Three-way RNA junctions with remote tertiary contacts: A recurrent and highly versatile fold

Marcos de la Peña, David Dufour and José Gallego

*RNA* 2009 15: 1949-1964 originally published online September 9, 2009
Access the most recent version at doi:10.1261/rna.1889509

References
This article cites 86 articles, 31 of which can be accessed free at:
http://rnajournal.cshlp.org/content/15/11/1949.full.html#ref-list-1

Article cited in:
http://rnajournal.cshlp.org/content/15/11/1949.full.html#related-urls

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here

To subscribe to *RNA* go to:
http://rnajournal.cshlp.org/subscriptions

Copyright © 2009 RNA Society
Three-way RNA junctions with remote tertiary contacts: A recurrent and highly versatile fold

MARCOS DE LA PEÑA,1 DAVID DUFOUR,2 and JOSÉ GALLEGO2
1Instituto de Biología Molecular y Celular de Plantas (UPV-CSIC), 46022 Valencia, Spain
2Centro de Investigación Príncipe Felipe, 46012 Valencia, Spain

ABSTRACT

Three-way junction RNAs adopt a recurrent Y shape when two of the helices form a coaxial stack and the third helix establishes one or more tertiary contacts several base pairs away from the junction. In this review, the structure, distribution, and functional relevance of these motifs are examined. Structurally, the folds exhibit conserved junction topologies, and the distal tertiary interactions play a crucial role in determining the final shape of the structures. The junctions and remote tertiary contacts behave as flexible hinge motifs that respond to changes in the other region, providing these folds with switching mechanisms that have been shown to be functionally useful in a variety of contexts. In addition, the juxtaposition of RNA domains at the junction and at the distal tertiary complexes enables the RNA helices to adopt unusual conformations that are frequently used by proteins, RNA molecules, and antibiotics as platforms for specific binding. As a consequence of these properties, Y-shaped junctions are widely distributed in all kingdoms of life, having been observed in small naked RNAs such as riboswitches and ribozymes or embedded in complex ribonucleoprotein systems like ribosomal RNAs, RNase P, or the signal recognition particle. In all cases, the folds were found to play an essential role for the functioning or assembly of the RNA or ribonucleoprotein systems that contain them.

Keywords: folding; hammerhead ribozyme; loop; ribonucleoprotein; ribosome; riboswitch; RNase P; signal recognition particle; structure; tertiary interaction; three-way junction

INTRODUCTION

In addition to the classical mRNA, tRNA, and rRNA molecules charged with transferring the genetic information from DNA to proteins, in recent years a plethora of new RNA molecules playing very diverse biological roles, alone or in association with proteins, have been described (Sharp 2009). In many cases, these RNAs fold into well-defined structures ranging in complexity from the simple hairpin or three-way junction folds found in some small ribozymes and riboswitches to the intricate architecture of ribosomal RNA or group II introns (Serganov and Patel 2007).

In the same way as protein architecture can be viewed as the assembly of two main types of structural units (α-helices and β-sheets) connected by loops and packaged by tertiary interactions in three-dimensional space, RNA architecture is essentially built through the assembly of double-stranded helices that pack in space side by side, orthogonally or on top of each other as continuous stacks. Key for organizing the three-dimensional packing of RNA helices are the junctions where three or more helices meet with each other, and tertiary contacts between different helical domains. Whereas the junctions impose topological constraints on three-dimensional helix organization (Lescoute and Westhof 2006; Laing and Schlick 2009), tertiary contacts are thought to modulate the final arrangement of stem–loops in space (Chauhan and Woodson 2008).

In this review we analyze simple RNA systems that are evolutionary and functionally very different, yet similar in terms of three-dimensional structure organization. These RNAs consist of three double helices that intersect at a three-way junction, adopting a recurrent-Y shape stabilized by coaxial stacking and remote tertiary contacts between the helices (Fig. 1). These folds are ubiquitously found in naked RNAs such as riboswitches or ribozymes, and ribonucleoprotein systems like ribosomal RNAs, RNase P, or the signal recognition particle (SRP). The juxtaposition of RNA domains at the junction and at the remote
tensory complexes enables the RNA helices of these motifs to adopt unusual shapes that are used as binding platforms by proteins and other RNA molecules. Moreover, the helical junction and the remote tertiary contact behave as flexible hinges whose structure is capable of responding to changes in the other region. Due to these properties, these motifs play essential functions in a variety of biological processes, representing a key building block in RNA architecture.

DESCRIPTION OF THE FOLD, AND FUNCTIONAL AND ARCHITECTURAL CONTEXTS

The RNA folds analyzed in this review comprise three double helices (domains) that, starting from the 5′ end of the RNA, will be identified as P1, P2, and P3 (Fig. 1). The three helices intersect at a three-way junction and adopt a “Y” shape, where the helix occupying the lower part of the “Y” coaxially stacks with one of the helices forming the two upper branches. Tertiary contacts between the two upper domains keep the remaining helix in parallel with the coaxial stack. These tertiary contacts are several base pairs away from the junction and usually involve loops, which will be named L or IL depending on whether they are apical or internal. In order to allow the parallel juxtaposition of the two upper domains, the junction strand (J) connecting these two helices is usually longer than the other two junction strands. In many cases, this strand forms a structured loop that establishes tertiary contacts with the lower helix. As a result of the sum of these tertiary interactions, all of the branches of the “Y” are parallel to each other, and for this reason these motifs will be hereafter identified as parallel-Y folds. These folds would fall within the C family of three-way junctions previously defined by Lescoute and Westhof (2006).

In many parallel-Y folds P1 and P3 are coaxially stacked, P2 and P3 occupy the upper branches of the “Y” and establish the remote tertiary contact, and J23 is the longest of the three junctions; an archetypal system of this type is shown in Figure 1. However, the coaxial stack can also be established between P1 and P2 or P2 and P3, and P1 may occupy one of the two upper branches rather than the lower part of the “Y” (Table 1).

The structural context in which these motifs have been observed is very diverse: they can be stable as isolated motifs (like hammerhead ribozymes) or embedded within larger RNA structures like RNase P or rRNA and can function either as naked RNA structures or forming complexes with proteins or metabolites (Table 1). Despite their global structural similarity (Figs. 2, 3), parallel-Y RNAs play very different functions across all kingdoms of life. This review first follows a functional criterion to group and to describe the structure and biological context of the motifs, and then compares all parallel-Y structures and discusses their mechanistic implications in the final sections of the article.

RIBOSWITCHES

Riboswitches are protein-free regulatory systems typically found embedded in the untranslated regions of bacterial metabolic genes. They generally consist of a 5′-sensing (aptamer) domain, which specifically binds metabolites when concentrations exceed a threshold, and a 3′-domain (expression platform) that controls gene expression. Depending on metabolite binding, the sensing domain can adopt alternative structures that influence the folding of the expression platform (Montange and Batey 2008; Batey 2009).

The adenine- and guanine-responsive riboswitches control the transcription of genes associated with purine metabolism in numerous bacterial species. The sensing domains of these riboswitches form similar 70-nucleotide (nt) parallel-Y folds (Fig. 2A; Batey et al. 2004; Serganov et al. 2004) and have strong sequence similarities, containing highly conserved nucleotides at the junction and at the L2 and L3 loops. These loops establish an extensive tertiary interaction across the major groove (Table 1), primarily comprising two Gs in the 3′ side of L2 and two Cs in the 5′ side of L3, which open toward the major groove to form two consecutive antiparallel Watson–Crick G·C pairs, a kissing loop-type interaction. Two noncanonical A·U and A·A pairs buttress the minor groove flanks of the G·C pairs, effectively forming with them two stacked tetrads (Fig. 2A). The ligand binding site is located at the three-way junction, completely encapsulated by the P1 helix and by J12, J31, and J23 nucleotides (Fig. 2A): it stacks above a P1
abolishes ligand binding (Batey et al. 2004), and single replacement of the wild-type loops with UUCG tetraloops for ligand binding by these riboswitches (Batey et al. 2004; Serganov et al. 2004). The L2–L3 interaction is essential for two types of purine riboswitches (Batey et al. 2004; Serganov et al. 2004), which also results in a further tightening of the junction structure and the stabilization of the nearby P1 helix containing the switching sequence (Fig. 2A,B); the sequestering of this sequence by P1 impedes alternative foldings of the expression platform (Montange and Batey 2008; Serganov 2009).

**HAMMERHEAD RIBOZYMES**

Hammerhead ribozymes are small catalytic RNAs initially discovered in viroids and virus satellite RNAs (Flores et al. 2001) and more recently identified in mrnas of plant (Przybilski et al. 2005) and animal (Martick and Scott 2008) species. They catalyze the specific cleavage of their own RNA backbone through a transesterification reaction, a step required for the replication of infectious circular RNA molecules.

Hammerhead ribozymes fold into typical parallel-Y structures (Fig. 2C). The catalytic center is located at the three-way junction and self-cleavage takes place at a J12 nucleotide, releasing 2',3'-cyclic phosphodiester and 5'-OH termini. Remote loop–loop interactions between domains P2 (domain I in the literature) and P3 (II) have been shown to greatly increase (from 100- to 1000-fold) the catalytic activity of these ribozymes (De la Peña et al. 2003; Khvorova et al. 2003) by promoting the adoption by the junction of a catalytically active conformation (Martick and Scott 2006). The L2–L3 interactions are established across the major groove of the helices and are quite extensive (Table 1), comprising a complex network of noncanonical base pairs as well as base interdigitations (Martick and Scott 2006; Chi et al. 2008) (Fig. 2C).

In contrast to the highly conserved sequence of the catalytic site nucleotides, there are many sequence and topology differences within the L2 loops and L3 loops of natural hammerheads (Fig. 4A; Dufour et al. 2009). Despite these differences, a conserved tertiary interaction motif embedded within the loop–loop complexes has been detected in at least four different hammerheads (Fig. 4; Dufour et al. 2009). The 5' U of L2 and the 3' A of L3 form a reverse Hoogsteen pair across the major groove in the X-ray structures of the *Schistosoma mansoni* (Martick and Scott 2006) and tobacco ring spot virus satellite RNA (Chi et al. 2008) (sTRSV) hammerheads, a base pair also supported by site-directed mutagenesis experiments in the hammerheads of the (+) and (−) polarity RNAs.
<table>
<thead>
<tr>
<th>Parallel-Y RNA</th>
<th>Organism</th>
<th>Coaxial stacking</th>
<th>J12/23/31 length (nt)</th>
<th>Main J structure</th>
<th>Main J tertiary interaction</th>
<th>Motifs involved</th>
<th>Motif side</th>
<th>Motif groove</th>
<th>Main feature</th>
<th>$\Delta$ASA (Å$^2$)</th>
<th>nhb</th>
<th>Protein binding</th>
<th>PDB code (ref)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purine riboswitch</td>
<td><em>Bacillus subtilis</em></td>
<td>P1–P3</td>
<td>3/4/2</td>
<td>G•C pair (J23)</td>
<td>J12–J23, J23–P1 (base triples, minor)</td>
<td>L2–L13,</td>
<td>3′–5′</td>
<td>Major</td>
<td>Kissing G•Cs</td>
<td>924 (66)</td>
<td>18</td>
<td>No</td>
<td>1JBD (Batey et al. 2004), 1Y27, 1Y26 (Serganov et al. 2004)</td>
</tr>
<tr>
<td>TPP riboswitch</td>
<td><em>Escherichia coli</em></td>
<td>P1–P2</td>
<td>0/5/2</td>
<td>None</td>
<td>J23–J31–P2 (base quads, minor)</td>
<td>IL2–L3,</td>
<td>5′–3′</td>
<td>Minor</td>
<td>Min-maj</td>
<td>860 (66)</td>
<td>8</td>
<td>No</td>
<td>2GDI (Serganov et al. 2006)</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td><em>Arabidopsis thaliana</em></td>
<td>P1–P2</td>
<td>0/5/2</td>
<td>None</td>
<td>J23–J31–P2 (base quads, minor)</td>
<td>IL2–L3,</td>
<td>3′</td>
<td>Minor</td>
<td>Min-maj</td>
<td>506 (63)</td>
<td>7</td>
<td>No</td>
<td>2CKY (Thore et al. 2006)</td>
</tr>
<tr>
<td>Hammerhead ribozyme</td>
<td><em>Struvite</em> (E+)</td>
<td>P1–P3</td>
<td>1/2/3</td>
<td>U-turn hxlp</td>
<td>J23–P1 (minor)</td>
<td>IL2–L3,</td>
<td>5′–3′</td>
<td>Major</td>
<td>Reverse U•A$^{\alpha}$</td>
<td>806 (62)</td>
<td>7</td>
<td>No</td>
<td>2QLS (Chi et al. 2008)</td>
</tr>
<tr>
<td>Schistosoma mansoni</td>
<td><em>Schistosoma mansoni</em></td>
<td>P1–P3</td>
<td>1/2/3</td>
<td>U-turn hxlp</td>
<td>J23–P1 (minor)</td>
<td>IL2–L3,</td>
<td>5′–3′</td>
<td>Major</td>
<td>Reverse U•A$^{\alpha}$</td>
<td>1190 (92)</td>
<td>8</td>
<td>No</td>
<td>2GQZ (Martick and Scott 2006)</td>
</tr>
<tr>
<td>RNase P (B-type)</td>
<td><em>B. subtilis</em></td>
<td>P1–P3</td>
<td>1/2/0</td>
<td>T-loop</td>
<td>J23–P1 (A-minor)</td>
<td>IL2–L3,</td>
<td>3′–5′</td>
<td>Minor</td>
<td>Receptor-GAAA Exh. G-A stack</td>
<td>588 (65)</td>
<td>6</td>
<td>No</td>
<td>1NBS (Krasilnikov et al. 2003)</td>
</tr>
<tr>
<td>SRP (S)</td>
<td><em>Methanococcus jannaschii</em></td>
<td>P1–P3</td>
<td>2/4/1</td>
<td>No</td>
<td>J23–P1 (base-minor)</td>
<td>IL2–L3,</td>
<td>3′–3′</td>
<td>Minor</td>
<td>GNAG-GNRA A-minor</td>
<td>535 (59)</td>
<td>7</td>
<td>SRP19, SRP54</td>
<td>1NGC (Hainzl et al. 2002), 1Z43 (Hainzl et al. 2005)</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td><em>Homo sapiens</em></td>
<td>P1–P3</td>
<td>0/8/0</td>
<td>T-loop</td>
<td>J23–P1 (A-minor)</td>
<td>IL2–L3,</td>
<td>3′–3′</td>
<td>Minor</td>
<td>GNAG-GNRA Minor-A</td>
<td>632 (57)</td>
<td>9</td>
<td>SRP19, SRP54</td>
<td>1L9A (Oubridge et al. 2002), 1MQO (Kuglstatter et al. 2002)</td>
</tr>
<tr>
<td>SRP (Ala)</td>
<td><em>H. sapiens</em></td>
<td>P1–P3</td>
<td>0/4/0</td>
<td>U-turn</td>
<td>J23–P1 (middle)</td>
<td>L2–L13,</td>
<td>3′–5′</td>
<td>Major</td>
<td>Kissing G•Cs</td>
<td>820 (63)</td>
<td>10</td>
<td>SRP9/14</td>
<td>1EBO (Weißenrieder et al. 2000)</td>
</tr>
<tr>
<td>23S rRNA H42-43-44</td>
<td><em>E. coli, Haloarcula marismortui</em></td>
<td>P1–P3</td>
<td>1/2/1</td>
<td>T-loop</td>
<td>J23–P1 (A-minor)</td>
<td>IL2–L3,</td>
<td>3′–3′</td>
<td>Major</td>
<td>Ribose-ribose Base triples Reverse U•A$^{\alpha}$</td>
<td>1581 (66)$^	ext{m}$</td>
<td>14$m$</td>
<td>L11</td>
<td>1MAMS (Wimberly et al. 1999), 1QA6 (Conn et al. 1999), 1S72 (Klein et al. 2004)</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Parallel-Y RNA</th>
<th>Organism</th>
<th>Coaxial stacking</th>
<th>Three-way junction</th>
<th>Remote tertiary interaction</th>
<th>Motif(s) involved</th>
<th>Motif side</th>
<th>Motif groove</th>
<th>Main feature</th>
<th>∆ASA / nc</th>
<th>nhb</th>
<th>Protein binding</th>
<th>PDB code</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA h35-36-37</td>
<td><em>Thermus thermophilus</em></td>
<td>P1–P2</td>
<td>0/2/2</td>
<td>None</td>
<td>J23–P1 (base triples, major)</td>
<td>IL1–L3</td>
<td>3’–3’</td>
<td>Major</td>
<td>Base-backbone</td>
<td>658 (66)</td>
<td>6</td>
<td>S2, S3, S5</td>
<td>1JSE (Wimberly et al. 2000)</td>
</tr>
</tbody>
</table>

*The junction strand connecting the helices forming the two upper branches of the “Y” is underlined.

bThe groove mainly used for establishing the tertiary interaction is also indicated.

cL is apical loop, IL is internal loop, H is double helix, and the distance in base pairs from the Ls or ILs to the junction or to the next relevant tertiary interaction is indicated in subscripts.

dSide (5’ or 3’) of the L or IL mainly used for establishing the tertiary interaction.

eSee the main text for more details.

fSolvent-accessible surface area (ASA) (Connolly 1983) buried by the interaction. The areas were computed with the program MOE 2007.09 (CCG, Inc.) after adding hydrogen atoms to the X-ray structures and using a radius of 1.4 Å for the water probe. A normalized ∆ASA / nc value is also shown in parentheses, where nc is the number of contacting residues. These were defined as those nucleotides containing any atom at a distance of 3.5 Å or less from the other domain. ∆ASA / nc indicates how much surface is buried by each nucleotide directly involved in the interaction; high values typically indicate inter-domain base stacking.

gTotal number of hydrogen bonds contained in the interaction. These were calculated with cutoffs of 3.5 Å for the donor-acceptor distance and 85° and 55° for the lone-pair and out-of-plane angles, respectively, using MOE 2007.09.

hWhen several PDB structures were available, the underlined one was used to compute ∆ASA, ∆ASA / nc, and nhb.

iHexaloop containing a U-turn and closed by a single Watson–Crick G=C pair.

jT-loop: pentaloop containing a U-turn and closed by a single base pair (Nagaswamy and Fox 2002; Krasilnikov and Mondragon 2003; Lee et al. 2003).

kPentaloop closed by a single Watson–Crick G=C pair.

lA is extrahelical and stacks into domain 2.

mThese tertiary contacts are adjacent to each other and had to be analyzed collectively.
of chrysanthemum chlorotic mottle viroid (CChMVd) (Dufour et al. 2009). In addition, the penultimate U of L2 in the sTRSV hammerhead forms a Watson–Crick pair with the 3' A of L3, so that a base triple is formed by these two nucleotides and the L2 5' U (Chi et al. 2008); this base-triple interaction is also consistent with the mutagenesis analyses of CchMVd(+) and CchMVd(-) hammerheads (Dufour et al. 2009). In the Schistosoma hammerhead this U:A:U base triple is not observed, but the last pyrimidine of L2 is extrahelical and protrudes into L3, facilitating the displacement of the 3' A of L3 to form the reverse Hoogsteen pair with the 5' U of L2 across the major groove and, in some cases [i.e., in sTRSV(+)] and CChMVd(+) and CChMVd(-) hammerheads], a base triple with the extrahelical L2 pyrimidine (Fig. 4C). The L2 and L3 nucleotides involved in this tertiary interaction motif are conserved among natural hammerheads (Fig. 4A), and mutations affecting these residues produce large reductions of catalytic activity (Khvorova et al. 2003; Dufour et al. 2009).

The substantial opening toward the minor groove (about 180°) of the last two nucleotides of L3 to interact with L2 across the major groove is highly unusual: L3s include, for example, GNRA loops (Fig. 4B), which commonly establish tertiary interactions across the minor groove (Cate de la Peña et al. 1954 RNA, Vol. 15, No. 11)
et al. 1996a). This unusual L2–L3 interaction likely contributes to the observed overwinding of helix P3 and to the severe bending and unwinding of helix P2, which in turn produce crucial conformational changes at the catalytic junction located 5–6 base pairs (bp) away from the interaction (Fig. 5A,B). Namely, the J23 junction becomes more structured, gaining one Watson–Crick base pair and moving away from the P1 helix, and the J12 cleavage nucleotide shifts to the J31 side of the junction, becoming properly oriented for in-line attack by J31 residues (Martick and Scott 2006).

SIGNAL RECOGNITION PARTICLE (S DOMAIN)

The SRP is a universally conserved ribonucleoprotein that mediates the cotranslational targeting of membrane proteins to the endoplasmic reticulum or the bacterial plasma membrane (Egea et al. 2005; Halic and Beckmann 2005). In eukaryotic and archaeal cells, the SRP RNA is composed of z300 nt arranged in two domains, named S and Alu. In mammals, six different proteins bind to the SRP RNA, whereas in archaea only SRP19 and SRP54 homologs have been identified so far. In eubacteria the SRP comprises a smaller RNA bound to an SRP54 homolog, the only SRP protein component present in all organisms. SRP54 plays a major role in SRP function, since it binds the signal sequence of the nascent peptides and directly interacts with the SRP receptor at the membrane.

The S domain of eukaryotic and archaeal SRPs is z100-nt long and adopts a typical parallel-Y fold (Fig. 3A). The loops capping helices P2 (helix 6 in the literature) and P3 (helix 8) are conserved GNAG and GNRA tetraloops,
respectively, and establish tertiary interactions across the minor groove that mainly consist of side-by-side base–base and base–backbone hydrogen-bonding contacts, with no base interdigitations (Table 1; Hainzl et al. 2002; Kuglstatter et al. 2002; Oubridge et al. 2002; Hainzl et al. 2005). Both tetraloops adopt similar conformations in their undocked form, with the last three nucleotides stacked and opened toward the minor groove (Jucker et al. 1996; Sakamoto et al. 2002). In fact, GNRA loops usually establish this type of tertiary interactions across the minor groove (Cate et al. 1996a), which is quite different from the one observed in hammerhead ribozymes (see above). The SRP19 protein binds to this L2–L3 complex from the major groove side (Fig. 3A), contributing to clamp together helices P2 and P3 (Hainzl et al. 2002; Kuglstatter et al. 2002; Oubridge et al. 2002). SRP19 needs a preformed L2–L3 complex to bind, since mutations affecting nucleotides involved in important RNA tertiary contacts destabilize the protein–RNA complex (Zwieb 1992).

Further down the parallel-Y, helix P2 consists of an almost regular A-form stem, but helix P3 contains two highly conserved internal loops that comprise the binding site of the SRP54 protein (Fig. 3A). The internal loop closest to the tip of P3 is symmetric and likely to be rigid because of the mostly continuous stacking of base pairs (Hainzl et al. 2002; Oubridge et al. 2002), whereas the other one (IL3) is asymmetric and provides a further flexible hinge region (Schmitz et al. 1999) in addition to the three-way junction (Fig. 3A). IL3 undergoes a dramatic structural rearrangement upon SRP54 binding, exposing some bases outside of the RNA helix for interaction with the protein (Batey et al. 2000; Kuglstatter et al. 2002; Oubridge et al. 2002). The main function of helix P2 and of the apical L2–L3 tertiary interaction is the stabilization of this protein–RNA interaction, which is essential for SRP function. In the archaeal SRP, two consecutive A’s loop out of helix P2 and form A-minor motif interactions (Doherty et al. 2001; Nissen et al. 2001) with base pairs flanking IL3, providing another point of attachment between the two helices—in addition to the apical L2–L3 interaction—which promotes a compact side-by-side association of the two helices (Yin et al. 2004). Binding of SRP19 stabilizes the L2–L3 interaction and causes a substantial repositioning of helix P2 relative to helix P3 that, transmitted through the A-minor P2–P3 interaction, results in the splaying out of the bases of IL3 (Hainzl et al. 2002, 2005), which can now interact with SRP54. However, spontaneous fluctuations of the two helices in the absence of SRP19 likely generate similar favorable conformations for SRP54 binding (Hainzl et al. 2005) and this probably accounts for the observed intrinsic affinity of free archaeal SRP RNA for SRP54 (Bhuiyan et al. 2000; Diener and Wilson 2000). In mammalian SRP RNA the additional A-minor P2–P3 interaction does not exist in the absence of SRP54 binding, so that due to the flexibility provided by the three-way junction and asymmetric IL3 (Schmitz et al. 1999) hinge regions, the relatively weak L2–L3 apical interaction does not bring the two helices side-by-side as efficiently. The clamping of the apical loops by SRP19 induces a closer side-by-side interaction of P2 and P3, so that when SRP54 binds to P3 and collapses the asymmetric loop, two A’s in the short strand of IL3 flip out of the helix to form A-minor motif interactions with helix P2, stabilizing the SRP54-P3–P2 ternary complex. The structure of the SRP54 bound-P3 would not be stable in the absence of a firm juxtaposition of helices P2 and P3 (Kuglstatter et al. 2002; Oubridge et al. 2002; Schmitz et al. 2002; Hainzl et al. 2005).
RIBOSOMAL RNA (23S rRNA COMPLEX WITH L11)

The highly conserved 58-nt parallel-Y motif formed by helices 42, 43, and 44 of 23S rRNA (Fig. 3B) was one of the first structures of this type to be described (Conn et al. 1999; Wimberly et al. 1999). The two upper domains of the “Y” (H43 and H44) bind to ribosomal protein L11, and the complex forms a knob protruding above the peptideyl transferase cleft of the large ribosomal subunit (Ban et al. 2000). This complex interacts with the EF-G and EF-Tu elongation factors during protein synthesis. The parallel-Y fold is characterized by severely distorted and unwound P2 (H43) and P3 (H44) domains. P2 comprises an internal loop (IL2) and a large and bent 9-nt apical loop (L2), whereas P3 has an internal loop (IL3) just below two G·C pairs leading to its apical loop (Fig. 3B). The fold is quite compact because the two unwound helices establish extensive tertiary interactions along their major grooves (Table 1; Conn et al. 1999; Wimberly et al. 1999). Two G and C nucleotides protruding out of the 3′ side of L2 bind to the major groove of the P3 G·C pairs, forming two consecutive triples that rest above a third triple formed by IL3 nucleotides, which acts as a stacking platform reminiscent of the A platforms observed in group I introns (Cate et al. 1996b). Just before this triple, the essential 5′ A of IL3 (A1088 in E. coli) protrudes out of the P3 domain in a syn conformation and stacks within the P2 helix, forming a reverse Hoogsteen pair with the RNA backbone, L11 also establishes direct hydrogen bonds with the 5′ U of IL2 and the 5′A of IL3 (A1088) forming the reverse Hoogsteen pair (Conn et al. 1999; Wimberly et al. 1999). Thus, L11 indirectly and directly promotes the P2–P3 tertiary interaction by stabilizing the P2 conformation and by locking into place the P3–P2 A·U base pair.

OTHER PARALLEL-Y RNAs

RNase P

A parallel-Y fold can also be found embedded in the specificity (S) domain of B-type bacterial RNase P (Krasilnikov et al. 2003). RNase P is a universal ribonucleoprotein that processes the 5′ end of pre-tRNA molecules (Kazantsev and Pace 2006; Torres-Larios et al. 2006). The main tertiary interaction between domains P2 and P3 (P10.1 and P12 in the literature) is a GAAA tetraloop–receptor interaction (Cate et al. 1996b) established across the minor groove of the helices (Fig. 2D). Further down the “Y,” the P3 helix contains a large and highly structured internal loop that provides additional tertiary contacts with an internal loop in helix P2 via stacking of bulged-out A and G nucleotides. The nucleotides of the J23 junction form a T-type pentaloop (Nagaswamy and Fox 2002; Krasilnikov and Mondragon 2003; Lee et al. 2003), and two A′s protruding from this loop buttress the minor groove of the P1 helix via ribose–ribose (Cate et al. 1996a) and A-minor (Doherty et al. 2001; Nissen et al. 2001) interactions. The nucleotides involved in the P2–P3 and J23–P1 tertiary interactions described above are conserved across all phylogenetic domains, and mutations that destabilize the Y fold above a certain threshold (like A1088U) abolish L11 binding and are lethal in vivo (Lu and Draper 1994; Conn et al. 1998; Maeder et al. 2006). The Ct domain of L11 recognizes the flat surface provided by the minor groove side of the bent and unwound P2 domain. Although the interaction mainly involves contacts with the RNA backbone, L11 also establishes direct hydrogen bonds with the 5′ U of IL2 and the 5′A of IL3 (A1088) forming the reverse Hoogsteen pair (Conn et al. 1999; Wimberly et al. 1999). Thus, L11 indirectly and directly promotes the P2–P3 tertiary interaction by stabilizing the P2 conformation and by locking into place the P3–P2 A·U base pair.
Thus, multiple tertiary interactions within the S domain may modulate tRNA recognition in addition to RNase P folding.

Alu domain of the SRP

In mammals, this 100-nt domain contains the 5′ and 3′ ends of the SRP RNA and, together with the SRP9/14 protein heterodimer, delays the elongation of the nascent protein chain targeted by the SRP (Egea et al. 2005; Halic and Beckmann 2005). The 5′ region of the domain forms a parallel-Y fold stabilized by a L2–L3 tertiary interaction involving three antiparallel G·C Watson–Crick pairs (Weichenrieder et al. 2000). The junction J23 nucleotides are highly conserved and form a U-turn, to which the SRP14 component of SRP9/14 binds from the side opposite P1 (Fig. 3C). It has been suggested that the P1 helix, which links the SRP Alu and S domains, folds back on itself and binds the SRP9 side of the heterodimer (Weichenrieder et al. 2000). This involves a sharp turn that would nevertheless be facilitated by the presence of the SRP RNA 5′- and 3′-ends within the P1 helix, 3 bp away from the junction. The function of the parallel-Y motif formed by the Alu RNA appears to be, in this case, the formation of an appropriate platform for SRP9/14 binding: mutations disrupting the loop–loop interaction abolish SRP9/14-RNA binding and SRP assembly (Huck et al. 2004).

Additional ribosomal motifs

In addition to the H42–H43–H44 complex with L11, the 23S rRNA H18–H19–H20 and H90–H91–H92 junctions adopt parallel-Y folds (Ban et al. 2000). The first is located within 23S rRNA domain I, relatively close to the rRNA 5′ end, is contacted from opposite sides by ribosomal proteins L4P and L24P (Klein et al. 2004) (Fig. 3D), and plays a critical role in the assembly of the large subunit (Skinner et al. 1985). The second junction belongs to domain V of 23S rRNA. In this motif, the coaxial stacking is established between P1 and P2 rather than P1 and P3; P1 and P3 are the interacting domains occupying the two upper branches and P2 forms the lower part of the “Y” (Fig. 3E; Table 1).
The solvent-exposed side of this parallel-Y (called loop A in the literature) forms the A-site of the peptidyl transferase cleft, whereas the opposite side is bound by proteins L3P and L14P (Fig. 3E). In contrast to the H42–H43–H44 junction, this parallel-Y is inverted relative to the surface of the ribosome. The P1 (H90)–P3 (H92) complex is at the bottom of the subunit cleft. Two G and U solvent-exposed bases at the tip of L3 have been shown to form antiparallel base pairs with the 3′ CCA tail of aminoacyl tRNA analogs (Schmeing et al. 2005). The P1 internal loop that establishes the remote contact with L3 accommodates the tRNA aminoacyl moiety (Schmeing et al. 2005) and is close to the catalytic nucleotides of the ribosome at the base of helix 89. At the other side of the “Y,” the third domain (P2 or H91) protrudes out of the subunit cleft and is packed between the upper portions of H89 and H95. In particular, at the top of helix 95 is the sarcin-ricin loop (SRL), which rests against the tip of P2 (Ban et al. 2000) and is essential for EF-G and EF-Tu binding (Moazed et al. 1988; Wool et al. 1992). The 120-nt 5S RNA of the 50S subunit adopts a parallel-Y conformation, where the coaxial stacking is established between domains P2 and P3 (Fig. 3F; Table 1). This Y fold is bound by ribosomal proteins L5P, L18P, and L21e, which contribute to glue the 5S RNA molecule to the rest of the subunit (Ban et al. 2000). At the head of the small ribosomal subunit, helix 35 (P1) also forms a parallel-Y fold with h36 (P2) and h37 (P3), which is bound by ribosomal proteins S2, S3, and S5 (Fig. 3G; Wimberly et al. 2000).

STRUCTURE–FUNCTION RELATIONSHIPS

Three-way junction

A number of structural features are shared by most parallel-Y folds studied in this review (Table 1). In order to allow for the parallel juxtaposition of the coaxial stack with the third helix, the junction strand connecting the two upper branches of the “Y” usually contains more nucleotides than the other two junction strands (Table 1; Lescoute and Westhof 2006) and makes a sharp turn when traveling from one helix to the next helix (Figs. 1, 2, 3). In addition, the topology of the three-way junction is remarkably conserved: one of the strands of the helix occupying the lower branch of the “Y” always travels in a 5′-to-3′ direction across the coaxial stack and toward the third helix (Figs. 2, 3), leaving the minor groove of the lower helix available for interactions with nucleotides of the long junction (Table 1; Lescoute and Westhof 2006). Most junctions maintain this topology regardless of the P1–P3 distribution among the three branches of the “Y.” In the H90–H91–H92 and h35–h36–h37 ribosomal junctions, for example, the coaxial stack is established between P1 and P2, P1 and P3 occupy the upper branches of the “Y,” and the junction strand making the sharp turn is J12 (Fig. 3E,G); in 5S RNA there is P2–P3 coaxial stacking and P1 and P2 occupy the upper branches of the “Y” connected by the J12 junction strand (Fig. 3F). The topology of these junctions, however, does not vary relative to that of classical folds with P1–P3 stacking (Figs. 1–3). The fold contained in the TPP riboswitch represents the only example of a parallel-Y with an unusual junction topology. In this system P2 and P3 occupy the two upper branches of the “Y,” but the coaxial stack is between P1 and P2 instead of P1 and P3, and the direction of the strand connecting the third helix with the lower domain of the coaxial stack is inverted (Fig. 2B). The global shape of this fold, however, is remarkably similar to that of parallel-Y’s with standard topologies (Figs. 2, 3).

Regardless of junction topology, the long and bent junction strand connecting the two upper helices of the “Y” often forms a U-turn structure or a more complex loop structure containing a U-turn, like T-type pentaloops (Nagaswamy and Fox 2002; Krasilnikov and Mondragon 2003; Lee et al. 2003) or U-turn-containing hexaloops closed by a single base pair (Table 1). From these T- or T-like loops, one or two bases often contact the minor groove of the lower helix through A-minor (Doherty et al. 2001; Nissen et al. 2001) or equivalent interactions. In several cases, however, the long junction does not form a structured loop, but participates in a network of interactions (usually involving base triples or quadruples) with other junction strands or helical domains. These unusually compact junctions are observed in riboswitches and ribosomal junctions (Table 1) and are connected to the switching and folding roles played by these systems.

Distal tertiary contacts

The remote tertiary interactions between the domains occupying the upper branches of the Y are more variable, since each system has evolved in a different way by varying the nature of the contacts, the number of base pairs separating them from the junction or other relevant regions of the motif, and the presence or absence of internal loops distorting the stems (Table 1). On average, these interactions take place at about half a helix turn (5 bp) from the junction (Table 1). A quantification of the surface areas and hydrogen bonds involved in the contacts (Table 1) indicates that, in general, tertiary interactions established across the major groove tend to be more extensive and contain more interdomain base stacks than those established across the minor groove, which often consist of side-by-side contacts without extensive base cross-stacking (Table 1). This is a logical consequence of the groove shapes: whereas the bases and sugars are more accessible in the shallow (minor) groove, greater distortions are required to make the nucleoside groups available for tertiary interactions across the deep (major) groove. For example, the apical regions of the parallel-Y motifs contained in the TPP riboswitch, the S domain of the SRP, and RNase P interact across the minor groove, but these interactions are always...
reinforced by additional tertiary interactions further down the fold, or by protein binding to the RNA tertiary complex (Table 1). In contrast, interactions across the major groove usually comprise extensive antiparallel kissing interactions with multiple hydrogen bonds, as observed in purine riboswitches, or involve the intercalation of key nucleotides within the base stack of the other domain, as observed in hammerhead ribozymes and the H42–H43–H44 23S rRNA junction (Table 1).

A comparison of the parallel-Y folds formed by a minimal hammerhead construct lacking the distal tertiary interaction (Fig. 5A; Scott et al. 1996), the sTRSV full-length hammerhead (Fig. 5B; Chi et al. 2008), and the aptamer domain of purine riboswitches (Fig. 5C; Batey et al. 2004; Serganov et al. 2004) provides clues about the effects of the distal tertiary contact and the length of the stems separating it from the junction. Both the sTRSV hammerhead and the purine riboswitch aptamer form naked Y-folds containing a single P2–P3 tertiary interaction. This interaction is established across the major groove of two apical loops separated from the junction by mostly standard base-paired stems (Table 1). Mainly due to a shorter P3 stem and to the unusual opening and inter-domain stacking of the conserved 3′ A of L3 to interact with the 5′ U of L2 in the tertiary complex (Fig. 4C), P2 is bent and P3 is partially overwound in the full-length hammerhead (Fig. 5B) compared with the minimal hammerhead (Fig. 5A) and the purine riboswitch (Fig. 5C) folds. In the purine riboswitch fold, the P2 and P3 stems have similar lengths, and the main tertiary interaction involves different loop sides (the 3′ side of L2 and 5′ side of L3) relative to the hammerhead fold. As a consequence, P2 and P3 are not overwound or bent, and a tight binding site can be formed for the purine ligand at the junction. The bending of P2 and overwinding of P3 in full-length hammerheads, on the other hand, allows the junction to adopt a catalytically competent conformation (Martick and Scott 2006). In fact, different natural hammerhead ribozymes may have tuned their self-cleaving activities by varying the length and irregularities of their P2 and P3 stems (Fig. 4A).

The parallel-Y fold formed by the Alu domain of SRP RNA (Weichenrieder et al. 2000) also contains a single tertiary interaction established across the major groove of two apical loops and provides another example of how the distal tertiary contact determines the shape of the Y structure. As in hammerhead ribozymes, P2 is longer than P3, but in the Alu fold, the P3 domain is bent and unwound instead of overwound (Fig. 5, cf. B and D). This is because the main tertiary interactions take place between the 3′ side of L2 and the 5′ side of L3; in this case, P3 needs to unwind to bring into plane the bases forming the two L2–L3 kissing G-C pairs (Fig. 5D).

Distortions in the P2 and P3 domains relative to regular A-form geometry also have an important impact in the conformation of the motif as illustrated by the 23S rRNA–L11 complex (Fig. 3B; Conn et al. 1999; Wimberly et al. 1999). In this case the motif is unusually compact because the P2 and P3 helices are distorted by IL2 and IL3 internal loops and a large and twisted L2 apical loop. These deformations unwind the helices and allow both domains to contact each other across the major groove along their entire length (Table 1).

Folding

The folding mechanism of these RNA motifs is relevant because some of these structures exert their function in a dynamic context, as they are transcribed. From the observations gathered in this review, the most important factor governing the shape of these motifs appears to be the presence of a junction strand that has a sufficient number of nucleotides to allow the parallel arrangement of the helices occupying the two upper branches of the “Y.” In fact, the most highly conserved structural motif in parallel-Y’s is a U-turn within this long junction strand (Table 1), allowing it to make a sharp turn when traveling from the first to the second upper domains. The position of this junction in the nucleotide sequence governs the distribution of the P1, P2, and P3 domains among the three branches of the Y (Table 1). However, once this condition is fulfilled, the remote tertiary contacts established between the two upper domains play a major role in fixing the global shape of the parallel-Y (Fig. 5), which is a likely requirement for function. This distal interaction usually involves apical or internal loops that have either conserved sequences or conserved three-dimensional structures (Fig. 4). As long as these contacts are sufficiently stable, parallel-Y’s can keep a globally similar three-dimensional shape while exhibiting an unusual junction topology, as observed in the TPP riboswitch (Fig. 2B). In this regard, tertiary interactions between helices have been recently shown to determine the accuracy of group I intron RNA folding, biasing the ensemble of secondary structures toward native-like conformations (Chauhan and Woodson 2008), and the conserved undocked conformations of the interacting loops have been proposed to play a key role in triggering the specific loop–loop contact leading to catalytically active hammerhead ribozymes (Fig. 4; Dufour et al. 2009).

The side-by-side packing of double-helices observed in parallel-Y RNAs is surprisingly common among RNA structures (Murthy and Rose 2000), showing that tertiary interactions and ion-binding effects can offset the electrostatic repulsion between the negatively charged phosphate backbones (Bassi et al. 1997; Murthy and Rose 2000; Penedo et al. 2004).

Dynamic effects

An important property of a number of parallel-Y folds is that both the junction and the remote tertiary interaction
have been shown to behave as flexible hinges that can respond to distal conformational changes in the other region. These changes may be caused by the binding of metabolites, auxiliary proteins, or even by the folding of the 3′ domain (P3) during transcription. The changes are transmitted through the regular base-paired stems and provide these motifs with the possibility of behaving as molecular switches, a property often connected with their biological function. The clearest examples in this respect are provided by riboswitches and hammerhead ribozymes together with the SRP S RNA. In the sensing domain of TPP riboswitches, the bridging by the TPP ligand of the P2 and P3 internal loops likely induces an allosteric stabilization of the three-way junction via tertiary interactions that sequesters the switching strand contained within the P1 helix (Serganov et al. 2006; Thore et al. 2006). In purine riboswitches, the L2–L3 interaction is necessary to preorganize the junction for purine binding (Batey et al. 2004; Lemay et al. 2006; Buck et al. 2007; Gilbert et al. 2007; Noeske et al. 2007; Ottink et al. 2007; Greenleaf et al. 2008; Stoddard et al. 2008), and ligand binding at the junction also tightens the remote L2–L3 complex (Buck et al. 2007). In hammerhead ribozymes, the L2–L3 interactions have been shown to reorganize the helical junction (Fig. 5A,B; Martick and Scott 2006), which becomes catalytically competent only upon the establishment of the remote loop–loop contacts (De la Peña et al. 2003; Khvorova et al. 2003; Martick and Scott 2006). These allosteric mechanisms are also observed in parallel-Y motifs contained in ribonucleoprotein systems. In the S domain of the mammalian SRP, the L2–L3 contact (stabilized by the SRP19 protein) brings the P2 and P3 helices together and enables a conformational change in a P3 internal loop located 10-bp away from the interaction that is necessary for SRP54 binding, since the protein–RNA interaction is stable in the context of the SRP54–P3–P2 ternary complex only (Batey et al. 2000; Kuglstatter et al. 2002; Oubridge et al. 2002). In addition, X-ray experiments have revealed that the J23 junction becomes ordered in the presence of SRP54 (Kuglstatter et al. 2002; Oubridge et al. 2002).

**Protein and RNA recognition**

In addition to their dynamic properties, the Y folds contained in ribonucleoprotein systems provide platforms for specific protein binding. This is because the juxtaposition of RNA domains at the junction and at the remote tertiary complexes enables the RNA helices to adopt unusual shapes that would not be stable in the absence of the RNA partner. For example, in the large ribosomal subunit, L11 recognizes the unusual shape of the P2 (H43) minor groove (Fig. 3B), whose conformation would not be possible in the absence of the H43–H44 tertiary interactions taking place in the parallel-Y fold (Conn et al. 1999; Wimberly et al. 1999). In the SRP S domain, the side-by-side arrangement of helices P2 and P3 induced by tertiary L2–L3 interactions (stabilized by SRP19) creates a binding platform for the essential SRP54 protein (Fig. 3A; Kuglstatter et al. 2002). In other cases, the proteins bind to the three-way junction: the SRP9/14 complex with the Alu junction (Weichenrieder et al. 2000) or ribosomal junction H18–H19–H20 bound to proteins L4P and L24P (Ban et al. 2000) are two representative examples of this type of complex (Fig. 3C,D). In most instances, protein binding typically has a stabilization role, because the RNA tertiary complexes are usually preformed in the absence of the protein partner (Zwieb 1992; Lu and Draper 1994; Conn et al. 1998; Huck et al. 2004; Maeder et al. 2006). Ribosomal three- and four-way RNA junctions and their protein complexes are believed to play an important role in rRNA folding and ribosome assembly (Ban et al. 2000; Wimberly et al. 2000). In this respect, the H18–H19–H20 parallel-Y, which is relatively close to the 5′ end of 23S rRNA, has been shown to play a critical role in the assembly of the large subunit (Skinner et al. 1985). Remarkably, the H18–H19–H20 and H90–H91–H92 parallel-Y’s are contacted by L24P and L3P, respectively, the only proteins capable of initiating 50S assembly in vitro (Nowotny and Nierhaus 1982). Likewise, the complex between the Alu SRP junction and the SRP9/14 protein heterodimer contributes to pack the P1 helix containing the SRP RNA 5′ and 3′ ends and is critical for SRP folding and assembly (Weichenrieder et al. 2000; Huck et al. 2004).

At least two of these motifs are also involved in the recognition of other RNA molecules. In the parallel-Y embedded within the S-domain of B-type RNase P (Fig. 2D), the large internal loop of domain P3 (P12) contributes to the recognition of the TΨC arm of the pre-tRNA molecule (LaGrandeur et al. 1994; Odell et al. 1998; Krasilnikov et al. 2003). In the H90–H91–H92 junction forming the A-site of the large ribosomal subunit, exposed bases within the tertiary complex formed by the apical loop of domain P3 (H92) and the internal loop of domain P1 (H90) participate in the recognition of the 3′ CCA tail of aminoacyl tRNA molecules (Schmeing et al. 2005).

**PARALLEL-Y JUNCTIONS AS TARGETS FOR DRUG ACTION**

Several parallel-Y RNAs have been shown to be the target of antibiotic action, a fact that adds further support to their functional relevance. The phosphorylated form of the antibiotic pyrithiamine exerts its activity by directly interacting with TPP riboswitches (Sudarsan et al. 2005). The thiopentapeptide antibiotics thiostrepton and micrococcin interact with the H42–H43–H44 junction of the large ribosomal subunit. They bind to a cavity formed by the tip of the L2 (H43)–L3 (H44) tertiary complex and the Nt domain of L11, interfering with the interaction of elongation factor EF-G with the ribosome (Harms et al. 2008).
Junction H90–H91–H92 forms the A-cavity of the peptidyl transferase center and is also the binding site of numerous antibiotics: puromycin (Nissen et al. 2000), chloramphenicol (Schlunzen et al. 2001), clindamycin (Schlunzen et al. 2001), anisomycin (Hansen et al. 2003), the pleuromutilin tiamulin (Schlunzen et al. 2004), and the recently developed oxazolidinone linezolid (Wilson et al. 2008) bind to the internal loop at the base of H90 and exert their action by interfering with the placement of the tRNA aminoacyl moiety in the A-site.

CONCLUDING REMARKS

Three-way junction RNAs adopt a recurrent parallel-Y shape when two of the helices form a coaxial stack and the third helix establishes one or more tertiary contacts several base pairs away from the junction. Structurally, these motifs are characterized by a conserved junction topology, where one of the strands of the lower helix crosses the coaxial stack in a 5′-to-3′ direction toward the third helix. In addition, the junction strand connecting the two upper helices is usually longer than the other two junction strands and often forms a U-turn structure. The remote tertiary contacts have a strong influence on the three-dimensional shape of the parallel-Y folds and usually involve apical or internal loops with conserved sequence or conserved three-dimensional structure. In fact, the helical junctions and the distal tertiary complexes have been shown to behave as flexible hinges that respond to conformational changes in the other region, a property that is often connected with the biological roles played by these motifs. In addition, the juxtaposition of RNA domains at the junction and at the remote tertiary complexes enables the RNA helices to adopt unusual shapes that are frequently used by proteins or other RNA molecules for specific binding. Not surprisingly, these motifs are targets for antibiotic action and play essential functions in a variety of biological processes including regulation of gene expression, translation, and ribonucleoprotein assembly.

ACKNOWLEDGMENTS

This work was supported by the Ministerio de Educación y Ciencia of Spain (grant SAF2007-60243 to J.G. and Ramón y Cajal contracts to M.d.l.P. and J.G.), Generalitat Valenciana (GV06/206 to M.d.l.P. and ACOMP2009/075 to J.G.), and the CIPF.

REFERENCES


www.rnajournal.org 1963


