

SCIENTIFIC REPORTS

OPEN

Modelling the three-dimensional structure of the right-terminal domain of pospiviroids

Gerhard Steger 

Viroids, the smallest known plant pathogens, consist solely of a circular, single-stranded, non-coding RNA. Thus for all of their biological functions, like replication, processing, and transport, they have to present sequence or structural features to exploit host proteins. Viroid binding protein 1 (Virp1) is indispensable for replication of pospiviroids, the largest genus of viroids, in a host plant as well as in protoplasts. Virp1 is known to bind at two sites in the terminal right (TR) domain of pospiviroids; each site consists of a purine- (R-) and a pyrimidine- (Y-)rich motif that are partially base-paired to each other. Here we model the important structural features of the domain and show that it contains an internal loop of two Y·Y *cis* Watson-Crick/Watson-Crick (cWW) pairs, an asymmetric internal loop including a cWW and a *trans* Watson/Hoogsteen pair, and a thermodynamically quite stable hairpin loop with several stacking interactions. These features are discussed in connection to the known biological functions of the TR domain.

Viroids, the smallest RNA pathogens known, consist of a circular, non-coding, single-stranded RNA (ssRNA) with a length of about 250–400 nucleotides depending on the viroid species^{1,2}. They are replicated in their plant hosts in an RNA-to-RNA rolling circle mechanism using either an asymmetric or a symmetric pathway³. In either pathway, replication includes a processing step of cleaving oligomeric replication intermediates to molecules of unit length. This step proceeds either by a viroid-internal ribozyme or by proteinaceous nucleases of the host. Based on this feature, viroids are classified into two families⁴: members of *Pospiviroidae* possess a thermodynamically stable rod-like secondary structure (Fig. 1a), replicate in the nucleus, and do not self-cleave; members of *Avsunviroidae* replicate in the chloroplast and self-cleave *via* a hammerhead ribozyme. The families are named after their respective type members potato spindle tuber viroid (PSTVd) and avocado sun blotch viroid. *Pospiviroidae* are divided into the genera *Pospi-*, *Hostu-*, *Cocad-*, *Apsca-*, and *Coleviroids*, based on sequence identity and biological properties⁴. All information for replication by cellular enzymes is included in structural features of the circular viroid RNA, which is abundant in the plant, and the less abundant oligomeric linear ssRNA replication intermediates of (+) and (–) polarity.

Only a few proteins or enzymes involved in the replication cycle of viroids are known. For example, members of *Pospiviroidae* and *Avsunviroidae* are transcribed by DNA-dependent RNA polymerase II and chloroplastic nuclear-encoded polymerase, respectively^{5,6}. Note that both enzymes are redirected to use ssRNA as template. Ligation of unit-length monomers to circles is supported by DNA ligase in case of PSTVd⁷ and by chloroplastic plant tRNA ligase for members of the *Avsunviroidae* family⁸, respectively.

Virp1 (Viroid binding protein 1) was isolated by screening a tomato cDNA expression library with PSTVd RNA⁹. The protein binds specifically (+) sense RNA of *Pospi-* and *Hostuviroid* members, but not of *Apscaviroid* and *Avsunviroidae* members¹⁰. Using electrophoretic mobility shift assays (EMSA) and the three-hybrid interaction system, the PSTVd region binding to the protein was determined: this is the terminal right (TR) domain of the rod like structure¹¹ (Fig. 1a,b). Two binding sites, named RY motifs (Fig. 1), composed of the sequence elements 5' ACAGG and 5' CCUUCUC were found. In the RY motif close to the terminal right hairpin loop, the 3' part of the purine-rich R sequence is paired to the 5' part of the pyrimidine-rich Y sequence¹²; the same pairing for the inner RY motif is only present in suboptimal foldings. The binding affinity of Virp1 was about fivefold stronger to the terminal RY motif than to the inner motif. Furthermore, Virp1 was shown to be indispensable for replication of PSTVd and citrus exocortis viroid (CEVd), a further *Pospiviroid*, in *Nicotiana benthamiana* host plants as well as in protoplasts¹³. Similarly, the nuclear import of the satellite RNA of cucumber mosaic virus

Institut für Physikalische Biologie, Heinrich-Heine-University Düsseldorf, 40225, Düsseldorf, Germany. Correspondence and requests for materials should be addressed to G.S. (email: steger@biophys.uni-duesseldorf.de)

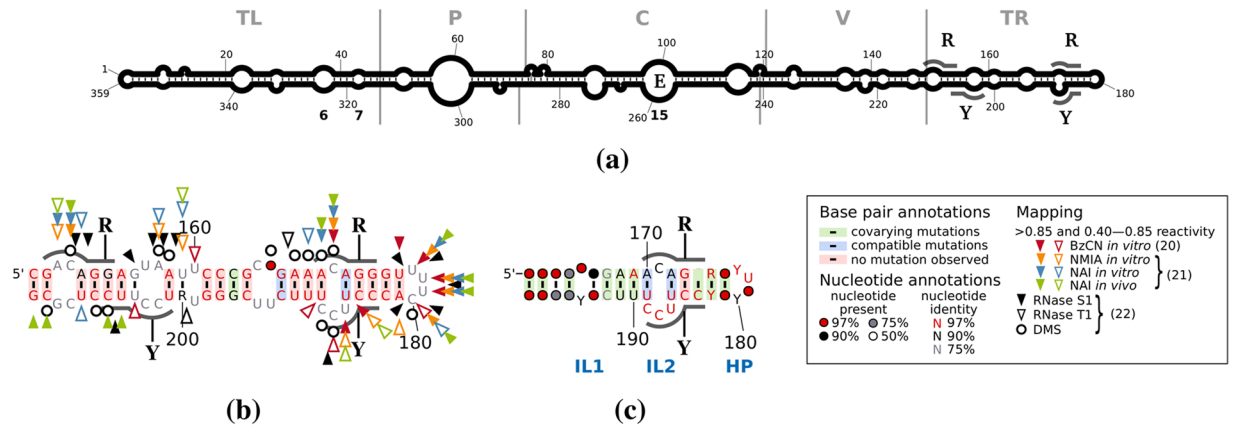


Figure 1. Secondary structure of pospiviroids. The consensus sequence and structure presented were predicted for alignments with MAFFT⁶⁶, optimized in ConStruct⁶⁷ at 37 °C, excluding lonely base pairs, and drawn with R2R⁶⁸. (a) Structure of circular PSTVd. Borders of the five domains⁷⁴ (TL, terminal left; P, pathogenicity-related; C, central; V, variable; TR, terminal right) are marked by gray lines. Loops 6, 7, and 15 (E) are marked. RY motifs critical for binding of Virp1¹² are outlined and marked by R and Y. (b) TR domain based on an alignment of 234 full-length unique PSTVd sequences. High and low SHAPE reactivity is marked by filled and open triangles, respectively [red⁶²; orange, blue and green⁷⁵]; reactivity with RNase S1 (black triangle), RNase T1 (open triangle), and dimethyl sulfate (DMS, circle) is marked⁷⁶. (c) Terminal right stem-loop based on an alignment (see Supplementary Fig. S2) of unique TR sequences of 95 pospiviroids (14 PSTVd, 25 CEVd, 21 columnea latent viroid (CLVd), 8 chrysanthemum stunt viroid (CSVd), 3 iredine viroid (Irvd), 3 Mexican papaya viroid (MPVd), 10 pepper chat fruit viroid (PCFVd), 4 tomato apical stunt viroid (TASVd), 7 tomato chlorotic dwarf viroid (TCDVd)). The internal loops are marked by IL1 and IL2, respectively, and the hairpin loop by HP. The color code used in (b) and (c) for the annotation of nucleotides, base pairs and mapping is given in the right box.

was shown to depend on Virp1¹⁴. Furthermore, the transcription factor TFIIA binds to the TR in the region G₁₆₇–U₁₇₆¹⁵.

Virp1 (see Supplementary Fig. S1) is a member of the bromodomain(s) and extraterminal domain (BET) protein family including a nuclear localization signal. BET proteins are known to be involved in chromatin biology and transcriptional regulation^{16,17}. The bromodomain recognizes acetyl-lysine residues in histones and other proteins^{18,19}. The ET domain consists of three regions; the conserved N-terminal ET (NET) region has an acidic patch that may interact with other proteins or nucleic acids²⁰. The RNA binding domain of Virp1 has been localized to the C-terminal half of the protein^{9,12}.

The knowledge on the three-dimensional structure of viroids is sparse. At least for PSTVd the presence of a tertiary structure (in the sense of tRNA-L-conformation) was excluded with certainty²¹. The left terminal part of the TL domain was shown by nuclear magnetic resonance (NMR) spectroscopy to be rod-like but not bifurcated²²; that is, this study supported secondary structure predictions but delivered no three-dimensional details. Loop 6 (Fig. 1a) has the sequence 5' G₃₆AC 3'/5' C₃₂₂GA 3' flanked by a U:G wobble pair and a *cis* Watson-Crick/Watson-Crick (cWW) G:C pair; according to structural modelling, the loop motif is a set of three stacked non-WC pairs²³. This motif mediates trafficking from palisade to spongy mesophyll in *N. benthamiana*. Loop 7 (Fig. 1a) is required to traffic from nonvascular into the vascular tissue phloem to initiate systemic infection. According to modelling, the loop nucleotides U₄₃·C₃₁₇ are a water-inserted cWW base pair flanked by canonical cWW base pairs²⁴. Loop 15 of PSTVd (Fig. 1a) is similar to loop E (sarcin/ricin motif) of eukaryotic 5 S rRNA^{25–27}. A detailed analysis of this loop, mainly based on isostericity matrices, allowed Zhong *et al.*²⁸ to design disruptive as well as compensatory mutations that are critical in rolling-circle replication^{29–31}.

We created three-dimensional RNA models of the terminal right domain of several circular viroids, using a variety of recently published web services, programs and databases, for example FARFAR/Rosetta^{32–34}, RNA Bricks³⁵, FRABase (RNA FRAGments search engine & dataBASE)³⁶, based on sequence and two-dimensional structural similarity to known three-dimensional RNA structures. These programs were selected for use mainly based on their complementarity (assembly of 3-nt fragments plus refinement versus search for similar structural elements using sequence and structure patterns) from a palette of available tools^{37–42}. The predicted models encompassed three loops: a symmetrical basepaired 2 × 2 loop, an asymmetrical internal loop, and a hairpin loop both stabilized by hydrogen bonding and stacking interactions. On the one hand this predicted structure should provide an extraordinary platform for protein and/or nucleic acid binding *in trans*, on the other this model should facilitate further experimental and structural studies of the TR interaction with Virp1 and TFIIA.

Results

Secondary structure prediction of TR domain. To get an overview on available TR sequences from members of *Pospiviroid* and their secondary structures, we assembled alignments of full-length sequences for each species of *Pospiviroid* and extracted the TR domain. For example, all unique sequences of PSTVd were aligned and a consensus structure was predicted; Fig. 1b details the TR domain of the PSTVd consensus structure.

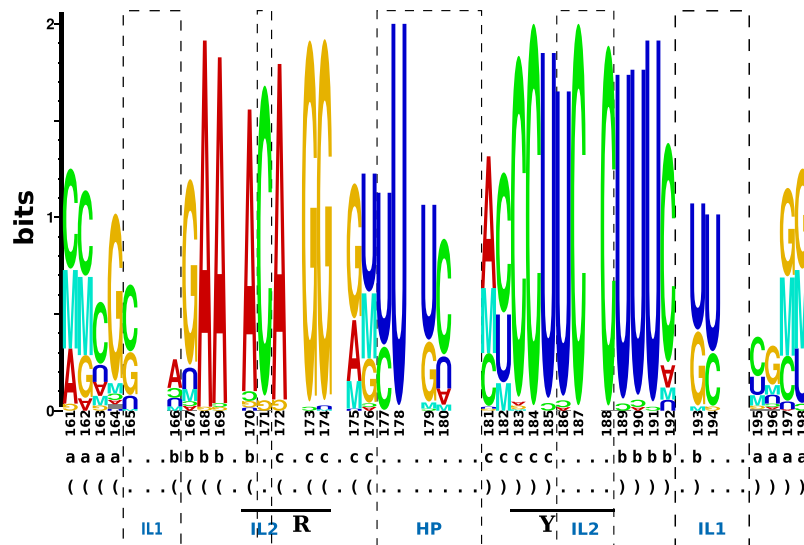


Figure 2. Structure logo of terminal right stem-loop. This logo⁴⁶ is based on the same alignment (see Supplementary Fig. S2) of 95 pospiviroids as used for Fig. 1c. The height of nucleotide characters is proportional to the observed frequency of nucleotides in the alignment and the expected (a priori) frequency $p = 0.25$; nucleotide characters that appear less than expected are displayed up-side-down. The mutual information (marked by ‘M’) between base-paired positions, as indicated in the two bottom lines, is calculated as the relative entropy between the fractions of complementary bases and the number of basepairs one would expect by chance from the distribution of nucleotides at the involved positions. The nucleotide numbering is as in Fig. 1c; missing numbers denote that the respective column of the alignment contains mostly gaps. The RY motif critical for binding of Virp1 is outlined and marked by R and Y. The sequences of internal loops are marked by IL1 and IL2, respectively, and the hairpin loop sequence by HP.

The TR stem-loop consisted of six well-defined helices but the sequences of loops more distant from the hairpin loop were difficult to align, especially if sequences of further *Pospiviroid* members were included. This is of concern only to the Virp1 binding site of lower affinity (nts 149–153/202–206), while the binding site of higher affinity (nts 170–174/183–187) is located in the more easily alignable region of the terminal stem-loop (nts 161–198; Fig. 1c). We will concentrate in the following on this terminal stem-loop that consists of three helices separated by two internal loops (IL1 and IL2) and the hairpin loop (HP). The final alignment of terminal stem-loop sequences (see Supplementary Fig. S2), on which the consensus structure (Fig. 1c) and the structure logo (Fig. 2) are based, contained 95 different sequences of pospiviroids.

The nucleotides in the two helices close to the HP are well conserved in sequence (marked by b and c in Fig. 2) but show some covariation marked by green and blue background in Fig. 1c. The sequence of the terminal HP loop is 5' UUUC in most pospiviroids, with the exception of CEVd and IrVd that have 5' CUCGW (W is A or U). The sequence of IL1 is 5' GA/GC in CEVd and 5' CW/UU in PCFVd, PSTVd, and most CLVd. The sequence of IL2 is 5' C/UCC, conserved in most pospiviroids. This degree of conservation in sequence as well as in secondary structure of the TR stem-loop gave some confidence as a basis for modelling.

Results from Rosetta modelling. We predicted models of IL1 (Fig. 3), IL2 (Fig. 4), and HP (Fig. 5) within the TR stem-loop (Fig. 1c) with the Rosetta web server (ROSIE/FARFAR/RNA De Novo). Fragment assembly for the full TR stem-loop showed only a limited convergence; with shorter structural elements, like the single loops, 1,000 generated structures and 10,000 MC cycles were usually sufficient to generate reasonable best-scoring structures. As input to the web server, we used the respective parts of the consensus structure (in bracket-dot notation) but sequences from individual viroid variants (for examples see Fig. 3, top row).

Internal loop 1 (IL1). The IL1 loop typical for PSTVd is a symmetric internal loop of 2×2 nucleotides ($10 \times$ CC/UU; Fig. 1b; Fig. 3a, top row). Other loop nucleotides also occur frequently (Fig. 2): $15 \times$ CU/UU (Fig. 3c); $2 \times$ UU/UU; a 1×1 loop of C/U followed by an A:U pair ($13 \times$; Fig. 3b); a G:C pair followed by 1×1 loop of A/G is typical for CEVd ($19 \times$). The loop is closed by G:C cWW pairs but also other closing pairs appear quite often. The loop nucleotides form standard cWW pairs in all models of these sequence variants (Fig. 3, middle row). Thus other cWW pairs should be able to substitute: U·C, G·A, A·G, and A·A are isosterically close to C·U; G·U, A·C, and C·C are isosterically close to U·U^{43,44}. These isosteric pairs, however, do not occur in viroids. Only a restricted set of pair combinations are able to retain the same backbone geometry according to RNAredesign^{45,46}; this is obvious from the logos shown in Fig. 3 bottom. For example, CC/UU might be replaced by UU/CC or CU/UU by UU/UC. Thus, the viroid-specific combinations of nucleotides might be either essential for interaction with Virp1, or the reduced stability of IL1—in comparison to an IL1 composed of purines with their increased stacking—is critical.

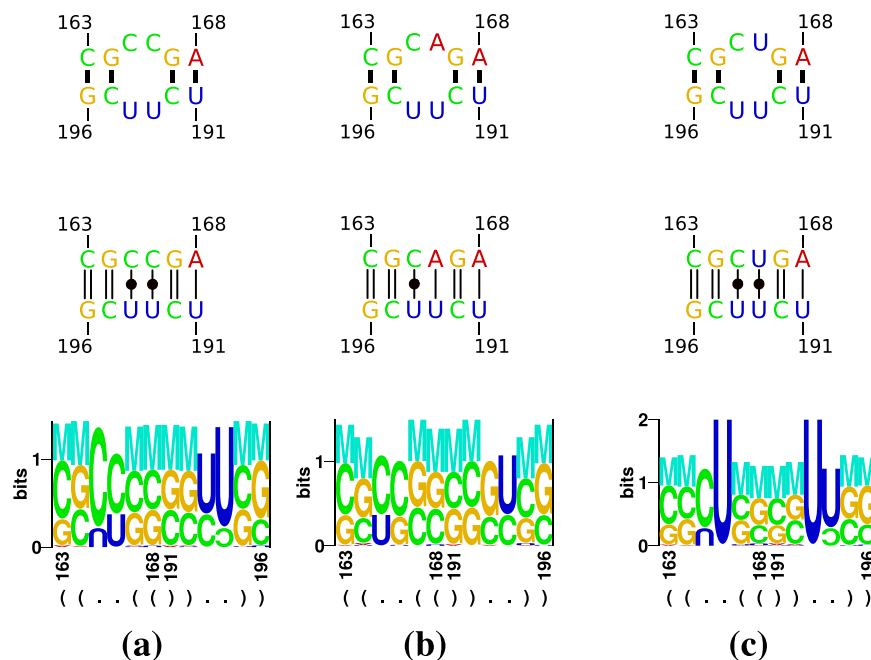


Figure 3. Results of Rosetta modelling of IL1 variants. Results of Rosetta modelling are shown for three typical loop variants: CC/UU (a), CA/UU (b), and CU/UU (c). Top: secondary structure of IL1 used as input to Rosetta modelling; that is, the input structure was ((..((+))..)) for all three sequences. Middle: interactions identified by rnaview in Rosetta's top model; standard cWW base pairs are marked by double lines; all "loop" nucleotides are paired in cWW conformation; for the basepairing nomenclature of Leontis & Westhof⁷⁷ see Supplementary Fig. S1. Bottom: Logos produced by RNAredesign for Rosetta's top model.

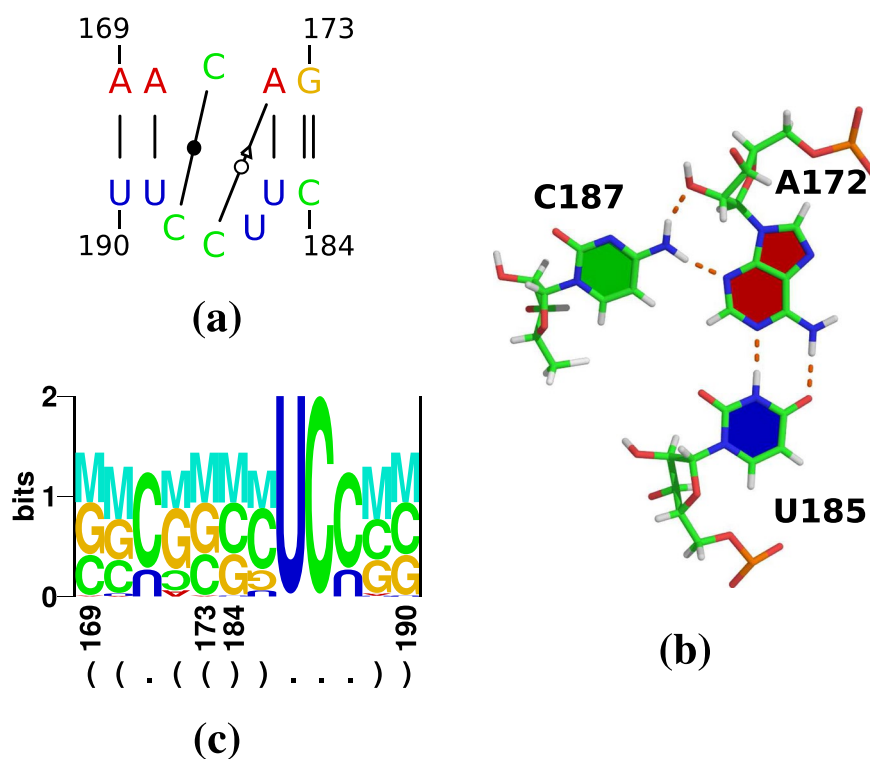


Figure 4. Results of Rosetta modelling of IL2 loop. (a) Interactions identified by rnaview in Rosetta's top model. Loop nucleotides were highly connected by a complex network of base/base hydrogen bonds and stacking interactions. (b) Triple pair C₁₈₇·A₁₇₂·U₁₈₅ (cWW A₁₇₂·U₁₈₅, *trans* Watson/Sugar edge (tWS) C₁₈₇·A₁₇₂). (c) Logo produced by RNAredesign for Rosetta's top model.

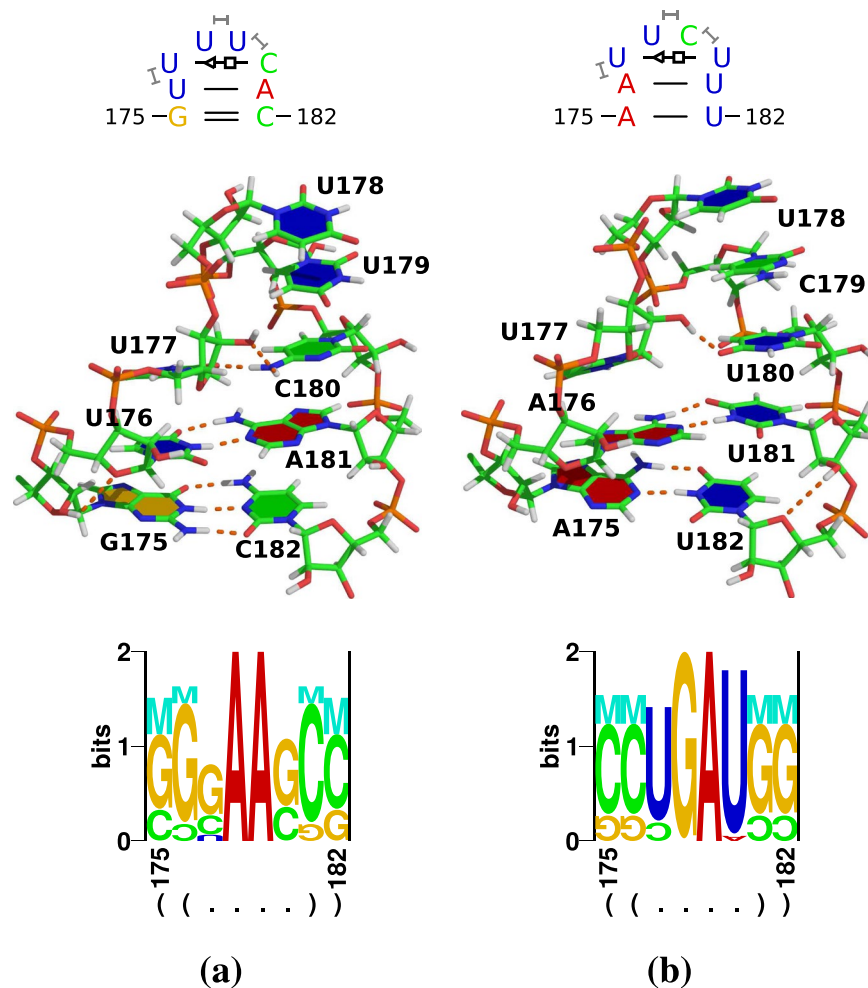


Figure 5. Results of Rosetta modelling of HP loop. Results for a typical loop of most pospiroviruses is shown in (a); an exceptional loop of TCDVd is shown in (b). Top: Interactions identified by rnaview and x3dna-dssr. Middle: 3D view produced by PyMOL. Bottom: Logos produced by RNAredesign.

The stability of such tandem “mismatches” is partially known^{47,48}: the tandem U·U stabilizes a helix when flanked by G·C pairs, as in case of viroids, but is less stable than symmetric tandem G·U or A·G pairs. The pyrimidine base-pairs led to a contraction of the sugar phosphate backbone of the opposing strands in comparison to standard cWW pairs; this might be of importance for protein binding.

Internal loop 2 (IL2). Predicted IL2 loop nucleotides were highly connected by a complex network of base/base hydrogen bonds and stacking interactions (Fig. 4a). An example of these interactions is shown in Fig. 4b: C₁₈₇·A₁₇₂·U₁₈₅ form a triple pair consisting of a cWW A:U and a *trans* Watson/Sugar edge (tWS) C·A.

According to RNAredesign, U₁₈₆ and C₁₈₇ are critical for the backbone geometry of the loop (Fig. 4c). The cWW pair C₁₇₁·C₁₈₈ is replaced by a U·U in 22% of structures, which is the only near-isosteric pair to C·C⁴⁴.

Hairpin loop (HP). The most common HP loop is UUUC closed by an U:A pair (Fig. 5a); the HP loop of CEVd and IrVd is CUCGW closed by a G:C pair (see Supplementary Fig. S2). The predicted HP loop structures were characterized by internal interactions of loop nucleotides. For the most common HP loop, Rosetta predicted a conformation with a *trans* Sugar/Hoogsteen (tSH) pair U₁₇₇:C₁₈₀ next to the loop-closing cWW U:A pair and stacking of the two last loop nucleotides (Fig. 5a). The exceptional HP loop of a TCDVd (Fig. 5b) was very similar with a tSH pair U₁₇₇:C₁₈₀ next to the loop-closing A:U pair and stacking of the two last loop nucleotides. For the HP loop of CEVd, Rosetta predicted a quite similar structure, despite the five loop nucleotides instead of the tetra-loops in most other cases: the fifth nucleotide of the loop was bulged out (see Supplementary Fig. S3b).

RNAredesign suggested quite different sequences for all loop models shown in Fig. 5 and S3: the optimized sequences contained more purines than the natural sequences, which might point on the one hand to increased thermodynamic stability of the optimized sequences due to increased stacking interactions, and on the other to functional constraints of the natural sequences. Such constraints might be an adaptation to stability requirements and restrictions induced by binding to other molecules.

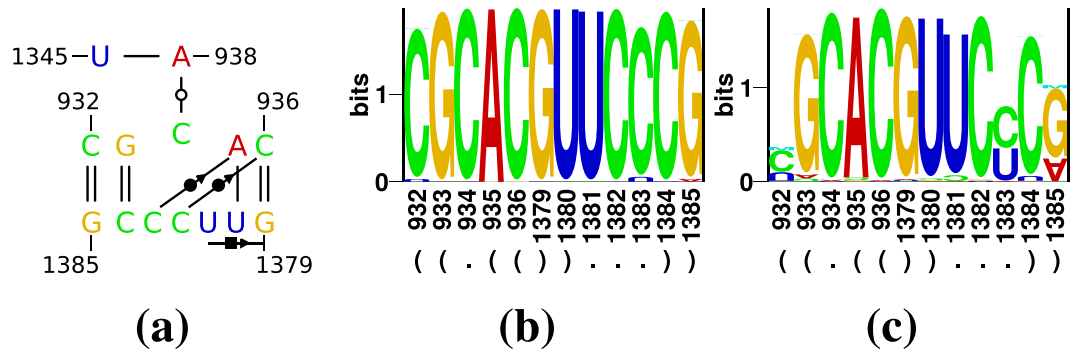


Figure 6. A motif from *Thermus thermophilus* 16S ribosomal RNA. (a) Nucleotides 932–936 and 1379–1385 are similar in sequence to IL2 of PSTVd; interactions are extracted by x3dna-dssr from PDB 1FJG⁷⁸. (b) A logo of the motif (a) from an alignment of 23,500 bacterial 16S ribosomal RNAs⁵⁴. (c) A logo of the motif (a) for 1850 sequences from the same alignment as used in (b), but sequences identical to the consensus 5' CGCAC/GUCCCCG 3' were removed. Columns with less than 1% of nucleotides and sequences not fully spanning the motif or containing nucleotides other than [AUCG] were removed from the alignments used for the logos.

Structural similarities in RNABricks and FRABase. The Rosetta models, as depicted above, are based on assembly of short fragments from known structures. These fragments might include nucleotides that interact with additional ligands like a protein or another RNA. That is, “unusual” conformations of nucleotides in the final models might be due to absent ligands. In contrast, if a larger fragment of the target is found in a known (template) structure including ligands, one might get insight on the basis of such “unusual” conformations and requirements for interactions. So we searched for structural elements similar to the terminal stem-loop in RNABricks³⁵ and FRABase³⁶.

IL1. An internal loop with a 5' CU/5' UU motif, frequent in pospirovirids, is described as a non-canonical tandem base-pair (i. e. two cWW pairs) in the 3'-UTR of poliovirus-like enteroviruses⁴⁹. Remarkably, this loop is conserved but replacement of the C·U and U·U pairs by a C·G and U·A pairs, respectively, had no effect on virus viability and growth in cell culture. This is in contrast with viroids: replacement of IL1 by two A·U pairs diminished replication efficiency and abolished systemic infection⁵⁰ (see below and Supplementary Table S4).

The intron 1 of human cellular nucleic acid-binding protein (also named zinc finger protein 9) contains a (CCUG)_n repeat; expansion of this repeat causes myotonic dystrophy type 2⁵¹. The repeat folds into a stem-loop with 5' CU/5' UC tandem cWW pairs flanked by GC pairs that binds and inactivates a splicing regulator⁵². This symmetric tandem mismatch, however, occurs only in CLVd JF446928.

IL2. RNA Bricks found an internal loop with some similarity to IL2 in PDB structures of 16S ribosomal RNA; for an example see Fig. 6a. The internal loop has the sequence C₉₃₄/U₁₃₈₁UCC₁₃₈₃ with a 3' closing A:U pair as in PSTVd (Fig. 4) but with a 5' G:C pair. C₉₃₄ is bulged out from the motif and makes a tWW pair with A₉₃₈:U₁₃₄₅ (see motif IL_40845.1 in the RNA 3D Motif Atlas)⁵³. C₁₃₈₂ and C₁₃₈₃ pair with C₉₃₆ and A₉₃₅ in cWS conformation, respectively. U₁₃₈₁ pairs with G₁₃₇₉ in cHS conformation as annotated by x3dna-dssr; rnaview strangely determines this interaction as a cSW pair. A protein contact seems not relevant for this motif. The sequence of this loop is highly conserved according to an alignment of bacterial 16S ribosomal RNAs⁵⁴, a corresponding logo is shown in Fig. 6b. The most variant loop nucleotide obviously is C₁₃₈₃ (Fig. 6b and c) that is replaced by an U in ~3% of sequences; this replacement did not occur in IL2.

HP. Hairpin loops with the four nucleotides and closing base pair common in PSTVd's HP (Fig. 5a) are present in the histone mRNA stem-loop (PDB entries 1JU7⁵⁵ and 1KKS⁵⁶), which binds to the stem-loop binding protein (SLBP) and the histone mRNA 3'-exonuclease 1 (Exo) for 3' end processing. These hairpin structures—without corresponding proteins—were determined by NMR spectroscopy. An overlay of both hairpin loops is shown in Fig. 7a; for this overlay, the terminal C₁:G₈ pairs were matched in chimera. Note the stackings of U₂, U₃ and U₄, and of U₅ on A₇ in 1JU7 (red), while U₃, U₅ and A₇ stack in 1KKS (cyan). In 1JU7 the C₆ is bulged out to the major groove, whereas in 1KKS the C₆ is bulged out to the minor groove. According to FRABase, further PDB entries (1ZBH, Cheng & Patel, unpublished; 4L8R and 4QOZ)⁵⁷ are similar to HP; each of these contains two of the histone stem-loops bound by protein (for example see Fig. 7b). These loop structures are very close, but not identical to 1JU7 and 1KKS, respectively, due to the interactions with protein, like stacking of Tyr₁₄₄(Exo)/U₃ and His₁₉₅(Exo)/C₆ and numerous hydrogen bondings between protein side chains and RNA backbone and bases. The sequence of the histone mRNA stem-loop is highly conserved (Fig. 7c,d), which might be forced by interactions with the proteins. The most variable position is C₆ (24% A, 68% C, 1% G, 7% U) as in HP (C₁₈₉; Fig. 2). The differences between the sequence logos of the histone mRNA stem-loop and HP might be due to additional requirements of HP like stability in the (–) strand, but might also point to a clear biological difference (see Discussion).

We further searched for HP-loop variants from other viroids. The PDB entry 4V7E⁵⁸, a model of the *Triticum aestivum* 80S ribosome based on cryo-electron microscopy with resolution of <5 Å, contains in the 28S rRNA a hairpin loop with sequence ₇₆₇UUCU₇₇₀ closed by a C:G pair instead of the A:U pair in a TCDVd (ID GQ169709). The first and last loop nucleotide stabilize the loop by a cWW interaction, and U₇₆₉ and U₇₇₀ stack on each other

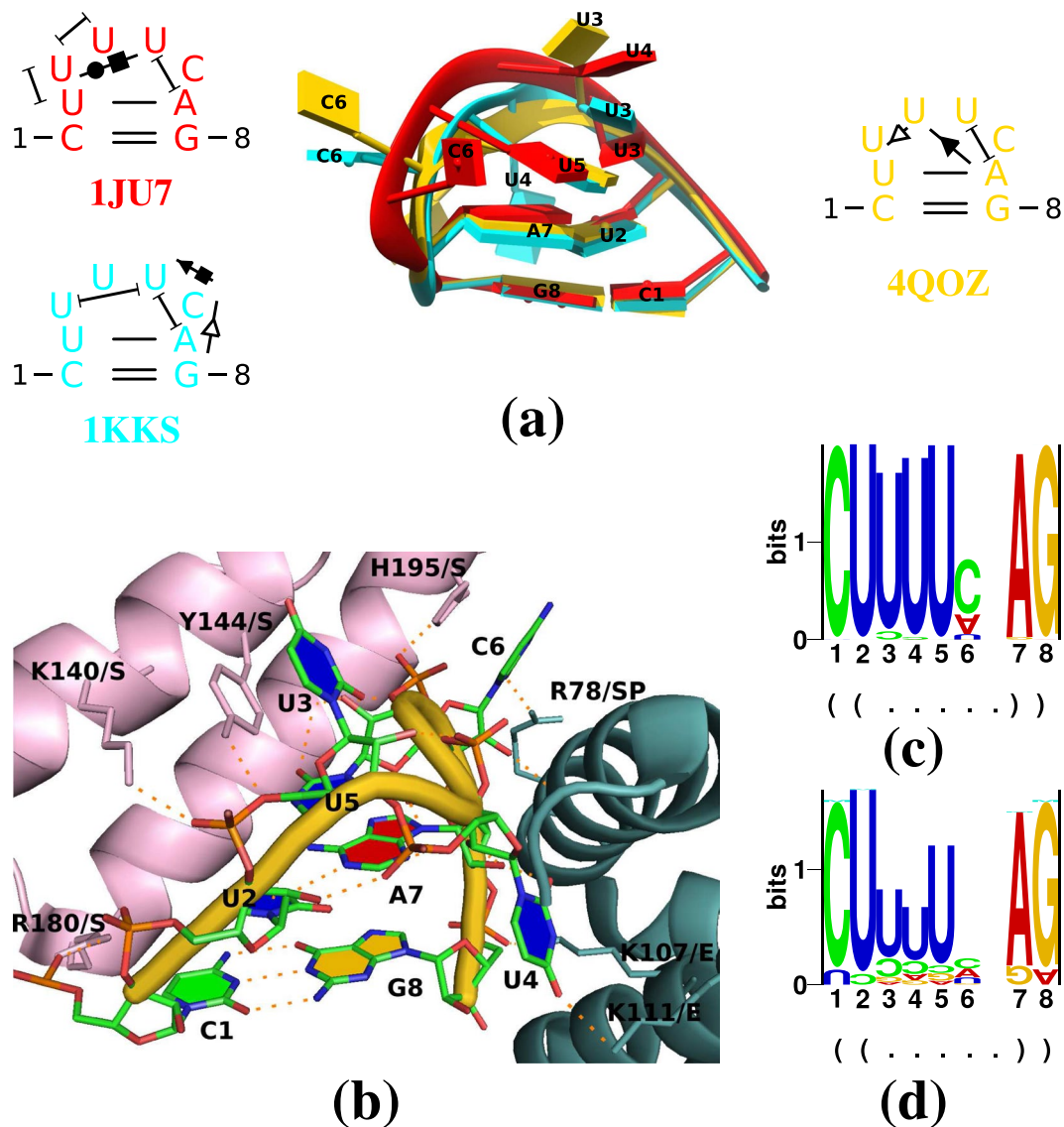


Figure 7. Histone mRNA stem-loop. This hairpin loop including the closing base pair is identical in sequence to PSTVd's HP (Fig. 5a). (a) Comparison of hairpin loops of PDB entries 1JU7 (red), 1KKS (cyan), and 4QOZ (gold); model 5 of 1JU7 and model 8 of 1KKS are shown, which are annotated as best structures in their respective PDB files. Views from different angles are shown in Supplementary Fig. S4. (b) Hairpin loop (cartoon in gold) from 4QOZ, bound by Exo (cartoon in lightpink) and SLBP (cartoon in darkgreen); a few amino acid side chains, which interact with RNA backbone and bases, are labelled. (c) Logo of 11,099 histone mRNA loops from Rfam-11⁷⁹ (ID: Histone3); three sequences have an additional nucleotide after position 6. (d) Logo of 36 non-redundant histone mRNA loops from Rfam-11.

(see Supplementary Fig. S5); both interactions might be independent from the loop-closing pair. The PDB entries 3J3V and 3J3W⁵⁹, which contain structures of the *Bacillus subtilis* 50S ribosome subunit based on cryo-electron microscopy and usage of ribosome structures of *Escherichia coli* and *Thermus thermophilus* as templates, show in the 23S rRNA a hairpin loop, similar to that of CEVd (see Supplementary Fig. S3b), that is involved in a kissing loop interaction (Fig. 8). Most interactions in the rRNA loop are due to the kissing interactions; however, the last loop nucleotide, in the rRNA as well as in the predicted CEVd loop, makes a cSH interaction with the 3' nucleotide of the loop-closing basepair.

Synthetic mutants in the TR domain. A synthetic mutant of HP and its progeny. Hammond⁶⁰ described a mutant PSTVd construct R+ (with the replacement $UU_{177} \rightarrow GAA$ in HP; see Supplementary Table S2) that was not infectious as transcribed RNA in contrast to the wild-type (WT) variant. After *Agrobacterium tumefaciens*-mediated inoculation of tomato plants with a corresponding plasmid containing a R+ cDNA insertion in (+) orientation, viroid progeny with secondary mutations appeared in galls and roots, and in 10% of infected plants also in leaves. The only variant (1 in Supplementary Table S2) appearing in leaves had, in comparison to WT, an inversion of the loop closing pair from $U_{176}:A_{181}$ in the WT to A:U, similar to the already

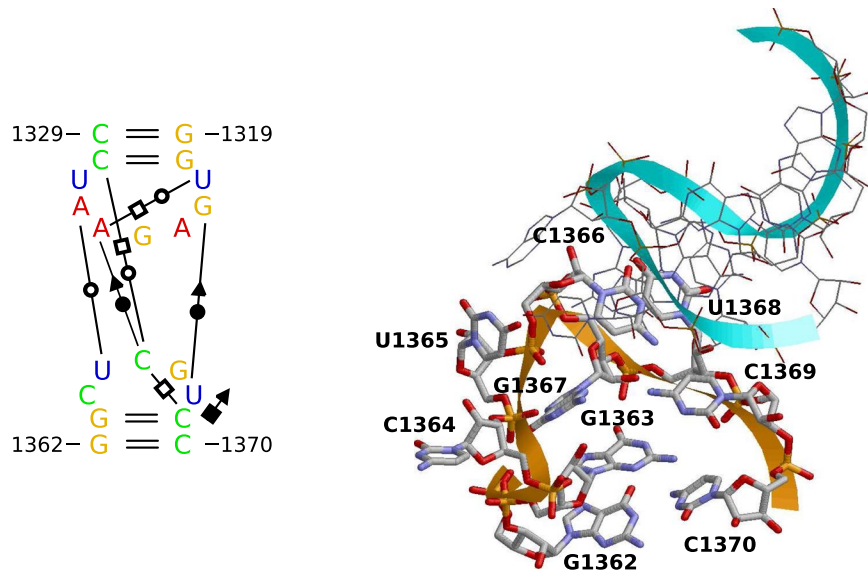


Figure 8. View on kissing-loop interaction. The nucleotides 1319–1329 and 1362–1370 from PDB entry 3J3W are visualized by rnaview (left) and by RasMol (right). The hairpin loop 1363–1369 is identical in sequence to the standard hairpin loop of CEVd.

mentioned TCDVd variant (Fig. 5b), and an additional U in the HP loop. The original variant R+, which was never recovered, and a further evolved variant (2 in see Supplementary Table S2) might be less viable due to a possible structural rearrangement influencing IL2 and the RY motif and thus did not point to critical loop nucleotides. The loop-closing pair and the loop size, however, seemed uncritical for replication and trafficking.

Binding affinity and infectivity of IL2 mutants. Gozmanova *et al.*¹² analyzed the binding of a Virp1 fragment to mutants of stem-loops, which were similar to the hairpin depicted in Fig. 1b. That is, their TR fragment contained both RY motifs, but only the terminal motif (as in Fig. 1c) was mutated. Infectivity of longer-than-unit-length viroid transcripts and relative binding affinity of stem-loops are given in Supplementary Table S3. From these data and further biophysical experiments, the authors concluded that the binding to the terminal RY motif was about fivefold stronger than to the internal motif.

Rosetta modelling of the synthetic IL2 mutants, which include single nucleotide replacements or a reverse complement of the loop sequence, predicted that these mutants possess at least one non-WC interaction different from but also additional to WT. Both modified interaction types might lead to an increased stability, rigidity and/or bending of the mutant loops, explaining the reduced binding of mutants to Virp1.

Inhibition of replication and trafficking by mutations in the TR domain. Zhong *et al.*⁵⁰ analyzed PSTVd loop motifs with respect to replication in protoplasts and to (replication plus) trafficking in plants. A replacement of the cWW pairs of IL1 by standard A·U WC pairs reduced replication efficiency to 29% of WT and abolished systemic infection. This experimental result fits to the findings from Rosetta modelling: either the lower thermodynamic stability of the WT loop is critical for replication competence, or the cWW loop pairs—in comparison to a loop without pairs—“protect” against endoribonucleases, or the backbone contraction in the WT loop is critical for protein binding. A replacement of the IL2 loop by three standard WC pairs reduced replication efficiency to 46% of WT and abolished systemic infection. A mutation of UU₁₇₆ → AA, which opens at least the HP-closing U·A pair, reduced replication efficiency to 53% of WT and abolished systemic infection. An additional mutation, which occurred in one plant, restored the loop-closing pair by an A·U pair that allowed this PSTVd mutant for systemic infection. This analysis points to a lower importance of loops in the TR domain for replication than loops in the TL and central domain but shows that these are most critical for trafficking⁵⁰.

Discussion

In the following we will discuss selected features from the predictions and compare several results from Rosetta modelling with findings from the structure databases.

Relevance of the two binding sites in TR. Figure 1b only shows the terminal Virp1 binding site with pairing of the R to Y motif, whereas the second, internal binding site is masked by different pairings. A structure with both Virp1 binding sites¹², is inferior to the one shown here by about 10 kJ/mol (mean $\Delta\Delta G_{37^\circ\text{C}}$ of 234 PSTVd variants) or a ratio of both structure ensembles (with two versus one perfect binding site) of about 50. This fits roughly to the different binding of Virp1 to both sites¹². However, only five of the 234 PSTVd variants miss the internal RY motif (ACs HQ452399–401, HQ452411, HQ452412)⁶¹, which points to the *in vivo* importance of a twofold RY motif.

IL1 is closed by two cWW pairs. The majority of IL1 is a 2×2 C·U forming cWW pairs; alternatives are a single cWW plus a standard WC pair (for example 5' CA/UU 3', 5' GA/GC 3'). These interactions fit to low BzCN reactivity of IL1⁶². A further “loop” stabilization by two A·U pairs with standard A-RNA geometry is prohibitive for viroid replication and abolishes trafficking⁵⁰.

Base-pairing in the asymmetric IL2 loop. According to the Rosetta model, only U₁₈₆ is bulged out while the three other nucleotides are paired. In contrast, the C₁₇₁ is bulged out in the 16S rRNA loop and makes contact to a distant base pair. In both cases, the logos for IL2 (Fig. 4c) and for the rRNA loop (Fig. 6b and c) do not point to compatible mutations of each motif. This fits to the higher degree of sequence conservation of IL2 than of IL1.

The HP structure allows for interaction despite its stability. According to DeJong *et al.*⁵⁵, the “stacking of loop nucleotides [in the histone mRNA stem-loop that is quite similar to HP] has the net effect of extending the helix by an additional base pair step”, which accounts for the relatively high thermodynamic stability of HP. However, the mRNA hairpin still allows for interaction with an RNA binding protein (Fig. 7) or for a kissing-loop interaction (Fig. 8).

In our opinion, however, the similarity between HP and histone mRNA stem-loop should not be taken as a hint towards biologically relevant interactions of viroid RNA with SLBP and Exo: (i) sequences similar to SLBP are not found by BLAST in higher plants; (ii) sequences similar to Exo are present in plant genomes, but these miss the SAP domain responsible for RNA binding; and (iii) histone 3' end processing is believed to be restricted to metazoa and green algae^{63,64}. Thus the detected similarity might be a pure consequence of the recurrence of evolutionary unrelated RNA motifs⁶⁵.

This work was inspired by the numerous tools for 3D RNA structure prediction that all promise to be manageable by non-experts. Indeed, our results point to an overall TR structure with a nice combination of rigidity and flexibility allowing the TR for binding of proteins. We have, however, not found a reason for the proposed TFIIIA binding to the TR region. Taken together, the predicted structure is encouraging for further exploration by mutational analysis and experimental structure determination.

Methods

Viroid sequences and consensus structures. For each of the 10 species of genus *Pospiviroid*⁴ a single sequence was used to search with nucleotide BLAST for somewhat similar sequences in GenBank. For sequences from each species, a preliminary alignment was produced with MAFFT/G-INS-i⁶⁶. According to these alignments, partial sequences were removed and the remaining sequences were adjusted for same strand orientation and similar start/end points. After a second alignment, redundant sequences were removed; that is, sequences from each genus had to differ from each other by at least one mutation. The final alignments were optimized in ConStruct⁶⁷ at 37 °C, excluding lonely base pairs, and drawn with R2R⁶⁸. An example is the consensus structure of PSTVd (Fig. 1a) which was based on 234 unique sequences. From the PSTVd alignment, the TR domain was cut out, redundant sequences were removed, aligned with MAFFT/X-INS-i⁶⁹, and optimized in ConStruct (Fig. 1b). From all individual alignments, the range of the terminal stem-loop of three helices was cut out, pooled, redundant sequences were removed, and aligned with MAFFT/X-INS-i. This alignment is shown in Supplementary Fig. S2, the corresponding R2R drawing in Fig. 1c.

Rosetta modelling. We used the Rosetta web server (ROSIE (Rosetta Online Server that Includes Everyone)/FARFAR (Fragment Assembly of RNA with Full Atom Refinement)/RNA De Novo) to predict 3D structures³³. FARFAR³⁴ assembles RNA fragments with length ≤ 3 nucleotides known from x-ray structures; these fragment sequences match partial sequences of the target RNA. The assembly is a Monte Carlo (MC) process guided by a knowledge-based energy function. Resulting structure models are then refined in an all-atom potential that is thought to discriminate native-like from non-native conformations. We used the following options: full run; variation of bond lengths and angles; optimization after fragment assembly; allowance for bulges; usage of updated (2012) force field; 1,000 structures; 10,000 MC cycles. From the predicted 3D structure(s), we extracted information on types of base pairs and further interactions by rnaview⁷⁰ and x3dna-dssr⁷¹.

RNAredesign⁴⁵ samples the possible sequence space that stabilizes a given 3D structure with a fixed structure backbone.

Database searches. The databases RNABricks³⁵ and FRABase³⁶ store short RNA motifs extracted from experimentally determined RNA 3D structures. Both databases allow a user to search for sequences similar to a query. Additionally, RNABricks includes a 3D search engine; as input we used 3D models predicted by ROSIE.

Further software. chimera v1.10.1⁷², PyMOL (<https://www.pymol.org/>) with nuccyl (<http://www.biosci.ki.se/groups/ljo/software/nuccyl.html>), and RasMol v2.7.5⁷³ were used to visualize 3D structures.

References

- Palukaitis, P. What has been happening with viroids? *Virus Genes* **49**, 175–184, doi:10.1007/s11262-014-1110-8 (2014).
- Navarro, B. *et al.* Viroids: How to infect a host and cause disease without encoding proteins. *Biochimie* **94**, 1474–1480, doi:10.1016/j.biochi.2012.02.020 (2012).
- Flores, R. *et al.* Rolling-circle replication of viroids, viroid-like satellite RNAs and hepatitis delta virus: variations on a theme. *RNA Biol.* **8**, 200–206, doi:10.4161/rna.8.2.14238 (2011).
- Di Serio, F. *et al.* Current status of viroid taxonomy. *Arch. Virol.* **159**, 3467–3478, doi:10.1007/s00705-014-2200-6 (2014).
- Schindler, I.-M. & Mühlbach, H.-P. Involvement of nuclear DNA-dependent RNA polymerases in potato spindle tuber viroid replication: a reevaluation. *Plant Sci.* **84**, 221–229, doi:10.1016/0168-9452(92)90138-C (1992).
- Navarro, J., Vera, A. & Flores, R. A chloroplastic RNA polymerase resistant to tagetitoxin is involved in replication of avocado sunblotch viroid. *Virology* **268**, 218–225, doi:10.1006/viro.1999.0161 (2000).

7. Nohales, M.-Á., Flores, R. & Darós, J. Viroid RNA redirects host DNA ligase 1 to act as an RNA ligase. *P. Natl. Acad. Sci. USA* **109**, 13805–13810, doi:[10.1073/pnas.1206187109](https://doi.org/10.1073/pnas.1206187109) (2012).
8. Nohales, M.-Á., Molina-Serrano, D., Flores, R. & Darós, J. Involvement of the chloroplastic isoform of tRNA ligase in the replication of viroids belonging to the family *Avsunviroidae*. *J. Virol.* **86**, 8269–8276, doi:[10.1128/JVI.00629-12](https://doi.org/10.1128/JVI.00629-12) (2012).
9. Martínez de Alba, A., Sägeser, R., Tabler, M. & Tsagris, M. A bromodomain-containing protein from tomato specifically binds potato spindle tuber viroid RNA *in vitro* and *in vivo*. *J. Virol.* **77**, 9685–9694, doi:[10.1128/JVI.77.17.9685-9694.2003](https://doi.org/10.1128/JVI.77.17.9685-9694.2003) (2003).
10. Martínez de Alba, A. *Isolation and characterisation of viroid-binding proteins*. Thesis, Universidad del País Vasco, Bilbao, Spain (2000).
11. Maniataki, E., Martínez de Alba, A., Sägeser, R., Tabler, M. & Tsagris, M. Viroid RNA systemic spread may depend on the interaction of a 71-nucleotide bulged hairpin with the host protein VirP1. *RNA* **9**, 346–354, doi:[10.1261/rna.2162203](https://doi.org/10.1261/rna.2162203) (2003).
12. Gozmanova, M., Denti, M., Minkov, I., Tsagris, M. & Tabler, M. Characterization of the RNA motif responsible for the specific interaction of potato spindle tuber viroid RNA (PSTVd) and the tomato protein Virp1. *Nucleic Acids Res.* **31**, 5534–5543, doi:[10.1093/nar/gkg777](https://doi.org/10.1093/nar/gkg777) (2003).
13. Kalantidis, K. *et al.* Virp1 is a host protein with a major role in *Potato spindle tuber viroid* infection in *Nicotiana* plants. *J. Virol.* **81**, 12872–12880, doi:[10.1128/JVI.00974-07](https://doi.org/10.1128/JVI.00974-07) (2007).
14. Chaturvedi, S., Kalantidis, K. & Rao, A. A bromodomain-containing host protein mediates the nuclear importation of a satellite RNA of *Cucumber mosaic virus*. *J. Virol.* **88**, 1890–1896, doi:[10.1128/JVI.03082-13](https://doi.org/10.1128/JVI.03082-13) (2013).
15. Wang, Y. *et al.* A land plant-specific transcription factor directly enhances transcription of a pathogenic noncoding RNA template by DNA-dependent RNA polymerase II. *Plant Cell* **28**, 1094–1107, doi:[10.1105/tpc.16.00100](https://doi.org/10.1105/tpc.16.00100) (2016).
16. Wu, S. & Chiang, C. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J. Biol. Chem.* **282**, 13141–13145, doi:[10.1074/jbc.R700001200](https://doi.org/10.1074/jbc.R700001200) (2007).
17. Sanchez, R. & Zhou, M. The role of human bromodomains in chromatin biology and gene transcription. *Curr. Opin. Drug Discov. Devel.* **12**, 659–665 (2009).
18. Mujtaba, S., Zeng, L. & Zhou, M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26**, 5521–5527, doi:[10.1038/sj.onc.1210618](https://doi.org/10.1038/sj.onc.1210618) (2007).
19. Rao, R., Thelen, J. & Miernyk, J. In silico analysis of protein Lys-N^ε-acetylation in plants. *Front Plant Sci.* **5**, 381, doi:[10.3389/fpls.2014.00381](https://doi.org/10.3389/fpls.2014.00381) (2014).
20. Lin, Y. *et al.* Solution structure of the extraterminal domain of the bromodomain-containing protein BRD4. *Protein Sci.* **17**, 2174–2179, doi:[10.1110/ps.037580.108](https://doi.org/10.1110/ps.037580.108) (2008).
21. Riesner, D., Kaper, J. & Randles, J. Stiffness of viroids and viroid-like RNA in solution. *Nucleic Acids Res.* **10**, 5587–5598, doi:[10.1093/nar/10.18.5587](https://doi.org/10.1093/nar/10.18.5587) (1982).
22. Dingley, A., Steger, G., Esters, B., Riesner, D. & Grzesiek, S. Structural characterization of the 69 nucleotide potato spindle tuber viroid left-terminal domain by NMR and thermodynamic analysis. *J. Mol. Biol.* **334**, 751–767, doi:[10.1016/j.jmb.2003.10.015](https://doi.org/10.1016/j.jmb.2003.10.015) (2003).
23. Takeda, R., Petrov, A., Leontis, N. & Ding, B. A three-dimensional RNA motif in *Potato spindle tuber viroid* mediates trafficking from palisade mesophyll to spongy mesophyll in *Nicotiana benthamiana*. *Plant Cell* **23**, 258–272, doi:[10.1105/tpc.110.081414](https://doi.org/10.1105/tpc.110.081414) (2011).
24. Zhong, X., Tao, X., Stombaugh, J., Leontis, N. & Ding, B. Tertiary structure and function of an RNA motif required for plant vascular entry to initiate systemic trafficking. *EMBO J.* **26**, 3836–3846, doi:[10.1038/sj.emboj.7601812](https://doi.org/10.1038/sj.emboj.7601812) (2007).
25. Branch, A., Benenfeld, B. & Robertson, H. Ultraviolet light-induced crosslinking reveals a unique region of local tertiary structure in potato spindle tuber viroid and HeLa 5 S RNA. *P. Natl. Acad. Sci. USA* **82**, 6590–6594, doi:[10.1073/pnas.82.19.6590](https://doi.org/10.1073/pnas.82.19.6590) (1985).
26. Wang, Y., Zhong, X., Itaya, A. & Ding, B. Evidence for the existence of the loop E motif of *Potato spindle tuber viroid in vivo*. *J. Virol.* **81**, 2074–2077, doi:[10.1128/JVI.01781-06](https://doi.org/10.1128/JVI.01781-06) (2007).
27. Eiras, M., Kitajima, E., Flores, R. & Darós, J. Existence *in vivo* of the loop E motif in potato spindle tuber viroid RNA. *Arch. Virol.* **152**, 1389–1393, doi:[10.1007/s00705-007-0952-y](https://doi.org/10.1007/s00705-007-0952-y) (2007).
28. Zhong, X. *et al.* Tertiary structural and functional analyses of a viroid RNA motif by isostericity matrix and mutagenesis reveal its essential role in replication. *J. Virol.* **80**, 8566–8581, doi:[10.1128/JVI.00837-06](https://doi.org/10.1128/JVI.00837-06) (2006).
29. Baumstark, T., Schröder, A. & Riesner, D. Viroid processing: switch from cleavage to ligation is driven by a change from a tetraloop to a loop E conformation. *EMBO J.* **16**, 599–610, doi:[10.1093/emboj/16.3.599](https://doi.org/10.1093/emboj/16.3.599) (1997).
30. Schrader, O., Baumstark, T. & Riesner, D. A mini-RNA containing the tetraloop, wobble-pair and loop E motifs of the central conserved region of potato spindle tuber viroid is processed into a minicircle. *Nucleic Acids Res.* **31**, 988–998, doi:[10.1093/nar/gkg193](https://doi.org/10.1093/nar/gkg193) (2003).
31. Gas, M., Hernández, C., Flores, R. & Darós, J. Processing of nuclear viroids *in vivo*: an interplay between RNA conformations. *PLoS Pathog.* **3**, e182, doi:[10.1371/journal.ppat.0030182](https://doi.org/10.1371/journal.ppat.0030182) (2007).
32. Kladwang, W., VanLang, C., Cordero, P. & Das, R. A two-dimensional mutate-and-map strategy for non-coding RNA structure. *Nat. Chem.* **3**, 954–962, doi:[10.1038/nchem.1176](https://doi.org/10.1038/nchem.1176) (2011).
33. Lyskov, S. *et al.* Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). *PLoS One* **8**, e63906, doi:[10.1371/journal.pone.0063906](https://doi.org/10.1371/journal.pone.0063906) (2013).
34. Das, R., Karanicolas, J. & Baker, D. Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* **7**, 291–294, doi:[10.1038/nmeth.1433](https://doi.org/10.1038/nmeth.1433) (2010).
35. Chojnowski, G., Walen, T. & Bujnicki, J. RNA Bricks—a database of RNA 3D motifs and their interactions. *Nucleic Acids Res.* **42**, D123–31, doi:[10.1093/nar/gkt1084](https://doi.org/10.1093/nar/gkt1084) (2013).
36. Popeno, M. *et al.* RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinformatics* **11**, 231, doi:[10.1186/1471-2105-11-231](https://doi.org/10.1186/1471-2105-11-231) (2010).
37. Miao, Z. *et al.* RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA*, doi:[10.1261/rna.060368.116](https://doi.org/10.1261/rna.060368.116) (2017).
38. Miao, Z. *et al.* RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures. *RNA* **21**, 1066–1084, doi:[10.1261/rna.049502.114](https://doi.org/10.1261/rna.049502.114) (2015).
39. Cruz, J. *et al.* RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction. *RNA* **18**, 610–625, doi:[10.1261/rna.031054.111](https://doi.org/10.1261/rna.031054.111) (2012).
40. Dawson, W. & Bujnicki, J. Computational modeling of RNA 3D structures and interactions. *Curr. Opin. Struct. Biol.* **37**, 22–28, doi:[10.1016/j.sbi.2015.11.007](https://doi.org/10.1016/j.sbi.2015.11.007) (2015).
41. Magnus, M. *et al.* Computational modeling of RNA 3D structures, with the aid of experimental restraints. *RNA Biol* **11**, 522–536, doi:[10.4161/rna.28826](https://doi.org/10.4161/rna.28826) (2014).
42. Laing, C. & Schlick, T. Computational approaches to 3D modeling of RNA. *J. Phys. Condens. Matter* **22**, 283101, doi:[10.1088/0953-8984/22/28/283101](https://doi.org/10.1088/0953-8984/22/28/283101) (2010).
43. Leontis, N., Stombaugh, J. & Westhof, E. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* **30**, 3497–3531, doi:[10.1093/nar/gkf481](https://doi.org/10.1093/nar/gkf481) (2002).
44. Stombaugh, J., Zirbel, C., Westhof, E. & Leontis, N. Frequency and isostericity of RNA base pairs. *Nucleic Acids Res.* **37**, 2294–2312, doi:[10.1093/nar/gkp011](https://doi.org/10.1093/nar/gkp011) (2009).
45. Yesselman, J. & Das, R. RNA-Redesign: a web server for fixed-backbone 3D design of RNA. *Nucleic Acids Res.* **43**, W498–501, doi:[10.1093/nar/gkv465](https://doi.org/10.1093/nar/gkv465) (2015).

46. Gorodkin, J., Heyer, L., Brunak, S. & Stormo, G. Displaying the information contents of structural RNA alignments: the structure logos. *Comp. Appl. Biosci./Bioinformatics* **13**, 583–586, doi:10.1093/bioinformatics/13.6.583 (1997).
47. Wu, M., McDowell, J. & Turner, D. A periodic table of symmetric tandem mismatches in RNA. *Biochemistry* **34**, 3204–3211, doi:10.1021/bi00010a009 (1995).
48. Xia, T., McDowell, J. & Turner, D. Thermodynamics of nonsymmetric tandem mismatches adjacent to G-C base pairs in RNA. *Biochemistry* **36**, 12486–12497, doi:10.1021/bi971069v (1997).
49. Lescrinier, E. *et al.* Structure of the pyrimidine-rich internal loop in the poliovirus 3'-UTR: the importance of maintaining pseudo-2-fold symmetry in RNA helices containing two adjacent non-canonical base-pairs. *J. Mol. Biol.* **331**, 759–769, doi:10.1016/S0022-2836(03)00787-3 (2003).
50. Zhong, X., Archual, A., Amin, A. & Ding, B. A genomic map of viroid RNA motifs critical for replication and systemic trafficking. *Plant Cell* **20**, 35–47, doi:10.1105/tpc.107.056606 (2008).
51. Liquori, C. *et al.* Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**, 864–867, doi:10.1126/science.1062125 (2001).
52. Childs-Disney, J. *et al.* Structure of the myotonic dystrophy type 2 RNA and designed small molecules that reduce toxicity. *ACS Chem. Biol.* **9**, 538–550, doi:10.1021/cb4007387 (2013).
53. Zirbel, C. *et al.* Identifying novel sequence variants of RNA 3D motifs. *Nucleic Acids Res.* **43**, 7504–7520, doi:10.1093/nar/gkv651 (2015).
54. Cannone, J. *et al.* The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**, 2, doi:10.1186/1471-2105-3-2 (2002).
55. DeJong, E., Marzluff, W. & Nikonowicz, E. NMR structure and dynamics of the RNA-binding site for the histone mRNA stem-loop binding protein. *RNA* **8**, 83–96, doi:10.1017/S1355838202013869 (2002).
56. Zanier, K. *et al.* Structure of the histone mRNA hairpin required for cell cycle regulation of histone gene expression. *RNA* **8**, 29–46, doi:10.1017/S1355838202014061 (2002).
57. Tan, D., Marzluff, W., Dominski, Z. & Tong, L. Structure of histone mRNA stem-loop, human stem-loop binding protein, and 3'hExo ternary complex. *Science* **339**, 318–321, doi:10.1126/science.1228705 (2013).
58. Gogala, M. *et al.* Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion. *Nature* **506**, 107–110, doi:10.1038/nature12950 (2014).
59. Li, N. *et al.* Cryo-EM structures of the late-stage assembly intermediates of the bacterial 50S ribosomal subunit. *Nucleic Acids Res.* **41**, 7073–7083, doi:10.1093/nar/gkt423 (2013).
60. Hammond, R. *Agrobacterium*-mediated inoculation of PSTVd cDNAs onto tomato reveals the biological effect of apparently lethal mutations. *Virology* **201**, 36–45, doi:10.1006/viro.1994.1263 (1994).
61. Luigi, M., Luison, D., Tomassoli, L. & Faggioli, F. Natural spread and molecular analysis of pospiviroids infecting ornamentals in Italy. *J. Plant Path.* **93**, 491–495 (2011).
62. Giguère, T., Adkar-Purushothama, C. & Perreault, J. Comprehensive secondary structure elucidation of four genera of the family Pospiviroidae. *PLoS One* **9**, e98655, doi:10.1371/journal.pone.0098655 (2014).
63. Dávila López, M. & Samuelsson, T. Early evolution of histone mRNA 3' end processing. *RNA* **14**, 1–10, doi:10.1261/rna.782308 (2007).
64. Marzluff, W., Wagner, E. & Duronio, R. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat. Rev. Genet.* **9**, 843–854, doi:10.1038/nrg2438 (2008).
65. Lescoute, A., Leontis, N., Massire, C. & Westhof, E. Recurrent structural RNA motifs, Isostericity Matrices and sequence alignments. *Nucleic Acids Res.* **33**, 2395–2409, doi:10.1093/nar/gki535 (2005).
66. Katoh, K., Kuma, K., Miyata, T. & Toh, H. Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome Inform. Ser.* **16**, 22–33 (2005).
67. Wilm, A., Linnenbrink, K. & Steger, G. ConStruct: Improved construction of RNA consensus structures. *BMC Bioinformatics* **9**, 219, doi:10.1186/1471-2105-9-219 (2008).
68. Weinberg, Z. & Breaker, R. R2R – software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics* **12**, 3, doi:10.1186/1471-2105-12-3 (2011).
69. Katoh, K. & Toh, H. Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics* **9**, 212, doi:10.1186/1471-2105-9-212 (2008).
70. Yang, H. *et al.* Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res.* **31**, 3450–3460, doi:10.1093/nar/gkg529 (2003).
71. Lu, X. & Olson, W. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.* **3**, 1213–1227, doi:10.1038/nprot.2008.104 (2008).
72. Pettersen, E. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612, doi:10.1002/jcc.20084 (2004).
73. Bernstein, H. Recent changes to RasMol, recombining the variants. *Trends Biochem. Sci.* **25**, 453–455, doi:10.1016/S0968-0004(00)01606-6 (2000).
74. Keese, P. & Symons, R. Domains in viroids: Evidence of intermolecular RNA rearrangement and their contribution to viroid evolution. *P. Natl. Acad. Sci. USA* **82**, 4582–4586, doi:10.1073/pnas.82.14.4582 (1985).
75. López-Carrasco, A. & Flores, R. Dissecting the secondary structure of the circular RNA of a nuclear viroid *in vivo*: a “naked” rod-like conformation similar but not identical to that observed *in vitro*. *RNA Biol.* doi:10.1080/15476286.2016.1223005 (2016).
76. Gast, F., Kempe, D., Spieker, R. & Sänger, H. Secondary structure probing of potato spindle tuber viroid (PSTVd) and sequence comparison with other small pathogenic RNA replicons provides evidence for central non-canonical base-pairs, large A-rich loops, and a terminal branch. *J. Mol. Biol.* **262**, 652–670, doi:10.1006/jmbi.1996.0543 (1996).
77. Leontis, N. & Westhof, E. Geometric nomenclature and classification of RNA base pairs. *RNA* **7**, 499–512, doi:10.1017/S1355838201002515 (2001).
78. Carter, A. *et al.* Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**, 340–348, doi:10.1038/35030019 (2000).
79. Burge, S. *et al.* Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.* **41**, D226–32, doi:10.1093/nar/gks1005 (2012).

Acknowledgements

I thank Prof. Mina Tsagris (Dep. Biology, Voutes University Campus, Heraklion, Crete, Greece) who initiated the project.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-00764-x

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017