RNA as a Flexible Scaffold for Proteins: Yeast Telomerase and Beyond

D.C. ZAPPULLA AND T.R. CECH
Howard Hughes Medical Institute, Department of Chemistry and Biochemistry,
University of Colorado, Boulder, Colorado 80309

Yeast telomerase, the enzyme that adds a repeated DNA sequence to the ends of the chromosomes, consists of a 1157-nucleotide RNA (TLC1) plus several protein subunits: the telomerase reverse transcriptase Est2p, the regulatory subunit Est1p, the nonhomologous end-joining heterodimer Ku, and the seven Sm proteins involved in ribonucleoprotein (RNP) maturation. The RNA subunit provides the template for telomeric DNA synthesis. In addition, we have reported evidence that it serves as a flexible scaffold to tether the proteins into the complex. More generally, we consider the possibility that RNPs may be considered in three structural categories: (1) those that have specific structures determined in large part by the RNA, including RNase P, other ribozyme–protein complexes, and the ribosome; (2) those that have specific structures determined in large part by proteins, including many small nuclear RNPs (snRNPs) and small nucleolar RNPs (snoRNPs); and (3) flexible scaffolds, with no specific structure of the RNP as a whole, as exemplified by yeast telomerase. Other candidates for flexible scaffold structures are other telomerases, viral IRES (internal ribosome entry site) elements, tmRNA (transfer-messenger RNA), the SRP (signal recognition particle), and Xist and rol1 RNAs that alter chromatin structure to achieve dosage compensation.

The ends of linear chromosomes cannot be replicated completely by the same enzymatic machinery that replicates the internal portions. In the absence of a special replication mechanism, chromosomes literally shrink from their ends. This DNA end-replication problem is overcome in most eukaryotes by the enzyme telomerase.

Telomerase is an RNP enzyme, and the function of one portion of its RNA subunit is gratifyingly easy to understand: It provides the template for extension of telomeric DNA (Greider and Blackburn 1989; Yu et al. 1990). The RNA itself performs other functions, the best-established example being the formation of the “template boundary,” a short intramolecular helix that defines the stopping point for reverse transcription (Tzfati et al. 2000). Finally, the RNA binds a number of proteins, foremost among these being the telomerase reverse transcriptase (TERT, or Est2p in Saccharomyces cerevisiae). The RNA also binds a number of accessory proteins, but these are more species-specific. In yeast, they consist of Est1p, which recruits or activates telomerase at the chromosome end; the Ku heterodimer, the DNA-repair protein that has a special role in recruiting telomerase to broken chromosome ends for de novo telomere addition; and the Sm proteins, which also contribute to biogenesis of the snRNPs. Yet the function, or absence of function, of the majority of the telomerase RNA remains unknown.

At an even more fundamental level, we still do not even know how to envision the structure of the telomerase RNP. Is it a more-or-less specific three-dimensional structure formed by a precise series of RNA–RNA, RNA–protein, and protein–protein interactions, along the lines of the ribosome or RNase P, such that it could someday be crystallized? Alternatively, is it possibly a loose collection of proteins held together by RNA “strings” such that no two complexes have an identical overall shape? In the case of yeast telomerase, we describe evidence that supports the latter model: RNA as a flexible scaffold for proteins. Our interrogation of yeast telomerase has stimulated us to think more generally about the structure–function relationships in RNPs, especially those involved in catalysis, and we present those thoughts in the Conclusions.

YEAST TELOMERASE RNA SECONDARY STRUCTURE

For noncoding RNAs, knowing the nucleotide sequence is of very limited value by itself, but having a good secondary structure model provides a useful framework for investigating structure–function relationships. In the case of telomerase RNAs, secondary structures were determined for ciliated protozoan examples and showed a single-stranded template region, a nearby pseudoknot, and a long “handle” (Stem IV) whose terminal loop is essential for function (Romero and Blackburn 1991; ten Dam et al. 1991; Lingner et al. 1994; Sperger and Cech 2001). Vertebrate telomerase RNAs folded into a structure somewhat similar to that of the ciliates, followed by a 3′-terminal pair of stem-loop structures characteristic of box H/ACA snoRNAs (Mitchell et al. 1999; Chen et al. 2000).

Budding yeast telomerase RNA structures were difficult to decipher due to their unusual length (typically >1000 nucleotides) and rapid divergence, which frustrates sequence alignments and makes it difficult to apply comparative sequence analysis. The Blackburn laboratory was able to model several helices of telomerase RNAs from Kluyveromyces species (Tzfati et al. 2003). Furthermore, for S. cerevisiae telomerase RNA (TLC1), several local structures involved in binding of the Ku and Est1 proteins...
were established (Peterson et al. 2001; Seto et al. 2002), and a single-stranded U-rich sequence involved in Sm protein binding was identified (Seto et al. 1999). Finally, in 2004, two laboratories converged on the same secondary structure model for TLC1 RNA, as shown in Figure 1A (Dandjinou et al. 2004; Zappulla and Cech 2004).

The three long arms of this secondary structure model are unusual features for a noncoding RNA; overall, the structure looks completely different from canonical structures of rRNA, group I and group II introns, RNase P RNA, tmRNA, etc., which are highly branched structures. Furthermore, each of the long RNA arms has near its end a binding site for one of the known accessory proteins: Ku, Est1p, or Sm. The three long arms consist of double-helical regions interrupted by internal loops and bulges, features known to provide bendable joints for RNA in solution. The elements more directly involved in catalysis—Est2p, the RNA template, and the template boundary element—are brought together in the central core of the structure. When the protein and RNA components are represented more or less according to predicted scale, telomerase appears as small protein beads separated by long RNA arms (Fig. 1B).

**FLEXIBLE SCAFFOLD MODEL**

The model that the yeast telomerase RNA provides a flexible scaffold for protein binding was derived from several observations. First, large stretches of the RNA could be deleted without significant loss of function in vivo (Livengood et al. 2002). Second, most of the RNA is evolving very rapidly, as evidenced by the divergent sequences of telomerase RNAs from species in the genus Kluyveromyces (Tzfati et al. 2000; Seto et al. 2002) or the genus Saccharomyces. In fact, TLC1 RNA has only 43% sequence identity between four Saccharomyces species examined, compared to 82–99% for other noncoding RNAs (U1, RNase P, and 18S rRNA) and 92% for the actin mRNA open reading frame (Zappulla and Cech 2004). These observations indicate that the RNA is “flexible,” meaning that it readily accommodates structural perturbations. It is likely that the RNA is also “flexible” in the sense that its long quasi-helical arms are free to twist, turn, and move in many directions, but this sort of structural flexibility is not directly addressed by our data.

**TESTING THE FLEXIBLE SCAFFOLD MODEL**

To provide a strong test of the flexible scaffold model, we mutated the natural binding site for the Est1p regulatory subunit on the RNA and then introduced a functional Est1p-binding element (nucleotides 524–704) at position 220, 450, or 1033 (Fig. 2A). Since mutation of the Est1p-binding site in TLC1 leads to senescence, if the repositioned Est1p site restores the essential function of Est1p in telomerase, the cell should become viable. Not only did these variant TLC1 RNAs with repositioned Est1p sites provide robust cell growth, but they maintained near wild-type telomere lengths (Zappulla and Cech 2004). Among other controls, a mutated version of the Est1p-binding site inserted at any of the same three unnatural locations did not rescue telomerase activity. Thus, it appears that as long as Est1p is tethered somewhere to the RNA, it is able to perform its function in recruiting telomerase to its site of action at the very end of the chromosome.

Given that yeast telomerase appeared to consist of proteins tethered by flexible RNA arms, we next asked whether the arm length was important. To avoid creating misfolded RNAs, the RNA secondary structure model was used to design pairs of deletions that would remove both sides of an RNA stem. First, a major portion of the terminal arm was deleted, and then deletions were made in all three major arms; in each case, the shortened “Mini-T” RNA (Fig. 2B) complemented a tlc1Δ mutant and telomere length was maintained at a stably shorter length. However, either the length or some other feature of the RNA arms was important for normal levels of RNA accumulation, as the Mini-T strains had substantially reduced...

**Figure 1.** Secondary structure model of yeast telomerase RNA including protein-binding sites. (A) RNA secondary structure model of Zappulla and Cech (2004), based on computer-generated RNA stability calculations and limited comparative sequence analysis and mutagenesis. Dandjinou et al. (2004) independently proposed a highly similar model. (B) Schematic model of the RNP with the RNA (green) shown approximately to scale with the proteins whose structures have been determined. Est1p and Est2p, whose detailed structures are independently proposed a highly similar model. (B) Schematic model of the RNP with the RNA (green) shown approximately to scale with the proteins whose structures have been determined.
levels of TLC1 RNA. When Mini-T(500), the 500-nucleotide version of Mini-T, was integrated into its normal locus in the yeast genome, the RNA expression level was only about 5% of the wild-type level.

Mixing the Mini-T(500) yeast cells with wild type and carrying out competitive growth in liquid culture revealed a large fitness disadvantage of Mini-T, about 22% per generation (Fig. 3). However, when the same RNA was expressed from a low-copy-number centromere-containing plasmid, which gives an expression level 11–16% of wild type, the fitness was much better: only about 3% selective disadvantage per generation (Zappulla et al. 2005). Thus, it appears that most if not all of the selective disadvantage of Mini-T(500) is due to its reduced expression level rather than some functional defect. It may be that the long arms of the natural TLC1 RNA provide binding sites for general RNA-binding proteins that promote nuclear retention or RNA stability while not contributing directly to telomerase activity.

Recent work by our colleague Dr. Amy Mozdy helps illuminate why reduced TLC1 expression has such a major effect on telomere length. Wild-type haploid yeast contain a steady-state number of about 30 TLC1 RNA molecules (Mozdy and Cech 2006). This can be compared to 32 telomeres per cell at the beginning of S phase and 64 telomeres per cell in late S phase when telomeres have been replicated. Thus, telomeres outnumber telomerase even in wild-type cells. The 5% expression level of integrated Mini-T(500) would correspond to about one molecule of TLC1 per cell. Thus, it is not surprising that the telomeres are maintained at an average length much shorter than wild type (Fig. 3A).

CONCLUSIONS

Yeast Telomerase

We have presented evidence that yeast telomerase RNA provides a flexible scaffold for tethering proteins into the complex. By “flexible,” we mean that there are only loose constraints on the length, sequence, and relative orientation of the RNA arms that tether the Est1p, Ku, and Sm protein components. Given the structure of the RNA arms—short helices separated by internal loops and
bulges (Fig. 1A)—they must also be flexible in the sense of allowing bending, rotation, and other dynamics. We have not tested the importance of this sort of flexibility, which could probably be done by site-specific mutagenesis, to eliminate the internal loops and bulges, followed by tests for activity (see, e.g., Nakamura et al. 1995).

We speculate that functional yeast telomerase might not even require the arms to be made of RNA. If the only purpose of the arms is to tether the proteins in a flexible manner, then the RNA arms might be replaced by protein α helices or loops, or even by polyethylene glycol. An interesting comparison concerns the scaffold proteins involved in cell signaling, which bind multiple protein kinases. Wendell Lim has asked exactly the same question of these scaffold proteins as we have asked of the telomerase RNA: “Do they simply tether components, or do they precisely orient and activate them?” (Park et al. 2003). These authors found that the yeast MAP kinase scaffold protein Ste5 survived major rearrangements, supporting the former model: “Scaffolds are highly flexible organizing factors.”

Even regarding yeast telomerase RNA, much remains unknown. The “flexible scaffold” model may break down near the active site, where it would not be surprising if the Est2p catalytic subunit had to be oriented in a specific manner relative to the RNA template that it reverse-transcribes. However, even in the central core, Chappell and Lundblad (2004) have described a collection of mutants that do not disrupt function, and they even replaced the yeast sequences in the core with the RNA pseudoknot from the ciliated protozoan Oxytricha nova and obtained a functional RNA. Thus, the flexible scaffold model may extend into the active-site region.

One possible explanation for the telomerase RNA of Saccharomyces having a structural organization differ-

Figure 3. Yeast with Mini-T(500) integrated in the genome, replacing the wild-type TLC1 gene, maintain shorter telomeres and show reduced fitness upon competitive growth. (A) Southern hybridization with a telomeric DNA probe shows the length distribution of wild-type (WT) and Mini-T telomeres. About half (17 of 32) of the yeast telomeres contain the repeated Y’ sequence just internal to the telomeric repeats, such that they are cleaved by XhoI restriction endonuclease to give a low-molecular-weight set of fragments. Non-Y’ telomeres form individual bands at higher molecular weights; one example is highlighted by a closed triangle (wild-type) and an open triangle (Mini-T). Upon continued growth in liquid medium, the cells containing WT TLC1 outcompete the Mini-T cells and take over the culture. (B) Northern hybridization with a probe to the region of TLC1 RNA that is shared between the WT RNA and Mini-T(500). Again, upon continued growth, the cells containing WT TLC1 outcompete the Mini-T cells. (C) Quantitation of the data in panel A and a repeat experiment (primary data not shown) indicate a fitness disadvantage of 22% per generation for Mini-T cells. Most of this fitness defect was due to reduced levels of the Mini-T(500) RNA (see text).
RNA AS FLEXIBLE SCAFFOLD FOR PROTEINS

Three Classes of RNPs

In considering RNPs more generally, we quickly come to the realization that the existing data are insufficient to prove or disprove many of them as being flexible scaffold structures, in the sense that much of their RNA framework can accommodate insertions, deletions, and base changes without disruption of function. In some cases, the “accommodativeness” or malleability of the RNA can be inferred from comparative sequence analysis of functionally equivalent RNAs from divergent species. In addition, in many cases, we know something about the structural flexibility of the RNP—which each particle has more or less the same arrangement of protein and RNA components, albeit with some conformational switching, or whether, on the other hand, proteins are loosely tethered into the complex. Here, we make the provisional assumption that structural flexibility of the RNP correlates with the ability of the RNA to accommodate insertions, deletions, and base changes in regions that serve as tethers; but in doing so, we realize that there may be cases where these features diverge. On the basis of these assumptions, we propose that RNPs may fit into three general structural categories, as follows.

1. RNA-determined RNP structures. Specific structure (or family of structures), determined in large part by the folded RNA.

   Examples: Ribosome, RNase P, and other ribozyme–protein complexes such as the yeast mitochondrial group I intron b15 bound to Cbp2 protein, or a number of Neurospora group I introns bound to CYT18 protein.

   Evidence: The RNP, and in some cases the RNA, can be crystallized. Individual ribosomal proteins often have disordered tails or internal loops, and these become ordered only upon RNP formation, so at least for the ribosome, the protein does not provide a pre-ordered template upon which the RNA folds (Brodersen et al. 2002; Klein et al. 2004). Not only do the ribosomal RNAs have conserved secondary structures, but there are bases throughout the structures that are universally conserved among the corresponding RNAs of all three primary kingdoms of life (Noller 1993). In Saccharomyces species, the sequence conservation of the rRNA far exceeds that of the telomerase RNA (99% and 43% sequence identity, respectively) (Zappulla and Cech 2004).

2. Protein-determined RNP structures. Specific structure (or family of structures), determined in large part by protein–protein interactions.

   Examples: U1, U2, U11/U12, and U5 snRNPs, box C/D, and box H/ACA snoRNPs.


   Evidence for structure being protein-based: The majority of U2-specific proteins form stable heteromeric complexes in absence of U2 snRNA (Will and Luhrmann 2006), U5 proteins assemble in absence of snRNA (Achsel et al. 1998), and the crystal structure of Cbf5–Nop10–Gar1 complex has been determined in the absence of snoRNA (Rashid et al. 2006).

3. Flexible scaffold RNPs. Proteins are tethered into the complex by a flexible RNA scaffold, with no specific structure of the RNP as a whole.

   Example: Yeast telomerase, other candidates discussed below.

   Evidence for flexibility: Phylogenetic variation in sequence even among closely related Saccharomyces species, ability to delete large portions of stems and still retain function, and ability to transplant Est1p-binding element to new locations with function.

   Evidence against specific structure: The RNA does not have the properties one would expect for a specific tertiary structure, so structure would have to come from the protein; but protein–protein interactions are not strong, unlike snRNPs and snoRNPs above.

   Prediction: The RNP would not crystallize, and would be heterogeneous by EM.

Telomerases from humans and from ciliated protozoa have been extensively studied, their RNA secondary structures are well established (Romero and Blackburn 1991; Chen et al. 2000), and a specific tertiary structure has been proposed for the core of the human RNA (Theimer et al. 2005). We find it attractive to think that these much smaller telomerase RNAs may be natural Mini-T RNAs, with the protein-binding arms reduced to a minimal length so that it is now difficult to even recognize the arms. If this is the case, it should be possible to add back flexible RNA arms to separate various functional elements of these other telomerase RNAs without perturbing activity. On the other hand, these other telomerases may be assembled with protein–protein interactions having replaced some of the RNA–protein interactions seen in yeast telomerase, in which case, the RNP may have a defined three-dimensional structure.

IRES elements that permit 5'-cap-independent initiation of protein synthesis of certain viral and cellular
mRNAs provide an interesting case. They bind multiple protein complexes, including the 12-protein initiation factor eIF3, as well as the small ribosomal subunit. The complex between the hepatitis C virus IRES RNA and eIF3 has some conformational flexibility, as determined by cryo-EM reconstructions (Siridechadilok et al. 2005). Both the IRES RNA and the ribosomal subunits undergo further conformational changes as translation is initiated and commences (Spahn et al. 2001; Boehringer et al. 2005). Thus, the hepatitis C IRES may be a flexible collection of proteins that becomes conformationally fixed upon binding to the ribosome. All IRESs are not identical, however. In the case of the cricket paralysis virus IRES, the RNA by itself forms a highly compact structure that interacts directly with the ribosome to achieve factor-independent translational initiation (Batey 2006).

If the RNA is responsible for tran-
scriptional repression and association with chromatin, the latter accomplished by functionally redundant sequences dispersed through the RNA (Wutz et al. 2002). Deletion analysis has shown that multiple 10% segments of roX2 RNA are dispensable for function, consistent with the RNA providing a flexible scaffold to bind multiple MSL (male-specific lethal) protein complexes (Stuckenholz et al. 2003).

Another intriguing candidate for a flexible scaffold is the NRON (noncoding repressor of NFAT) recently identified as a repressor of a transcription factor, NFAT (nuclear factor of activated T cells) (Willingham et al. 2005). NRON binds a number of proteins, including members of the importin β family, and appears to act as a repressor of the transcription factor NFAT. It will be important to understand something of its structure and the spatial arrangement of its protein-binding sites.

It may seem surprising that we have placed yeast telomerase into a different category from the snoRNPs. Superficially at least, they have similar features: In all cases, the RNA provides a “guide sequence” that base-pairs to the site of a chemical reaction in a nucleic acid substrate. The reaction is 2′-O-methylation in the case of the box C/D snoRNPs (Kiss-Laszlo et al. 1996), conversion of U to pseudouracil in the case of the box H/ACA snoRNPs (Ganot et al. 1997), and nucleotide addition to single-stranded DNA in the case of telomerase. In all cases, the RNA also binds a protein enzyme that catalyzes the reaction: fibrillarin or Nop1p in the case of 2′-O-methylation (Tollervey et al. 1993; Wang et al. 2000; Omer et al. 2002), NAP57/dyskerin or Chl5p in the case of pseudouridylation (Zebbar biased et al. 1999), and TERT/Es2p in the case of telomerase (Lingner et al. 1997). Certainly, there is some flexibility in the box H/ACA snoRNP system: The protein trimer can accommodate about 100 different snoRNAs (Meier 2006). We place the snoRNPs in a different category from yeast telomerase because the former form a specific protein structure in the absence of the RNA, whereas stable protein–protein interactions have not been observed between Es2p and Est1p, Ku, or Sm in yeast. Time will tell whether this is a useful distinction or whether the apparent differences arise in large part because of the primitive state of our current knowledge.

ACKNOWLEDGMENTS

We thank Karen Goodrich for excellent technical assistance; Rob Batey, Quentin Vicens, and Art Zaug for helpful discussions; and Anne Stellwagen (Boston College) for an ongoing collaboration on the Ku protein. This work was supported in part by a grant from the National Institutes of Health.

REFERENCES

Chappell A.S. and Lundblad V. 2004. Structural elements
Igel A.H. and Ares M., Jr. 1988. Internal sequences that distin-
Greider C.W. and Blackburn E.H. 1989. A telomeric sequence
Dandjinou A.T., Levesque N., Larose S., Lucier J.F., Abou Elela
Kiss-Laszlo Z., Henry Y., Bachellerie J.P., Caizergues-Ferrer
Klein D.J., Moore P.B., and Steitz T.A. 2004. The roles of ribo-
Mura C., Cascio D., Sawaya M.R., and Eisenberg D.S. 2001. The crystal structure of a heptamer of Sm protein: Implications for the eukaryotic snRNPs.


RNA as a Flexible Scaffold for Proteins: Yeast Telomerase and Beyond

D.C. ZAPPULLA and T.R. CECH

Access the most recent version at doi:10.1101/sqb.2006.71.011

License

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 Cold Spring Harbor Symposia on Quantitative Biology go to:
http://symposium.cshlp.org/subscriptions

Copyright 2006, Cold Spring Harbor Laboratory Press