BioMol Concepts, Vol. 2 (2011), pp. 171–181 • Copyright © by Walter de Gruyter • Berlin • Boston. DOI 10.1515/BMC.2011.016

Review

It's a loop world – single strands in RNA as structural and functional elements

Christian Schudoma

Bioinformatics Group, Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, D-14476 Potsdam/Golm, Germany

e-mail: schudoma@mpimp-golm.mpg.de

Abstract

Unpaired regions in RNA molecules – loops – are centrally involved in defining the characteristic three-dimensional (3D) architecture of RNAs and are of high interest in RNA engineering and design. Loops adopt diverse, but specific conformations stabilised by complex tertiary structural interactions that provide structural flexibility to RNA structures that would otherwise not be possible if they only consisted of the rigid A-helical shapes usually formed by canonical base pairing. By participating in sequence-non-local contacts, they furthermore contribute to stabilising the overall fold of RNA molecules. Interactions between RNAs and other nucleic acids, proteins, or small molecules are also generally mediated by RNA loop structures. Therefore, the function of an RNA molecule is generally dependent on its loops. Examples include intermolecular interactions between RNAs as part of the microRNA processing pathways, ribozymatic activity, or riboswitch-ligand interactions. Bioinformatics approaches have been successfully applied to the identification of novel RNA structural motifs including loops, local and global RNA 3D structure prediction, and structural and conformational analysis of RNAs and have contributed to a better understanding of the sequence-structure-function relationships in RNA loops.

Keywords: RNA bioinformatics; RNA function; RNA interactions; RNA loops; RNA structure.

Introduction

Almost three decades have passed since the discovery of catalytically active RNA molecules in the late 1970s/early 1980s (1, 2). No longer was RNA function known to be limited to information-storage (mRNA, viral RNA), amino acid transport (tRNA), and 'structural' component of the ribosome (rRNA). These 'new' RNAs – the ribozymes – RNA molecules capable of enzymatic activity, the domain of proteins, heralded a biological paradigm shift resulting in a revision of the central dogma of molecular biology. They also posed supporting evidence to the theory of an ancient RNA world (3). Since then, a large number of RNA 3D structures have been experimentally determined [1884 RNA-containing entries in the NDB (4) as of January 2011], mainly by X-ray and NMR technologies, shedding light on the mechanism of RNA function. A usually single-stranded molecule, RNA is capable of folding into complex hierarchical 3D structures. By forming Watson-Crick base pairs between purine and pyrimidine bases, the molecule adopts its secondary structure, resulting in alternating regions of paired and unpaired bases. The subsequent formation of a complex network of base pairs within and between unpaired regions, the tertiary structure, gives rise to the three-dimensional conformation of the RNA molecule. While the majority of bases in an RNA belong to the paired regions (stems), it is the unpaired regions, or loops, that harbour the functionality of most RNAs or serve as important structural features. This review aims to be an overview over the field of RNA loops, taking into account different loop types, characteristic structural motifs, and the roles of RNA loops in RNA function and structure, e.g., in current RNA research areas such as riboswitches, synthetic aptamers, and micro RNAs. Additionally, we provide a short excursion to the field of (RNA) bioinformatics, since RNA loops are a central part of motif detection, e.g., in the search for regulatory elements in genomic sequences. Furthermore, understanding loop structures is essential for RNA 3D structure prediction and modeling.

Loops in RNA architecture

The 3D structure of RNA molecules is based on a three-tiered (primary, secondary, and tertiary structure) hierarchical architecture [e.g., (5)]. The Watson-Crick paired regions of the secondary structure usually form A-form helical structures, resulting in thermodynamically stable and rigid regions in the RNA molecule. In contrast, loop regions play a role in the spatial arrangement or 'packing' of paired regions (6) by introducing both flexibility and additional stability (through the formation of tertiary structure base pairs) into the RNA architecture, thus exerting control on the molecule's global fold (Figure 1). Intrinsically being parts of the 3D structure, the different types of loops (loop motifs) can nevertheless be distinguished, albeit in a coarse-grained fashion, at the secondary structure level. At this level, RNA architecture follows a simple stem-loop rule, i.e., each base-paired region (stem) ultimately has to be connected to an unpaired region of one of the three types hairpin loop, internal loop/

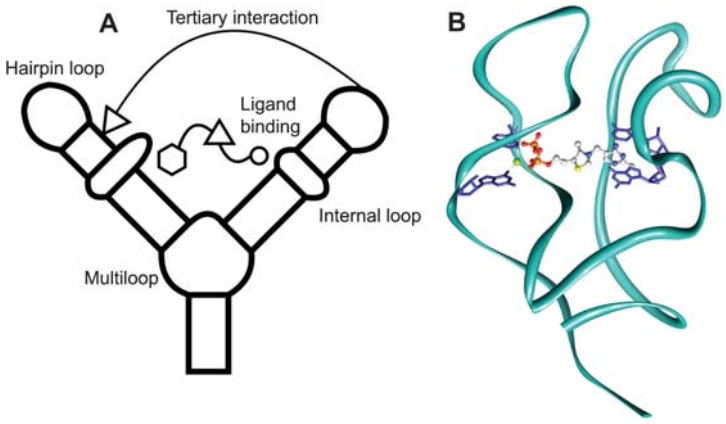

Figure 1 Loops in RNA structure.
(A) Interactions and structural elements in a TPP-riboswitch aptamer domain (98). A-helical base-paired regions are given as rectangles, loops as curve drawings. (B) Ribbon-visualisation of the 3D sugar-phosphate backbone trace (turquoise) of the TPP-riboswitch aptamer domain. Bases involved in ligand binding are depicted in dark blue, the TPP-ligand is coloured by atom type. Molecular graphics were generated with Chimera (143).

bulge, and multiloop. The tertiary structure level provides enough information (i.e., base pair pattern, base stacking pattern) to obtain a fine-grained loop classification and even to infer 3D structure. While in fact certain sequence motifs (e.g., GNRA hairpin loops) describing individual loop types exist, the small nucleotide alphabet size implies high sequence ambiguity (in contrast to proteins) and thus makes it difficult to infer specific RNA loops from the primary structure of a single RNA alone.

Hairpin loops

Hairpin loops arise as a consequence of the RNA strand folding back unto itself, therefore 'terminating' or 'capping' base-paired regions. They promote both intra- and extra-molecular interactions and thus are essential elements for RNA structure and function. Interactions with other hairpin loops (so-called 'kissing hairpins') constrain 3D conformational space for an RNA molecule and often give rise to characteristic molecular shapes, e.g., the L-fold of tRNA or the group II intron 'wishbone'. Furthermore, hairpin loops are important in RNA recognition [e.g., the tRNA anticodon (deca-) loop] and RNA-protein interactions (5, 7, 8). The most prominent examples of hairpin loops are tetraloops [>50% of all hairpins (9, 10)], which are frequently found and well-conserved especially in ribosomal RNAs (11, 12). Nevertheless, they are also present across a large variety of different RNAs such as the hammerhead ribozyme (13, 14) and the P4-P6 domain of group I self-splicing introns (15). Tetraloops have been extensively studied, both structurally and thermodynamically and characteristic thermodynamic parameters have been established [e.g., (16)].

Most ribosomal tetraloops confer to well-known sequence patterns such as GNRA, UNCG, and rarely CUUG with generally conserved but not completely invariant structural patterns. These tetraloops are known to be extraordinarily thermodynamically stable [e.g., (7)].

While GNRA tetraloops commonly contain U-Turn motifs [(13), s. below], other conformations with different stacking patterns or bulged out single bases are known (17). CUUG tetraloops are especially interesting, since their bases C1 and G4 can form a Watson-Crick base pair (18, 19), resulting in the formation of a di-loop structure violating the usual hairpin loop minimum size rule of three nucleotides. Interestingly, also five-membered hairpin loops (pentaloops) can conform to the architecture of GNRA or UNCG tetraloops. In such cases, the extra base (often at position 4) is bulged out of the loop, while the remaining 4 bases adopt e.g., a GNRA-like fold (20, 21). Additionally, GNRA-like folds can occur in internal loop environments (22).

Two similar structural motifs, U-turn and T-loop, are commonly found in hairpin loops involved in RNA-RNA (both intra- and extra-molecular) or RNA-protein interactions. Either of these motifs induce a conformational change in the sugar-phosphate backbone, leading a number of bases in the loop to 'bulge out' and make them accessible for long-range interactions.

The U-turn The U-turn (or Uridine Turn) motif was first discovered in the TΨC loop of $tRNA^{phe}$ (23). U-turns are also present in tRNA anticodon loops (24) and additionally occur in other RNA species, e.g., in the hammerhead ribozyme (13), 23S rRNA (25–27), U2 snRNA (28), and HIV RNA (29). The presence of a U-turn induces a sharp directional change of the RNA backbone, mediated by a local base pairing interaction network involving a Uridine, Pseudouridine, or Guanosine (30) residue. U-turns are important structural elements for both intra- and intermolecular interactions. In the case of the anticodon loop, the U-turn-mediated change in backbone direction forces the three anticodon bases to bulge out and therefore makes them accessible for interaction with the codon triplet and the ribosomal P-Site (31, 32). Gutell and coworkers used comparative sequence analysis to predict occurrences of U-turns in 16S and 23S

ribosomal RNA, identified a U-turn consensus structure, and established 10 distinct U-turn families according to their sequence/structural context (33).

The T-loop Like the U-turn, the T-loop motif was first observed in the TΨC loop of $tRNA^{phe}$ (23). It has since then found in tmRNA (34), in the TYMV genome (35), and in bacterial ribosomal subunits of Thermus thermophilus (30S) and Haloarcula marismortui (50S) (36). Two types of T-loop are known, both following a consensus structure based on a *trans* WC/H U-A base pair (37) stacked on a canonical WC base pair. T-loop motifs allow for the formation of loop-loop interactions such as the D-loop/T-loop interaction in tRNA. T-loop-mediated loop-loop interactions are also frequently observed in rRNA, including T-loop-T-loop interactions involving different T-loop types. Furthermore, interactions between T-loop motifs in rRNA hairpin loops and ribosomal proteins L4, L15, L23, L24, S14, S19, and S20 have been observed (36).

Internal loops

Internal loops are unpaired regions connecting exactly two stems. An internal loop is 'symmetric' if both of its strands are of equal length and 'asymmetric' if they are not. Bulges are special asymmetric internal loops with only one unpaired strand. Bases in bulged regions are either stacked between the two flanking stems or extrude from the stem. In the former case, a kink is introduced into the structure between the two subsequent stems. In the latter case, the stems can form a virtually uninterrupted A-form helix [e.g., (38)]. Furthermore, extruding bases can be packed into one of the helix grooves or function as a 'flap' closing ligand binding sites (39).

The A-minor motif In ribosomal RNAs, Adenosine residues were found to be overrepresented in bulged regions and at the same time being only minimally exposed to solvent. The bulged A's were observed to participate in interactions with the minor groove of base paired regions, i.e., in loop-stem interactions (40). Facing the minor groove of the target base pair with the Hoogsteen-edge (37) of its base, the Adenosine contacts the Sugar-edges of both paired bases (type I) or with the ribose-O2′ atom of the nearer of the bases (type II). Additionally, there is one variant where the Adenosine pairs with a ribose-O2′ atom via its Watson-Crick edge (type III) and another – rare – variant with the Adenosine placing its ribose into the minor groove (type 0). A composite motif, the A-patch, is formed by multiple stacking Adenosines participating in A-minor motifs and takes part in RNA-protein interactions.

The A-minor (Adenine – minor groove) motif has been found across different RNA species and is now seen as an, if not one of the most, essential structural building block(s) for RNA 3D structure formation. The A-minor motif has also been found playing a role in intermolecular interactions, for instance in the translational decoding recognition process (41).

Kink-turns Kink- or K-turns (42) are asymmetric internal loops first observed in the 50S rRNA of Haloarcula marismortui. One has recently been discovered in the Bacillus subtilis yitJ SAM-I Riboswitch (43). The motif induces a kink in the sugar-phosphate backbone, bending the axis between two flanking helical regions by approximately 120° and thus bringing the two minor grooves into proximity. Kink-turns have been observed interacting with various ribosomal proteins of the large (L4, L7Ae, L10, L15e, L19e, L24, L29, and L37Ae) and small (S11 and S17) subunits as well as promoting RNA tertiary structure interactions. Furthermore, they are believed to play a role in the transport of RNA in neuronal and glial cells (44). One of the two flanking helical regions, the C-Stem ('canonical stem'), consists of canonical base pairs. The second flanking helical region, the NC-Stem ('non-canonical stem') starts with two non-canonical [usually G-A sheared (37)] base pairs. These G-A/A-G base pairs appear to be essential for the kink-turn formation and structures with exchanged base pairs will not adopt the kinked conformation. An exception to that rule is Kt-23 in the Thermus thermophilus 30S ribosomal subunit which has an A-U pair replacing the bulge-distal A-G pair and is still capable of forming a kink-turn structure *in vitro* (45).

A reverse kink-turn (46) is an internal loop motif first discovered in an Azoarcus group I intron (47). It contains a similar sharp bend as observed in kink-turns, albeit with a curvature of approximately 90° and into the opposite direction, leading to a juxtaposition of the major grooves. The high sequence similarity between the kink-turn consensus and the J9/9.0 reverse kink-turn connecting helices J9 and J9.0 in the Azoarcus group I intron poses the question as to why the two motifs bend differently. A possible explanation partially supported by fluorescence studies is that internal loops following the kink-turn consensus exist in a dynamic 3-state equilibrium (unbent state, kinked, reverse-kinked) (48). In fact, it is known that at least the kinked state exists in such a dynamic equilibrium between tightly and loosely kinked. The different known conformational states of kink-turns make them interesting targets for molecular dynamics studies [e.g., (49)]. The tightly kinked state is generally dependent on the presence of divalent metal ions (48). Another possibility would be the dependence on the presence of external factors, such as interactions with proteins or through RNA tertiary structure elements [as confirmed by mutation studies (50)]. In such a case, kink-turns would not qualify as primary building blocks of RNA structure (48).

C-loops C-loops (21) are asymmetric internal loops involved in RNA-protein interactions. They have been observed in 16S and 23S rRNAs and in threonyl-tRNA-synthetase (thrRS) mRNA (51). In the latter case the C-loop facilitates the interaction between the mRNA and thrRS and therefore allows thrRS to repress the translation of its own mRNA. The base pair patterns of C-loops have been thoroughly analysed (52).

Sarcin/ricin loops The motif of the sarcin/ricin asymmetric internal loop is universally conserved in the 5S, 16S,

and 23S rRNA of Archaea, bacteria, and eukaryotes, as well as the hairpin ribozyme loop B and the PSTV conserved central domain (53, 54). Sarcin/ricin loops are platforms for intra- (e.g., interdomain packing), as well as intermolecular interactions with other RNAs, proteins (e.g., ribosomal protein L15E), and small compounds.

UA_handles The UA_handle (55) is a bulge motif generally following the consensus sequence (5′XU/ANnX3′). An n-nucleotide-long bulge is flanked by a U-A Watson-Crick/Hoogsteen (37) and a canonical Watson-Crick base pair (X-X), with the flanking pairs stacking on each other. Often a directional backbone change is observed at the Adenosine residue of the U-A pair. Two major types of the motif can be distinguished. Instances of type I have a bulge of one, three, or more nucleotides and the Watson-Crick pair is usually C-G, while instances of type II have a bulge of length 2 with a G-C Watson-Crick base pair. UA_handles appear to be present in a wide range of structures and are involved in the formation of tertiary structure, serving as a 'handle' for long-range contacts.

Multiloops

Multi (-branched) loops (or junctions), i.e., unpaired regions connect three or more stems. They play a central role in RNA architecture (56). The structural complexity of an RNA molecule increases with the presence of multiloops. The simplest possible RNA secondary structure is a stem with terminating hairpin loop. Adding extra stems, interspersed by internal loops and/or bulges keeps the original stem-loop helical 3D structure generally intact, possibly slightly bent via one or more of the internal loops. In contrast, the base pair interactions within multiloops may exert direct influence on the global conformation of an RNA molecule. They either cause the individual stems to branch into different directions or to adopt a structure where multiple stems are stacked coaxially, therefore appearing as one long uninterrupted stem. Additionally, the number of hairpins in an RNA secondary structure is directly related to the number of multiloops. Each multiloop gives rise to a number of stems and according to the stem-loop structure each stem has to be terminated by a loop. Ultimately, each stem ends in one or more hairpin loops, allowing for more loop-loop interactions to be formed and thus additionally increasing structural complexity. The topologies and conformations of the most common types of multiloops [three-branched (57) and four-branched (58) junctions] have been studied and tertiary structure motifs have been found in higher-order junctions (59).

Intramolecular interactions

In an RNA molecule, tertiary interactions can form between any two loops. Prominent examples are interactions between the D-Loop and T-Loop in tRNA (23) and tmRNA (34, 60–63). The function of riboswitches, i.e., the conformational change ('switch', s. below) depends on the formation of long-range tertiary structure contacts [e.g., (64)]. Another example is the D5 bulge A376-C377 in the Oceanobacillus ihyensis group-IIC intron (65, 66), which takes part in tertiary structural interactions important for the catalytical function of the intron while forming an unusual backbone conformation with geometrical similarities to peptide α-helices. A376 is involved in the orientation of the 5′-terminus via a stacking interaction and its backbone moieties coordinate a divalent metal ion implicated in catalysis. C377 serves as the terminal stack of the catalytic triplex between domain D5 and the J2/3 linker via a base triple interaction with C360 and G383 (66). Intramolecular interactions in RNAs have been reviewed e.g., in (67) and the complex tertiary structure interaction networks that can arise in RNAs have been examined e.g., in (68).

RNA loop functionality

RNA loops guide cleavage processes

Loops in RNA molecules play essential roles in regulatory processes, such as RNA cleavage mediated by ribonucleases of the RNAse III family (69). While the exact molecular mechanisms have not yet been solved in 3D, bioanalytic approaches have found evidence for the importance of RNA loop regions for these processes. Generally spoken, RNA loops appear to serve some kind of landmark-function for enzymatic RNA cleavage with ribonuclease-binding and –activity being dependent on the presence of certain structural and sequence features. The substrates of RNAse III enzymes usually are short stem-hairpin structures, which are selectively recognised. In a T7 R1.1 substrate of *Escherichia coli* RNase III, a single cleavage event occurs within an internal loop region (70). Single cleavage sites in the 5-strand of an asymmetric internal loop have been observed in other RNase III substrates as well (71). An interesting example is RNA cleavage mediated by the yeast RNase III Rnt1p. Targets of Rnt1p generally contain an AGNN-tetraloop hairpin with cleavage occurring 13–16 base pairs away from the terminal loop. The hairpin is believed to be essential for substrate recognition (72, 73). It was observed that in addition to the hairpin sequence, the sequences of the first two base-pairs adjacent to the loop exert strong influence on Rnt1p-binding and – activity (74). NMR-structures for two different Rnt1p targets (9 base pairs capped by an AGUC or AGAA tetraloop) showed a common fold of the loop regions, with the Guanine and Adenine stacking, and the 5′-Adenine forming a non-canonical base pair with the 3′-Adenine respectively Cytosine. In both loops, the Guanine is in *syn* conformation mediated by base-phosphate contacts. The contact with Rnt1p is likely to be made via the Hoogsteen edges (37) of the 5′-A and G, as well as the Watson-Crick edge of the 5′-A. The similar folds suggest a shape-specific recognition event independent of sequence for Rnt1p targets (75).

Loop structures in miRNA processing pathways

The cleavage of microRNA transcripts (pri-miRNAs) and microRNA precursors (pre-miRNAs) by the Drosha, or

respectively, Dicer enzymes has been found to depend on the terminal loop region (a stem-internal-loop-stem-hairpin structure), which is believed to be structurally flexible (76). Dicer-like 1 (DCL1) cleavage often occurs at a distance of about 15 nucleotides away from an unpaired region (77, 78). Deletion of the terminal loop completely abrogates the accumulation of Arabidopsis thaliana miR172a (79). *In vitro*, larger hairpin loops facilitate both Drosha and Dicer cleavage of human microRNAs miR-16, miR-30, and miR-31 (while structures with more base pairs result in decreased cleavage activity), suggesting that the accessibility of bases plays a role in this otherwise unknown structural mechanism (76). A bit further down the miRNA-mediated silencing pathway, loading of miRNAs into the RNA-induced silencing complex (RISC) is influenced by the presence or absence of bulges within the miRNA duplex (80). Furthermore, 'quaternary structure' bulges in the central region of imperfect animal miRNA-mRNA hybrids have been found to inhibit translation or promote mRNA decay (81).

RNA loops drive RNA-ligand interactions

Interactions with target molecules, such as small molecular compounds as well as larger biomolecules, are an important aspect of RNA function. The 'binding platforms' for these interaction events are usually located in internal loop [e.g., the Tat and Rev regulatory proteins involved in HIV replication (82) or aminoglycoside antibiotics such as paromomycin (83)] or multiloop environments [e.g., purine riboswitches (84, 85)]. The potential of RNA as a drug target has been extensively studied and was reviewed e.g., in (86). About two decades ago, the so-called selection techniques of '*In vitro* selection' (87) and 'systematic evolution of ligands by exponential enrichment' (SELEX) (88) were first presented. They made it possible to synthesise nucleic acid sequences ('aptamers') that could bind to a target molecule (organic dyes in the former case, bacteriophage T4 DNA Polymerase in the latter) with very high specificity. Aptamers have been extensively studied in order to assess their therapeutic and diagnostic potential [e.g., (89–91)]. NMR structures of aptamers in complex with different ligands (such as cofactors, drugs, amino acids, and aminoglycid antibiotics) were reviewed e.g., in (92). A current example of *in vitro* selected RNA is a flexizyme that, due to containing a phenylalanine-binding site, is capable of specific $tRNA^{phe}$ aminoacylation (93).

Later, actual natural examples of these high-specificity binding-platforms as part of highly sophisticated molecular switches ('riboswitches') controlling gene expression by direct and specific sensing of metabolite levels were discovered in bacterial mRNA molecules (94–97). A riboswitch is generally a two-platform system, containing an aptamer platform for binding a respective metabolite and an effector or expression platform which, via conformational change after ligand binding to the aptamer platform, promotes or represses the expression of the gene encoded by respective mRNA on the transcriptional, post-transcriptional, or translational level. The bases of the aptamer domain must be accessible (generally by being part of a loop) for metabolite-binding in order for the switch-mechanism to work, rendering gene expression control via riboswitches loop-dependent. Crystal structures are available for a number of riboswitch-ligand complexes [reviewed e.g., in (64)], allowing the mode and location of ligand-binding to be assessed. The G- and A-bacterial purine riboswitches (84, 85) bind their target in a buried and therefore solvent-inaccessible multiloop environment. The binding pocket of the eukaryotic and bacterial thiamine pyrophosphate (TPP)-sensing riboswitches (98, 99) is formed by two internal loops. One of these loops (the pyrimidine-sensor) adopts a T-loop like fold and is responsible to bind the pyrimidine residue of TPP while the other one (the phosphate-sensor) coordinates the negatively charged pyrophosphate-residue via two divalent metal ions. S-Adenosylmethionine (SAM) is partially bound via an internal loop and buried between two stems of the SAM-I riboswitch (100–103) and the glmS ribozyme-riboswitch (104, 105) has a glucosamine-6-phosphate (GlcN6P) binding pocket located in the linker-environment between two stems.

RNA-protein interactions [reviewed e.g., in (106, 107)] are a driving force of cellular mechanisms. In addition to the above mentioned C-loop interaction site, anticodon-like hairpin loops play a role in RNA-protein interaction in autotranslational control by threonyl-tRNA synthetase (51).

Loop bioinformatics

Bioinformatics approaches dealing with RNA loops or structural motifs in general usually aim at deriving sequence constraints of known motifs from 3D data and finding new motif instances within RNA sequence data. While methods such as comparative sequence analysis and *ab initio* predictions based on thermodynamics and statistical mechanics are core areas in RNA motif bioinformatics, this review focuses on methods that are based on experimentally determined three-dimensional structure information (and therefore can use actual tertiary structural information) and are applied in order to detect and analyse RNA loop structural motifs.

A recent study, determining the distribution of 3-, 4-, and 5-mer 3D motifs in 23S rRNA of the Haloarcula marismortui large ribosomal subunit was able to predict most of the 43 previously known tetraloop hairpins (108). Another recent approach based on dynamic programming using base pair patterns derived from 3D structure, RNAMotifScan, identified a high number of instances of five known loop motifs (from ribosomal RNA) in a set of 1445 RNA structures from the PDB. Interestingly, the numbers of discovered instances were significantly higher than the currently known motifs, despite 'rather stringent' cutoffs. Comparisons between newly discovered instances and the previously known motifs yielded sequence identities as low as 66%, somewhat questioning the applicability of purely sequence-based approaches. Furthermore, the motifs were found in non-ribosomal RNAs, supporting previous assumptions of universal RNA building blocks (109). Investigations of tertiary structure networks using graph-grammars, resulted in sequence constraints for sarcin/ricin loops (110). Detection of novel

structural motifs in RNA is performed by searching common structural patterns in the 3D data of different RNA molecules. Software tools such as FR3D ['Find RNA 3D' (111)], MC-Annotate (112), RNAVIEW (113), and 3DNA (114) apply geometrical [or network-theoretical (MC-Annotate)] procedures to detect base pairs and thus allow to find loop regions both on the secondary and tertiary structure level in RNA 3D structures. Another approach is the COMPADRES algorithm, using a reduced RNA backbone representation (the η/θ pseudotorsions). This approach has been successfully applied in the identification of new motifs, such as the π-turn, the Ω-turn, and the α-loop (115).

Three major bioinformatics-based studies on RNA loop structures have been published in the recent years. UPGMA cluster analysis was applied on the pairwise structural distances of RNA tetraloops and found major clusters corresponding to the GNRA and UNCG tetraloops, as well relationships between other sequence motifs and structural conformations (116). Sequence-structure relationships were discovered analyzing the tertiary structure networks of triloop hairpins (117) and methods from both studies were combined, discovering further sequence-structure relationships within hairpin loops and unpaired regions in general (118).

Databases

A number of public databases have been established to provide access to RNA structural motif data. The at the time available databases are SCOR ['Structural Classification of RNA' (119)], storing a fully hand-curated motif annotation of RNA structures in the PDB (last updated in 2004), RNA-Junction (120), a database holding information about multiloops ('junctions'), internal loops, and kissing hairpin motifs (last updated in 2008), RLooM ['RNA Loop Modeling' (118, 121)] storing 3D data for all kinds of secondary structure loop motifs (last updated in 2009), and FRABASE (frequent update schedule) (122) storing motif annotations for all structures in the PDB. The most recent addition to RNA structural databases is a database specifically dedicated to kink-turn-related data (123).

Structure prediction

Approaches on *ab initio* structure prediction of RNA loops date back at least into the early 1990s but appear to have become less popular nowadays. Using the constraint-satisfaction-based approach MC-SYM, tRNA hairpin loop models were modeled within 2–3 Å all-atom RMSD to the experimentally verified structures (124, 125). Loop modeling based on a genetic algorithm for the conformational search of tRNA hairpin loop structures, achieved models within 1.8 Å RMSD, albeit not among the energetically fittest structure models (126). The adaptation of a method from protein loop structure prediction using bond scaling and relaxation was applied for the prediction not only of tRNA and sarcin/ricin hairpin loops but also of different tRNA variable loops (a substructure of the central three-branched tRNA multiloop), thus expanding RNA loop structure prediction attempts beyond the rather constrained hairpin loops (127).

Traditional secondary structure prediction approaches based on stacking energies and partition functions [e.g., (128)] can be used in order to distinguish between stem and loop regions in RNA structures and have been applied for the prediction of coaxial stacking of stems connected by the same multiloop (129). Nowadays, RNA structure prediction approaches focus on modeling global structure, rather than local structural motifs. Nevertheless, all current methods should in theory be capable of predicting loop structures, since the underlying folding mechanisms are the same. In addition to the aforementioned constraint-satisfaction-based software tool MC-SYM [which in combination with the software MC-Fold has been applied to reproduce a series of experimentally determined RNA three-dimensional structures from sequence (130)], there are several other RNA structure prediction tools available. FARNA (131) is an *ab initio*/knowledge-based hybrid using potentials derived from ribosome structures. Its recent extension FARFAR (132) was successfully applied to the modeling of RNA structural motifs, yielding structural models between 1 and 2 Å RMSD. iFoldRNA (133, 134) is an *ab initio* approach based on a 3-bead-string model utilising discrete molecular dynamics simulations. NAST/C2S (135, 136) is a pipeline for the geometrical modeling of RNA backbone structures allowing for the incorporation of structural constraints (e.g., secondary structure, tertiary structure, small angle X-ray scattering data, etc.) in order to augment the prediction (NAST) and subsequential addition of atomic details (C2S). MANIP (137), S2S (138), and ASSEMBLE (139) are tools for homology modeling of RNA structures, available under the PARADISE web service (http://paradise-ibmc.u-strasbg.fr/). The most recent tool, ModeRNA (140) is an easy-to-use software package, designed for RNA homology modeling and modification of RNA structures, including RNA loops. The only current purely loop-focused approach is RLooM (118, 121). RLooM is based on a loop template database derived from experimentally verified structures from the NDB and uses a sequence/structure search in combination with geometrical fitting of loop structures into specified potential loop sites of RNA structures. As of now, homology-/geometry-based approaches seem to be the gold standard for RNA loop modeling, albeit being somewhat limited in their performance by the number of available experimentally verified template structures.

Expert opinion

Loops in RNA molecules are essential structural and functional building blocks and as such key to understanding both RNA folding mechanisms and structure-function relationships.

Finding exact sequence and base pair patterns in order to describe the different types of RNA loops and their associ-

ated structural elements are of high importance for solving the RNA 3D structure prediction problem and are a good starting point for identifying yet unknown RNA structural motifs.

In order to facilitate cooperation between different research groups in the field and avoid misunderstandings resulting from using different terminology [Is a 'hairpin' the whole construct of stem and loop or only the loop? What about a 'hairpin loop'? Are bulges simply special cases of internal (interior?) loops or does the term also describe the individual unpaired 'linker' regions of multi-(branched) loops (junctions?)? A 'helix' is a set of consecutive canonical *cis* Watson-Crick base pairs, but does it or does it not include internal loop regions mimicking the A-helical conformation?], an unambiguous description of RNA structural motifs is essential. Initiatives to integrate general RNA-related data as aimed for by the RNA Ontology (RNAO) Consortium [ROC (141)] will help the field greatly. The identification and analysis of novel motifs as building blocks has contributed greatly and will continue to contribute to the understanding of global RNA folding mechanisms as well as how RNA molecules function. We now know that microRNA precursor processing is loop-dependent, as are metabolite-binding by natural and synthetic aptamer structures, or certain RNA-(protein/RNA) recognition events. The RNA world is (and highly likely has been) a world of loops that, as interaction platforms to other molecules, drive important regulatory processes within the cell.

Outlook

The understanding of RNA loop structures is tightly linked to advances in global RNA structure prediction. With steadily increasing numbers and variability of available experimentally solved RNA 3D structures, we may also see an increase of novel RNA loop types with direct consequences for RNA 3D structure prediction and structure design. More structural data will lead to better knowledge-based potentials and force field calculations. However, it is also possible that we already know all (or most) there is. The facts that different structural motifs have been observed across RNA species and that the conformational space of the RNA backbone is limited, supports the notion that the fold space of RNA molecules and their loops might indeed be restricted to a small number of 'folds' or motifs. Thus, it seems rather unlikely that there are many more folds and loop types yet to be found. Nevertheless, understanding the role of RNA loops e.g., in aptamers may play an essential role for the biomedical field in general, especially in the research and design of novel RNA-based drugs.

Emerging algorithmic approaches together with recent advances in RNA structure probing [e.g., (142)] will allow accurate RNA structure predictions, possibly allowing the identification of certain loop motifs purely from sequence and probing data.

Highlights

- RNA loops are structural elements essential for the formation of RNA 3D structure.
- RNA loops adopt specific conformations (motifs) with different structural and functional attributes.
- RNA loops drive interactions with other biopolymers.
- A wide range of RNA function is loop-dependent, such as metabolite binding.
- RNA loops are important structural features in microRNA pathways.
- Bioinformatics approaches are successful in identifying novel loop motifs and understanding their sequence-structure-function relationships.

Acknowledgements

I wish to thank Dirk Walther for reading the manuscript and supporting the work on this review. Furthermore, I apologise to all fellow RNA scientists whose work contributed to the field of RNA loops, but which was not referenced here.

Conflict of interest statement

None declared.

References

1. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 1982; 31: 147–57.
2. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 1983; 35: 849–57.
3. Gilbert W. The RNA World. Nature 1986; 319: 618.
4. Berman HM, Olson WK, Beveridge DL, Westbrook J, Gelbin A, Demeny T, Hsieh SH, Srinivasan AR, Schneider B. The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. Biophys J 1992; 63: 751–9.
5. Tinoco I, Bustamante C. How RNA folds. J Mol Biol 1999; 293: 271–81.
6. Doudna JA, Doherty EA. Emerging themes in RNA folding. Fold Des 1997; 2: R65–70.
7. Tuerk C, Gauss P, Thermes C, Groebe DR, Gayle M, Guild N, Stormo G, d'Aubenton-Carafa Y, Uhlenbeck OC, Tinoco, JR I. CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes. Proc Natl Acad Sci USA 1988; 85: 1364–8.
8. Varani G. Exceptionally stable nucleic acid hairpins. Annu Rev Biophys Biomol Struct 1995; 24: 379–404.
9. Antao VP, Tinoco I. Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. Nucleic Acids Res 1992; 20: 819–24.
10. Wolters J. The nature of preferred hairpin structures in 16S-like rRNA variable regions. Nucleic Acids Res 1992; 20: 1843–50.

11. Woese CR, Gutell R, Gupta R, Noller HF. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol Rev 1983; 47: 621–69.
12. Woese CR, Winker S, Gutell RR. Architecture of ribosomal RNA: constraints on the sequence of 'tetra-loops'. Proc Natl Acad Sci USA 1990; 87: 8467–71.
13. Pley HW, Flaherty KM, McKay DB. Three-dimensional structure of a hammerhead ribozyme. Nature 1994; 372: 68–74.
14. Scott WG, Finch JT, Klug A. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. Cell 1995; 81: 991–1002.
15. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. Crystal structure of a group I ribozyme domain: principles of RNA packing. Science 1996; 273: 1678–85.
16. Sheehy JP, Davis AR, Znosko BM. Thermodynamic characterization of naturally occurring RNA tetraloops. RNA 2010; 16: 417–29.
17. Klosterman PS, Hendrix DK, Tamura M, Holbrook SR, Brenner SE. Three-dimensional motifs from the SCOR, structural classification of RNA database: extruded strands, base triples, tetraloops and U-turns. Nucleic Acids Res 2004; 32: 2342–52.
18. Jucker FM, Pardi A. Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif. Biochemistry 1995; 34: 14416–27.
19. Moore PB. Structural motifs in RNA. Annu Rev Biochem 1999; 68: 287–300.
20. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 2000; 289: 905–20.
21. Leontis NB, Westhof E. Analysis of RNA motifs. Curr Opin Struct Biol 2003; 13: 300–8.
22. Lemieux S, Major F. Automated extraction and classification of RNA tertiary structure cyclic motifs. Nucleic Acids Res 2006; 34: 2340–6.
23. Quigley GJ, Rich A. Structural domains of transfer RNA molecules. Science 1976; 194: 796–806.
24. Sussman JL, Kim SH. Three-dimensional structure of a transfer RNA in two crystal forms. Science 1976; 192: 853–8.
25. Huang S, Wang YX, Draper DE. Structure of a hexanucleotide RNA hairpin loop conserved in ribosomal RNAs. J Mol Biol 1996; 258: 308–21.
26. Conn GL, Draper DE, Lattman EE, Gittis AG. Crystal structure of a conserved ribosomal protein-RNA complex. Science 1999; 284: 1171–4.
27. Culver GM, Cate JH, Yusupova GZ, Yusupov MM, Noller HF. Identification of an RNA-protein bridge spanning the ribosomal subunit interface. Science 1999; 285: 2133–6.
28. Stallings SC, Moore PB. The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. Structure 1997; 5: 1173–85.
29. Puglisi EV, Puglisi JD. HIV-1 A-rich RNA loop mimics the tRNA anticodon structure. Nat Struct Biol 1998; 5: 1033–6.
30. Jucker FM, Pardi A. GNRA tetraloops make a U-turn. RNA 1995; 1: 219–22.
31. Prince JB, Taylor BH, Thurlow DL, Ofengand J, Zimmermann RA. Covalent crosslinking of tRNA1Val to 16S RNA at the ribosomal P site: identification of crosslinked residues. Proc Natl Acad Sci USA 1982; 79: 5450–4.
32. Cate JH, Yusupov MM, Yusupova GZ, Earnest TN, Noller HF. X-ray crystal structures of 70S ribosome functional complexes. Science 1999; 285: 2095–104.
33. Gutell RR, Cannone JJ, Konings D, Gautheret D. Predicting U-turns in ribosomal RNA with comparative sequence analysis. J Mol Biol 2000; 300: 791–803.
34. Barends S, Björk K, Gultyaev AP, De Smit MH, Pleij CWA, Kraal B. Functional evidence for D- and T-loop interactions in tmRNA. FEBS Lett 2002; 514: 78–83.
35. Fechter P, Rudinger-Thirion J, Florentz C, Giegé R. Novel features in the tRNA-like world of plant viral RNAs. Cell Mol Life Sci 2001; 58: 1547–61.
36. Nagaswamy U, Fox GE. Frequent occurrence of the T-loop RNA folding motif in ribosomal RNAs. RNA 2002; 8: 1112–9.
37. Leontis NB, Stombaugh J, Westhof E. The non-Watson–Crick base pairs and their associated isostericity matrices. Nucleic Acids Res 2002; 30: 3497–531.
38. Noller HF. RNA structure: reading the ribosome. Science 2005; 309: 1508–14.
39. Hermann T, Patel DJ. RNA bulges as architectural and recognition motifs. Structure 2000; 8: R47–54.
40. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. Proc Natl Acad Sci USA 2001; 98: 4899–903.
41. Lescoute A, Westhof E. The A-minor motifs in the decoding recognition process. Biochimie 2006; 88: 993–9.
42. Klein DJ, Schmeing TM, Moore PB, Steitz TA. The kink-turn: a new RNA secondary structure motif. EMBO J 2001; 20: 4214–21.
43. Lu C, Ding F, Chowdhury A, Pradhan V, Tomsic J, Holmes WM, Henkin TM, Ke A. SAM recognition and conformational switching mechanism in the Bacillus subtilis yitJ S Box/SAM-I riboswitch. J Mol Biol 2010; 404: 803–18.
44. Tiedge H. K-turn motifs in spatial RNA coding. RNA Biology 2006; 3: 133–9.
45. Schroeder KT, Lilley DMJ. Ion-induced folding of a kink turn that departs from the conventional sequence. Nucleic Acids Res 2009; 37: 7281–9.
46. Strobel SA, Adams PL, Stahley MR, Wang J. RNA kink turns to the left and to the right. RNA 2004; 10: 1852–4.
47. Adams PL, Stahley MR, Gill ML, Kosek AB, Wang J, Strobel SA. Crystal structure of a group I intron splicing intermediate. RNA 2004; 10: 1867–87.
48. Goody TA, Melcher SE, Norman DG, Lilley DMJ. The kink-turn motif in RNA is dimorphic, and metal ion-dependent. RNA 2004; 10: 254–64.
49. Spackova N, Reblova K, Sponer J. Structural dynamics of the box C/D RNA kink-turn and its complex with proteins: the role of the A-minor 0 interaction, long-residency water bridges, and structural ion-binding sites revealed by molecular simulations. J Phys Chem B 2010; 114: 10581–93.
50. Antonioli AH, Cochrane JC, Lipchock SV, Strobel SA. Plasticity of the RNA kink turn structural motif. RNA 2010; 16: 762–8.
51. Torres-Larios A, Dock-Bregeon AC, Romby P, Rees B, Sankaranarayanan R, Caillet J, Springer M, Ehresmann C, Ehresmann B, Moras D. Structural basis of translational control by Escherichia coli threonyl tRNA synthetase. Nat Struct Biol 2002; 9: 343–7.
52. Lescoute A, Leontis NB, Massire C, Westhof E. Recurrent structural RNA motifs, isostericity matrices and sequence alignments. Nucleic Acids Res 2005; 33: 2395–409.
53. Leontis NB, Westhof E. A common motif organizes the structure of multi-helix loops in 16 S and 23 S ribosomal RNAs. J Mol Biol 1998; 283: 571–83.
54. Leontis NB, Stombaugh J, Westhof E. Motif prediction in ribosomal RNAs lessons and prospects for automated motif predic-

tion in homologous RNA molecules. Biochimie 2002; 84: 961–73.
55. Jaeger L, Verzemnieks EJ, Geary C. The UA_handle: a versatile submotif in stable RNA architectures. Nucleic Acids Res 2009; 37: 215–30.
56. Cruz JA, Westhof E. The dynamic landscapes of RNA architecture. Cell 2009; 136: 604–9.
57. Lescoute A, Westhof E. Topology of three-way junctions in folded RNAs. RNA 2006; 12: 83–93.
58. Laing C, Schlick T. Analysis of four-way junctions in RNA structures. J Mol Biol 2009; 390: 547–59.
59. Laing C, Jung S, Iqbal A, Schlick T. Tertiary motifs revealed in analyses of higher-order RNA junctions. J Mol Biol 2009; 393: 67–82.
60. Ushida C, Himeno H, Watanabe T, Muto A. tRNA-like structures in 10Sa RNAs of Mycoplasma capricolum and Bacillus subtilis. Nucleic Acids Res 1994; 22: 3392–6.
61. Felden B, Hanawa K, Atkins JF, Himeno H, Muto A, Gesteland RF, McCloskey JA, Crain PF. Presence and location of modified nucleotides in Escherichia coli tmRNA: structural mimicry with tRNA acceptor branches. EMBO J 1998; 17: 3188–96.
62. Stagg SM, Frazer-Abel AA, Hagerman PJ, Harvey SC. Structural studies of the tRNA domain of tmRNA. J Mol Biol 2001; 309: 727–35.
63. Zwieb C, Guven SA, Wower IK, Wower J. Three-dimensional folding of the tRNA-like domain of Escherichia coli tmRNA. Biochemistry 2001; 40: 9587–95.
64. Edwards TE, Klein DJ, Ferré-D'Amaré AR. Riboswitches: small-molecule recognition by gene regulatory RNAs. Curr Opin Struct Biol 2007; 17: 273–9.
65. Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. Science 2008: 320: 77–82.
66. Keating KS, Toor N, Perlman PS, Pyle AM. A structural analysis of the group II intron active site and implications for the spliceosome. RNA 2010; 16: 1–9.
67. Hermann T, Patel DJ. Stitching together RNA tertiary architectures. J Mol Biol 1999; 294: 829–49.
68. Lescoute A, Westhof E. The interaction networks of structured RNAs. Nucleic Acids Res 2006; 34: 6587–604.
69. Gan J, Shaw G, Tropea JE, Waugh DS, Court DL, Ji X. A stepwise model for double-stranded RNA processing by ribonuclease III. Mol Microbiol 2008; 67: 143–54.
70. Zhang K, Nicholson AW. Regulation of ribonuclease III processing by double-helical sequence antideterminants. Proc Natl Acad Sci USA 1997; 94: 13437–41.
71. Nicholson AW. Function, mechanism and regulation of bacterial ribonucleases. FEMS Microbiol Rev 1999; 23: 371–90.
72. Chanfreau G, Buckle M, Jacquier A. Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. Proc Natl Acad Sci USA 2000; 97: 3142–7.
73. Nagel R, Ares M. Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5′-AGNN-3′ tetraloop. RNA 2000; 6: 1142–56.
74. Lamontagne B, Ghazal G, Lebars I, Yoshizawa S, Fourmy D, Elela SA. Sequence dependence of substrate recognition and cleavage by yeast RNase III. J Mol Biol 2003; 327: 985–1000.
75. Lebars I, Lamontagne B, Yoshizawa S, Abou Elela S, Fourmy D. Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing. EMBO J 2001; 20: 7250–8.
76. Zhang X, Zeng Y. The terminal loop region controls microRNA processing by Drosha and Dicer. Nucleic Acids Res 2010; 38: 1–9.
77. Song L, Axtell MJ, Fedoroff NV. RNA secondary structural determinants of miRNA precursor processing in Arabidopsis. Curr Biol 2010; 20: 37–41.
78. Werner S, Wollmann H, Schneeberger K, Weigel D. Structure determinants for accurate processing of miR172a in Arabidopsis thaliana. Curr Biol 2010; 20: 42–8.
79. Mateos JL, Bologna NG, Chorostecki U, Palatnik JF. Identification of microRNA processing determinants by random mutagenesis of Arabidopsis MIR172a precursor. Curr Biol 2010; 20: 49–54.
80. Kawamata T, Seitz H, Tomari Y. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat Struct Mol Biol 2009; 16: 953–60.
81. Brodersen P, Voinnet O. Revisiting the principles of microRNA target recognition and mode of action. Nat Rev Mol Cell Biol 2009; 10: 141–8.
82. Gait MJ, Karn J. RNA recognition by the human immunodeficiency virus Tat and Rev proteins. Trends Biochem Sci 1993; 18: 255–9.
83. Fourmy D, Recht MI, Blanchard SC, Puglisi JD. Structure of the A site of Escherichia coli 16S ribosomal RNA complexed with an aminoglycoside antibiotic. Science 1996; 274: 1367–71.
84. Batey RT, Gilbert SD, Montange RK. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. Nature 2004; 432: 411–5.
85. Serganov A, Yuan Y-R, Pikovskaya O, Polonskaia A, Malinina L, Phan AT, et al. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. Chem Biol 2004; 11: 1729–41.
86. Hermann T, Westhof E. RNA as a drug target: chemical, modelling, and evolutionary tools. Curr Opin Biotechnol 1998; 9: 66–73.
87. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. Nature 1990; 346: 818–22.
88. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 1990; 249: 505–10.
89. Brody EN, Gold L. Aptamers as therapeutic and diagnostic agents. J Biotechnol 2000; 74: 5–13.
90. Hesselberth J, Robertson MP, Jhaveri S, Ellington AD. In vitro selection of nucleic acids for diagnostic applications. J Biotechnol 2000; 74: 15–25.
91. Kusser W. Chemically modified nucleic acid aptamers for in vitro selections: evolving evolution. J Biotechnol 2000; 74: 27–38.
92. Patel DJ, Suri AK. Structure, recognition and discrimination in RNA aptamer complexes with cofactors, amino acids, drugs and aminoglycoside antibiotics. J Biotechnol 2000; 74: 39–60.
93. Xiao H, Murakami H, Suga H, Ferré-D'Amaré AR. Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. Nature 2008; 454: 358–61.
94. Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR. Genetic control by a metabolite binding mRNA. Chem Biol 2002; 9: 1043–9.
95. Winkler WC, Cohen-Chalamish S, Breaker RR. An mRNA structure that controls gene expression by binding FMN. Proc Natl Acad Sci USA 2002; 99: 15908–13.
96. Winkler W, Nahvi A, Breaker RR. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature 200; 419: 952–6.
97. Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, Kreneva RA, Perumov DA, Nudler E. Sensing small molecules

by nascent RNA: a mechanism to control transcription in bacteria. Cell 2002; 111: 747–56.
98. Thore S, Leibundgut M, Ban N. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. Science 2006; 312: 1208–11.
99. Serganov A, Polonskaia A, Phan AT, Breaker RR, Patel DJ. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. Nature 2006; 441: 1167–71.
100. Winkler WC, Nahvi A, Sudarsan N, Barrick JE, Breaker RR. An mRNA structure that controls gene expression by binding S-adenosylmethionine. Nat Struct Biol 2003; 10: 701–7.
101. Corbino KA, Barrick JE, Lim J, Welz R, Tucker BJ, Puskarz I, Mandal M, Rudnick ND, Breaker RR. Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in α-proteobacteria. Genome Biol 2005; 6: R70.
102. Fuchs RT, Grundy FJ, Henkin TM. The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase. Nat Struct Mol Biol 2006; 13: 226–33.
103. Montange RK, Batey RT. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. Nature 2006; 441: 1172–5.
104. Klein DJ, Ferré-D'Amaré AR. Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. Science 2006; 313: 1752–6.
105. Cochrane JC, Lipchock SV, Strobel SA. Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. Chem Biol 2007; 14: 97–105.
106. Serganov A, Patel DJ. Towards deciphering the principles underlying an mRNA recognition code. Curr Opin Struct Biol 2008; 18: 120–9.
107. Burd CG, Dreyfuss G. Conserved structures and diversity of functions of RNA-binding proteins. Science 1994; 265: 615–21.
108. Sargsyan K, Lim C. Arrangement of 3D structural motifs in ribosomal RNA. Nucleic Acids Res 2010; 38: 3512–22.
109. Zhong C, Tang H, Zhang S. RNAMotifScan: automatic identification of RNA structural motifs using secondary structural alignment. Nucleic Acids Res 2010; 38: 1–11.
110. St-Onge K, Thibault P, Hamel S, Major F. Modeling RNA tertiary structure motifs by graph-grammars. Nucleic Acids Res 2007; 35: 1726–36.
111. Sarver M, Zirbel CL, Stombaugh J, Mokdad A, Leontis NB. FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. J Math Biol 2008; 56: 215–52.
112. Gendron P, Lemieux S, Major F. Quantitative analysis of nucleic acid three-dimensional structures. J Mol Biol 2001; 308: 919–36.
113. Yang H, Jossinet F, Leontis N, Chen L, Westbrook J, Berman H, Westhof E. Tools for the automatic identification and classification of RNA base pairs. Nucleic Acids Res 2003; 31: 3450–60.
114. Lu X-J, Olson WK. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Res 2003; 31: 5108–21.
115. Wadley LM, Pyle AM. The identification of novel RNA structural motifs using COMPADRES: an automated approach to structural discovery. Nucleic Acids Res 2004; 32: 6650–9.
116. Huang H-C, Nagaswamy U, Fox GE. The application of cluster analysis in the intercomparison of loop structures in RNA. RNA 2005; 11: 412–23.
117. Lisi V, Major F. A comparative analysis of the triloops in all high-resolution RNA structures reveals sequence-structure relationships. RNA 2007; 13: 1537–45.
118. Schudoma C, May P, Nikiforova V, Walther D. Sequence-structure relationships in RNA loops: establishing the basis for loop homology modeling. Nucleic Acids Res 2010; 38: 970–80.
119. Tamura M, Hendrix DK, Klosterman PS, Schimmelman NRB, Brenner SE, Holbrook SR. SCOR: Structural Classification of RNA, version 2.0. Nucleic Acids Res 2004; 32(Database issue): D182-4.
120. Bindewald E, Hayes R, Yingling YG, Kasprzak W, Shapiro BA. RNA junction: a database of RNA junctions and kissing loops for three-dimensional structural analysis and nanodesign. Nucleic Acids Res 2008; 36(Database issue): D392–7.
121. Schudoma C, May P, Walther D. Modeling RNA loops using sequence homology and geometric constraints. Bioinformatics 2010; 26: 1671–2.
122. Popenda M, Szachniuk M, Blazewicz M, Wasik S, Burke EK, Blazewicz J, Adamiak RW. RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures BMC Bioinformatics 2010; 11: 231.
123. Schroeder KT, McPhee SA, Ouellet J, Lilley DMJ. A structural database for k-turn motifs in RNA. RNA 2010; 16: 1463–8.
124. Major F, Turcotte M, Gautheret D, Lapalme G, Fillion E, Cedergren R. The combination of symbolic and numerical computation for three-dimensional modeling of RNA. Science 1991; 253: 1255–60.
125. Gautheret D, Major F, Cedergren R. Modeling the three-dimensional structure of RNA using discrete nucleotide conformational sets. J Mol Biol 1993; 229: 1049–64.
126. Ogata H, Akiyama Y, Kanehisa M. A genetic algorithm based molecular modeling technique for RNA stem-loop structures. Nucleic Acids Res 1995; 23: 419–26.
127. Frederic T, Rosenfeld R, Cantor CR, DeLisi C. RNA loop structure prediction via bond scaling and relaxation. Biopolymers 1996; 38: 769–79.
128. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P. Fast folding and comparison of RNA secondary structures. Monatsh Chem 1994; 125: 167–188.
129. Tyagi R, Mathews DH. Predicting helical coaxial stacking in RNA multibranch loops. RNA 2007; 13: 939–51.
130. Parisien M. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. Nature 2008; 452: 51–5.
131. Das R, Baker D. Automated de novo prediction of native-like RNA tertiary structures. Proc Natl Acad Sci USA 2007; 104: 14664–9.
132. Das R, Karanicolas J, Baker D. Atomic accuracy in predicting and designing noncanonical RNA structure. Nat Methods 2010; 7: 291–4.
133. Ding F, Sharma S, Chalasani P, Demidov VV, Broude NE, Dokholyan NV. Ab initio RNA folding by discrete molecular dynamics: from structure prediction to folding mechanisms. RNA 2008; 14: 1164–73.
134. Sharma S, Ding F, Dokholyan NV. iFoldRNA: three-dimensional RNA structure prediction and folding. Bioinformatics 2008; 24: 1951–2.
135. Jonikas MA, Radmer RJ, Laederach A, Das R, Pearlman S, Herschlag D, et al. Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. RNA 2009; 15: 189–99.
136. Jonikas MA, Radmer RJ, Altman RB. Knowledge-based instantiation of full atomic detail into coarse-grain RNA 3D structural models. Bioinformatics 2009; 25: 3259–66.

137. Massire C, Westhof E. MANIP: an interactive tool for modelling RNA. J Mol Graph Model 1998; 16: 197–205, 255–7.
138. Jossinet F, Westhof E. Sequence to structure (S2S): display, manipulate and interconnect RNA data from sequence to structure. Bioinformatics 2005; 21: 3320–1.
139. Jossinet F, Ludwig TE, Westhof E. Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. Bioinformatics 2010; 26: 2057–9.
140. Rother M, Rother K, Puton T, Bujnicki JM. ModeRNA: a tool for comparative modeling of RNA 3D structure. Nucleic Acids Res 2011; Epub ahead of print; DOI 10.1093/nar/gkq1320.
141. Leontis NB, Altman RB, Berman HM, Brenner SE, Brown JW, Engelke DR, Harvey SC, Holbrook SR, Jossinet F, Lewis SE, Major F, Mathews DH, Richardson JS, Williamson JR, Westhof E. The RNA ontology consortium: an open invitation to the RNA community. RNA 2006; 12: 533–41.
142. Weeks KM. Advances in RNA structure analysis by chemical probing. Curr Opin Struct Biol 2010; 20: 295–304.
143. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera – a visualization system for exploratory research and analysis. J Comput Chem 2004; 25: 1605–12.

Received February 4, 2011; accepted March 25, 2011