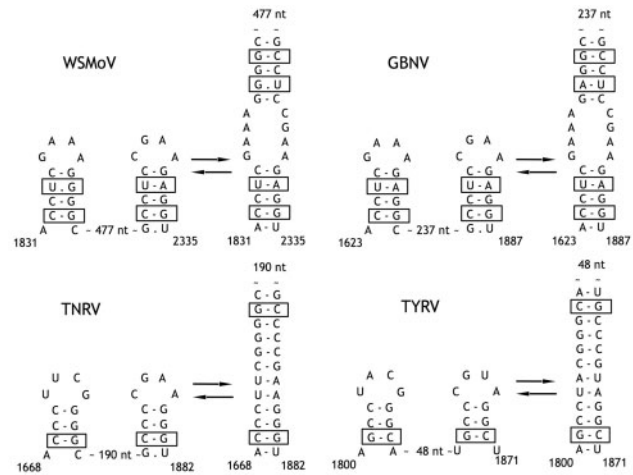


Fig. 2. Coupled covariations (boxed) in the S segment vRNA structures

These constraints suggest fast cotranscriptional formation of the metastable branched configurations that may delay the folding of more stable extended alternative stem-loops. The energy barrier for this transition is determined by simultaneous disruption of two hairpins, with total stacking free energy of ~15–20 kcal/mol in 8–10 bp capped by stable tetraloops. Thus, the transition times may be considerable and effectively trap tospovirus IGRs in branched configurations. Such kinetic trapping in a metastable conformation is supported by the IGR folding simulations and the experimental study of alternative structures (Supplementary Fig. S5). The coupled covariation pattern involving more than two bases (Fig. 2) is an indicator of a fine-tuned conformational transition (Gulyaev *et al.*, 2000), and multiple



examples of functionally important cotranscriptional RNA folding are known (Lai *et al.*, 2013).

Functional interchangeability of tetraloop types is usually explained by their exceptional stabilities (Sahu *et al.*, 2012; Selinger *et al.*, 1993; Uhlenbeck, 1990; Woese *et al.*, 1990). However, the majority of tospovirus M segment GAAA tetraloops are closed by a U-A base pair (Table 1), which neither confers the exceptional stability nor is a frequent GAAA closing pair (Woese *et al.*, 1990). Furthermore, the GAAA tetraloops fold significantly slower than YNMG ones and have folding kinetics comparable with the loops with an arbitrary sequence (Menger *et al.*, 2000; Nagel *et al.*, 2006; Proctor *et al.*, 2004). Thus, it seems that certain features other than just thermodynamic stability and folding rate are important in the tospovirus RNA tetraloop replacements.

Both GNRA and YNMG tetraloops are able to form tertiary interactions with various receptors in RNA (e.g. Proctor *et al.*, 2002; Wu *et al.*, 2012) and also can be recognized by proteins (Du *et al.*, 2004; Legault *et al.*, 1998; Melchers *et al.*, 2006). For instance, the less stable uGAAA motif, present in many tospovirus M segments (Table 1), specifically binds the phage  $\lambda$  N peptide (Legault *et al.*, 1998). Furthermore, by analogy to the behavior of the phage  $\lambda$  N peptide targets, the GTV vRNA 5'-proximal loop GAAAA (Table 1), containing the extruding adenine from the GAAA-like fold, is expected to retain comparable binding properties.

The tospovirus IGR structures were suggested to stimulate translation by assisting in circularization of mRNA (Geerts-Dimitriadou *et al.*, 2012). Their Y-shape conformation resembles those of some of the plant virus translational enhancers such as the so-called BYDV-like cap-independent 3'UTR translational enhancers (Meulewaeter *et al.*, 2004; Shen and Miller, 2004). Remarkably, the BYDV-like cap-independent 3'UTR translational enhancer left arms have GNRNA and GNRA loops that have been suggested to determine the binding of the translation factor eIF4G and recruitment of eIF4E or other yet unidentified proteins (Kraft *et al.*, 2013; Shen and Miller, 2004; Wang *et al.*, 2010). The tospovirus S RNA structures have been shown to act in concert with the viral proteins N and NSs, presumably involving NSs binding to RNA (Geerts-Dimitriadou *et al.*, 2012). One could speculate that the tetraloop motifs in the tospovirus IGR structures determine the binding of translation factors and/or other proteins.

A transition between two RNA structures may determine switching between different processes such as translation, transcription and replication. Replication enhancers may also adopt various Y-shape or unbranched conformations with diverse loop motifs (An *et al.*, 2010; Baumstark and Ahlquist, 2001; Olsthoorn and Bol, 2002).

Structurally diverse IGR domains are also folded in the ambisense segments of some viruses of the genera *Phlebovirus* and *Arenavirus* (Emery and Bishop, 1987; Lopez and Franze-Fernandez, 2007; Simons and Pettersson, 1991), but these structures are not conserved. This may be explained by differences in transcription termination mechanisms (Geerts-Dimitriadou *et al.*, 2012). Further research is needed to elucidate the roles of the tospovirus ambisense IGR structures and the function of the diverse tetraloop motifs.

## ACKNOWLEDGEMENTS

The authors thank Maarten de Smit for fruitful discussions.

*Funding:* Leiden University.

*Conflict of Interest:* none declared.

## REFERENCES

- An, M. *et al.* (2010) A Y-shaped structure in the 3' untranslated region together with the trans-activator and core promoter of Red clover necrotic mosaic virus RNA2 is required for its negative-strand RNA synthesis. *Virology*, **405**, 100–109.
- Baumstark, T. and Ahlquist, P. (2001) The brome mosaic virus RNA3 intergenic replication enhancer folds to mimic a T $\psi$ C-stem loop and is modified *in vivo*. *RNA*, **7**, 1652–1670.
- Bernhart, S.H. *et al.* (2008) RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics*, **9**, 474.
- Chen, T.C. *et al.* (2013) Molecular characterization of the full-length L and M RNAs of Tomato yellow ring virus, a member of the genus *Tospovirus*. *Virus Genes*, **46**, 487–495.
- Cheong, C. *et al.* (1990) Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature*, **346**, 680–682.
- Ciuffo, M. *et al.* (2008) A member of a new Tospovirus species isolated in Italy from wild buckwheat. *Arch. Virol.*, **153**, 2059–2068.
- de Haan, P. *et al.* (1990) The S RNA segment of tomato spotted wilt virus has an ambisense character. *J. Gen. Virol.*, **71**, 1001–1007.
- de Oliveira, A.S. *et al.* (2012) Characterization of Bean necrotic mosaic virus: a member of a novel evolutionary lineage within the genus *Tospovirus*. *PLoS One*, **7**, e38634.
- Du, Z. *et al.* (2004) Solution structure of a consensus stem-loop D RNA domain that plays important roles in regulating translation and replication in enteroviruses and rhinoviruses. *Biochemistry*, **43**, 11959–11972.
- Emery, V.C. and Bishop, D.H.L. (1987) Characterization of Punta Toro S mRNA species and identification of an inverted complementary sequence in the intergenic region of Punta Toro phlebovirus ambisense S RNA that is involved in mRNA transcription termination. *Virology*, **156**, 1–11.
- Gardner, D.P. *et al.* (2011) Statistical potentials for hairpin and internal loops improve the accuracy of the predicted RNA structure. *J. Mol. Biol.*, **413**, 473–483.
- Geerts-Dimitriadou, C. *et al.* (2012) Analysis of the tomato spotted wilt virus ambisense S RNA—encoded hairpin structure in translation. *PLoS One*, **7**, e31013.
- Gulyaev, A.P. *et al.* (1995) The computer simulation of RNA folding pathways using a genetic algorithm. *J. Mol. Biol.*, **413**, 473–483.
- Gulyaev, A.P. *et al.* (2000) Coupled covariations reveal dynamic RNA interaction patterns. *RNA*, **6**, 1483–1491.
- Heus, H.A. and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science*, **253**, 191–194.
- Knierim, D. *et al.* (2006) The complete nucleotide sequence of a capsicum chlorosis virus isolate from *Lycopersicon esculentum* in Thailand. *Arch. Virol.*, **151**, 1761–1782.
- Kraft, J.J. *et al.* (2013) Cation-dependent folding of 3' cap-independent translation elements facilitates interaction of a 17-nucleotide conserved sequence with eIF4G. *Nucleic Acids Res.*, **41**, 3398–3413.
- Lai, D. *et al.* (2013) On the importance of cotranscriptional RNA structure formation. *RNA*, **19**, 1461–1473.
- Legault, P. *et al.* (1998) NMR structure of the bacteriophage  $\lambda$  N peptide/box B RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell*, **93**, 289–299.
- Lemieux, S. and Major, F. (2006) Automated extraction and classification of RNA tertiary structure cyclic motifs. *Nucleic Acids Res.*, **34**, 2340–2346.
- Lopez, N. and Franze-Fernandez, M.T. (2007) A single stem-loop structure in Tacaribe arenavirus intergenic region is essential for transcription termination but it is not required for a correct initiation of transcription and replication. *Virus Res.*, **124**, 237–244.
- Melchers, W.J.G. *et al.* (2006) A GCUA tetranucleotide loop found in the poliovirus oriL by *in vivo* SELEX (un)expectedly forms a YNMG-like structure: extending the YNMG family with GYYA. *RNA*, **12**, 1671–1682.
- Menger, M. *et al.* (2000) Dynamics of the RNA hairpin GNRA tetraloop. *Biochemistry*, **39**, 4500–4507.

- Meulewaeter, F. et al. (2004) Conservation of RNA structures enables TNV and BYDV 5' and 3' elements to cooperate synergistically in cap-independent translation. *Nucleic Acids Res.*, **32**, 1721–1730.
- Nagel, J.H. et al. (2006) Structural parameters affecting the kinetics of RNA hairpin formation. *Nucleic Acids Res.*, **34**, 3568–3576.
- NCBI Resource Coordinators. (2013) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **41**, D8–D20.
- Nguyen, M. and Haenni, A.L. (2003) Expression strategies of ambisense viruses. *Virus Res.*, **93**, 141–150.
- Olsthoorn, R.C. and Bol, J.F. (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.
- Proctor, D.J. et al. (2002) Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions. *Biochemistry*, **41**, 12062–12075.
- Proctor, D.J. et al. (2004) Folding thermodynamics and kinetics of YNMG RNA hairpins: specific incorporation of 8-bromoguanosine leads to stabilisation by enhancement of the folding rate. *Biochemistry*, **43**, 14004–14014.
- Sahu, B. et al. (2012) Functional replacement of two highly conserved tetraloops in the bacterial ribosome. *Biochemistry*, **51**, 7618–7626.
- Selinger, D. et al. (1993) Functional interchangeability of the structurally similar tetra nucleotide loops GAAA and UUCG in fission yeast signal recognition particle RNA. *Proc Natl Acad. Sci. USA*, **90**, 5409–5413.
- Shen, R. and Miller, W.A. (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.
- Shu, Z. and Bevilacqua, P.C. (1999) Isolation and characterization of thermodynamically stable and unstable RNA hairpins from a triloop combinatorial library. *Biochemistry*, **38**, 15369–15379.
- Simons, J.F. and Pettersson, R.F. (1991) Host-derived 5' ends and overlapping complementary 3' ends of the two mRNAs transcribed from the ambisense S segment of Uukuniemi virus. *J. Virol.*, **65**, 4741–4748.
- Thulasi, P. et al. (2010) Thermodynamic characterization of RNA triloops. *Biochemistry*, **49**, 9058–9062.
- Tuerk, C. et al. (1988) CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes. *Proc. Natl Acad. Sci. USA*, **85**, 1364–1368.
- Uhlenbeck, O.C. (1990) Tetraloops and RNA folding. *Nature*, **346**, 613–614.
- van Knippenberg, I. et al. (2005) Tomato spotted wilt virus S-segment mRNAs have overlapping 3'-ends containing a predicted stem-loop structure and conserved sequence motif. *Virus Res.*, **110**, 125–131.
- Wang, Z. et al. (2010) Structural plasticity of Barley yellow dwarf virus-like cap-independent translation elements in four genera of plant viral RNAs. *Virology*, **402**, 177–186.
- Woese, C.R. et al. (1990) Architecture of ribosomal RNA: constraints on the sequence of tetra-loops. *Proc. Natl Acad. Sci. USA*, **87**, 8467–8471.
- Wu, L. et al. (2012) Structural variation and uniformity among tetraloop-receptor interactions and other loop-helix interactions in RNA crystal structures. *PLoS One*, **7**, e49225.
- Xayaphoummine, A. et al. (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.*, **33**, 605–610.
- Zhao, Q. et al. (2012) UNAC tetraloops: to what extent do they mimic GNRA tetraloops? *Biopolymers*, **97**, 617–628.
- Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridisation prediction. *Nucleic Acids Res.*, **31**, 3406–3415.