Published in final edited form as:

Wiley Interdiscip Rev RNA. 2013 September ; 4(5): 507–522. doi:10.1002/wrna.1175.

Structure and function of the T-loop structural motif in non-coding RNAs

Clarence W Chan^{1,*}, Bhaskar Chetnani^{1,*}, and Alfonso Mondragón¹

¹Department of Molecular Biosciences, Northwestern University, Evanston, IL60208, USA

Abstract

The T-loop is a frequently occurring five-nucleotide motif found in the structure of non-coding RNAs where it is commonly assumed to play an important role in stabilizing the tertiary RNA structure by facilitating long-range interactions between different regions of the molecule. T-loops were first identified in tRNA^{Phe} and a formal consensus sequence for this motif was formulated and later revised based on analyses of the crystal structures of prokaryotic ribosomal RNAs and RNase P and the corresponding primary sequence of their orthologues. In the past decade, several new structures of large RNA molecules have been added to the RCSB Protein Data Bank, including the eukaryotic ribosome, a self-splicing group II intron, numerous synthetases in complex with their cognate tRNAs, tmRNA in complex with SmpB, several riboswitches, and a complex of bacterial RNase P bound to its tRNA substrate. In this review, the search for T-loops is extended to these new RNA molecules based on the previously established structure-based criteria. The review highlights and discusses the function and additional roles of T-loops in four broad categories of RNA molecules, namely tRNAs, rRNAs, P RNAs, and RNA genetic elements. Additionally, the potential application for T-loops as interaction modules is also discussed.

In recent years, discoveries of non-coding RNAs^{1–3} combined with the growing realization that their diverse functions are determined in large part by their tertiary fold^{4, 5} have refueled the impetus to characterize and examine recurrent RNA structural motifs^{6–10}. Many of these motifs were initially identified in efforts to catalogue the common building blocks of folded RNA^{11, 12} when near atomic-resolution RNA structures first became available. To date, numerous RNA structural motifs have been classified and surveyed, including A-minor motifs¹³, A–A platforms¹⁴, kink-turns^{15, 16}, kissing loops¹⁷, pseudoknots¹⁸, S-turns¹⁹, T-loops^{20, 21}, tetraloops and their receptors^{22, 23}, as well as a large variety of non-canonical base-pairs and base-triples²⁴. These motifs collectively represent a tremendous range of structural diversity and share the common function of stabilizing tertiary conformation in folded RNA molecules. Whereas basic structural interactions such as base-pairing and base-stacking lead to coaxial stacking and hairpin formation, structural motifs bring these secondary structural elements together through intramolecular interactions to complete the tertiary fold^{6–10, 25}.

T-loops are unique among RNA structural motifs for several reasons. First, T-loops are one of the first structural motifs discovered, originally in transfer RNA (tRNA)^{26, 27}, and have subsequently been identified in a variety of non-coding RNA molecules, including transfer-messenger RNA (tmRNA)^{28, 29}, ribonuclease P (RNase P)^{30–32}, the ribosome^{33–45}, and a variety of riboswitches^{46–53}. Second, T-loops act to facilitate not only long-range intramolecular interactions, but also intermolecular RNA-RNA⁵⁴ and RNA-ligand interactions^{50–53} in naturally-occurring molecules, and represent the only RNA structural

Corresponding author: Alfonso Mondragón (a-mondragon@northwestern.edu).
* contributed equally as co-first authors

motif known to date to have this wide range of functions. Finally, the definition of a T-loop has evolved from partially sequence-based^{21, 55} to entirely structure-based²⁰ as additional examples have become available in recent RNA crystal structures.

Although a T-loop was first observed in a crystal structure of *Saccharomyces cerevisiae* tRNA^{Phe}, determined in 1974^{26, 27}, the first formal definition of a T-loop motif was not formulated until 2002 when crystal structures of the *Thermus thermophilus* 30S⁴³ and *Haloarcula marismortui* 50S⁴⁰ ribosome subunits were reported, which substantially increased the number of available T-loop-containing structures known at the time. An analysis of the nine T-loops present in the *T. thermophilus* 16S and *H. marismortui* 23S ribosomal RNAs (rRNAs) resulted in a partially sequence-based consensus, which defined the T-loop as a reverse Hoogsteen U·A base-pair stacked upon a Watson-Crick base-pair to form a compact U-turn-like pentaloop structure²¹. By expanding the T-loop search to the 23S rRNA in the *Deinococcus radiodurans* 50S ribosome subunit crystal structure³⁸, a subsequent study revealed that the reverse Hoogsteen U·A base-pair can infrequently be replaced by canonical Watson-Crick G·C or U·A base-pairs⁵⁵. While these initial studies were limited to the T-loop structures of prokaryotic rRNAs and various *Escherichia coli* and *S. cerevisiae* tRNAs, the importance of T-loops in facilitating tertiary interactions in non-coding RNAs was clearly established.

In 2003, following the structural determination of a folded domain of the *T. thermophilus* RNase P RNA subunit (P RNA) in which two T-loops were identified³², a systematic search for T-loops in RNA structures was conducted using a strictly structure-based approach²⁰. By searching for structural similarity rather than recurring sequence patterns, additional T-loops were found in previously examined rRNA crystal structures. A comparison of T-loop sequences in more than 4,000 orthologous RNAs (tRNAs, rRNAs, and P RNAs) revealed a relatively broad sequence consensus that is fundamentally dictated by structural conservation (Fig. 1)²⁰.

In the last decade, numerous RNA and RNA-protein crystal structures have been added to the RCSB Protein Data Bank (PDB, www.pdb.org)⁵⁶, including structures of tRNAs in complex with various synthetases^{57–59}; examples of tmRNA bound to its protein cofactor, SmpB^{28, 29}; the prokaryotic ribosome captured in different stages of its catalytic cycle in association with various protein factors and antibiotics^{60–66}; the *S. cerevisiae* ribosome^{33, 34}; the intact bacterial P RNA from both A and B type phylogenetic groups^{30, 67}, as well as an A type bacterial RNase P holoenzyme in complex with tRNA⁵⁴; self-splicing group I^{68–70} and group II⁷¹ introns; and an assorted array of riboswitches⁷². In this review, a systematic search for additional T-loops was performed using the previously established structure-based criteria²⁰. The results demonstrate that T-loops are indeed present not only in tRNAs, rRNAs, and P RNAs, but also, as have already been noted, in numerous gene regulatory RNAs such as riboswitches^{46–53} and in a self-splicing group II intron⁷¹. More importantly, T-loops are shown to assume diverse functions, ranging from stabilizing local tertiary structure to mediating long-range tertiary interactions to acting as a receptor for ligand binding, making the T-loop an essential and versatile widespread RNA structural motif.

Frequent occurrence of T-loops in known RNA structures

A total of 5,085 RNA chains from 2,398 RNA-containing X-ray, NMR, and electron microscopy structures currently deposited in the PDB were surveyed for T-loops using a structure-based criterion, as described previously²⁰. In brief, the T-loop in the high-resolution crystal structure of *S. cerevisiae* tRNA^{Phe} (PDB ID IEHZ, residues 54 to 58)⁷³ was superposed on every possible contiguous five-nucleotide segment in each RNA chain using the general atomic superposition program *LSQKAB* from the CCP4 package⁷⁴. In the

superposition, every atom was treated equally. Structural similarity was then assessed according to the backbone root mean square deviation (RMSD) for the superposed atoms. To make the search process sequence independent, only the backbone atoms were used for the superpositions. An identical search was also conducted using 692 RNA-containing crystal structures (all determined to or better than 4 Å resolution) from a non-redundant database⁷⁵. The searches uncovered the presence of over 105 unique T-loops. In order to minimize false negatives produced by the search algorithm, a generous RMSD cutoff (1.0 Å and 1.5 Å for all backbone atoms in the non-redundant and complete datasets, respectively) was used; however, all potential matches were manually verified to eliminate any apparent false positives. A comparable search conducted using only the phosphorus atoms in the backbone of the search model yielded similar results, demonstrating that, at minimum, the 3D coordinates of only these five atoms are generally sufficient for systematically locating T-loops in RNA tertiary structures. A visual and statistical summary of the final results is provided in Tables I and II, respectively.

Collectively, T-loop-containing RNAs make up a structurally and functionally diverse set of folded non-coding RNA molecules. Although most examples of T-loops are derived from rRNAs exceeding several kilobases in length, T-loops can also be found in small RNA molecules such as tRNAs and certain riboswitches. T-loops are not limited to any particular structural context, as they occur in both proximal and distal regions of RNA molecules, cap both short and long stems, and can be involved in both intramolecular and intermolecular tertiary interactions^{6–10, 25}. Since the basic T-loop architecture can accommodate a wide range of base substitutions, T-loops exhibit a relatively broad sequence consensus²⁰, which likely accounts for their capacity to act as receptors for both nucleobases and other ligands.

The RNA T-loop structural motif

Based on the T-loops identified in the 2,398 RNA-containing structures, a clear structure-based definition emerges with the following characteristics (Fig. 1 and Table I):

- A T-loop consists of five consecutive nucleotides assuming a compact U-turn-like loop structure, which can be identified objectively based on backbone RMSD measurements from a reference structure.
- T-loops are closed by a base-pair formed by nucleotides at positions 1 and 5. This base-pair is often a non-canonical base-pair but conventional Watson-Crick base-pairing are also supported by the T-loop architecture.
- The nucleobase at position 2 stacks with the nucleobase at position 1.
- The nucleobase at position 4 either base-pairs with the nucleobase at position 2 (Fig. 1D, 1E, and 1F) or sandwiches an inserted nucleobase or ligand with the nucleobase at position 5, leading to the formation of a continuous base-stack (Fig. 1A, 1B, and 1C). When an inserted nucleobase is present, the inserted nucleobase will interact with the nucleobase at position 2 through hydrogen bonding.
- The nucleobase at position 3 is unstacked and often free to facilitate tertiary interactions with a distant region of the RNA molecule either through base-pairing and/or base-stacking interactions.

The classical tRNA T-loop

In protein synthesis, tRNAs decode genetic information encoded in messenger RNAs (mRNAs) by acting as adapter molecules that supply nascent polypeptides of translating ribosomes with amino acids specified by the codons of their corresponding mRNA sequence. Structurally, tRNAs are small (~70–100 nucleotides) folded RNAs consisting of

at least three stem-loops (D arm, anti-codon arm, T Ψ C arm, and an optional variable arm) and a closing stem formed by their 5' and 3' ends, which acts as the amino acid acceptor region. On the tertiary level, tRNAs assume an L-shaped structure in which the loop regions of the D and T Ψ C arms closely associate to form a near-90° kink⁷³.

The long-range intramolecular docking of two conserved tRNA D arm guanosines into the T-loop of the T Ψ C arm constitutes an essential feature of the tRNA tertiary structure that is conserved across all tRNAs⁷⁶. As T-loops were first identified in a crystal structure of yeast tRNA^{Phe}, the tRNA T-loop represents the prototypic T-loop structural motif formed by five consecutive nucleotides in loop regions of RNA molecules (Fig. 2A). Closure of the tRNA T-loop is achieved by a non-Watson-Crick, reverse Hoogsteen U-A base-pair formed by a conserved uridine and adenosine at positions 1 and 5 (U_{T1} and A_{T5}, respectively) of the T-loop. A pyrimidine at position 2 base-stacks with U_{T1} and hydrogen bonds with a D-loop guanosine (G_{D1}), which in turn is inserted between A_{T5} and a purine at position 4 of the T-loop, forming a three-nucleotide continuous base-stack. The center nucleotide at position 3 of the tRNA T-loop is a conserved cytidine (C_{T3}), which forms a Watson-Crick base-pair with a conserved guanosine (G_{D2}) adjacent to G_{D1}, further stabilizing the tertiary association of the D and T Ψ C-loops.

Although T-loops have been observed in a wide variety of structured RNAs, crystal structures of tRNAs purified from their native source provide unique examples of how natural post-transcriptional RNA modifications can affect the T-loop structure. In yeast tRNA^{Phe}, the U_{T1} and A_{T5} residues forming the U-A base-pair closing the T-loop are replaced by modified nucleotides 5-methyluridine (ribothymidine) and 6-hydro-1-methyladenosine, respectively. In addition, the pyrimidine at position 2 is a pseudouridine instead of a uridine²⁴. Although it remains unclear how these nucleobase modifications affect the T-loop structure since they do not appear to significantly alter the hydrogen bonding network within the T-loop, the replacement of U_{T2} with a pseudouridine does result in the formation of a hydrogen bond with G_{D1}. Whether this single hydrogen bond accounts for a more stable T-loop, and hence a more stable tRNA, awaits further investigation. However, previous thermodynamic studies have demonstrated that the presence of these modified nucleotides not only increases the melting temperature of the entire tRNA molecule⁷⁷, but also specifically enhances the binding affinity between the D and T Ψ C-loops⁷⁸.

Structural mimicry of the tRNA T-loop in tmRNA

In bacteria, the accumulation of ribosomes stalled on a string of rare codons or on an aberrant mRNA lacking a stop codon (nonstop mRNA) can be extremely detrimental to growth and survival. tmRNA is a chimeric RNA molecule consisting of both a tRNA-like domain (TLD) (Fig. 2B) and a mRNA-like domain (MLD) that, when associated with a small protein subunit, SmpB, acts to rescue stalled ribosomes through a process of *trans*-translation. The tmRNA/SmpB complex hijacks a stalled ribosome and switches its translation template to that of a degradation signal encoded *in trans* by the tmRNA-MLD. Consequently, tmRNA/SmpB recycles an otherwise stalled ribosome by simultaneously ejecting the abnormal mRNA and tagging the partially translated polypeptide for targeted degradation⁷⁹.

Proper function of the tmRNA/SmpB system relies largely on its ability to deceive the entire protein expression machinery by structurally mimicking key features of both tRNA and mRNA, as well as of the cognate tRNA-mRNA codon interaction in a single RNP complex⁶³. In particular, by resembling tRNA^{Ala}, the tmRNA-TLD is charged by alanyl-tRNA synthetase and recognized subsequently by the elongation factor EF-Tu⁸⁰. However,

notable structural differences between the tmRNA-TLD/SmpB complex and tRNA^{Ala} also exist. Whereas tRNA^{Ala} folds into an L-shaped structure, the tmRNA-TLD anti-codon-like stem-loop protrudes out at an obtuse angle reminiscent of the variable loop in other tRNAs and is nearly anti-parallel to the coaxial helical stack formed by the T Ψ C and acceptor stems^{28, 29}. In the tmRNA-TLD, the D arm is present but without a stem; instead, the D-loop is stabilized by SmpB, which also acts as a structural and functional substitute to the tRNA anti-codon stem in tmRNA²⁸ (Fig. 2C).

Despite the structural differences between tRNA^{Ala} and the tmRNA-TLD, the long-range docking of two D arm guanosines into the T-loop of the T Ψ C arm is a feature of tRNAs that is preserved identically in the tmRNA-TLD structure^{28, 29}. The T-loop in both molecules closes with a reverse Hoogsteen U·A base-pair (a 5-methyluridine-6-hydro-1-methyladenosine pairing when the bases are modified), which flanks a uridine/pseudouridine, cytidine, and guanosine in the middle positions 2, 3, and 4, respectively. The same stabilizing interactions, both within the T-loop and between the T-loop and D-loop guanosines, G_{D1} and G_{D2}, are present in tRNA and tmRNA. This striking similarity demonstrates that T-loops are not only important for long-range interactions, but also a recurrent structural motif that can be found in two different RNA molecules that serve a similar function (Fig. 2D).

T-loops help stabilize the tertiary fold in ribosomal RNA (rRNA)

Ribosomes are large cytoplasmic RNP particles that translate mRNAs into protein. Both the 70S prokaryotic and 80S eukaryotic ribosome consist of large (50S and 60S, respectively) and small (30S and 40S, respectively) subunits that total a molecular weight up to several megadaltons. Whereas the 50S is formed by a 23S and 5S rRNA in association with 34 protein subunits, the 60S is a complex consisting of a 28S/25–26S, 5.8S, and 5S rRNA with nearly 50 proteins. Similarly, the 30S includes the 16S rRNA and 21 proteins and the 40S is made up of the 18S rRNA and over 30 proteins⁸¹. In addition to its size and complexity, a notable feature of the ribosome is that it contains a RNA-based active site and functions as a true RNP enzyme (ribozyme) capable of catalyzing polypeptide synthesis in a multiple-turnover fashion⁸².

Despite containing numerous proteins, the rRNA subunits make up a significant mass percentage of both the prokaryotic and eukaryotic ribosome and act effectively as a core scaffold to which the ribosomal proteins bind. Several T-loops are present in rRNAs, and they have been examined and catalogued in detail previously^{20, 21, 55}. In summary, rRNA T-loops assume a variety of structural roles, ranging from forming noninteracting stem-capping pentaloops to pentaloops that participate in tertiary interactions involving either base-pairing or base-stacking via their second, third, and fourth nucleotides. In some cases, even the first and fifth base-pairing T-loop nucleotides form long-range nucleobase or backbone contacts.

The key role of rRNA T-loops is to support the ribosome tertiary structure by stabilizing loop junctions and sharp turns and by making long-range contacts between stem-loops that are distant in primary sequence but adjacent in physical space. Most rRNA T-loops occur in phylogenetically conserved regions⁵⁵. In fact, a comparison of a recent crystal structure of the *S. cerevisiae* ribosome³³ with the prokaryotic ribosome⁶⁶ structure does reveal that most T-loop structures and interactions are conserved from bacteria to eukaryotes (Fig. 3).

Interlocked T-loops in RNase P

RNase P is an RNP endonuclease that catalyzes the removal of a leading segment from the 5'-terminus of tRNA precursors. Like the ribosome, RNase P is a true multiple-turnover

enzyme that utilizes an RNA-based active site. The catalytic RNA subunit of RNase P is structurally organized into two independently-folded domains, namely, the specificity (S) and catalytic (C) domains⁸³. Based on crystal structures of the bacterial P RNA, the S and C domains form a pseudo-planar surface to which tRNA precursor substrates bind, primarily as a result of shape-complementarity, but also by base-stacking and base-pairing interactions that enhance the specificity of the interaction (Fig. 4A)⁵⁴. To date, several T-loops have been observed in both phylogenetic A-type and B-type bacterial P RNAs, including those that are unique to expansion regions in various B-type P RNAs^{32, 67, 84}.

A unique feature of the P RNA is the presence of two conserved T-loops in the J11/12 and J12/11 regions of the S domain that interlock tightly to form a rigid structural module (Fig. 4B)⁸³. Whereas the J11/12 T-loop is always closed by a non-Watson-Crick A·A (N7-amino) base-pair, the J12/11 T-loop is predominantly closed by a reverse Hoogsteen U·A base-pair, with non-Watson-Crick C·A and U·C pairings occurring as rare exceptions^{83, 85}. In both T-loops of the bacterial S domain structure, the nucleotide at position 2 is a conserved guanosine, which interacts with the 5'-phosphate of the closing pair nucleotide at position 5 through hydrogen bonding⁸⁶ and base-stacks with the nucleotide at position 1. Adjacent to these conserved guanosines, the nucleotide at position 4 contains exclusively a purine base that participates in a Hoogsteen/sugar edge interaction with the 2'-hydroxyl of the ribose sugar group at position 2. In the case of the bacterial S domain T-loops, the center nucleotide at position 3 base-stacks with the position 4 purine base, allowing for additional backbone-backbone contacts and thus an ostensibly tighter and more stable U-turn configuration in the overall T-loop structure^{30–32, 54}.

Two salient features of the bacterial S domain J11/12 and J12/11 T-loops allow the formation of a particularly stable module. First, the anti-parallel juxtaposition and close proximity of the closing base-pair nucleotides and conserved position 2 guanosine (G₂) result in two coaxial base triples. In one layer, the A·A base-pair associates with the G₂ of the other T-loop to form an A·A·G (amino-N7, N1-amino, amino-N3, N1-amino) base triple. Similarly, Watson-Crick/sugar edge and Watson-Crick/Hoogsteen interactions between the other G₂ and closing base-pair form an adjacent base triple. As the two base triples form the structural core of the double T-loop module, a continuous 6-nucleotide base-stack incorporates the remaining T-loop nucleotides and serves as the second key stabilizing feature (Fig. 4B).

The numerous and intricate non-covalent interactions between the J11/12 and J12/11 T-loops not only make them structurally unique, but also serve a functional role in assisting RNase P substrate recognition. Across all domains of life, the J11/12 and J12/11 P RNA loops each contain a cluster of universally conserved nucleotides. Recent structural studies^{30–32, 54} have shown that in bacterial P RNA these conserved regions give rise to the double T-loop structural module, and thus, it is likely that this structural module also exists in the P RNA of higher organisms. The role of the T-loops in RNase P extends beyond a purely stabilizing one in that the formation of the double T-loop structural module results in unstacked bases that form direct intermolecular contacts with conserved regions in both the D and TΨC-loops of tRNA in the RNase P/tRNA complex. The center position 3 nucleotide of the J12/11 T-loop base-stacks directly with C_{T3} of the T-loop in the TΨC arm of tRNA, and the J11/12 T-loop helps position a base-stacking interaction between an upstream nucleotide at the -2 position with a conserved D-loop guanosine (G_{D2}) (Fig. 4B)⁵⁴.

T-loops in genetic elements: self-splicing introns and riboswitches

Genetic elements in untranslated segments of mRNAs, such as self-splicing introns and riboswitches, encompass a diverse class of non-coding RNAs⁸⁷. Self-splicing introns are

spliceosomal-independent catalytic RNAs that are capable of carrying out self-excision and the subsequent re-ligation of flanking exons. All self-splicing introns share a two-step catalytic mechanism involving two consecutive transesterification reactions. However, self-splicing introns are classified into groups I, II, and III primarily based on differences in their domain architecture and on whether the branch-site nucleophile in the first transesterification reaction is a guanosine or an adenosine. To date, several crystal structures of group I and group II self-splicing introns in various catalytic states have been determined^{68–71, 88, 89}. The *Oceanobacillus iheyensis* group II intron is a large, folded RNA (> 400 nucleotides) consisting of six structural domains (I–VI). Making up nearly half of the intron, domain I consists of several subdomains and contains the binding site for the branch-site nucleophile (an adenosine in domain VI), as well as for the flanking 5' and 3' exons⁷¹. A wide variety of long-range tertiary interactions are needed to stabilize this intron structure, including a T-loop (5'-UGAGA-3') in subdomain IA that serves as a receptor for an intercalating unpaired adenosine in subdomain ID1 (Fig. 5). In addition, a sharp turn in the backbone directly 3' of the T-loop mediates a continuous base-stacking interaction between the guanosine and uridine nucleobases directly following the T-loop with another nearby adenosine in subdomain ID1. Together, these T-loop mediated interactions help stabilize a corner region of the intron where several domain I stems meet.

Riboswitches are untranslated, structured regions upstream of mRNA coding sequences that bind highly specifically to an associated ligand and thereby regulate gene expression. Riboswitches are broadly classified into type I and II based on the architecture of their binding pockets and the degree of conformational change that is induced by ligand binding⁷². However, as crystal structures of different riboswitches have become available, there is mounting evidence that most ligand binding pockets are formed at junctions and in loop regions between stems that in turn are stabilized by a range of structural motifs, even though each ligand and respective binding pocket is unique²⁵.

The adenosylcobalamin-sensing (coenzyme B₁₂) riboswitch is a relatively large riboswitch (>200 nucleotides) whose binding pocket is formed by a four-way junction between the P3, P4, P5, and P6/7–11 stems. An essential and conserved feature of the coenzyme B₁₂ riboswitch is that a long-range contact is made between the P4 and P6 arms mediated by a T-loop in the P4 arm, as observed in the recent B₁₂ riboswitch crystal structures from *Symbiobacterium thermophilum* (Fig. 6C)⁴⁶, *Thermoanaerobacter tengcongensis* (Fig. 6E), and another ocean surface microorganism (Fig. 6G)⁴⁷. Despite slight variation in sequence (5'-UGCAA-3', 5'-UGAGA-3', and 5'-UGUAA-3', respectively), the T-loops in all three riboswitches are structurally identical and form a tertiary contact with the P6/7–11 arm by acting as a receptor for an intercalating adenosine from the P6/7–11 arm (Fig. 6). Interestingly, in an entirely different context but in similar fashion, two T-loops (each associated with an intercalated adenosine) help bring together two short stem-loop pairs (i.e. P2:P6 and P3:P5 arms) in the flavin mononucleotide (FMN) riboswitch to form the ligand-binding pocket located at the core of its six-way junction of stems P1–P6 (Fig. 7)^{48, 49}.

Even more remarkably, recent structures of a prokaryotic and eukaryotic coenzyme thiamine pyrophosphate-sensing (THI-box) riboswitch from *E. coli*^{51, 52} and *Arabidopsis thaliana*^{50, 53} demonstrated that T-loops can also be involved in ligand binding. Thiamine pyrophosphates consist of three chemical moieties: a 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP) and a pyrophosphate (PP) separated by a thiazole group. THI-box riboswitches assume a pseudo-planar structure formed largely by two anti-parallel coaxial helical stacks. A loop region within each helical stack creates a binding pocket that is highly specific for both the HMP and PP of a thiamine pyrophosphate. Whereas specificity for the PP is achieved through nucleobase contacts and hydrogen bonding with water-coordinated

magnesium ions, HMP binds to the adjacent helical stack by inserting into a conserved T-loop (5'-UGAGA-3') reminiscent of an intercalated base (Fig. 8)⁵².

Natural functions of T-loops and their potential applications

T-loops are essential structural elements in many RNA molecules owing to their capacity to facilitate long-range intermolecular and intramolecular tertiary contacts and thereby to stabilize key regions of the overall RNA fold. As a result, T-loops occur frequently in phylogenetically-conserved areas of an RNA molecule and tend either to assemble into non-helical structural junctions or act as receptors in the distal portion of a solvent exposed stem-loop²⁵. Whereas the structural properties of T-loops are clearly evident and conserved in homologous tRNAs, rRNAs, and P RNAs, additional RNA structures are needed to determine the true prevalence of this important structural motif.

From a functional perspective, T-loops are versatile structural motifs due to the fact that a majority of the interactions stabilizing the T-loop are preserved regardless of whether an intercalating nucleobase or ligand is inserted between the T-loop nucleotides at positions 4 and 5. As a consequence, T-loops are extremely stable pentaloops that can act naturally as receptors for long-range docking with another RNA molecule or for recognizing an exogenous ligand. In fact, there are several RNAs that crystallize in part due to the lattice contacts formed by T-loop-mediated RNA-RNA interactions between symmetry-related molecules (Fig. 9)^{49, 71, 73}, thus suggesting that the use of T-loops can be potentially extended beyond its physiological function. An obvious application of this phenomenon is the intentional use of T-loops as a module for facilitating crystallization of RNA or RNA-protein complexes. By using a strategy that has already been demonstrated successfully with the GAAA tetraloop-tetraloop receptor interaction⁹⁰, kissing-loops⁹¹, the U1A spliceosomal protein⁹², and synthetic antibodies⁹³, perhaps T-loops too can be strategically positioned on peripheral stem regions to facilitate crystallization of RNA molecules.

Conclusion

In recent years, the structural and functional significance of T-loops in both small and large non-coding RNAs has become increasingly apparent as more RNA structures were solved. T-loops assume a sharp U-turn-like conformation and require specific stabilizing interactions to maintain their characteristic fold; the result is a motif strictly definable at the structural level but with a variable sequence signature. These structural considerations therefore make the accurate prediction of T-loops from sequence data alone difficult, even when sequence-based criteria are used in conjunction with secondary structural information. In tRNAs, a single T-loop helps bring together two highly-conserved regions in the TΨC and D loops and thus mediates the formation of the characteristic tRNA tertiary fold⁷³. The structural importance of T-loops is also demonstrated in the substantially larger rRNAs, in which multiple T-loops are utilized to interconnect key regions in rRNAs^{33, 66}. In RNase P, two interlocked T-loops in the P RNA S domain assume not only a structural role, but also a functional role by mediating specific base-stacking interactions with its precursor tRNA substrates⁵⁴. Recently solved crystal structures of a group II self-splicing intron^{71, 88, 89, 94} and riboswitches^{46–53, 95} reveal that T-loops are indeed present in a wide range of non-coding RNAs serving important functions. Although numerous RNA structural motifs have been characterized to date, more will likely be identified in coming years as new RNA structures become available. It remains to be seen whether closely related variants of the T-loop exist, whether T-loops can assist in binding to other small molecules like thiamine pyrophosphate^{50–53}, and whether T-loops can be utilized as an effective RNA crystallization module.

Acknowledgments

Funding for this work was provided by NIH grant GM58443 (A.M.). C.W.C. is a Takeda Global Research and Development Scholar of the ARCS Foundation, Chicago Chapter.

Bibliography

1. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*. 2012; 338:1435–1439. [PubMed: 23239728]
2. Wery M, Kwapisz M, Morillon A. Noncoding RNAs in gene regulation. *Wiley Interdiscip Rev Syst Biol Med*. 2011; 3:728–738. [PubMed: 21381218]
3. Eddy SR. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet*. 2001; 2:919–929. [PubMed: 11733745]
4. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*. 2012; 81:145–166. [PubMed: 22663078]
5. Hogg JR, Collins K. Structured non-coding RNAs and the RNP Renaissance. *Curr Opin Chem Biol*. 2008; 12:684–689. [PubMed: 18950732]
6. Holbrook SR. Structural principles from large RNAs. *Annu Rev Biophys*. 2008; 37:445–464. [PubMed: 18573090]
7. Leontis NB, Lescoute A, Westhof E. The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol*. 2006; 16:279–287. [PubMed: 16713707]
8. Hendrix DK, Brenner SE, Holbrook SR. RNA structural motifs: building blocks of a modular biomolecule. *Q Rev Biophys*. 2005; 38:221–243. [PubMed: 16817983]
9. Holbrook SR. RNA structure: the long and the short of it. *Curr Opin Struct Biol*. 2005; 15:302–308. [PubMed: 15963891]
10. Leontis NB, Westhof E. Analysis of RNA motifs. *Curr Opin Struct Biol*. 2003; 13:300–308. [PubMed: 12831880]
11. Batey RT, Rambo RP, Doudna JA. Tertiary Motifs in RNA Structure and Folding. *Angew Chem Int Ed Engl*. 1999; 38:2326–2343. [PubMed: 10458781]
12. Moore PB. Structural motifs in RNA. *Annu Rev Biochem*. 1999; 68:287–300. [PubMed: 10872451]
13. 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 U S A*. 2001; 98:4899–4903. [PubMed: 11296253]
14. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Szewczak AA, Kundrot CE, Cech TR, Doudna JA. RNA tertiary structure mediation by adenosine platforms. *Science*. 1996; 273:1696–1699. [PubMed: 8781229]
15. Schroeder KT, McPhee SA, Ouellet J, Lilley DM. A structural database for k-turn motifs in RNA. *RNA*. 2010; 16:1463–1468. [PubMed: 20562215]
16. Klein DJ, Schmeing TM, Moore PB, Steitz TA. The kink-turn: a new RNA secondary structure motif. *EMBO J*. 2001; 20:4214–4221. [PubMed: 11483524]
17. Mujeeb A, Clever JL, Billeci TM, James TL, Parslow TG. Structure of the dimer initiation complex of HIV-1 genomic RNA. *Nat Struct Biol*. 1998; 5:432–436. [PubMed: 9628479]
18. Wyatt JR, Puglisi JD, Tinoco I Jr. RNA folding: pseudoknots, loops and bulges. *Bioessays*. 1989; 11:100–106. [PubMed: 2695075]
19. Szewczak AA, Moore PB, Chang YL, Wool IG. The conformation of the sarcin/ricin loop from 28S ribosomal RNA. *Proc Natl Acad Sci U S A*. 1993; 90:9581–9585. [PubMed: 8415744]
20. Krasilnikov AS, Mondragon A. On the occurrence of the T-loop RNA folding motif in large RNA molecules. *RNA*. 2003; 9:640–643. [PubMed: 12756321]
21. Nagaswamy U, Fox GE. Frequent occurrence of the T-loop RNA folding motif in ribosomal RNAs. *RNA*. 2002; 8:1112–1119. [PubMed: 12358430]
22. Costa M, Michel F. Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J*. 1995; 14:1276–1285. [PubMed: 7720718]

23. Pley HW, Flaherty KM, McKay DB. Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature*. 1994; 372:111–113. [PubMed: 7526219]
24. Abu Almakarem AS, Petrov AI, Stombaugh J, Zirbel CL, Leontis NB. Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res*. 2012; 40:1407–1423. [PubMed: 22053086]
25. Reiter NJ, Chan CW, Mondragon A. Emerging structural themes in large RNA molecules. *Curr Opin Struct Biol*. 2011; 21:319–326. [PubMed: 21474301]
26. Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BF, Klug A. Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature*. 1974; 250:546–551. [PubMed: 4602655]
27. Kim SH, Suddath FL, Quigley GJ, McPherson A, Sussman JL, Wang AH, Seeman NC, Rich A. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science*. 1974; 185:435–440. [PubMed: 4601792]
28. Bessho Y, Shibata R, Sekine S, Murayama K, Higashijima K, Hori-Takemoto C, Shirouzu M, Kuramitsu S, Yokoyama S. Structural basis for functional mimicry of long-variable-arm tRNA by transfer-messenger RNA. *Proc Natl Acad Sci U S A*. 2007; 104:8293–8298. [PubMed: 17488812]
29. Gutmann S, Haebel PW, Metzinger L, Sutter M, Felden B, Ban N. Crystal structure of the transfer-RNA domain of transfer-messenger RNA in complex with SmpB. *Nature*. 2003; 424:699–703. [PubMed: 12904796]
30. Torres-Larios A, Swinger KK, Krasilnikov AS, Pan T, Mondragon A. Crystal structure of the RNA component of bacterial ribonuclease P. *Nature*. 2005; 437:584–587. [PubMed: 16113684]
31. Krasilnikov AS, Xiao Y, Pan T, Mondragon A. Basis for structural diversity in homologous RNAs. *Science*. 2004; 306:104–107. [PubMed: 15459389]
32. Krasilnikov AS, Yang X, Pan T, Mondragon A. Crystal structure of the specificity domain of ribonuclease P. *Nature*. 2003; 421:760–764. [PubMed: 12610630]
33. Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, Yusupov M. The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science*. 2011; 334:1524–1529. [PubMed: 22096102]
34. Ben-Shem A, Jenner L, Yusupova G, Yusupov M. Crystal structure of the eukaryotic ribosome. *Science*. 2010; 330:1203–1209. [PubMed: 21109664]
35. Brodersen DE, Clemons WM Jr, Carter AP, Wimberly BT, Ramakrishnan V. Crystal structure of the 30 S ribosomal subunit from *Thermus thermophilus*: structure of the proteins and their interactions with 16 S RNA. *J Mol Biol*. 2002; 316:725–768. [PubMed: 11866529]
36. Carter AP, Clemons WM Jr, Brodersen DE, Morgan-Warren RJ, Hartsch T, Wimberly BT, Ramakrishnan V. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science*. 2001; 291:498–501. [PubMed: 11228145]
37. Clemons WM Jr, Brodersen DE, McCutcheon JP, May JL, Carter AP, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. Crystal structure of the 30 S ribosomal subunit from *Thermus thermophilus*: purification, crystallization and structure determination. *J Mol Biol*. 2001; 310:827–843. [PubMed: 11453691]
38. Harms J, Schluenzen F, Zarivach R, Bashan A, Gat S, Agmon I, Bartels H, Franceschi F, Yonath A. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell*. 2001; 107:679–688. [PubMed: 11733066]
39. Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH, Noller HF. Crystal structure of the ribosome at 5.5 Å resolution. *Science*. 2001; 292:883–896. [PubMed: 11283358]
40. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*. 2000; 289:905–920. [PubMed: 10937989]
41. Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*. 2000; 407:340–348. [PubMed: 11014183]
42. Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon I, Franceschi F, et al. Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell*. 2000; 102:615–623. [PubMed: 11007480]

43. Wimberly BT, Brodersen DE, Clemons WM Jr, Morgan-Warren RJ, Carter AP, Vonnrhein C, Hartsch T, Ramakrishnan V. Structure of the 30S ribosomal subunit. *Nature*. 2000; 407:327–339. [PubMed: 11014182]
44. Clemons WM Jr, May JL, Wimberly BT, McCutcheon JP, Capel MS, Ramakrishnan V. Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature*. 1999; 400:833–840. [PubMed: 10476960]
45. Cate JH, Yusupov MM, Yusupova GZ, Earnest TN, Noller HF. X-ray crystal structures of 70S ribosome functional complexes. *Science*. 1999; 285:2095–2104. [PubMed: 10497122]
46. Peselis A, Serganov A. Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch. *Nat Struct Mol Biol*. 2012; 19:1182–1184. [PubMed: 23064646]
47. Johnson JE Jr, Reyes FE, Polaski JT, Batey RT. B12 cofactors directly stabilize an mRNA regulatory switch. *Nature*. 2012; 492:133–137. [PubMed: 23064232]
48. Vicens Q, Mondragon E, Batey RT. Molecular sensing by the aptamer domain of the FMN riboswitch: a general model for ligand binding by conformational selection. *Nucleic Acids Res*. 2011; 39:8586–8598. [PubMed: 21745821]
49. Serganov A, Huang L, Patel DJ. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature*. 2009; 458:233–237. [PubMed: 19169240]
50. Thore S, Frick C, Ban N. Structural basis of thiamine pyrophosphate analogues binding to the eukaryotic riboswitch. *J Am Chem Soc*. 2008; 130:8116–8117. [PubMed: 18533652]
51. Edwards TE, Ferre-D'Amare AR. Crystal structures of the thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. *Structure*. 2006; 14:1459–1468. [PubMed: 16962976]
52. Serganov A, Polonskaia A, Phan AT, Breaker RR, Patel DJ. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature*. 2006; 441:1167–1171. [PubMed: 16728979]
53. Thore S, Leibundgut M, Ban N. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science*. 2006; 312:1208–1211. [PubMed: 16675665]
54. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragon A. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*. 2010; 468:784–789. [PubMed: 21076397]
55. Lee JC, Cannone JJ, Gutell RR. The lonepair triloop: a new motif in RNA structure. *J Mol Biol*. 2003; 325:65–83. [PubMed: 12473452]
56. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res*. 2000; 28:235–242. [PubMed: 10592235]
57. Ito T, Yokoyama S. Two enzymes bound to one transfer RNA assume alternative conformations for consecutive reactions. *Nature*. 2010; 467:612–616. [PubMed: 20882017]
58. Nakanishi K, Bonnefond L, Kimura S, Suzuki T, Ishitani R, Nureki O. Structural basis for translational fidelity ensured by transfer RNA lysidine synthetase. *Nature*. 2009; 461:1144–1148. [PubMed: 19847269]
59. Palencia A, Crepin T, Vu MT, Lincecum TL Jr, Martinis SA, Cusack S. Structural dynamics of the aminoacylation and proofreading functional cycle of bacterial leucyl-tRNA synthetase. *Nat Struct Mol Biol*. 2012; 19:677–684. [PubMed: 22683997]
60. Gao YG, Selmer M, Dunham CM, Weixlbaumer A, Kelley AC, Ramakrishnan V. The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science*. 2009; 326:694–699. [PubMed: 19833919]
61. Jin H, Kelley AC, Loakes D, Ramakrishnan V. Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. *Proc Natl Acad Sci U S A*. 2010; 107:8593–8598. [PubMed: 20421507]
62. Korostelev A, Trakhanov S, Laurberg M, Noller HF. Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell*. 2006; 126:1065–1077. [PubMed: 16962654]
63. Neubauer C, Gillet R, Kelley AC, Ramakrishnan V. Decoding in the absence of a codon by tmRNA and SmpB in the ribosome. *Science*. 2012; 335:1366–1369. [PubMed: 22422985]

64. Petry S, Brodersen DE, Murphy FVt, Dunham CM, Selmer M, Tarry MJ, Kelley AC, Ramakrishnan V. Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. *Cell*. 2005; 123:1255–1266. [PubMed: 16377566]
65. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FVt, Weir JR, Ramakrishnan V. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science*. 2009; 326:688–694. [PubMed: 19833920]
66. Selmer M, Dunham CM, Murphy FVt, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science*. 2006; 313:1935–1942. [PubMed: 16959973]
67. Kazantsev AV, Krivenko AA, Harrington DJ, Holbrook SR, Adams PD, Pace NR. Crystal structure of a bacterial ribonuclease P RNA. *Proc Natl Acad Sci U S A*. 2005; 102:13392–13397. [PubMed: 16157868]
68. Adams PL, Stahley MR, Kosek AB, Wang J, Strobel SA. Crystal structure of a self-splicing group I intron with both exons. *Nature*. 2004; 430:45–50. [PubMed: 15175762]
69. Golden BL, Kim H, Chase E. Crystal structure of a phage Twort group I ribozyme-product complex. *Nat Struct Mol Biol*. 2005; 12:82–89. [PubMed: 15580277]
70. Guo F, Gooding AR, Cech TR. Structure of the Tetrahymena ribozyme: base triple sandwich and metal ion at the active site. *Mol Cell*. 2004; 16:351–362. [PubMed: 15525509]
71. Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. *Science*. 2008; 320:77–82. [PubMed: 18388288]
72. Montange RK, Batey RT. Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys*. 2008; 37:117–133. [PubMed: 18573075]
73. Shi H, Moore PB. The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA*. 2000; 6:1091–1105. [PubMed: 10943889]
74. Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*. 1994; 50:760–763. [PubMed: 15299374]
75. Leontis N.; Westhof E., editors. *RNA 3D Structure Analysis and Prediction*. Springer; Berlin Heidelberg: 2012. p. 281–298.
76. Abe T, Ikemura T, Sugahara J, Kanai A, Ohara Y, Uehara H, Kinouchi M, Kanaya S, Yamada Y, Muto A, et al. tRNADB-CE 2011: tRNA gene database curated manually by experts. *Nucleic Acids Res*. 2011; 39:D210–213. [PubMed: 21071414]
77. Sampson JR, Uhlenbeck OC. Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. *Proc Natl Acad Sci U S A*. 1988; 85:1033–1037. [PubMed: 3277187]
78. Nobles KN, Yarian CS, Liu G, Guenther RH, Agris PF. Highly conserved modified nucleosides influence Mg²⁺-dependent tRNA folding. *Nucleic Acids Res*. 2002; 30:4751–4760. [PubMed: 12409466]
79. Moore SD, Sauer RT. The tmRNA system for translational surveillance and ribosome rescue. *Annu Rev Biochem*. 2007; 76:101–124. [PubMed: 17291191]
80. Gillet R, Felden B. Transfer RNA(Ala) recognizes transfer-messenger RNA with specificity; a functional complex prior to entering the ribosome? *EMBO J*. 2001; 20:2966–2976. [PubMed: 11387229]
81. Moore PB. The three-dimensional structure of the ribosome and its components. *Annu Rev Biophys Biomol Struct*. 1998; 27:35–58. [PubMed: 9646861]
82. Moore PB, Steitz TA. The roles of RNA in the synthesis of protein. *Cold Spring Harb Perspect Biol*. 2011; 3:a003780. [PubMed: 21068149]
83. Torres-Larios A, Swinger KK, Pan T, Mondragon A. Structure of ribonuclease P--a universal ribozyme. *Curr Opin Struct Biol*. 2006; 16:327–335. [PubMed: 16650980]
84. Kazantsev AV, Krivenko AA, Pace NR. Mapping metal-binding sites in the catalytic domain of bacterial RNase P RNA. *RNA*. 2009; 15:266–276. [PubMed: 19095619]
85. Brown JW. The Ribonuclease P Database. *Nucleic Acids Res*. 1999; 27:314. [PubMed: 9847214]

86. Lancelot G, Helene C. Phosphate-guanosine interactions. A model for the involvement of guanine derivatives in autocatalytic reactions of ribonucleic acids. *J Biol Chem.* 1984; 259:15046–15050. [PubMed: 6210285]
87. Serganov A, Patel DJ. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet.* 2007; 8:776–790. [PubMed: 17846637]
88. Chan RT, Robart AR, Rajashankar KR, Pyle AM, Toor N. Crystal structure of a group II intron in the pre-catalytic state. *Nat Struct Mol Biol.* 2012; 19:555–557. [PubMed: 22484319]
89. Toor N, Rajashankar K, Keating KS, Pyle AM. Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol.* 2008; 15:1221–1222. [PubMed: 18953333]
90. Zhang L, Doudna JA. Structural insights into group II intron catalysis and branch-site selection. *Science.* 2002; 295:2084–2088. [PubMed: 11859154]
91. Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature.* 2009; 458:475–480. [PubMed: 19325628]
92. Ferre-D'Amare AR, Zhou K, Doudna JA. Crystal structure of a hepatitis delta virus ribozyme. *Nature.* 1998; 395:567–574. [PubMed: 9783582]
93. Koldobskaya Y, Duguid EM, Shechner DM, Suslov NB, Ye J, Sidhu SS, Bartel DP, Koide S, Kossiakoff AA, Piccirilli JA. A portable RNA sequence whose recognition by a synthetic antibody facilitates structural determination. *Nat Struct Mol Biol.* 2011; 18:100–106. [PubMed: 21151117]
94. Toor N, Keating KS, Fedorova O, Rajashankar K, Wang J, Pyle AM. Tertiary architecture of the *Oceanobacillus iheyensis* group II intron. *RNA.* 2010; 16:57–69. [PubMed: 19952115]
95. Serganov A, Patel DJ. Metabolite recognition principles and molecular mechanisms underlying riboswitch function. *Annu Rev Biophys.* 2012; 41:343–370. [PubMed: 22577823]

Figure 1. Basic architecture of the T-loop RNA motif both with and without an intercalating nucleobase

In Panels A and B, the 2D and 3D representations of the prototypic T-loop of tRNA (**PDB ID: 1EHZ**) are shown in yellow. The T-loop acts as an intramolecular receptor for an intercalating nucleobase (IB) stacked between the fourth and fifth T-loop positions. In Panel C, the T-loop is shown with cross-eyed stereoview with a ball-and-stick representation and its nucleotides are numbered 1 to 5 in the 5' to 3' direction. The intercalating guanine nucleobase is depicted in red. In panels D and E, the 2D and 3D representations of one of two T-loops identified in the THI-box (thiamine pyrophosphate-sensing) riboswitch (**PDB ID: 2HOP**), which acts as a receptor for a small molecule ligand, are shown in grey. In Panel F, the T-loop is shown with a cross-eyed stereoview with a ball and stick representation.

Figure 2. Functional and structural similarities between the tRNA and tmRNA T-loops

In panel A, the overall tRNA structure (**PDB ID: 1EHZ**) is shown in purple as cartoon representation, whereas the T-loop in its T Ψ C-loop is shown in yellow with ball-and-stick representation. In panel B, the tRNA-like domain of tmRNA (tmRNA-TLD) bound to SmpB (**PDB ID: 2CZJ**) is shown in green, whereas the T-loop in its T Ψ C-like-loop is shown in yellow with ball and stick representation. Panel C depicts the two superposed molecules aligned using only the backbone atoms of the T-loops. In Panel D, the superposed T-loops from tRNA (in purple) and tmRNA (in green) are shown as ball and sticks, illustrating the high degree of structural similarity.

Figure 3. Structural importance of T-loops in prokaryotic and eukaryotic rRNAs

In panels A and B, the overall structure of the prokaryotic 16S (PDB ID: 3V2C) and eukaryotic 18S (PDB ID: 2XZM) small subunit rRNAs are shown in grey with T-loops colored in red. Similarly, in panels C and D, the overall structure of the prokaryotic 23S (PDB ID: 2ZJR) and eukaryotic 28S (PDB ID: 3IZ9) large subunit rRNAs are shown in grey with T-loops colored in red. T-loops appear frequently in rRNAs and serve important structural roles primarily by mediating intramolecular tertiary interactions between loop regions that are distant in primary sequence. Furthermore, most rRNA T-loops are structurally conserved across prokaryotic and eukaryotic rRNAs.

Figure 4. T-loops are involved in substrate recognition and tertiary interactions in bacterial RNase P

In panel A, three interacting T-loops in the ternary complex structure of a bacterial RNase P holoenzyme bound to a tRNA substrate (PDB ID: 3Q1Q) are depicted in ball-and-stick representation. The two interlocked T-loops of the P RNA are shown in magenta and green, whereas the T-loop of the tRNA is shown in yellow. Panel B depicts the key interactions that stabilize the two interlocked T-loops of the P RNA. The nucleobase at position 2 (marked with * and **) in of the P RNA T-loops plays a key stabilizing role by forming a co-planar base triple with the closing base-pair (at positions 1 and 5) of the adjacent P RNA T-loop. These position 2 nucleobases also participate in hydrogen bonding interactions with the phosphate oxygen of the fifth nucleotide of the same T-loop. The nucleobases in positions 3, 4, and 5 of both P RNA T-loops form a contiguous, crescent-shaped base-stack that stacks on the position 3 nucleobase of the tRNA T-loop.

Figure 5. A structural T-loop mediates a long-range tertiary interaction in a self-splicing group II intron

In panel A, the secondary structure diagram of the *Oceanobacillus iheyensis* group II intron is shown with subdomains IA and IB and subdomain ID1 colored in purple and in green, respectively. The T-loop (5'-UGAGA-3') in subdomain IA mediates a long-range tertiary contact with a proximal loop region of subdomain ID1 (depicted by arrow) by serving as a receptor for an intercalating adenosine from ID1. The tertiary structure of this interaction in the context of the overall group II intron structure (**PDB ID: 3EOG**) is shown in panel B, and a close up view is shown in panel C. A sharp turn in the backbone directly 3' of the T-loop mediates a continuous base-stacking interaction between the guanosine and uridine nucleobases directly following the T-loop with another nearby adenosine (not shown) in subdomain ID1.

Figure 6. A conserved T-loop mediates a long-range tertiary interaction in two classes of B12-sensing riboswitches

In panels A and B, the secondary structure diagrams of two classes of B12-sensing riboswitches are shown with the P4 and P6 arms colored in purple and in green, respectively. A conserved T-loop is located in the loop region of the P4 arm and is colored in yellow. Arrows depict a structurally-important, tertiary interaction between the P4 and P6 arms. The overall tertiary structure of the B12-sensing riboswitches from an ocean surface microorganism (**PDB ID: 4FRN**), *Thermoanaerobacter tengcongensis* (**PDB ID: 4GMA**), and *Symbiobacterium thermophilum* (**PDB ID: 4GXY**) are depicted in panels C, E, and G, with the corresponding close up views of the T-loop mediated interaction in panels D, F, and H.

Figure 7. T-loops mediate two important long-range tertiary interaction in a flavin mononucleotide-sensing riboswitch

In panel A, the secondary structure diagram of a flavin mononucleotide (FMN)-sensing riboswitch is shown with the P2, P3, P5, and P6 arms colored in green, blue, magenta, and purple, respectively. T-loops located in the loop region of the P2 and P5 arms are colored in yellow. Both T-loops associate with an intercalated adenosine from the loop region of the P6 and P3 arms, respectively, to help bring together the P2:P6 and P3:P5 arms as seen in the overall tertiary structure of the FMN-sensing riboswitch depicted in panel B (**PDB ID: 3F2Q**). Close up views of the two T-loops and their key interactions are shown in panels C and D.

Recognition of thiamine pyrophosphate (TPP) by both prokaryotic (panel A; **PDB ID: 2GDI**) and eukaryotic (panel B; **PDB ID: 2CKY**) THI-box riboswitches involves base intercalation of the 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP) moiety of the TPP ligand between the fourth and fifth nucleotides of the T-loop, as well as hydrogen bonding interactions with the guanosine at the second T-loop position.

Figure 9. T-loops mediate the formation of crystal lattice contacts

Panels A, C, and E show the crystal packing of symmetry related molecules for a tRNA^{Phe} (PDB ID: 1EHZ), a FMN riboswitch (PDB ID: 3F2T), and a group II self-splicing intron (PDB ID: 3G78), respectively. The T-loops in all symmetry-related molecules are shown in yellow with ball-and-stick representation. Panels B, D, and F are the corresponding close up views showing the interactions between the T-loops and the symmetry-related molecules. In all three examples shown here, the third nucleotide of the T-loop stacks with a nucleobase in the adjacent molecule.

Table 1
Representative examples of T-loop-containing, non-coding RNAs from the Protein Data Bank

The T-loop in tRNA^{Phe} (PDB ID: 1EHZ) was used as a reference model and the T-loops in all RNA molecules were aligned using only the backbone atoms by *LSQKAB* of the CCP4 package. The resulting orientation of the RNA molecule was used to depict the overall structure in the table. The T-loop forming nucleotides are shown in red in the overall structure and in yellow with ball-and-stick representation in the close up view. The table also lists the T-loop sequence, the identity of the intercalating base (IB) stacked between the fourth and fifth T-loop nucleobases when applicable, and the backbone RMSD (Å) against the reference model.

PDB ID	Molecule	Overall structure	Close up view of T-loop	T-loop sequence (position 1 – 5)	IB	RMSD (Å)
1EHZ	tRNA ^{Phe}			UUCGA (54 – 58)	G	--
2CZJ	tmRNA-TLD			UUCGA (54 – 58)	G	0.77
3Q1Q	RNase P			AGAGA (114 – 118)	A	0.81

PDB ID	Molecule	Overall structure	Close up view of T-loop	T-loop sequence (position 1 – 5)	IB	MSD (Å)
3V2C	16S rRNA			GGAAG (1177 – 1181)	A	0.96
2XZM	18S rRNA			UGGAA (1524 – 1528)	A	0.80
2ZJR	23S rRNA			UGAAA (509 – 513)	A	0.66
3IZ9	28S rRNA			UGAAA (368 – 372)	A	0.78
3G78	Group II intron			UGAGA (31 – 35)	A	0.52

PDB ID	Molecule	Overall structure	Close up view of T-loop	T-loop sequence (position 1 – 5)	IB	MSD (Å)
4FRN	Cobalamin riboswitch			UGUAA (29 – 33)	A	0.63
3F2T	FMN riboswitch			UGAAA (18 – 22)	A	0.63
2HOP	THI-box riboswitch			UGAGA (39 – 43)	--	0.65

T-loop nucleotide frequencies and sequence consensuses

193 potential T-loops were initially identified using a non-redundant set of 692 RNA-containing X-ray crystal structures solved at or better than 4Å. Of these potential T-loops, 88 were subsequently determined to be false positives mostly due to artifacts created by truncations or disordered regions in the RNA molecule. The subgroup frequencies presented as percentages in the table were determined based on the resulting 105 visually-verified unique T-loops. The sequence consensuses are established based on the highest nucleotide frequency at the given position.

	Position 1	Position 2	Position 3	Position 4	Position 5
Transfer RNA (47 T-loops)	A = 2.1%	A = 0%	A = 2.1%	A = 36.2%	A = 97.9%
	G = 0%	G = 2.1%	G = 0%	G = 63.8%	G = 2.1%
	C = 0%	C = 0%	C = 97.9%	C = 0%	C = 0%
	U = 97.9%	U = 97.9%	U = 0%	U = 0%	U = 0%
Consensus sequence	U	U	C	G	A
Ribosomal RNA (39 T-loops)	A = 5.1 %	A = 0%	A = 59%	A = 69.2%	A = 87.2%
	G = 5.1%	G = 82.1%	G = 20.5%	G = 30.8%	G = 7.7%
	C = 15.4%	C = 0%	C = 5.1%	C = 0%	C = 0%
	U = 74.4%	U = 17.9%	U = 15.4%	U = 0%	U = 5.1%
Consensus sequence	U	G	A	A	A
Ribonuclease P (8 T-loops)	A = 37.5%	A = 0%	A = 62.5%	A = 37.5%	A = 100%
	G = 0%	G = 87.5%	G = 12.5%	G = 62.5%	G = 0%
	C = 0%	C = 0%	C = 12.5%	C = 0%	C = 0%
	U = 62.5%	U = 12.5%	U = 12.5%	U = 0%	U = 0%
Consensus sequence	U	G	A	G	A
Genetic elements: introns and riboswitches (11 T-loops)	A = 9.1%	A = 0%	A = 72.7%	A = 54.5%	A = 81.8%
	G = 0%	G = 90.9%	G = 0%	G = 45.5%	G = 9.1%
	C = 9.1%	C = 0%	C = 27.3%	C = 0%	C = 0%
	U = 81.8%	U = 9.1%	U = 0%	U = 0%	U = 9.1%
Consensus sequence	U	G	A	A/G	A
Average of subgroup frequencies	A = 13.5%	A = 0%	A = 49.1%	A = 49.3%	A = 91.7%

	Position 1	Position 2	Position 3	Position 4	Position 5
	G = 1.3% C = 6.1% U = 79.1%	G = 65.6% C = 0% U = 34.4%	G = 8.3% C = 35.7% U = 6.9%	G = 50.7% C = 0% U = 0%	G = 4.7% C = 0% U = 3.6%
Consensus sequence	U	G	A	G/A	A
Overall frequency	A = 6.7% G = 1.9% C = 6.7% U = 84.7%	A = 0% G = 47.6% C = 0% U = 52.4%	A = 35.2% G = 8.6% C = 49.5% U = 6.7%	A = 50.5% G = 49.5% C = 0% U = 0%	A = 92.4% G = 4.8% C = 0% U = 2.8%
Consensus sequence	U	U/G	C	A/G	A