Published in final edited form as:

Biochemistry. 2017 January 10; 56(1): 3–13. doi:10.1021/acs.biochem.6b01106.

Structural Roles of Noncoding RNAs in the Heart of Enzymatic Complexes

William J. Martin and Nicholas J. Reiter*

Department of Biochemistry, Vanderbilt University, Nashville, Tennessee 37232, United States

Abstract

Over billions of years of evolution, nature has embraced proteins as the major workhorse molecules of the cell. However, nearly every aspect of metabolism is dependent upon how structured RNAs interact with proteins, ligands, and other nucleic acids. Key processes, including telomere maintenance, RNA processing, and protein synthesis, require large RNAs that assemble into elaborate three-dimensional shapes. These RNAs can (i) act as flexible scaffolds for protein subunits, (ii) participate directly in substrate recognition, and (iii) serve as catalytic components. Here, we juxtapose the near atomic level interactions of three ribonucleoprotein complexes: ribonuclease P (involved in 5' pre-tRNA processing), the spliceosome (responsible for pre-mRNA splicing), and telomerase (an RNA-directed DNA polymerase that extends the ends of chromosomes). The focus of this perspective is profiling the structural and dynamic roles of RNAs at the core of these enzymes, highlighting how large RNAs contribute to molecular recognition and catalysis.

Graphical Abstract

Noncoding (nc) RNAs were instrumental in the evolution of the genetic code and likely served as key progenitor molecules for all life forms.^{1–3} Examining the compositions of a few of these ancient RNAs, from the analyses of sequence conservation to secondary and

*Corresponding Author: Vanderbilt University, 670 Robinson Research Building, Nashville, TN 37232. Phone: 615-343-6031. nick.reiter@vanderbilt.edu.

ORCID

Nicholas J. Reiter: 0000-0002-6095-6528

Author Contributions

W.J.M. and N.J.R. wrote the manuscript and prepared the figures.

Notes

The authors declare no competing financial interest.

tertiary structure determination, has yielded tremendous insight into our understanding of RNA biology. Detailed, biophysical studies of these RNAs in isolation or as part of a ribonucleoprotein complex have revealed a high degree of structural diversity, illuminating how large RNAs assemble into defined tertiary architectures and interact with their protein partners at the atomic level.⁴ These structural studies serve as paradigms for improving our understanding of the versatile roles of RNAs and the emerging functional roles of newly discovered ncRNAs in the modern world.

NcRNA possesses several evolutionary advantages in acting as a central regulatory molecule in the cell. Unlike proteins, ncRNAs can utilize both shape-based recognition and intermolecular base pair interactions with RNA and DNA to provide specificity.⁵ In an almost symbiotic relationship, ncRNAs can also act as scaffolds that bring proteins together and direct them to a nucleotide substrate while the proteins in turn stabilize RNA structures. This added stabilization by conserved protein–RNA interactions plays a central role throughout the RNP assembly process and fortifies key RNA conformations during the reaction cycle.

Here, we focus on large ncRNAs that contain structured domains and are intimately associated with protein components to execute a specific biological function. Ribonuclease (RNase) P, the spliceosome, and telomerase are three essential multidomain ribonucleoprotein (RNP) complexes that exhibit elaborate architectures and contain highly structured RNA domains. RNase P catalyzes the hydrolysis of a phosphodiester bond to generate a 5' end of mature tRNA (tRNA); the spliceosome removes noncoding pre-mRNA introns and ligates exonic sequences to generate a mature mRNA via two consecutive transesterification reactions, and telomerase serves as an RNA template-directed DNA polymerase that adds a species-dependent repetitive sequence at the ends of chromosomes.

While RNase P, spliceosome, and telomerase systems by no means represent a rigorous list of multidomain RNP complexes, and many comprehensive reviews exist for each RNP,^{6–10} general insights and commonalities have emerged from a comparison of recently determined structures derived from these full-length and intact holoenzyme complexes.^{11–16} In each case, large ncRNAs play an integral role in the overall reaction mechanism, from substrate recognition to the assembly of the active site.

At the Core: Conserved RNA Helices, Unpaired Nucleotides, and Base Triples

Functional features that are shared among three RNPs include well-defined RNA helical scaffolds and flexible internal-loop nucleotides. We are only starting to comprehend the role of RNA at the core of these enzymatic complexes, though it is likely that both structured RNA helices and local RNA motions represent conserved features of the reaction mechanism. For example, comparative sequence analyses within each of these RNAs show that assembling large RNA core regions requires highly conserved, short nucleotide stretches four to nine nucleotides in length.^{17–19} RNA sequences from RNase P, the spliceosome, and telomerase are clustered together to form a single helix or pseudoknot RNA element to function as the enzyme's central scaffold, located at or immediately adjacent to the active

site (Figure 1). These conserved structural elements at the RNP core help to explain how homologous RNAs can perform similar functions despite possessing a high overall degree of sequence variability.

Interestingly, the most highly conserved nucleotides within these core elements are often unpaired and highly dynamic. Single-nucleotide bulges, asymmetric internal loops, or catalytically important loop nucleotides are positioned within the central RNA helices, suggesting that local flexibility is an important feature in the assembly or orientation of an RNA-mediated enzyme. Whereas the conformations of base-paired nucleotides are restricted and limited by specific hydrogen bonds and/or stacking environments, bulged nucleotides offer a wide variety of spatial arrangements and can contribute significantly to the overall stability of the RNA fold. There are many advantages to having unpaired nucleotides near functionally important regions of an RNP complex, including (1) preserving continuous stacking interactions or enhancing hydrophobic packing between RNA regions, (2) subtly distorting or kinking the helical backbone to promote an RNA or protein contact, (3) stabilizing charge–charge repulsion forces of neighboring RNA backbones, (4) forming long-range intermolecular base pair interactions, and (5) potentially interacting directly with solvent or metal ions to activate catalysis.²⁰ Within the RNA core regions of RNase P, the spliceosome, and telomerase, unpaired and internal-loop nucleotides undergo dynamic motions, helping to preserve the functionality of the enzymatic core.

In addition, base triples are repeatedly employed as dynamic switches at key regions of RNP complexes to enhance intra- or intermolecular recognition. The enzymatic functions of RNase P, the spliceosome, and telomerase all require conserved base triple interactions, though it appears that each of these complexes utilizes base triples in distinct ways. RNase P and telomerase employ intramolecular base triples to form a preorganized RNA shape that is essential for substrate recognition and catalytic activity, respectively.^{7,10} However, in the spliceosome, stacked, intermolecular base triples form transiently to promote catalytic metal binding.²¹ Emerging structural studies of these RNPs begin to demonstrate how these stacked triples facilitate the organization of key molecular recognition events.

RNase P: Shape-Dependent Recognition by a Catalytic RNA

RNase P is an RNP complex composed of an essential RNA ribozyme subunit and one or more protein subunits required for its enzymatic function *in vivo*. Aside from the ribosome, RNase P constitutes the only known example of a multiple-turnover RNA enzyme that is required for cell viability in all domains of life.^{6,22} Newly transcribed precursor tRNA contains excess nucleotides at its 5′ and 3′ ends. The 5′ leader region is always removed by RNase P, whereas the 3′ tail extension can be trimmed or cleaved by various protein enzymes.⁶ Because of its vital role in cell metabolism, the bacterial RNase P RNP complex is a prime target for novel antibiotics.²³

Although our focus is on the RNase P RNP complex, a protein-only form of RNase P does exist in some eukaryotic organelles.^{24,25} It appears that the catalytic mechanisms of these two forms of RNase P enzymes are similar, where both utilize two metal ions to perform a concerted hydrolytic cleavage reaction.^{11,24} One metal (M1) serves to position and activate a

hydroxide nucleophile, while a second metal (M2) stabilizes the transition state and coordinates the oxyanion leaving group. Despite these mechanistic similarities, key features involving molecular recognition and assembly are unique, underscoring RNase P as a remarkable example of convergent evolution that involves distinct RNA-based and protein-only enzyme forms.

Numerous structural and biochemical studies have demonstrated that the large RNA component of RNase P can display versatility in terms of substrate recognition and that its functional complexity is comparable with that of multidomain protein enzymes.^{6,7} The RNase P RNP complex not only catalyzes the maturation of the 5' end of tRNA but also can exhibit broad specificity and acts on many different RNAs, including viral and phage RNA, mRNAs, noncoding RNAs, rRNA, and riboswitches.⁷ RNase P often targets substrates that mimic portions of the distinctive tRNA shape.⁶ For example, mammalian long ncRNAs that contain tRNA-like elements, such as MALAT1 and NEAT1, undergo processing and become activated via an RNase P-mediated mechanism.²⁶

Sequence and Tertiary Structural Conservation of the RNase P RNA

The catalytic RNA of RNase P (P RNA) varies greatly across phylogeny. It averages ~350–450 nucleotides in length but ranges from ~200 nucleotides (in archeal *Pyrobaculum*) to >1100 nucleotides (in the human pathogen *Candida glabrata*).^{27,28} Sequence analyses and structure mapping studies show that many architectural features of the P RNA are preserved in all domains of life, including a core secondary structure that is comprised of universally conserved nucleotides (Figure 2A).^{7,18} Crystallographic studies at near-atomic resolutions of different isoforms of RNase P show how key conserved regions are clustered together to organize the active site and participate in the recognition of the pre-tRNA substrate.^{11,29–32} Specifically, these universally conserved regions (CRs) are positioned near or within the catalytically important P4 helix (CR-I, -IV, and -V)⁷ or assemble as part of two interlocked T-loop tertiary structural motifs (CR-II and -III) that are critical for substrate recognition (Figure 2A).³³

RNA–RNA Shape-Dependent Recognition

The bacterial RNase P RNA recognizes at least two features of the pre-tRNA substrate: the minor groove helix of the acceptor stem and the TΨC and D stem–loop structures. The anticodon and acceptor stems of pre-tRNAs form separate coaxial helices. The resulting L-shaped “elbow” includes the TΨC- and D-loop regions, which adopt a distinctive T-loop motif that is commonly recognized by tRNA binding proteins and large RNAs.³⁴ Interestingly, this tRNA TΨC–D T-loop motif is recognized by two interlocked T-loop tertiary structural motifs (CR-II and -III) of RNase P, permitting a continuous hydrophobic stacking interaction between the P RNA and the TΨC–D interface of tRNA (Figure 2B).¹¹ Within these preorganized T-loop motifs, CR-II and -III nucleotides form two layers of stacked base triples, which serve to stabilize the structure and prime the adjacent P RNA unpaired nucleotides for hydrophobic stacking with the TΨC–D region of tRNA. Additional peripheral elements that help to facilitate this molecular recognition event include (1) a metal binding site (termed M4) that serves to stabilize the interlocked P RNA T loops and

(2) a conserved, unpaired adenosine that is wedged into the tRNA acceptor stem minor groove, forming an intermolecular A-minor motif. Taken together, these preorganized P RNA features allow for the accurate recognition of the substrate's shape, providing structural evidence for why biology has chosen to retain this architecture throughout evolution.

RNase P Protein Increases Catalytic Efficiency by Stabilizing RNA Conformations

While the RNA component is active on its own at high Mg^{2+} concentrations *in vitro*, RNase P is reliant on accessory protein components to stabilize the enzyme and aid in substrate recognition. The addition of a single protein domain in bacterial RNase P improves catalytic efficiency dramatically (>100-fold), and a synergistic coupling likely occurs between the bacterial RNase P RNA and protein to properly orient and position the substrate.^{35–39} Here, the P protein primarily binds the backbone of the catalytic RNA via basic amino acids and a bacterially conserved electropositive-rich α -helix.^{11,40} By stabilizing core scaffold regions of the catalytic P RNA, including CR-IV and -V and helix P3 (Figure 2C), the bacterial protein becomes integrated as part of the holoenzyme.

In higher organisms, the identified helix–bulge–helix P3 domain serves as an essential scaffold for assembly of multiple protein components.^{41,42} This P3 RNA scaffold likely permitted the evolutionary transition of RNase P to more protein-rich eukaryotic RNase P systems and is structurally analogous to the P3-like RNA scaffold that exists in yeast telomerase.^{41,43} This protein–RNA interface allows protein surfaces to aid in substrate selectivity by interacting directly with the 5' leader region of the pre-tRNA substrate. Protein–leader interactions also reflect a versatile and tractable network that contains both intrinsic specificity and nonspecific binding modes.^{44–46} Thus, in contrast to the P RNA–T Ψ C–D tRNA interaction that is largely preorganized and shape-dependent, the P protein–pre-tRNA leader recognition mode involves a conformational equilibrium, where the flexibility of the pre-tRNA leader and minor, local changes of the protein allow favorable overall binding interactions that give rise to an activated enzyme–substrate complex.

Beyond the simple bacterial RNase P RNP, there are five identified protein components in archaeal RNase P and 10 or more proteins in the eukaryotic holoenzyme.⁶ This increased protein content is indicative of a slightly reduced role for the functional RNAs and likely allows the enzyme to broaden its substrate binding potential in the complex cellular environment of eukaryotes. Interestingly, many of the protein components of eukaryotic RNase P also participate in the closely related mitochondrial RNA processing complex, or MRP. In eukaryotes, gene duplication of RNase P allowed the formation of the *RMRP* gene, which encodes the MRP RNA.⁴¹ The MRP RNP complex has localization and substrate specificity distinct from RNase P but retains the P3 helix, core P4 helix, and many of the same RNase P protein components.^{6,41,43,47} This supports the model in which functionally distinct RNP complexes can share and swap large protein modules to aid in the assembly and activation of the enzyme. Future studies will be required to discern whether these more complex enzymes employ a mode of substrate binding similar to that of bacterial RNase P.

RNase P Ribozyme Core: A largely Preorganized Active Site with a Dynamic Uridine Bulge

A comparison of the apo and tRNA-bound structures indicates that the RNase P holoenzyme does not undergo large conformational changes, suggesting an overall RNP complex that is largely preassembled to engage with its target substrate.⁷ Nonetheless, FRET studies and a bacterial RNase P holoenzyme–tRNA structure show that two catalytically important metal ions possess different binding affinities.^{11,48} Metal ions are sandwiched among the 5′ end of the tRNA, the phosphate oxygen-rich cleft of the P4 helix, and conserved region V (CR-V) (Figure 2D). It is likely that the identified RNase P active site is the same in all RNA-based RNase P complexes, where one metal site (M1) is nestled between the tRNA 5′ end, a universally conserved uridine (U52 in *Thermatoga maritima* numbering), and the phosphate oxygens of the P4 helix, while the other metal ion site (M2) is fully formed in the presence of the pre-tRNA leader substrate.¹¹ In the apo *Bacillus stearothermophilus* RNase P RNA structure without pre-tRNA, a metal site (termed M6) is in the same location as the M1 site in the tRNA-bound *T. maritima* holoenzyme structure.^{11,49} A cluster of phosphate oxygens from conserved nucleotides A50 (A50), A313 (A389), and A389 (A390) in the *T. maritima* RNase P RNA (*B. stearothermophilus* numbering in parentheses) help to create this M1 metal site, strongly suggesting that the P RNA has a partially preorganized metal binding location within the active site. However, the exact positions of the M1 site in the apo- and holoenzyme–tRNA structure likely differ, as the tRNA ligands are absent and U52 is positioned out of the helical stack in the apo structure (Figure 2D). In contrast, the M2 site is not preorganized and likely represents a relatively diffuse metal binding location that becomes stable only upon binding to the pre-tRNA substrate.

An additional component within the active site region that undergoes dynamic motions is the universally conserved U52 nucleotide. Here, the unpaired uridine within the P4 helix is flipped out relative to its neighboring nucleotides. In the apo crystal structure, this uridine is pointed out of the helical stack and resides in the proximity of the catalytically important M1 metal binding site.^{29,49} In the presence of tRNA, this U52 nucleobase appears to swivel toward the substrate, interacting directly with the catalytically important metal (M1) and helping to properly position the local tRNA structure at the active site (Figure 2D).¹¹ The central role of this uridine at the active site has been previously proposed, and single-point deletion or mutational studies of this bulged nucleotide show dramatically reduced catalytic efficiency of the enzyme.^{35,37,50,51} These studies provide a near-atomic resolution glimpse of the structural features of the ribozyme active site, yet higher-resolution information about this region will be required to understand how core RNA–metal ion interactions participate in the RNase P catalytic cycle.

Spliceosome: Small Nuclear RNAs Form an Ancient Catalytic Core

Studies of RNase P and the spliceosome demonstrate that RNA–metal-based catalysis is a fundamental process in the conversion of genetic information. However, unlike that of RNase P, global preorganization from a singular RNA structural scaffold does not occur in the spliceosome. Instead, large-scale dynamic rearrangements driven by ATP-dependent

U6 RNA Elements Prime Active Site Formation and Contain a Dynamic Uridine Bulge

The U6 snRNA is the most conserved RNA component in the spliceosome, and its tertiary structure plays a central role in the active site cavity.¹⁹ The U6 ISL⁶¹ forms the essential structural scaffold, and key nucleotides within U6 form metal ligands that function in both steps of splicing. Phosphorothioate metal rescue studies of key regions within U6 support a model in which both phosphoryl transfer steps of the splicing reactions are catalyzed by the identical, two-metal ion catalytic core, as opposed to two independent active sites.^{62,63} The U6 ISL within intact spliceosomes contains an asymmetrical internal loop, where conserved nucleotides (A79 and U80) are flipped out of the helical stack and orient the metalloenzyme active site. The highly conserved and bulged U80 nucleotide serves as a metal ligand in both steps of splicing, and its nucleobase forms a critical base triple with C61 of U6 and G21 of U2.^{12,21,62,63} In addition, the U80 nucleobase is also in the proximity of two highly conserved regions of Cwc15 (*Cwf15*) and Cef1 (*Cdc5*) proteins, and the entire ISL is surrounded by Prp8 (*Spp42*), revealing key protein residues involved in stabilizing the U6 ISL structure. Opposite to U80 within the U6 ISL helix, a single nucleotide (C66) is bulged out of the helix and protrudes into the interface of Cwc15 (*Cwf15*), Cef1 (*Cdc5*), and Clf1 (*Cwf4*).¹²

Many of the U6–protein interactions were identified by genetics and biochemical studies before being confirmed by structural studies, and it is clear that the internal bulge region of the U6 ISL serves multiple, highly interactive functions within the active site core.^{52,64,65} Specifically, the internal bulge provides catalytically important ligands for two metals (M1 and M2) that are directly involved in both splicing steps (Figure 3C,D) and also interacts with a conserved U2–U6 helical region (termed helix IB) to form a structurally conserved catalytic triplex (Figure 3E,F).^{21,59,63,66} The dual functions of this internal bulge of U6 suggest that this region of the spliceosome contains both ancient structures and conserved dynamic motions that are essential for properly orienting the substrate and activating catalysis of the pre-mRNA substrate.^{56,67}

Conserved RNA Structural Similarities to Group II Intron Ribozymes

In fact, the U6 ISL and its corresponding internal-loop region are structurally related to the core domain V (D5) structure within group II self-splicing introns.^{56,68–71} Both of these evolutionarily related helices are approximately of the same dimensions and contain two-nucleotide bulged internal loops at similar locations. Interestingly, both of these bulges interact directly with metal ions in a stereospecific manner and reside immediately adjacent to a catalytic triplex (Figure 3C,D). In fact, it has even been demonstrated that domain V can functionally replace the U6 ISL.^{72,73}

Likewise, the three stacked base triples comprising the catalytic triplex are also structurally conserved between the spliceosome and group II introns (Figure 3E,F).^{21,74} The multiple tertiary structures of group II introns at atomic resolution demonstrate how these base triples might prime the formation of the active site cavity,^{57,74,75} and it is likely that this is a universal mechanism shared between the spliceosome and group II introns. However, in

terms of the organization of the active site, it is important to emphasize that protein components that stabilize the U6 ISL and U2–U6 catalytic triplex in the spliceosome have usurped the role belonging to RNA helices in group II introns. For example, in the spliceosome, the catalytic triplex forms upon NTC–NTR binding, and it appears that Prp8 (*Spp42*) and both Cwc15 (*Cwf15*) and Cef1 (*Cdc5*) of the NTC–NTR complex are poised to stabilize the triplex during both steps of the splicing reaction.¹² In addition, structural studies of a minimal U2–U6 RNA-only spliceosome have demonstrated that the U2–U6 core, by itself, cannot spontaneously form a group II intron-like RNA tertiary structure in isolation.⁷⁶ Nonetheless, because the two transesterification reactions catalyzed by the spliceosome are identical to those of group II self-splicing intron ribozymes, it is likely that these shared RNA elements have been preserved to facilitate conformational changes and serve as catalytically important metal ion ligands at the active site.

The central U6 ISL scaffold and catalytic triplex highlight the importance of core RNA tertiary structures at the heart of the spliceosome. Although recent cryo-EM structures have revealed the very first structural insights into the fully assembled spliceosome architectures, much remains to be understood. Additional details involving higher-resolution structures and exploration of the conformational motions of the RNP components around the active site will be needed to fully understand how the spliceosome facilitates activation of the pre-mRNA splicing reaction, especially the mechanisms of substrate rearrangement and how ATP-dependent helicases influence the splicing reaction cycle.

Telomerase: A Multifunctional RNA Plays Essential Roles in Telomere Biology

While RNA catalyzes the enzymatic reaction in RNase P and the spliceosome, more often it regulates a catalytic protein. This is exemplified by the telomerase RNA (TR), which serves as an essential ncRNA component of telomerase and interacts extensively with the catalytic protein telomerase reverse transcriptase (TERT). During DNA replication, the lagging DNA strand is unable to fully copy the chromosome ends.^{10,77} This end replication problem results in a progressive shortening of telomeres that, if not countered, leads to cellular senescence. TR contains a single-stranded template region that encodes a telomeric repeat sequence for TERT to generate new telomeric repeats. Several conserved TR structural domains help to position the template region and regulate the enzymatic activity.

The telomerase complex is primarily active in embryonic stem cells but is also inappropriately expressed in a large majority of cancers, making it a prime therapeutic target.⁷⁸ Additionally, mutations in the protein or RNA components of telomerase are associated with the bone marrow failure syndrome dyskeratosis congenita, a marked decrease in telomerase activity, and shortened telomeres.⁷⁹ Thus, an improved structural understanding of the mechanics of telomerase holds significant therapeutic implications.

Sequence Conservation and Diverse Functional Elements of TR

TR varies greatly in size. In humans, TR (451 nucleotides) is only slightly longer than the RNase P ncRNA (341 nucleotides), but in some organisms, it reaches over 2 kb.⁸⁰ While

global structural information about the *Tetrahymena* telomerase holoenzyme is now available (Figure 1C), only individual domains of human telomerase have been resolved, and the actual RNP sequence and structural complexity across distinct evolutionary lineages are quite diverse.^{13,81} However, despite the wide variety in TR length and secondary structure, several components are well conserved (Figure 4). These conserved regions form the core components of the TR, including the template, template boundary element, pseudoknot, and conserved region 4 and 5 (CR 4/5).⁸¹

Template

Perhaps the most striking feature of the otherwise well-ordered TR is the unstructured template region. This stretch is located within conserved region 1 (CR-1) and must be able to both bind at the active site of TERT and base pair with the DNA substrate. Structural studies of the *Tribolium castaneum* TERT indicate that it binds the RNA template region nonspecifically via the RNA ribose groups and phosphate backbone.⁸² RNA oligonucleotides have been demonstrated to be capable of functioning *in trans* to complement truncated TR lacking the template sequence and can restore activity *in vitro*.^{83,84} The crystal structures of apo and substrate-bound TERT do not reveal large conformational changes in protein structure, indicating that, like bacterial RNase P, the telomerase active site is largely preformed.⁸²

Template Boundary Element

For telomerase to synthesize regular telomeric repeats, it must consistently terminate reverse transcription at the correct nucleotide. Incorporation of the wrong telomeric sequence can be lethal, as evidenced by the sharply increased mortality of cells expressing a TR containing mutant template sequences.⁸⁵ The precise mechanism varies between species; in the human form, a template boundary element (TBE) containing the P1 helix forms immediately 5' of the template (Figure 4A) and constrains transcription.⁸⁶ Disruption of the helix not only leads to inappropriate reverse transcription beyond the template region but also dramatically perturbs overall activity.

Interestingly, this 5' region contains conserved guanosine repeats that have a propensity to form a stable G-quadruplex structure (Figure 4A). The function of this G-quadruplex remains unclear, though it has been proposed to serve as a protective cap to limit RNA degradation.⁸⁷ As this competing RNA structure prevents formation of the P1 helix and influences catalytic activity, it must be removed. An RNA helicase, RHAU, likely remodels this region by recognizing and unwinding G-quadruplexes to allow proper formation of the template boundary helix. In removing the G-quadruplex RNA, RHAU may also remodel additional portions of RNA structure,⁸⁸ suggesting that this helicase plays an important mechanistic role in activating the telomerase holoenzyme.

Tertiary Structural Elements and Unpaired Bases within the TR Pseudoknot

Downstream of the template, the RNA forms a structurally conserved helix–bulge–helix motif. NMR analysis identifies highly dynamic motions within the motif, termed J2a/b, with

a five-nucleotide bulge forcing the RNA to form an $\sim 90^\circ$ angle.⁸⁹ In the middle of this bulge is an unpaired uridine, conserved in mammals, that is required for full telomerase function *in vitro*. This dynamic J2a/b motif may allow for conformational changes during the telomere elongation step.⁸⁹

Downstream of this hinge are highly conserved regions (CR2/3) that form a critical pseudoknot (Figure 4C). The pseudoknot forms a triple helix that is essential for telomerase function *in vitro* and *in vivo*.^{90,91} Pseudoknot mutations can disrupt enzyme processivity after the first round of DNA synthesis, indicating that dynamics in the triple helix may be involved in the translocation step of the reverse transcriptase catalytic cycle. However, the precise mechanistic and functional role of the pseudoknot remains unclear.⁹¹ Interestingly, the presence of an unpaired nucleotide (U177 in humans) within the pseudoknot is well-conserved in vertebrates.¹⁷ This bulged nucleotide may destabilize the triple helix to produce an equilibrium between the pseudoknot and an alternative simple hairpin structure *in vitro*.^{90,92} Recent electron microscopy models of the *Tetrahymena thermophila* telomerase place the pseudoknot apart from the active site, calling into question whether the pseudoknot directly interacts with the active site.¹⁰ Atomic-resolution information at different stages of the catalytic cycle will be needed to define the structural mechanisms of the pseudoknot within the intact holoenzyme.

Conserved Regions 4 and 5

Additional conserved regions 4 and 5 (CR4 and -5, respectively) adopt a defined RNA tertiary structure that involves two stem-loop structures, termed the P6 and P6.1 helices, which originate from a three-way junction (Figure 4D). This RNA region forms a flexible Y-shaped scaffold that interacts with the TERT RNA binding domain (TRBD).^{93–95} An NMR structure of this RNA region by itself confirms the three-way junction, but the RNA adopts an alternative conformation from the crystal structure of the complex. In the complex, the P6.1 arm forms a much larger angle with the P6 arm, adopting an almost orthogonal orientation.^{94,95} The TRBD-bound RNA also forms two base triples at the three-way junction that are absent in the apo structure. The triplexes form a base for the helical stacking of the P6 arm and are likely involved in the stabilization of the bound RNP conformation. Thus, while the protein structure is largely unchanged by RNA binding, the TR CR4/5 domain appears to undergo substantial conformational changes to bind TRBD.

Mutations within CR4/5 and small molecules that block its interaction with the TRBD result in a loss of telomerase activity.^{93–95} Most of the TRBD contacts are with the backbone of CR4/5 and are not sequence specific. However, important exceptions include two dynamic and unpaired nucleotides that make stacking interactions with specific TERT amino acids: a guanosine at the terminal loop of the larger P6 helix and a uridine in the middle of the P6 helix. These bases and/or their unpaired positions in the helix are well conserved in vertebrates and necessary for binding, although the uridine in the human form appears to be base-paired while an adjacent cytosine is bulged out and likely functionally substitutes for the uridine (Figure 4A).^{94,96}

Accessory Proteins Reveal Significant Diversity and a Unique, Shared RNA–Protein Scaffold

The eukaryotic TR interacts with an array of proteins beyond TERT. While these proteins have functional analogues from yeast to humans, much of the composition appears to vary significantly between eukaryotes.⁹⁷ For instance, the RNase P protein components POP6 and POP7 also regulate the telomerase complex and are necessary for telomerase activity in *S. cerevisiae*.⁴³ Here, the protein module recognizes a helix–bulge–helix region of TR that mimics the P3 helix of RNase P (Figure 2C). While this precise mechanism does not appear to be replicated in humans, it demonstrates that the RNase P protein components can participate in a number of RNP complexes.

Other accessory proteins also target specific regions of the TR. For example, four snoRNA-associated proteins (dyskerin, NHP2, NOP10, and GAR1) recognize a conserved hairpin–hinge–hairpin–ACA, or H/ACA, motif at the 3′ end of TR (CR6/8).⁹⁸ This asymmetrical stem–loop structure was likely coopted by telomerase to act as a scaffold for chaperones that stabilize and localize the TR during RNP assembly.

Structured ncRNAs Significantly Alter the Action of Telomerase

TR is not the only RNA that regulates telomerase. The repetitive telomeric DNA sequence encoded by TR can itself be transcribed into a structured RNA, termed telomerase repeat-containing RNA (TERRA).⁹⁹ TERRA is transcribed following the replication of DNA during the S phase and can form stable G-quadruplex structures *in vivo*. These RNA structures can interact with and influence epigenetic regulators, including the Suv39H1 methyltransferase and LSD1 demethylase,^{100–102} and also regulate telomere length by inhibiting telomerase.¹⁰³ The detailed mechanisms of these recognition events are unknown, though direct contacts between TERRA and the unstructured template region of TR dramatically influence telomerase activity.¹⁰⁴

In addition, a TR-mimicking lncRNA in *A. thaliana*, termed TER2, is highly homologous to TR.¹⁰⁵ TER2 appears to be upregulated upon excessive DNA damage to outcompete TR and inhibit telomerase activity, presumably by forming an inactive TER2–TERT RNP complex. These ncRNAs influence the action of telomerase and serve primarily as RNA mimics to regulate telomere biology. How ncRNA structures are regulated at telomeres and how alternative ncRNAs impact telomerase function must be addressed to improve our understanding of the role of lncRNAs at specific chromosomal locations.

SUMMARY

Atomic level understanding of large and dynamic RNP complexes is beginning to emerge, revealing how conserved structural features give rise to RNA-mediated recognition and catalysis. RNA helices, unpaired and internal-loop nucleotides, and triplexes represent fundamental motifs that are repeatedly used by evolution in various combinations to introduce diversity into RNA structures and perform specific tasks. In all cases, functional

protein modules contribute to stabilization of these RNA structures and enhance the molecular recognition properties of RNP complexes.

In RNase P and the spliceosome, highly conserved RNA helices form around the active sites and bulged internal-loop nucleotides are involved in coordinating catalytically important metal ions. At the core of telomerase, the stability of a pseudoknot structure and unpaired nucleotides play essential roles in contacting conserved protein domains and in facilitating enzyme activation. All three of these RNP complexes additionally require the formation of stacked base triple motifs to ensure substrate recognition and efficient catalysis. Base triples within conserved T-loop motifs of the RNase P are flanked by unpaired nucleotides that enhance intermolecular coaxial stacking, permitting shape specific molecular recognition of pre-tRNA substrates (Figure 2B).¹¹ A highly conserved catalytic triplex in the spliceosome involves three stacked triples that function to define the active site cavity, providing a structural mechanism for optimally positioning the pre-mRNA substrate with catalytic metals in both steps of the splicing reaction (Figure 3E,F).^{12,21} In the case of telomerase, two sets of conserved and stacked base triples operate in distinct functional domains: a three-way junction and an extended triple-helical pseudoknot structure (Figure 4C,D).^{90,91,94} Both telomerase components are required for RNP assembly and play essential roles in reverse transcriptase binding and catalytic activation. Additional molecular details demonstrating how these specific RNAs and RNA–protein interactions accurately recognize and reorganize their substrates during catalysis will be required to improve our understanding of the diverse modes of RNA-mediated recognition and RNP catalysis.

Future studies of these RNP structures will also be critical to gain insight into the emerging roles of ncRNAs in biology. For example, long ncRNA (lncRNA) transcripts are pervasively transcribed and form modular functional domains that can regulate processes such as gene transcription, chromatin structure, post-transcriptional processing, and protein translation.¹⁰⁶ Atomic level knowledge of the more traditional RNase P, spliceosome, and telomerase RNP assemblies will provide powerful insight into how other large ncRNA molecules have evolved and are integrated with proteins in the modern world. The advantages in using ncRNAs as centralized regulatory molecules include the ability to respond to a stimulus quickly upon transcription, the ability to evolve rapidly without generating missense mutations, and the opportunity to act as a scaffold that brings proteins together and localizes them to specific RNA or DNA targets. The roles of these ncRNAs are largely unexplored but are likely multifaceted and versatile, suggesting that these new RNAs are part of fluid, responsive RNP networks that are highly adaptable. On the basis of structural studies of RNase P, the spliceosome, and telomerase, it is clear that many of these newly discovered ncRNAs will not simply be passive platforms but may utilize specific tertiary structural motif combinations and conserved RNA–protein interactions to functionalize a dynamic RNP assembly.

Acknowledgments

Funding

This work is supported by Grant T32-GM008320 to W.J.M., start-up funds to N.J.R. from Vanderbilt University, and American Heart Association Grant GRNT20380334 to N.J.R.

We thank Juli Feigon for the pseudoatomic model of the catalytic core of *Tetrahymena* telomerase, Martin Egli, Melanie Ohi, and Amanda Erwin for critical reading of the manuscript, and Benjamin Brown for insightful discussions.

ABBREVIATIONS

A-minor	adenosine minor groove
CR	conserved region
EM	electron microscopy
FRET	Förster resonance energy transfer
G-quadruplex	guanosine quadruplex
ISL	internal stem-loop
LSD1	lysine specific histone demethylase 1
MALAT1	metastasis-associated lung adenocarcinoma transcript 1
NEAT1	nuclear paraspeckle assembly transcript 1
NTC	Nineteen Complex
NTR	Nineteen-Related Complex
Pop6/7	processing of precursor RNA protein 6/7
pre-	precursor
Prp8	pre-messenger processing factor 8
PDB	Protein Data Bank
RHAU	RNA helicase associated with AU-rich elements
RNA	ribonucleic acid
RNase P	ribonuclease P
RNP	ribonucleoprotein
lncRNA	long noncoding RNA
mRNA	messenger RNA
ncRNA	noncoding RNA
rRNA	ribosomal RNA
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
tRNA	transfer RNA

Sc	<i>S. cerevisiae</i>
Sp	<i>Sc. pombe</i>
SNP	single-nucleotide polymorphism
TR	telomerase RNA
TERT	telomerase reverse transcriptase protein
TER2	telomerase RNA component 2
TERRA	telomeric repeat-containing RNA
TBE	template boundary element

References

1. Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A*. 1977; 74:5088–5090. [PubMed: 270744]
2. Crick FH. The origin of the genetic code. *J Mol Biol*. 1968; 38:367–379. [PubMed: 4887876]
3. Orgel LE. Evolution of the genetic apparatus. *J Mol Biol*. 1968; 38:381–393. [PubMed: 5718557]
4. Butcher SE, Pyle AM. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc Chem Res*. 2011; 44:1302–1311. [PubMed: 21899297]
5. Reiter NJ, Chan CW, Mondragón A. Emerging structural themes in large RNA molecules. *Curr Opin Struct Biol*. 2011; 21:319–326. [PubMed: 21474301]
6. Kazantsev AV, Pace NR. Bacterial RNase P: a new view of an ancient enzyme. *Nat Rev Microbiol*. 2006; 4:729–740. [PubMed: 16980936]
7. Mondragón A. Structural Studies of RNase P. *Annu Rev Biophys*. 2013; 42:537–557. [PubMed: 23654306]
8. Chen W, Moore MJ. The spliceosome: disorder and dynamics defined. *Curr Opin Struct Biol*. 2014; 24:141–149. [PubMed: 24530854]
9. Nguyen THD, Galej WP, Fica SM, Lin PC, Newman AJ, Nagai K. CryoEM structures of two spliceosomal complexes: starter and dessert at the spliceosome feast. *Curr Opin Struct Biol*. 2016; 36:48–57. [PubMed: 26803803]
10. Feigon J, Chan H, Jiang J. Integrative structural biology of *Tetrahymena* telomerase - insights into catalytic mechanism and interaction at telomeres. *FEBS J*. 2016; 283:2044–2050. [PubMed: 26918633]
11. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragón A. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*. 2010; 468:784–789. [PubMed: 21076397]
12. Galej WP, Wilkinson ME, Fica SM, Oubridge C, Newman AJ, Nagai K. Cryo-EM structure of the spliceosome immediately after branching. *Nature*. 2016; 537:197–201. [PubMed: 27459055]
13. Jiang J, Chan H, Cash DD, Miracco EJ, Ogorzalek Loo RR, Upton HE, Cascio D, O'Brien Johnson R, Collins K, Loo JA, Zhou ZH, Feigon J. Structure of *Tetrahymena* telomerase reveals previously unknown subunits, functions, and interactions. *Science*. 2015; 350:aab4070. [PubMed: 26472759]
14. Wan R, Yan C, Bai R, Huang G, Shi Y. Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science*. 2016; 353:895–904. [PubMed: 27445308]
15. Yan C, Wan R, Bai R, Huang G, Shi Y. Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science*. 2016; 353:904–911. [PubMed: 27445306]
16. Rauhut R, Fabrizio P, Dybkov O, Hartmuth K, Pena V, Chari A, Kumar V, Lee CT, Urlaub H, Kastner B, Stark H, Lührmann R. Molecular architecture of the *Saccharomyces cerevisiae* activated spliceosome. *Science*. 2016; 353:1399–1405. [PubMed: 27562955]

17. Chen JL, Blasco MA, Greider CW. Secondary structure of vertebrate telomerase RNA. *Cell*. 2000; 100:503–514. [PubMed: 10721988]
18. Chen JL, Pace NR. Identification of the universally conserved core of ribonuclease P RNA. *RNA*. 1997; 3:557–560. [PubMed: 9174091]
19. Brow DA, Guthrie C. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature*. 1988; 334:213–218. [PubMed: 3041282]
20. Hermann T, Patel DJ. RNA bulges as architectural and recognition motifs. *Structure*. 2000; 8:R47–54. [PubMed: 10745015]
21. Fica SM, Mefford MA, Piccirilli JA, Staley JP. Evidence for a group II intron-like catalytic triplex in the spliceosome. *Nat Struct Mol Biol*. 2014; 21:464–471. [PubMed: 24747940]
22. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*. 1983; 35:849–857. [PubMed: 6197186]
23. Drainas D. Antibiotics and RNase P. *Antibiotics (Basel, Switz)*. 2016; 5:15.
24. Howard MJ, Lim WH, Fierke CA, Koutmos M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *Proc Natl Acad Sci U S A*. 2012; 109:16149–16154. [PubMed: 22991464]
25. Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmann W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*. 2008; 135:462–474. [PubMed: 18984158]
26. Wilusz JE, Sunwoo H, Spector DL. Long non-coding RNAs: functional surprises from the RNA world. *Genes Dev*. 2009; 23(13):1494–1504. [PubMed: 19571179]
27. Kachouri R, Stribinskis V, Zhu Y, Ramos KS, Westhof E, Li Y. A surprisingly large RNase P RNA in *Candida glabrata*. *RNA*. 2005; 11:1064–1072. [PubMed: 15987816]
28. Lai LB, Chan PP, Cozen AE, Bernick DL, Brown JW, Gopalan V, Lowe TM. Discovery of a minimal form of RNase P in *Pyrobaculum*. *Proc Natl Acad Sci U S A*. 2010; 107:22493–22498. [PubMed: 21135215]
29. 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]
30. Torres-Larios A, Swinger KK, Krasilnikov AS, Pan T, Mondragón 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, Mondragón A. Basis for structural diversity in homologous RNAs. *Science (Washington, DC, U S)*. 2004; 306:104–107.
32. Krasilnikov AS, Yang X, Pan T, Mondragón A. Crystal structure of the specificity domain of ribonuclease P. *Nature*. 2003; 421:760–764. [PubMed: 12610630]
33. Chan CW, Chetani B, Mondragón A. Structure and function of the T-loop structural motif in noncoding RNAs. *Wiley Interdiscip Rev: RNA*. 2013; 4:507–522. [PubMed: 23754657]
34. Zhang J, Ferré-D'Amaré AR. The tRNA Elbow in Structure, Recognition and Evolution. *Life (Basel, Switz)*. 2016; 6(1):3.
35. Christian EL, Smith KMJ, Perera N, Harris ME. The P4 metal binding site in RNase P RNA affects active site metal affinity through substrate positioning. *RNA*. 2006; 12:1463–1467. [PubMed: 16822954]
36. Koutmou KS, Zahler NH, Kurz JC, Campbell FE, Harris ME, Fierke CA. Protein-precursor tRNA contact leads to sequence-specific recognition of 5' leaders by bacterial ribonuclease P. *J Mol Biol*. 2010; 396:195–208. [PubMed: 19932118]
37. Reiter NJ, Osterman AK, Mondragón A. The bacterial ribonuclease P holoenzyme requires specific, conserved residues for efficient catalysis and substrate positioning. *Nucleic Acids Res*. 2012; 40:10384–10393. [PubMed: 22904083]
38. Buck AH, Dalby AB, Poole AW, Kazantsev AV, Pace NR. Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J*. 2005; 24:3360–3368. [PubMed: 16163391]
39. Crary SM, Niranjanakumari S, Fierke CA. The protein component of *Bacillus subtilis* ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNA^{Asp}. *Biochemistry*. 1998; 37:9409–9416. [PubMed: 9649323]

40. Stams T, Niranjanakumari S, Fierke CA, Christianson DW. Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science*. 1998; 280:752–755. [PubMed: 9563955]
41. Esakova O, Krasilnikov AS. Of proteins and RNA: the RNase P/MRP family. *RNA*. 2010; 16:1725–1747. [PubMed: 20627997]
42. Perederina A, Esakova O, Quan C, Khanova E, Krasilnikov AS. Eukaryotic ribonucleases P/MRP: the crystal structure of the P3 domain. *EMBO J*. 2010; 29:761–769. [PubMed: 20075859]
43. Lemieux B, Laterreur N, Perederina A, Noel JF, Dubois ML, Krasilnikov AS, Wellinger RJ. Active Yeast Telomerase Shares Subunits with Ribonucleoproteins RNase P and RNase MRP. *Cell*. 2016; 165:1171–1181. [PubMed: 27156450]
44. Niland CN, Zhao J, Lin HC, Anderson DR, Jankowsky E, Harris ME. Determination of the Specificity Landscape for Ribonuclease P Processing of Precursor tRNA 5' Leader Sequences. *ACS Chem Biol*. 2016; 11:2285–2292. [PubMed: 27336323]
45. Lin HC, Zhao J, Niland CN, Tran B, Jankowsky E, Harris ME. Analysis of the RNA Binding Specificity Landscape of C5 Protein Reveals Structure and Sequence Preferences that Direct RNase P Specificity. *Cell Chemical Biology*. 2016; 23:1271–1281. [PubMed: 27693057]
46. Guenther UP, Yandek LE, Niland CN, Campbell FE, Anderson D, Anderson VE, Harris ME, Jankowsky E. Hidden specificity in an apparently nonspecific RNA-binding protein. *Nature*. 2013; 502:385–388. [PubMed: 24056935]
47. Chamberlain JR, Lee Y, Lane WS, Engelke DR. Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes Dev*. 1998; 12:1678–1690. [PubMed: 9620854]
48. Hsieh J, Koutmou KS, Rueda D, Koutmos M, Walter NG, Fierke CA. A divalent cation stabilizes the active conformation of the *B. subtilis* RNase P x pre-tRNA complex: a role for an inner-sphere metal ion in RNase P. *J Mol Biol*. 2010; 400:38–51. [PubMed: 20434461]
49. 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]
50. Chen WY, Xu Y, Cho IM, Oruganti SV, Foster MP, Gopalan V. Cooperative RNP assembly: complementary rescue of structural defects by protein and RNA subunits of archaeal RNase P. *J Mol Biol*. 2011; 411:368–383. [PubMed: 21683084]
51. Kaye NM, Zahler NH, Christian EL, Harris ME. Conservation of helical structure contributes to functional metal ion interactions in the catalytic domain of ribonuclease P RNA. *J Mol Biol*. 2002; 324:429–442. [PubMed: 12445779]
52. Brow DA. Allosteric cascade of spliceosome activation. *Annu Rev Genet*. 2002; 36:333–360. [PubMed: 12429696]
53. Alcid EA, Jurica MS. A protein-based EM label for RNA identifies the location of exons in spliceosomes. *Nat Struct Mol Biol*. 2008; 15:213–215. [PubMed: 18223660]
54. Ohi MD, Ren L, Wall JS, Gould KL, Walz T. Structural characterization of the fission yeast U5. U2/U6 spliceosome complex. *Proc Natl Acad Sci U S A*. 2007; 104:3195–3200. [PubMed: 17360628]
55. Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harbor Perspect Biol*. 2011; 3(7):a003707.
56. Keating KS, Toor N, Perlman PS, Pyle AM. A structural analysis of the group II intron active site and implications for the spliceosome. *RNA*. 2010; 16:1–9. [PubMed: 19948765]
57. Marcia M, Pyle AM. Visualizing group II intron catalysis through the stages of splicing. *Cell*. 2012; 151:497–507. [PubMed: 23101623]
58. Manning KS, Cooper TA. The roles of RNA processing in translating genotype to phenotype. *Nat Rev Mol Cell Biol*. 2016; doi: 10.1038/nrm.2016.139
59. Madhani HD, Guthrie C. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell*. 1992; 71:803–817. [PubMed: 1423631]
60. O'Keefe RT, Norman C, Newman AJ. The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell*. 1996; 86:679–689. [PubMed: 8752221]

61. Huppler A, Nikstad LJ, Allmann AM, Brow DA, Butcher SE. Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat Struct Biol.* 2002; 9:431–435. [PubMed: 11992125]
62. Yean SL, Wuenschell G, Termini J, Lin RJ. Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature.* 2000; 408:881–884. [PubMed: 11130730]
63. Fica SM, Tuttle N, Novak T, Li NS, Lu J, Koodathingal P, Dai Q, Staley JP, Piccirilli JA. RNA catalyses nuclear pre-mRNA splicing. *Nature.* 2013; 503:229–234. [PubMed: 24196718]
64. Query CC, Konarska MM. CEF1/CDC5 alleles modulate transitions between catalytic conformations of the spliceosome. *RNA.* 2012; 18:1001–1013. [PubMed: 22408182]
65. Collier SE, Voehler M, Peng D, Ohi R, Gould KL, Reiter NJ, Ohi MD. Structural and functional insights into the N-terminus of *Schizosaccharomyces pombe* Cdc5. *Biochemistry.* 2014; 53:6439–6451. [PubMed: 25263959]
66. Yu YT, Maroney PA, Darzynkiwicz E, Nilsen TW. U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA.* 1995; 1:46–54. [PubMed: 7489488]
67. Sontheimer EJ, Gordon PM, Piccirilli JA. Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes Dev.* 1999; 13:1729–1741. [PubMed: 10398685]
68. Reiter NJ, Blad H, Abildgaard F, Butcher SE. Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change. *Biochemistry.* 2004; 43:13739–13747. [PubMed: 15504036]
69. Venditti V, Clos L, Niccolai N, Butcher SE. Minimum-energy path for a u6 RNA conformational change involving protonation, base-pair rearrangement and base flipping. *J Mol Biol.* 2009; 391:894–905. [PubMed: 19591840]
70. McManus CJ, Schwartz ML, Butcher SE, Brow DA. A dynamic bulge in the U6 RNA internal stem-loop functions in spliceosome assembly and activation. *RNA.* 2007; 13:2252–2265. [PubMed: 17925343]
71. Chanfreau G, Jacquier A. Catalytic site components common to both splicing steps of a group II intron. *Science.* 1994; 266:1383–1387. [PubMed: 7973729]
72. Butcher SE. The spliceosome and its metal ions. *Met Ions Life Sci.* 2011; 9:235–251. [PubMed: 22010274]
73. Shukla GC, Padgett RA. A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. *Mol Cell.* 2002; 9:1145–1150. [PubMed: 12049749]
74. Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. *Science.* 2008; 320:77–82. [PubMed: 18388288]
75. Robart AR, Chan RT, Peters JK, Rajashankar KR, Toor N. Crystal structure of a eukaryotic group II intron lariat. *Nature.* 2014; 514:193–197. [PubMed: 25252982]
76. Burke JE, Sashital DG, Zuo X, Wang YX, Butcher SE. Structure of the yeast U2/U6 snRNA complex. *RNA.* 2012; 18:673–683. [PubMed: 22328579]
77. Levy MZ, Allsopp RC, Fitcher AB, Greider CW, Harley CB. Telomere end-replication problem and cell aging. *J Mol Biol.* 1992; 225:951–960. [PubMed: 1613801]
78. Xu Y, Goldkorn A. Telomere and Telomerase Therapeutics in Cancer. *Genes.* 2016; 7:22.
79. Vulliamy TJ, Kirwan MJ, Beswick R, Hossain U, Baqai C, Ratcliffe A, Marsh J, Walne A, Dokal I. Differences in disease severity but similar telomere lengths in genetic subgroups of patients with telomerase and shelterin mutations. *PLoS One.* 2011; 6:e24383. [PubMed: 21931702]
80. Kachouri-Lafond R, Dujon B, Gilson E, Westhof E, Fairhead C, Teixeira MT. Large telomerase RNA, telomere length heterogeneity and escape from senescence in *Candida glabrata*. *FEBS Lett.* 2009; 583:3605–3610. [PubMed: 19840797]
81. Podlevsky JD, Chen JJ. Evolutionary perspectives of telomerase RNA structure and function. *RNA Biol.* 2016; 13:720–732. [PubMed: 27359343]
82. Mitchell M, Gillis A, Futahashi M, Fujiwara H, Skordalakes E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat Struct Mol Biol.* 2010; 17:513–518. [PubMed: 20357774]

83. Qi X, Xie M, Brown AF, Bley CJ, Podlevsky JD, Chen J. RNA/DNA hybrid binding affinity determines telomerase template-translocation efficiency. *EMBO J.* 2012; 31:150–161. [PubMed: 21989387]
84. Miller MC, Collins K. Telomerase recognizes its template by using an adjacent RNA motif. *Proc Natl Acad Sci U S A.* 2002; 99:6585–6590. [PubMed: 11997465]
85. Kim MM, Rivera MA, Botchkina IL, Shalaby R, Thor AD, Blackburn EH. A low threshold level of expression of mutant-template telomerase RNA inhibits human tumor cell proliferation. *Proc Natl Acad Sci U S A.* 2001; 98:7982–7987. [PubMed: 11438744]
86. Chen JL, Greider CW. Template boundary definition in mammalian telomerase. *Genes Dev.* 2003; 17:2747–2752. [PubMed: 14630939]
87. Sexton AN, Collins K. The 5' guanosine tracts of human telomerase RNA are recognized by the G-quadruplex binding domain of the RNA helicase DHX36 and function to increase RNA accumulation. *Mol Cell Biol.* 2011; 31:736–743. [PubMed: 21149580]
88. Booy EP, Meier M, Okun N, Novakowski SK, Xiong S, Stetefeld J, McKenna SA. The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. *Nucleic Acids Res.* 2012; 40:4110–4124. [PubMed: 22238380]
89. Zhang Q, Kim NK, Peterson RD, Wang Z, Feigon J. Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. *Proc Natl Acad Sci U S A.* 2010; 107:18761–18768. [PubMed: 20966348]
90. Theimer CA, Blois CA, Feigon J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell.* 2005; 17:671–682. [PubMed: 15749017]
91. Qiao F, Cech TR. Triple-helix structure in telomerase RNA contributes to catalysis. *Nat Struct Mol Biol.* 2008; 15:634–640. [PubMed: 18500353]
92. Theimer CA, Finger LD, Trantirek L, Feigon J. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci U S A.* 2003; 100:449–454. [PubMed: 12525685]
93. Bryan C, Rice C, Hoffman H, Harkisheimer M, Sweeney M, Skordalakes E. Structural Basis of Telomerase Inhibition by the Highly Specific BIBR1532. *Structure.* 2015; 23:1934–1942. [PubMed: 26365799]
94. Huang J, Brown AF, Wu J, Xue J, Bley CJ, Rand DP, Wu L, Zhang R, Chen J, Lei M. Structural basis for protein-RNA recognition in telomerase. *Nat Struct Mol Biol.* 2014; 21:507–512. [PubMed: 24793650]
95. Kim NK, Zhang Q, Feigon J. Structure and sequence elements of the CR4/5 domain of medaka telomerase RNA important for telomerase function. *Nucleic Acids Res.* 2014; 42:3395–3408. [PubMed: 24335084]
96. Leeper TC, Varani G. The structure of an enzyme-activating fragment of human telomerase RNA. *RNA.* 2005; 11:394–403. [PubMed: 15703438]
97. Snow BE, Erdmann N, Cruickshank J, Goldman H, Gill RM, Robinson MO, Harrington L. Functional conservation of the telomerase protein Est1p in humans. *Curr Biol.* 2003; 13:698–704. [PubMed: 12699629]
98. Egan ED, Collins K. Specificity and stoichiometry of subunit interactions in the human telomerase holoenzyme assembled in vivo. *Mol Cell Biol.* 2010; 30:2775–2786. [PubMed: 20351177]
99. Azzalin CM, Reichenbach P, Khoraiuli L, Giulotto E, Lingner J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science.* 2007; 318:798–801. [PubMed: 17916692]
100. Hirschi A, Martin WJ, Luka Z, Loukachevitch LV, Reiter NJ. G-quadruplex RNA binding and recognition by the lysine-specific histone demethylase-1 enzyme. *RNA.* 2016; 22:1250–1260. [PubMed: 27277658]
101. Porro A, Feuerhahn S, Delafontaine J, Riethman H, Rougemont J, Lingner J. Functional characterization of the TERRA transcriptome at damaged telomeres. *Nat Commun.* 2014; 5:5379. [PubMed: 25359189]

102. Porro A, Feuerhahn S, Lingner J. TERRA-reinforced association of LSD1 with MRE11 promotes processing of uncapped telomeres. *Cell Rep.* 2014; 6:765–776. [PubMed: 24529708]
103. Schoeftner S, Blasco MA. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol.* 2008; 10:228–236. [PubMed: 18157120]
104. Redon S, Reichenbach P, Lingner J. The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res.* 2010; 38:5797–5806. [PubMed: 20460456]
105. Cifuentes-Rojas C, Nelson ADL, Boltz KA, Kannan K, She X, Shippen DE. An alternative telomerase RNA in *Arabidopsis* modulates enzyme activity in response to DNA damage. *Genes Dev.* 2012; 26:2512–2523. [PubMed: 23109676]
106. Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics.* 2013; 193:651–669. [PubMed: 23463798]
107. Jiang J, Miracco EJ, Hong K, Eckert B, Chan H, Cash DD, Min B, Zhou ZH, Collins K, Feigon J. The architecture of *Tetrahymena* telomerase holoenzyme. *Nature.* 2013; 496:187–192. [PubMed: 23552895]
108. Blad H, Reiter NJ, Abildgaard F, Markley JL, Butcher SE. Dynamics and metal ion binding in the U6 RNA intramolecular stem-loop as analyzed by NMR. *J Mol Biol.* 2005; 353:540–555. [PubMed: 16181635]
109. Kim NK, Zhang Q, Zhou J, Theimer CA, Peterson RD, Feigon J. Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J Mol Biol.* 2008; 384:1249–1261. [PubMed: 18950640]

Figure 1.

Conserved RNA scaffolds at the core of RNP complexes. (A) Two conserved domains (blue and cyan) of the ribonuclease P (RNase P) enzyme are separated by 250–350 nucleotides but assemble to form the central scaffold, termed the P4 helix. A single-nucleotide bulge within the P4 helix is universally conserved. The crystal structure of a bacterial RNase P holoenzyme with tRNA (PDB entry 3Q1Q) is shown, with the P4 helix (blue and cyan), active site metal ion location (magenta spheres), surrounding P RNA structure (gray), and P protein (blue-white surface).¹¹ (B) An asymmetric internal stem-loop structure (green and lime) from the U6 small nuclear RNA (termed the U6 ISL) helps to form the activated spliceosome core. Cryo-EM structures^{12,14–16} confirm a wealth of genetic and biochemistry studies, revealing that this structurally conserved helix is central to the core of the spliceosome. The cryo-EM-derived minimal structure is shown, with U6 ISL (green/lime), active site metal ions (magenta), surrounding snRNAs (gray), and protein components (blue-white surface) (PDB entry 5LJ3).¹² (C) The telomerase RNA component (TR) utilizes a conserved pseudoknot, where a stem-loop structure (red) intercalates with a single-stranded region (pink). This structural element is central to the assembly of an activated telomerase holoenzyme and is located near the RNA template of the DNA polymerase reaction but is not directly involved in the reverse transcriptase protein (TERT) enzyme active site.^{13,107} A pseudoatomic experimental model of the catalytic core of the *Tetrahymena* telomerase¹⁰ illustrates the location of the pseudoknot (red/pink), the approximate active site metal (magenta), surrounding TR (gray), and core protein components (blue-white surface). Although the RNA from *Tetrahymena* holoenzyme differs from human TR in size and putative tertiary structure, both contain the RNA pseudoknot and three-way junction structures.^{13,81} In panels A–C, dashed lines represent base pairs and a summary of the reaction chemistry, evolutionarily related RNAs (paralogs), and primary substrate recognition modes are indicated.

Figure 2.

Core RNA features of the RNase P holoenzyme. (A) RNA secondary structural schematic and crystal structure of the bacterial RNase P holoenzyme with tRNA, highlighting five conserved regions (CR-I–V) in cyan and a universally conserved uridine nucleotide (U52, *Thermatoga maritima* numbering) that is located within helix P4. Long-range interactions are noted; four structural or catalytically important metal ion sites are identified (magenta spheres), and a dashed circle (red) indicates the location of an A-minor interaction between the P RNA and tRNA. The tertiary structure shows how the *T. maritima* RNase P RNA (blue) and P protein (green) globally engage with tRNA and a short oligonucleotide leader mimic (pink) (PDB entry 3Q1R). (B) CR-II and -III form two preorganized T loops that promote hydrophobic stacking interactions (dashed black) with the TΨC–D loop of tRNA. A structurally important metal ion (M4, pink sphere) and an A-minor interaction (dashed red circle) facilitate this interaction and enhance RNA–RNA shape complementarity. Two conserved, stacked base triples reside within the interlocked T loops of the P RNA (boxed). (C) The bacterial P protein increases catalytic efficiency by stabilizing conserved, nonhelical junctions (CR-IV and -V) proximal to the active site (denoted as M1 and M2 metal ions) and adjacent to helix P3. (D) RNA active site environment of RNase P. Universally conserved nucleotides (bold and underlined) surround the 5′ end of tRNA (pink) and provide ligands

to catalytically important metal ions (M1 and M2, spheres). Crystal structures of the bacterial apo form (orange) and holoenzyme product (cyan) show a largely preorganized (PDB entries 2A64 and 3Q1Q, respectively) active site, though the universally conserved and bulged uridine (U52) likely undergoes a dynamic motion (indicated by arrows), helping to contribute to metal ion coordination and active site stabilization upon substrate binding.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3.

Active site of the spliceosome that suggests an ancient architecture. (A) RNA secondary structural schematic showing RNA–RNA interactions immediately after branching and prior to exon ligation. Intermolecular base pairing [involving ACAGAGA, branch, and 5′ exon/U5 loop 1 helices (gray)], an internal stem–loop structure in U6 (U6 ISL), and stacked base triple interactions (triplex) are likely structurally conserved. Labeled nucleotides within the U2, U5, and U6 snRNAs (purple, pink, and green, respectively) and intron region (black) are invariant on the basis of sequence conservation. The branched adenosine, which contains the 2′-OH nucleophile in the first step of the reaction, is highlighted (yellow). The 5′ exon (orange) contains the 3′-OH nucleophile in the second step. Metal ions (M1 and M2, magenta spheres) facilitate both chemical steps of the reaction.⁶³ (B) Active site according to the *S. cerevisiae* cryo-EM structure (PDB entry 5LJ3).¹² RNA tertiary structural elements, including the U6 ISL and catalytic triplex (bold) and intermolecular helices (gray), help to configure the catalytic core and surround the catalytic metal ions (M1 and M2). All components as colored as in panel A. In addition, critical *S. cerevisiae* proteins that orchestrate these RNA–RNA interactions at this step include Prp8 (220K and Spp42), Clf1 (CRNKL1 and Cwf4), Cef1 (CDC5L and Cdc5), Yju2 (CCDC94 and Cwf16), Cwc25 (CWC25 and Cwf25), and Cwc15 (CWC15 and Cwf15). *Homo sapiens* and *Sc. pombe* names, respectively, are given in parentheses. (C) U6 ISL (green) overlay with domain 5 (D5) from group II self-splicing introns reveals an ancient and conserved stem–loop structure at the active site core (PDB entries 5LJ3, 4R0D, and 4E8K). (D) Overlay of the internal loop within the ISL reveals a conserved bulge structure that is shared between the spliceosome and group II self-splicing introns.^{12,57,75} This internal-loop region undergoes dynamic motions in activated spliceosomes.^{69,70,108} (E) Three stacked base triples are proximal to the enzyme active site and are structurally conserved between the spliceosome and group II introns.^{21,56} (F) Base triple interactions between U6 and U2 snRNAs are conserved in the spliceosome and function to define the 5′ splice site (*S. cerevisiae* numbering).

Figure 4.

Multifunctional structural features of the telomerase RNA (TR). (A) RNA secondary structure schematic of human telomerase RNA (TR), based on experimental data.¹⁷ The eight conserved regions (CR) are highlighted in yellow, and core features are annotated. These include guanosines clustered in the 5' region with the propensity to form an inhibitory G-quadruplex tertiary structure, the P1 helix that forms the telomerase boundary element (TBE), the J2a/b region with a five-nucleotide bulge that forms a critical bend, and the destabilizing, unpaired uridine (green) in the pseudoknot. CR4/5 forms a three-way junction and contains a bulged cytosine that likely interacts with reverse transcriptase, providing specificity. Near the 3' end of TR, CR6/8 forms an H/ACA domain that interacts with snoRNA-associated proteins. (B) Simplified schematic of the telomerase reverse transcriptase (TERT) complex, with the RNA colored red and the multidomain reverse transcriptase protein colored blue. TR regions are labeled as in panel A. The pseudoknot and P6 helix regions are boxed, corresponding to tertiary structures in panels C and D, respectively. (C) The pseudoknot region forms a kinked, well-ordered structure with tertiary interactions and extended triple helices.¹⁰⁹ The three strands of the triple helix are colored as in Figure 1C, with the bulged uridine colored green. Several stacked base triples are indicated (*H. sapiens*, PDB entry 2K95). (D) CR4/5 (red) in *O. latipes* TR forms a three-way

junction and binds the telomeric RNA binding domain (TRBD, blue).⁹⁴ Specificity is provided by the well-conserved, unpaired nucleotides uridine 182 (U182), which likely functionally substitutes for the bulged cytosine in humans, and guanosine 189 (G189). The boxed inset regions show that two stacked base triples at the crux of the triple helix form the base for the P6 helix (PDB entry 4O26).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript