UCLA
UCLA Previously Published Works

Title
Structure and folding of the Tetrahymena telomerase RNA pseudoknot

Permalink
https://escholarship.org/uc/item/1k283373

Journal
Nucleic Acids Research, 45(1)

ISSN
0305-1048

Authors
Cash, Darian D
Feigon, Juli

Publication Date
2017-01-09

DOI
10.1093/nar/gkw1153

Peer reviewed

Structure and folding of the *Tetrahymena* telomerase RNA pseudoknot

Darian D. Cash and Juli Feigon*

Department of Chemistry and Biochemistry, and Molecular Biology Institute, University of California, Los Angeles, CA 90095-1569, USA

ABSTRACT

Telomerase maintains telomere length at the ends of linear chromosomes using an integral telomerase RNA (TER) and telomerase reverse transcriptase (TERT). An essential part of TER is the template/pseudoknot domain (t/PK) which includes the template, for adding telomeric repeats, template boundary element (TBE), and pseudoknot, enclosed in a circle by stem 1. The *Tetrahymena* telomerase holoenzyme catalytic core (p65-TER-TERT) was recently modeled in our 9 Å resolution cryo-electron microscopy map by fitting protein and TER domains, including a solution NMR structure of the *Tetrahymena* pseudoknot. Here, we describe in detail the structure and folding of the isolated pseudoknot, which forms a compact structure with major groove U•A-U and novel C•G-A⁺ base triples. Base substitutions that disrupt the base triples reduce telomerase activity *in vitro*. NMR studies also reveal that the pseudoknot does not form in the context of full-length TER in the absence of TERT, due to formation of a competing structure that sequesters pseudoknot residues. The residues around the TBE remain unpaired, potentially providing access by TERT to this high affinity binding site during an early step in TERT-TER assembly. A model for the assembly pathway of the catalytic core is proposed.

INTRODUCTION

Telomerase is a ribonucleoprotein (RNP) complex composed of the catalytic telomerase reverse transcriptase (TERT), telomerase RNA (TER) and accessory proteins that vary with species [1,2]. Telomerase maintains the telomeric DNA at the 3′ ends of linear chromosomes by adding *de novo* telomeric DNA repeats [3,4]. Telomeres, composed of telomeric DNA and bound proteins, counteract DNA damage due to incomplete replication, degradation, chromosomal fusion and illicit DNA damage re-

pair [5–7]. TERT is generally well conserved ranging from ∼900 to 1100 amino acids (aa) among the commonly studied vertebrates, yeast and ciliates. It contains four conserved domains: the TERT essential N-terminus (TEN), telomerase RNA binding domain (TRBD), reverse transcriptase (RT) domain (palm and fingers) and C-terminal extension (CTE) (thumb) [8]. TER is more divergent, ranging in size from ∼150 nucleotides (nt) in ciliates to over 3000 nt in yeasts [9]. TER has a number of conserved elements, which are the template, pseudoknot (PK), template boundary element (TBE) and stem terminus element (STE) [1,10]. The template/pseudoknot domain (t/PK) (also called core domain) of TER includes the template that is used to copy the telomeric sequence and an adjacent pseudoknot. The pseudoknot is important for activity and has been proposed to have a direct role in catalysis, in template positioning, and/or in assembly [11–15]. The TBE, within the t/PK, is typically a stem or hairpin upstream of the template that binds tightly to the TRBD and helps prevent copying of nontemplate residues [16,17]. Ciliates also have a template recognition element (TRE) 3′ of the template, which TERT uses to direct efficient use of the template for reverse transcription [18]. The STE is distal from the t/PK and stimulates telomerase activity typically through TERT TRBD interactions [10]. Along with TERT, the t/PK and STE of TER are the minimal components necessary to reconstitute telomerase activity *in vitro* [1,19].

The structures of individual TER domains have been extensively studied in yeast, human and ciliates [15,20–29]. The NMR solution structures of the human and yeast *K. lactis* TER pseudoknots revealed a number of conserved features [15,21,29]. Despite a divergent sequence and secondary structure between the two pseudoknots, the 3D structures are very similar, indicating conserved functions or interactions. Both pseudoknots contain an extended triple helix where the loop residues form base triples with the Watson–Crick (WC) paired stem(s). Formation of the base triples was shown to be important for catalytic activity *in vitro* and *in vivo* [14,15,29–31]. Early biochemical studies have also linked the pseudoknot to telomerase assembly and telomere repeat addition processivity (RAP) [32,33].

---

*To whom correspondence should be addressed. Tel: +1 310 206 6922; Fax: +1 310 825 0982; Email: feigon@mbi.ucla.edu

A smaller pseudoknot was predicted to form in ciliate TERs (34). The telomere-rich ciliate *Tetrahymena* has served as a model organism from which telomeres and telomerase was discovered (35,36). Ciliate TERs are generally composed of four base paired regions (numbered 1–4 and depicted in Figure 1) where Stems 3A and 3B form the pseudoknot (34,37,38), Stem 2 and adjacent single-strand nucleotides are the TBE, and Stem 1 closes the t/PK circle. Distal stem–loop (SL) 4 is the ciliate STE, and is required along with the t/PK for activity. NMR structures of SL2 and SL4 have been determined (22,25,26,39). Stem 4 was crystallized in the presence of telomerase accessory protein p65 C-terminal domain (40). p65 binds Stem 4, inducing a 105° bend at the conserved GA bulge that positions Loop 4 to interact with the TRBD (40,41). The conformational change promotes hierarchical assembly of telomerase, with the p65-TER complex facilitating assembly with TERT (40–43). The ciliate pseudoknot has been particularly challenging to characterize, due to its conformational dynamics, which was revealed in early chemical probing studies (44). It has a much shorter Stem A and Loop A than those found in vertebrates (P2 and J2b/3) and yeasts. Interestingly, the short length of Stem A is conserved amongst ciliates (3–4 bp) suggesting a possible function for pseudoknot conformational flexibility (45). Based on phylogenetic analysis and modeling this pseudoknot was predicted to have 2 U•A-U triples in Stem B-Loop A (shown in Figure 1B) (45).

A number of studies have examined pseudoknot formation in the *Tetrahymena* TER. Chemical probing revealed that the isolated *Tetrahymena* pseudoknot (tetPK) can form a pseudoknot, while tetPK does not form in the context of full length (protein-free) TER (44). The authors concluded that in full-length TER the pseudoknot region is in equilibrium with alternate conformations. More recent single-molecule FRET experiments verified that tetPK forms in isolation but not in full-length TER (46). Furthermore, FRET values for free TER suggested that the pseudoknot residues were involved in competing interactions with other regions of TER (46). Based on SHAPE data, a model for this alternative secondary structure was proposed, in which residues from the pseudoknot region are base paired with residues from the template and TRE (47). Importantly, both SHAPE and FRET studies show that while tetPK does not form in free TER, it does form in the presence of TERT.

Recently, a structural model of the *Tetrahymena* telomerase holoenzyme was generated using a 9-Å resolution cryo-electron microscopy (cryo-EM) map and fitting with NMR and X-ray crystal domain structures and homology models of proteins and RNA (12). This study revealed new subunits and interactions, including two previously unidentified proteins: telomeric repeat binding (Teb)2 and Teb3, which interact with Teb1 to form TEB, a replication protein A (RPA) paralog, that enhances processivity through DNA binding. Furthermore, fitting of high-resolution structures of domains of TER, TERT and p65 into the cryo-EM map provided a detailed model of the RNP catalytic core. The TERT TRBD-RT-CTE forms a ring which is encircled by the TER t/PK. The template traverses the RT domain, while the PK is distal from the active site, on the other side of the TERT ring from the template near the CTE. p65 binds and bends Stem 4, and Loop 4 is inserted at the interface between TRBD and CTE, on the opposite side of TERT from the PK.

We determined the solution NMR structure of the *Tetrahymena* telomerase pseudoknot and used it in modeling the path of TER in our 9 Å cryo-EM map of the telomerase holoenzyme (12). Here we used NMR to study the structure and folding of the pseudoknot as an isolated domain and in the context of the full length t/PK core. We present details of the NMR data analysis and describe the solution NMR structure of the isolated pseudoknot, which we have further refined here. The isolated pseudoknot folds into a structure that has stacked stems and loops interacting in the respective major and minor grooves. The tetPK tertiary structure is more compact than the human and yeast pseudoknots, with only two stable major groove triples, an expected U•A-U and a novel C•G-A$^+$ triple with a protonated A and *syn* G, and two minor groove triples. Telomerase activity assays and NMR data confirm the importance of the C•G-A$^+$ triple for pseudoknot stability and telomerase activity. NMR data shows that the pseudoknot does not form in full length TER due to the competing interactions between the template-TRE and pseudoknot residues to form an alternative stem. Based on the determined secondary structure of the t/PK and the cryo-EM structure (12) a model for assembly of the catalytic core is proposed.

MATERIALS AND METHODS

RNA preparation for NMR studies

TetPK (and mutants tetPK-C75A, -C75U, -A80C, -G95U), Stem 1 (nts 5–12:100–107), Stem 2 (nts 15–40), Template-stem (nts 43–53:88–100), TRE-stem (nts 51–88), tet51-58:82–88, Stem 3alt (nts 42–100), t/PK (nts 5–107) and TER (nts 1–159) RNAs (see Figures 1 and 4, and Supplementary Figure S7 for sequences) were made as follows for NMR studies. RNAs were synthesized by *in vitro* transcription, using T7 RNA polymerase, from a synthetic DNA template as previously described (48,49). The DNA templates for t/PK and TER were cloned into the ApaI (5′) and XhoI (3′) sites of pRSF-1 vector. After linearization with XhoI enzyme, 0.5–1.0 mg of DNA was used for *in vitro* transcription with T7 RNA polymerase. The tetPK, t/PK and TER templates contain a minimal consensus hammerhead ribozyme sequence at the 3′ end to allow self-cleavage of each RNA molecule at a precise position (50,51). Briefly, purified T7 RNA polymerase (P266L mutant) (52) is added to a reaction containing 25–40 mM MgCl$_2$, 4–6 mM of each NTP, and 1 μM DNA template in 10 mM Tris–HCl pH 8, 1 mM spermidine, 2.5 mM DTT and 0.01% Triton. For tetPK and TRE-stem, uniformly ^{13}C,^{15}N labeled rNTPs were appropriately substituted in the reaction to make ^{13}C,^{15}N-(A,U) and ^{13}C,^{15}N-(G,C) double labeled samples. For tetPK, a uniformly ^{13}C,^{15}N-labeled sample was also made for RDC measurements. The reactions were incubated at 37°C for 3–4 h for transcription, and an additional 3–4 h for cleavage of hammerhead ribozyme constructs. The RNAs were subsequently purified by electrophoresis on a 10–15% denaturing polyacrylamide gel followed by electroelution of the excised RNA bands (53). The RNAs were washed, using Amicon centrifugal filters, once with high salt (1.5 M KCl), three times with water, and then
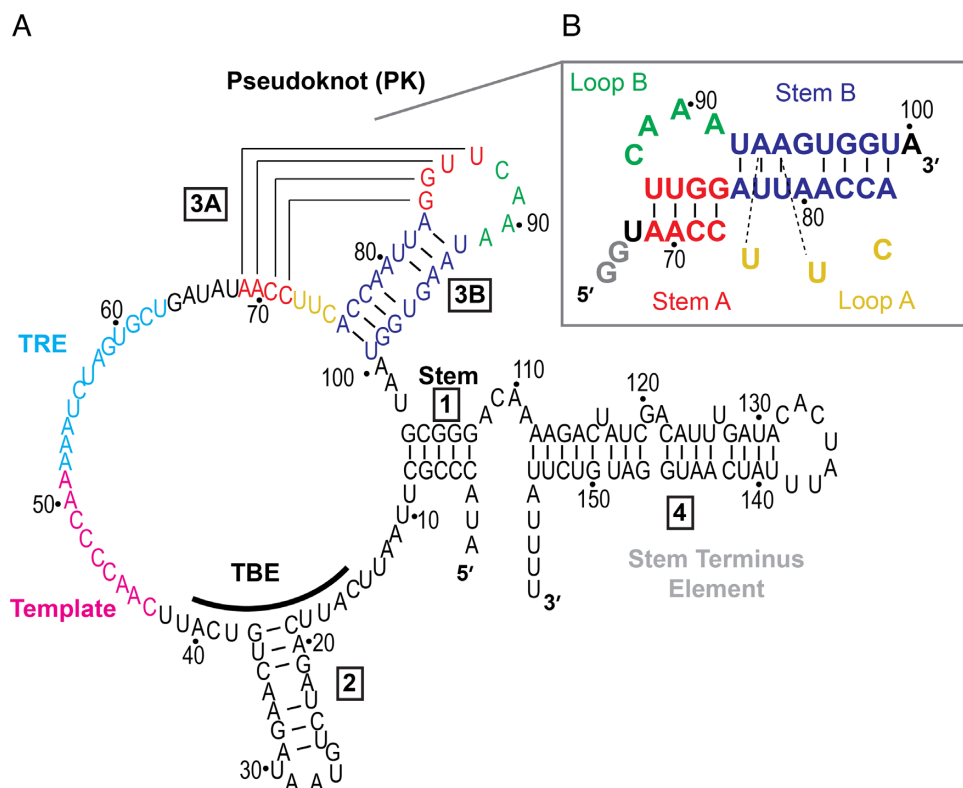
Figure 1. (**A**) Sequence and predicted secondary structure of *Tetrahymena* TER, with base paired helices labeled 1–4 (boxed). 3A and 3B form the pseudoknot (PK). Template (magenta), template boundary element (TBE, black), template recognition element (TRE, cyan), stem terminus element (STE, gray) and PK are labeled. (**B**) Sequence and predicted secondary structure of tetPK. The secondary elements are colored as Stem A (red), Loop A (gold), Stem B (blue) and Loop B (green). Two G's (gray) were added to enhance *in vitro* transcription.

NMR buffer (10 mM NaPO$_4$ pH 6.3, 50 mM KCl). The RNA solution was heated under dilute conditions (10–50 μM) at 95°C for 4 min and then snap cooled on ice. The RNAs were then concentrated to 0.5–1 mM for NMR studies.

NMR spectroscopy

The solution NMR structure of tetPK was previously reported (PDB 2N6Q) (12). Below, we describe the details for NMR assignments, dihedral angle restraints, and RDCs for structure calculation, and NMR assignments for tetPK and other RNAs used in this study. NMR spectra were recorded on Bruker DRX 500 MHz and 600 MHz and Avance 800MHz spectrometers equipped with HCN cryoprobes. Exchangeable proton resonances for RNA constructs were assigned at 283 K with samples in 90%H$_2$O/10%D$_2$O, using 1D and 2D NOESY spectra in H$_2$O, and additionally ^{1}H–^{15}N HSQC and JNN-COSY spectra for tetPK. Non-exchangeable proton resonances for tetPK and TRE-stem were assigned with samples in 100% D$_2$O, using 2D NOESY, 2D TOCSY, ^{1}H–^{13}C HSQC, 2D HCCH-COSY, 3D HCCH-TOCSY, and a suite of filtered/edited NOESY (F1fF2f, F2f, F1fF2e, F1eF2e) experiments at 283K (49,54–56) on ^{13}C,^{15}N-A,U and -G,C labeled samples. Proton attached ^{13}C and ^{15}N resonances were assigned with the HSQC experiments. The proton assignment strategy used sequential NOE connectivity as previously described

(21,27,56). NOESY experiments were acquired with 100, 200 and 300 ms mixing times to determine which mixing time best approximated a linear relationship between NOE crosspeak volume and inter-proton distance. For use in tetPK structure calculations, hydrogen bonds for Watson-Crick (WC) and Hoogsteen base pairs were confirmed by JNN-COSY (57,58). The imino proton resonances of the terminal base pairs of tetPK (A69-U87, A76-U99) were not detected due to rapid exchange with water, but these base pairs were confirmed by NOEs (in D$_2$O NOESY spectra). Non WC base pairs (A80-G95, A91-C72, A90-C71) were also confirmed by indicative NOEs (in D$_2$O NOESY spectra).

RDCs for tetPK were measured for C–H ($^1D_{C1'H1'}$, $^1D_{C2H2}$, $^1D_{C5H5}$, $^1D_{C6H6}$, $^1D_{C8H8}$) and N–H ($^1D_{N1H1}$, $^1D_{N3H3}$) using the uniformly ^{13}C,^{15}N-labeled sample on 800 MHz spectrometer at 283 K (59). The spectra were processed with Bruker TOPSPIN and analyzed with NMR-Draw. A total of 69 C–H and 10 N–H RDCs were determined by measuring the difference in J-coupling in the absence and presence of 15 mg/ml Pf1 phage.

New structure calculations using Xplor-NIH 2.42

For structure calculations, the NOE cross-peaks were integrated to generate distance restraints using the pyrimidine H5–H6 crosspeak as an internal standard (2.45 Å), using 200 ms mixing time NOESY spectra. NOEs were

categorized as very strong (2.5 Å), strong (3.5 Å), moderate (4.5 Å), weak (5.5 Å) or very weak (6.5 Å), with a range of ±1 Å (except 'very strong' was limited to van der Waals lower bound, 1.8 Å). Loose A-form dihedral angle restraints were used for the stem residues ($\alpha = -62° \pm 30°$, $\beta = -179° \pm 30°$, $\gamma = 47° \pm 30°$, $\zeta = -73° \pm 30°$, $\nu = 37° \pm 30°$). The ribose sugar pucker, correlated to δ angle, was determined based on the H1′-H2′ crosspeak intensity in the DQF-COSY: strong (C2′ endo, $\delta = 145° \pm 30°$), intermediate (C2′ endo, $\delta = 120° \pm 30°$), or no cross peak (C3′ endo, $\delta = 82° \pm 30°$). *Syn/anti* configuration is correlated to χ, and determined based on characteristic NOE pattern, where strong H8–H1′ crosspeak intensity indicates a *syn* conformation. Only G95 was *syn* ($\chi = 25° \pm 30°$), and all other residues were set as anti ($\chi = -160° \pm 30°$). Hydrogen bond restraints were used to restrain all experimentally determined base pairs, with two distance restraints per hydrogen bond.

The previously reported structure was refined with Xplor-NIH version 2.9.8 (60). The structure was recalculated here using Xplor-NIH 2.42, with a modified van der Waal radius of 1.10 (up from 0.9) (61). An initial 100 structures were calculated starting from an extended single strand, using NOE, hydrogen bond, weak planarity (weight = 300.0 for individual base and weight = 6.0 for base pairs), and dihedral angle restraints (21,60). This was followed by a second round of NOE refinement with a lower starting temperature (1000 K) and more cooling steps (40 000). The next step of refinement incorporates the 79 RDCs, in which a grid search produced optimal values for the axial (D_a) and rhombic (D_r) components of the alignment tensor: $D_a = -39.0$, $D_r = 0.13$ (62). In the last step, the database potentials are applied. The structural statistics for the lowest 10 (out of 100) energy structures are given in Supplementary Table S1. The structures were viewed and analyzed with MOLMOL, PYMOL and MolProbity (63).

Telomerase activity assays

Telomerase activity was measured by primer extension assay as previously described (64). TERT was *in vitro* translated in rabbit reticulocyte lysate (RRL) using the TNT transcription/translation kit (Promega). The 10 μl translation reaction contained 0.4 μl of PCR enhancer (0.5 M KCl, 12.5 mM Mg(OAc)$_2$), 0.2 μl of 1 mM methionine, 8 μl RRL mix and 150 ng TERT DNA plasmid (pCITE-TERT) (gift from Kathy Collins), and was incubated at 30°C for 1 h. The DNA templates of TER and variants TER-C75A, -C75U, -A80C, -G95U, -U73C, -U73-bp (A93G/U82C), -U73-T (U73C/A93G/U82C), -U74C, -U74-bp (A94G/U81C), -U74-T (U74C/A94G/U81C), -A91G and -U74C/C75U were made and the RNAs were transcribed and purified as described above, except they were exchanged into water instead of buffer for use in activity assays. 0.5–1.0 μM of the purified RNA was added to the RRL translated TERT (1 μl volume RNA per 10 μl RRL reaction) and incubated at 30°C for 30 min for RNP reconstitution. p65 was purified as described (65), and was added to 1 μM during RNP reconstitution when necessary.

A 20 μl telomerase primer extension reaction contains 50 mM Tris–HCl pH 7.0 or 8.3, 1.25 mM MgCl$_2$, 1 mM TCEP, 1 μM primer ((GT$_2$G$_3$)$_3$), 100 μM TTP, 9 μM nonradioactive dGTP, 0.4 μl of (α-^{32}P)dGTP at 3000 Ci/mmol and 6 μl of the reconstituted RNP. The reaction was incubated at 30°C for 60 min and terminated by addition of quench buffer (1 mM Tris–HCl, pH 8.0, 0.5 M EDTA) containing a 15-mer (α-^{32}P)5′-end labeled RNA recovery control (RC). The nucleic acid products and control were recovered by phenol/chloroform extraction and ethanol precipitation, and loaded on a 10% polyacrylamide sequencing gel (19:1 polyacrylamide, 7 M urea, 1× TBE buffer). The gel was electrophoresed at 50 W for 1 h, dried and exposed overnight to a phosphor imaging screen. The screen was scanned and analyzed with QuantumOne Software. The relative activity was determined by normalizing the integrated density of each lane relative to the RC, and comparing to the WT (as 100%).

Model of a step in catalytic core assembly

For help in visualizing the proposed assembly pathway (Figure 7) a model of an early step in assembly of the catalytic core was built as follows. *Tetrahymena* TERT (TEN, TRBD, RT and CTE), SL4 and p65 were modeled as described in (12), and positioned using Chimera. SL1 and Stem 3alt were generated by the online server RNAComposer (66). The structure and relative position of SL2 and TRBD were obtained from a recent crystal structure (PDB 5C9H) (67). To generate the full-length TER model, the remaining single-stranded regions of TER connecting the above subdomains (SL1, SL2, PK and SL4) were modeled by Pymol and Coot (68), and standard bond angles and lengths of backbone of these nucleotides were optimized by using the 'Regularize Zone' tool in Coot.

RESULTS

Folding of the *Tetrahymena* telomerase RNA pseudoknot

For structural studies of the isolated pseudoknot, an RNA construct was designed from the WT TER sequence (nt 69–100) with an additional two G's added at the 5′ end to enhance *in vitro* transcription (Figure 1B). The *in vitro* transcribed RNA includes a 3′ hammerhead ribozyme sequence which self-cleaves the RNA product precisely at A100, ensuring 3′ end homogeneity. TetPK has low stability compared to the TER pseudoknots of human and *K. lactis*, likely due to its short Stem A and Loop A. 1D imino and 2D TOCSY (H5–H6) data show that the pseudoknot begins to unfold (or pre-melt) at temperatures greater than 10°C or if the pH is increased to 8.0 (Supplementary Figure S1). The pseudoknot is in equilibrium with alternate conformations, which can include the hairpins of each stem and single-stranded RNA, as seen to a lesser extent for human TER PK, which is in equilibrium with a small amount of a Stem 1 (P2b; equivalent to Stem A in *Tetrahymena*) hairpin at 25°C (27). The alternate conformations of tetPK are evidenced by the appearance of additional TOCSY (H5–H6) crosspeaks of increasing intensity as the temperature is increased >10°C or the pH above 6.3 (with 10 mM sodium phosphate and 50 mM KCl) (Supplementary Figure S1A). The additional crosspeaks arise from separate unique structure(s) that are in slow exchange with the folded pseudo-

knot. As temperature and pH increase, the alternate conformations increase in population. Additional Mg^{2+} or KCl does not have a significant effect on this equilibrium. Increasing salt (50–200 mM KCl and 0–5 mM Mg^{2+}) also did not affect the pseudoknot fold, as evidenced by insignificant changes in the 1D imino spectra (Supplementary Figure S1B). For these reasons, the tetPK structure was solved at 10°C and pH 6.3, with 10 mM sodium phosphate and 50 mM KCl.

The base pairing of tetPK was determined by analyzing the imino region of 2D NOESY spectra and JNN-COSY spectra, which detects hydrogen bonds ([57,58]). The stems of the pseudoknot form the predicted Watson-Crick base pairs (shown in Figure 1B), with the expected NOE connectivities, including an NOE crosspeak between Stem A and Stem B residues (G84H1 to U92H3, respectively) indicating stacking of the two stems. In Stem B, A80 and G95 form an unusual base pair. G95 is in the *syn* conformation and its Hoogsteen edge faces the Watson-Crick edge of A80. A80N1 is protonated and hydrogen bonds with G95N7 to form a cWH base pair (Neocles and Westhof nomenclature ([69])). Although the A80 imino resonance was not observed (even at pH 5.0), there is substantial evidence that the A80N1 is protonated. N1 protonation causes the associated amino protons to shift significantly downfield and split further apart ([70]). This is observed with the A80 amino protons of tetPK (Supplementary Figure S2B). Similar chemical shift changes were observed for the aminos of the protonated cytosine in the C^+•G-C base triple formed in the *K. lactis* telomerase RNA pseudoknot ([29,71]). In addition, A80C2 is upfield shifted, another characteristic of N1 adenine protonation (Supplementary Figure S2A) ([72]). G95 was identified as *syn* by the strong NOE between its H8 and H1′, which is stronger than its H8 to H2′ NOE. In addition, there is an NOE between A80H2 and G95H8, which is a strong indication of the synG-A^+ base pair. This type of *syn*G-A^+ pair was previously observed in the crystal structure of an RNA 16-mer duplex with G-A mismatches ([73]). Of the two predicted U•A-U base triples (U73•A93–U82 and U74•A94–U81), only U74•A94–U81 (Figure 2D) could be directly confirmed by JNN-COSY. Surprisingly, the loop C75 interacts with the G-A base pair to form a unique C•G-A^+ base triple (C75•*syn*G95-A^+80) (Figure 2E) (cWHcWW nomenclature in ([69])). The *syn*G95 has its Watson-Crick edge in the major groove and can thus form a canonical Watson-Crick base pair with the loop C75.

Solution NMR structure of tetPK reveals a compact fold

The solution NMR structure of tetPK was previously calculated by simulated annealing ([60]) using Xplor-NIH version 2.9.8 ([12]). Here, we recalculated the structure using the updated Xplor-NIH version 2.42, with the previously reported restraints (414 NOE, 171 dihedral angle and 79 RDC restraints). The new calculation improved the RMSD (between lowest 10 energy structures) from 0.83 Å to 0.42 Å, resulting in a better defined structure (Figure 2B and Supplementary Table S1). The new structure is slightly more extended than the previous one, with an RMSD between lowest energy structures of the old and new calculations of 1.5 Å. There were improvements in steric contacts, confor-

mation and minimal energy (better fit to NOE and dihedral data, Supplementary Table S1).

In tetPK, the two stems form a quasi-continuous A-form helix, stacked on top of each other without any significant bend. Loops A and B bind in the major groove (of Stem B) and minor groove (of Stem A), respectively. The pseudoknot is compact, with a major groove width of 11.4 Å (defined as the phosphate distance between i and i+6 cross-strand residues). Loop A and Stem B interact to form a small triplex, composed of consecutive U•A-U and C•G-A^+ triples (discussed above). A second U•A-U triple (U73•A93–U82) was predicted based on phylogenetic analysis ([45]). The U73 is positioned in the major groove, but in most of the structures it is not within hydrogen bond distance for a Hoogsteen base pair to A83. The U73 imino resonance was not observed in any NMR spectra. In addition, the non-exchangeable resonances for U73 exhibited broader linewidths, indicative of dynamics, and the NOE pattern expected for this base triple, in which U73 would have crosspeaks to U92 and A93 (i.e. U73-H1′ to U92-H8, -H1′, -H2′ and U73-H1′ to A93-H8), was not present ([15]). We conclude that this base triple is conformationally dynamic in the free tetPK.

The adenine-rich Loop B interacts with the minor groove of Stem A of tetPK, to form A91•C72-G84 and A90•C71-G85 triples. An adenine-rich Loop B is common in vertebrate TER pseudoknots, as well as in ribosomal frameshifting viral pseudoknots ([74,75]). In these RNAs, the loop adenines form minor groove triples involving sugar 2′-OH and base protons ([15,75]). In tetPK, A90 and A91 are stacked on top of each other and pair with with C71 and C72, respectively, to form identical A•C-G triples. The adenine amino protons form two hydrogen bonds with the cytosine carbonyl and 2′-OH (shown in Figure 2F). This interaction is consistent with *in vivo* footprinting data which indicated that the Loop B—CAAA residues were protected (from dimethyl sulfate modification) ([76]).The human TER pseudoknot forms similar minor groove triples with its Stem 1 and Loop 2 adenine residues, further signifying the importance of these types of minor groove interactions on pseudoknot structure ([15]).

Pseudoknot folding is important for telomerase activity

The loop-stem tertiary interactions of TER pseudoknots have previously been shown to be important for telomerase activity in humans and yeast (*in vitro* and *in vivo*) ([14,15,29,31]). By direct primer extension assay, we tested the potential importance of tertiary interactions found in the tetPK structure. First, we examined in detail the novel C75•G95-A^+80 triple. In full-length TER, we made C75A, C75U, G95U and A80C mutations. TER variants were then reconstituted with TERT, which had been *in vitro* transcribed/translated in rabbit reticulocyte lysate (RRL), and assayed for telomerase activity (Figure 3B, Supplementary Figure S3). Since tetPK folding and stability is pH dependent due to the protonated A in the C•G-A^+ triple, these and other activity assays were done at pH 7.0 in addition to the standard pH 8.3 ([64,65]).

For the substitutions in the C75•G95-A^+80 base triple, TER-C75A has the biggest effect on telomerase activity, de-
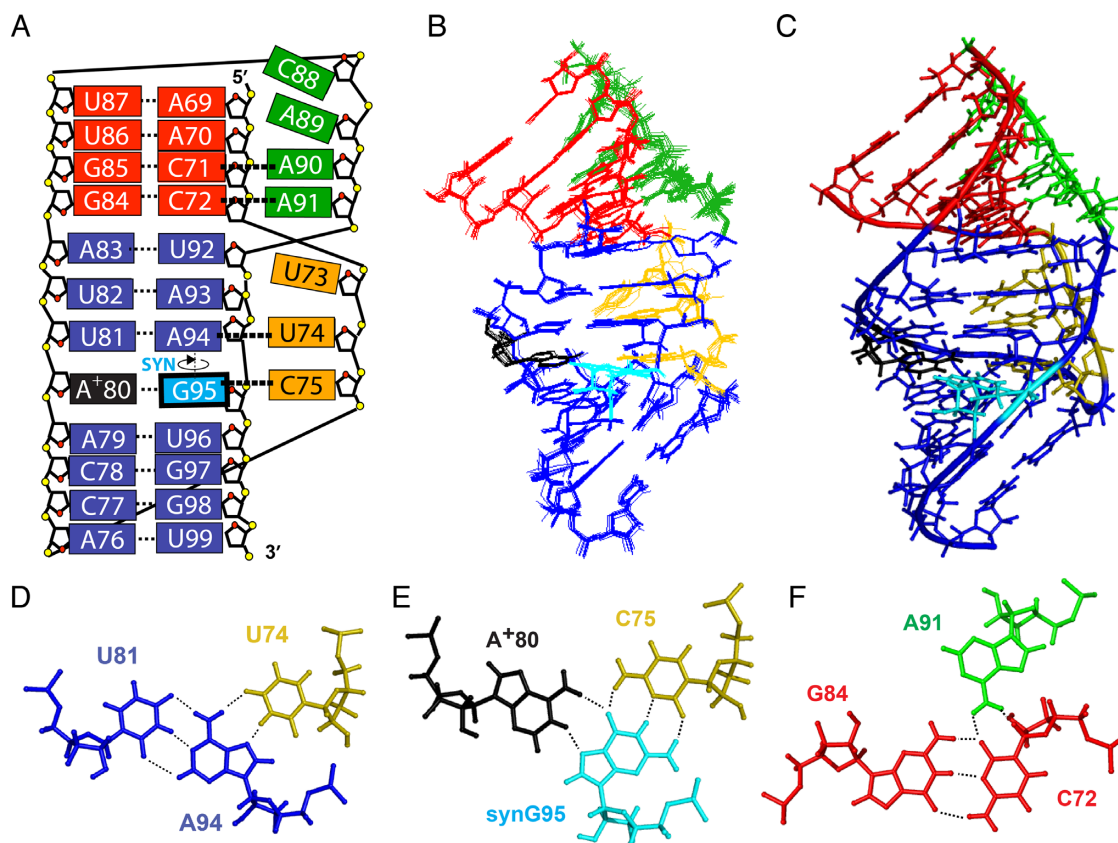
Figure 2. (**A**) Schematic of determined secondary structure of tetPK. Solid lines represent backbone connectivity, and dotted lines represent hydrogen bonds. (**B**) Family of lowest 10 energy NMR structures. (**C**) Lowest energy NMR structure. (**D**) U81-A94-U74 major groove triple, (**E**) A80–G95–C75 major groove triple, and (**F**) G84–C72–A91 minor groove triple. Hydrogen bonds are shown as dotted lines. Pseudoknot elements are colored as in Figure 1, except A80 is black and G95 is cyan, for clarity when viewing structures.

creasing activity by ∼50% compared to WT. This substitution would abolish the tertiary interaction (loop-stem C75–G95 bp) that stabilizes the unusual Stem B synG95-A$^+$80 bp. TER-G95U and TER-80C, which would be expected to replace the synG95-A$^+$80 bp with a U–A and G–C base pair, respectively, have 92% and 78% activity, respectively. For both of these substitutions, an alternative triple (C•U-A and C$^+$•G-C, respectively) could form with the loop nucleotide. TER-C75U has 78% of WT activity; U75 could pair with G95 similar to C75. Interestingly, decreasing pH to 7.0 from the usual activity assay conditions of pH 8.3 increases activity for each TER variant (13–37%) This increase is the smallest for the TER variants whose tertiary interactions are not pH dependent (C75A, A80C and C95U) and largest for WT and C75U (Figure 3B, Supplementary Figure S3).

To investigate the effect of these mutations on pseudoknot structure and stability, the C•G-A$^+$ base triple substitutions were made in the context of the isolated tetPK, and 1D and 2D imino NMR spectra were acquired to provide secondary structure information (Figure 3D, Supplementary Figure S4). Comparison of these spectra with that of WT reveals that only tetPK-C75A does not form a folded pseudoknot structure. The C75A substitution disrupts formation of Stem A, as indicated by the near disappearance of Stem A U86, G85 and G84 imino resonances (Figure

3D, Supplementary Figure S4). This explains why this TER variant has the lowest activity. The tetPK substitutions with higher activity, tetPK-C75U, tetPK-G95U and tetPK-A80C, each form a pseudoknot with Stem A and Stem B as shown in the 1D and 2D spectra, although based on the increase in imino proton linewidths all are somewhat less stable than the WT tetPK (Figure 3D, Supplementary Figure S4). As discussed above, tetPK-C75U and tetPK-A80C can potentially form U•G-A$^+$ and C$^+$•G-C base triples respectively. There is evidence for a U•G-A$^+$ base triple in tetPK-C75U, which folds into a pseudoknot with similar imino resonance chemical shifts and NOE crosspeak patterns as WT tetPK, with the same protonated adenine amino resonances (Figure 3D, Supplementary Figure S4). We were unable to confirm whether a C$^+$•G-C base triple forms in tetPK-A80C. TetPK-G95U replaces the relatively unstable synG95-A$^+$80 base pair with a Watson–Crick U-A base pair, where U95 could form a possible tertiary interaction with C75. In summary, all of the nucleotide substitutions for the C•G-A$^+$ triple except for C75A had only modest effects on activity, since an alternative stem base pair and/or loop interaction could form. The C75A mutation strongly destabilized the pseudoknot and showed a concomitant decrease in activity.

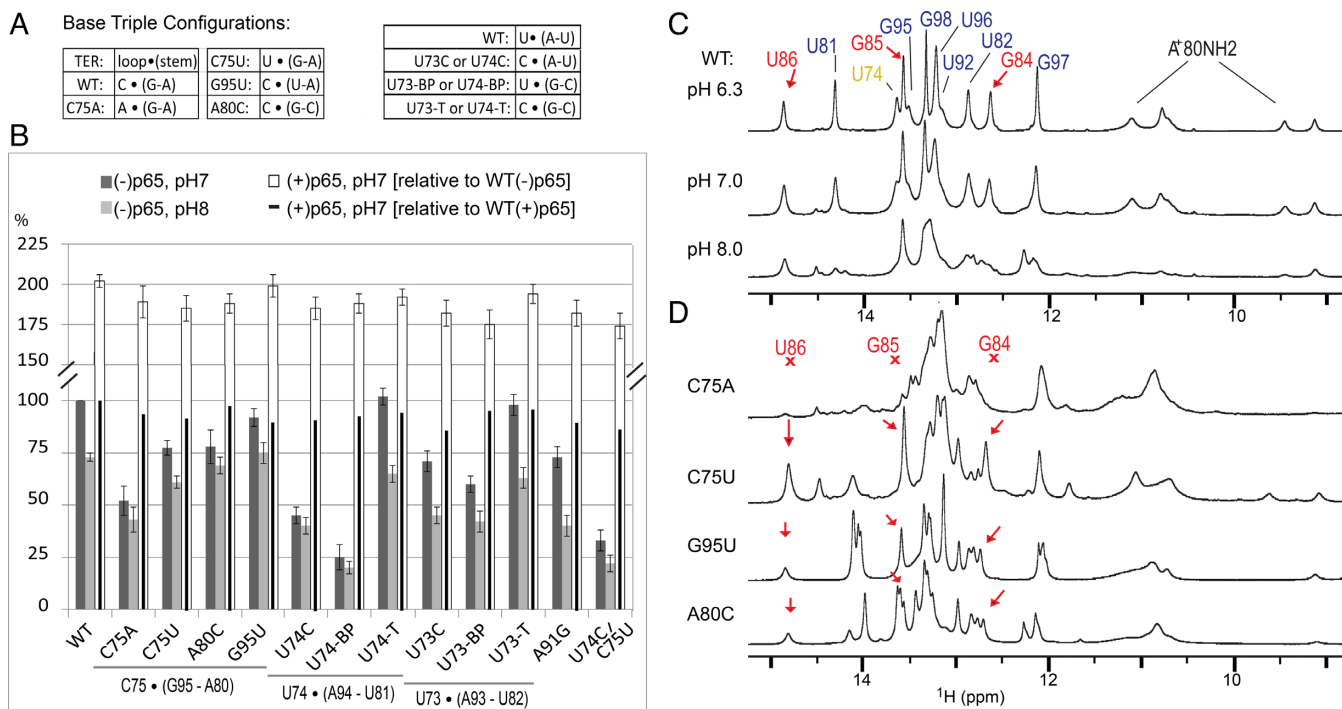We next examined the effects of nucleotide substitutions in the U74•A94–U81 and predicted U73•A93–U82 triples

Figure 3. (**A**) TER substitutions with associated stem–loop base triple configuration. (**B**) Relative telomerase activity of WT and mutant TER constructs at pH 7 (dark gray), pH 8 (light gray), and with p65 at pH 7 (white). The density of each lane was integrated from the assay images (in Supplementary Figure S3), and normalized to WT at pH7 (as 100%). For (+) p65, the density was also normalized relative to WT(+)p65 (black line). Error bars are from duplicate experiments. (**C**) 1D imino spectra of WT tetPK at pH 6.3, 7.0, and 8.0. Imino resonances are labeled and colored according to secondary structure elements as in Figure 1. (**D**) 1D imino spectra at pH 6.3 of tetPK variants C75A, C75U, G95U, A80C that change the A80-G95-C75 base triple. Stem 3A iminos are present (shown with arrows) in all of the constructs except tetPK-C75A (indicated by an 'X').

on telomerase activity. For each set of triples, base pair (AU to G–C, referred to as -bp), loop (U to C), and compensatory base triple (U•A-U to C$^+$•G-C, referred to as -T) substitutions were made. The compensatory triple requires protonation at CN3, so is also expected to be pH dependent. For the U74•A94–U81 triple, TER-U74C and TER-U74-bp (A94G/U81C) substitutions which disrupt the base triple have activity levels of 45% and 25% respectively, while the compensatory mutant U74C/A94G/U81C (TER-U74-T) restores activity to 102%. Replacing U•A-U with C$^+$•G-C base triples has previously been shown to (partially) restore activity in the human and yeast pseudoknots (15,29). For the predicted U73•A93–U82 triple, TER-U73C and TER-U73-bp (A93G/U82C) have a smaller but significant decrease in activity levels to 71% and 60% respectively, and TER-U73-T (U73C/A93G/U82C) restores activity to 98%. The activity assay results on substitutions in the predicted U73•A93–U82 triple indicate that at least in the context of assembly with TERT this triple is formed although perhaps without optimal hydrogen bond geometry. Finally, we tested the effect of disrupting the consecutive U•A–U and C•G–A$^+$ triples by changing the loop nucleotides (TER-U74C/C75U). This decreases activity to 33%, compared to 45% for TER-U74C and 78% for TER-C75U.

Last, we tested the importance of the minor groove interaction between Loop B A91 and Stem A base pair G84–C72. Substitution of A91G, which should abolish the tertiary interaction of A91 with the relatively unstable Stem A, decreases activity to 73%. As seen for the C•G–A$^+$ base

triple substitutions, all TER variants showed an increase in activity at pH 7 compared to pH 8.3, with most showing an increase of more than 50%.

Taken together, the activity assays and NMR data provide a direct correlation between stable pseudoknot folding and telomerase activity levels. Disruption of the base triples by changing the stem base pair or loop nucleotide results in a decrease in activity. Not surprisingly, the largest decrease in activity for substitutions in a single triple is seen for the central U74C•A94–U81 triple.

Telomerase activity was also assayed in the presence of holoenzyme assembly protein p65. p65 has been shown to rescue a number of TER mutations that affect assembly, including mutations within the pseudoknot region (65). For all of the nucleotide substitutions tested here, p65 increased the activity to near WT levels (90–98% relative to WT with p65) (Figure 3B, Supplementary Figure S3). This is consistent with a role for the pseudoknot in assembly rather than catalysis, as proposed based on the location of the pseudoknot in the pseudoatomic model of the catalytic core in the cryo-EM map (12).

Secondary structure of TER t/PK

As discussed above, previous studies have shown that in the absence of TERT, the *Tetrahymena* pseudoknot does not form in the context of full length TER (46,47). Based on SHAPE data, an alternate stem–loop structure with base pairing between template-TRE and pseudoknot residues
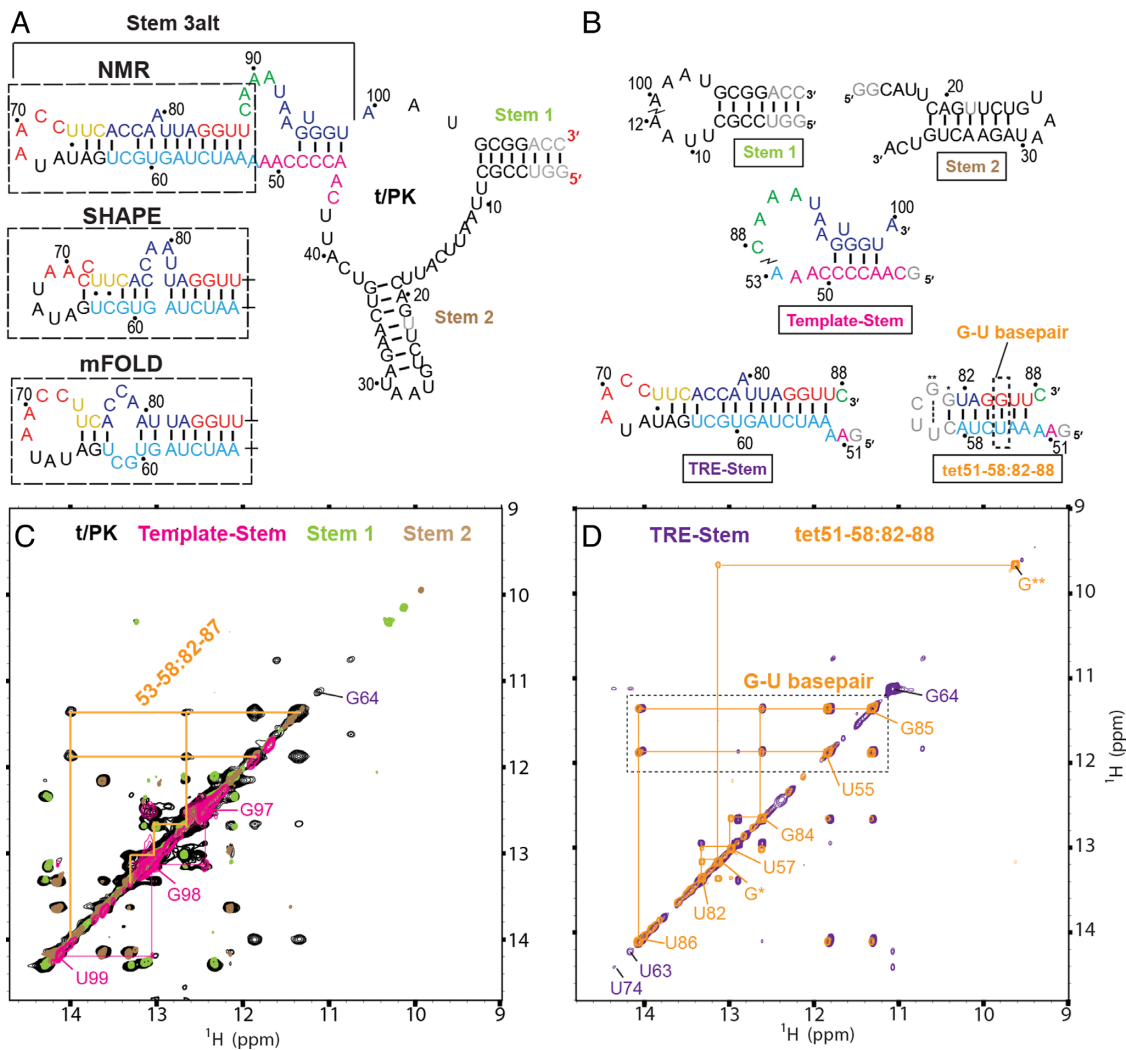
Figure 4. (**A**) Schematics and sequence of t/PK, with secondary structures determined by NMR and predicted by SHAPE and mFOLD for the alternative stem–loop 3 structure (boxed). (**B**) Sequences and determined secondary structures for RNA constructs (Stem 1, Stem 2, Template-Stem, TRE-stem and tet51-58:82–88) used for NMR studies to verify the t/PK secondary structure. Nucleotides are colored as in Figure 1, with gray representing non-native residues. (**C**) Overlay of 2D imino NOESY spectra from t/PK (black), template-stem (magenta), Stem 1 (lime), and Stem 2 (brown). NOE connectivities for t/PK corresponding to the template-stem (magenta lines) and stem 53–58:82–87 (orange lines) are shown on the right and left side of the diagonal respectively. (**D**) Overlay of 2D imino NOESY spectra from TRE-stem (purple) and tet51-58:82–88 (orange). NOE connectivities for tet51-58:82–88 are shown with orange lines. The cross peaks from the G–U base pair are boxed (dashed line).

was proposed for free TER (Figure 4A) (44,46,47). mFOLD predicted a similar alternate structure, but with a significant difference in the apical stem loop of the alternate stem–loop (nt 51–88) (Figure 4A). We therefore investigated the folding of free TER by NMR. The t/PK resonances in full length TER (nt 1–159) and t/PK alone (nt 5–107) have similar NOE crosspeak patterns, indicating that SL4 does not affect the structure of the t/PK (Supplementary Figure S5). Addition of MgCl$_2$ also had no apparent effect on the conformation of the t/PK (Supplementary Figure S5). We therefore focused our NMR studies on the t/PK alone (in 10mM phosphate, pH 6.3, 50 mM KCl), which has the best spectra quality. RNA constructs with sequences of Stem 1, Stem 2, and the predicted template-stem (nt 43–53:88–100) were also made for comparison of their NMR spectra to those of the t/PK (Figure 4A and B). Comparison of the chemical shifts and crosspeak patterns in 2D imino NOESY

spectra of these constructs with those of t/PK shows that Stem 1, Stem 2, and the template-stem form in the t/PK (Figure 4C). The lack of the tetPK NOE crosspeak pattern (spectra in Figure 4C compared to Supplementary Figure S7A) and the presence of the template-stem that sequesters pseudoknot residues, confirms that the pseudoknot does not form within the free t/PK. There are additional unassigned imino resonances in t/PK, which include a distinct G-U base pair pattern, which must arise from the alternative structure. Both the SHAPE and mFOLD models predict that the TRE is paired in an alternate stem–loop (TRE-stem); the exact secondary structures differ but both have an identical central stem (nt 53–58:82–87) that contains a G-U basepair. We made a construct of this central stem sequence capped with a UUCG tetraloop (tet51-58:82–88) for comparison to t/PK. Tet51-58:82–88 imino resonances were assigned, which confirmed that the predicted stem forms
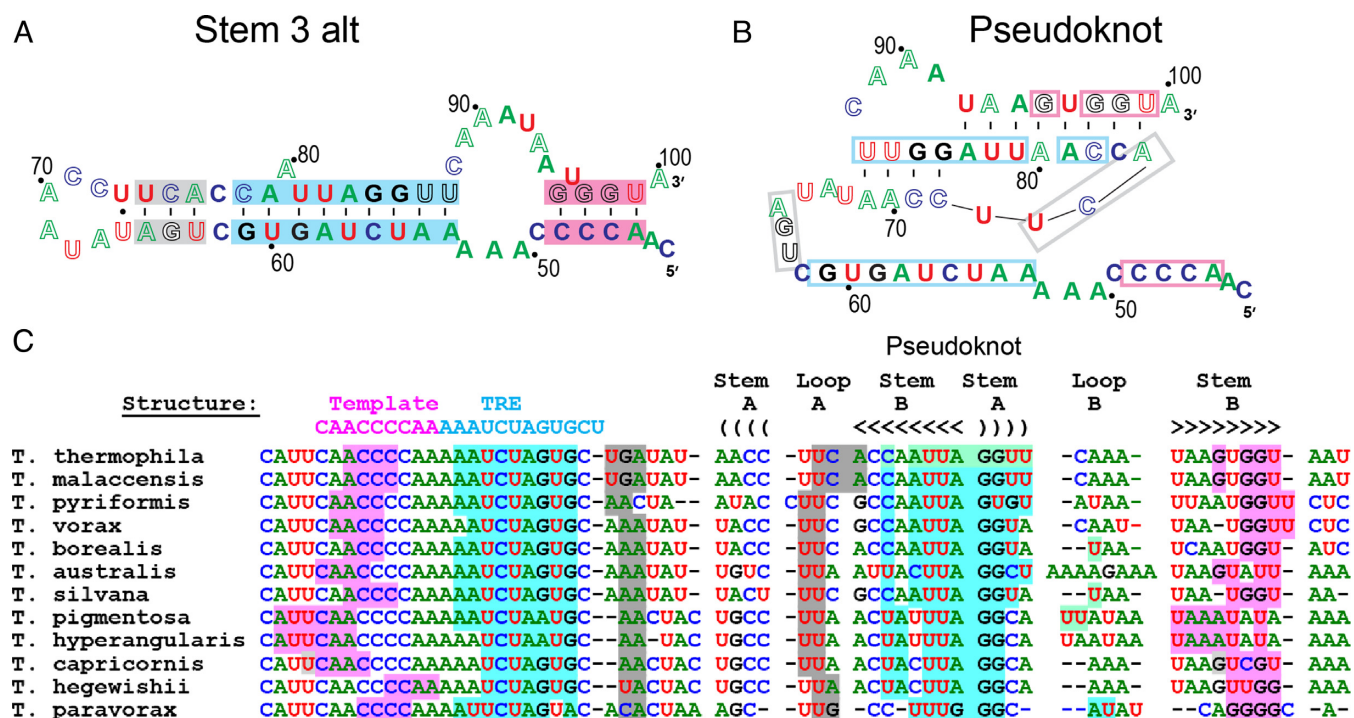
Figure 5. (**A**) NMR determined secondary structure of Stem 3alt with nucleotides colored according to identity: G (black), U (red), C (blue), A (green). Conserved residues (>10 out 12 *Tetrahymena* species) are in bold, while non-conserved are white-filled. Regions of Stem 3alt that are conserved to base pair are highlighted for the template, TRE and apical stems (magenta, cyan and gray respectively). (**B**) Pseudoknot colored and bolded as in (A), with residues that base-pair in the conserved Stem 3alt boxed. (**C**) Sequence alignment of template-TRE-PK region of *Tetrahymena* TERs with nucleotides and potential base-paired regions highlighted as in (A).

(Figure 4D, orange). Furthermore, comparison of the spectra of tet51-58:82–88 and t/PK confirmed that the central TRE-paired stem is present in the t/PK (Figure 4C, orange). Finally, we made an RNA containing residues 51–88 (TRE-stem) and sequential assignments were obtained from analysis of 2D imino NOESY (Figure 4D, purple) and D_2O NOESY spectra (Supplementary Figure S6). The AH2 proton is the only non-exchangeable resonance with cross-strand NOEs: (weak NOE) to the H1′ of its base-paired residue and (strong NOE) to the H1′ of its 3′ (or i – 1) cross-strand neighbor, depicted in Supplementary Figure S6. These NOEs were used to validate assignments and determine TRE-stem secondary structure, which is shown in Figures 4B and 5A. This secondary structure is a long helix, which contains a single bulge A(80) base, a U–U and C–C base pair, capped by a hexaloop. While the central stem (nt 53–58:82–87) is the same as the SHAPE and mFOLD models, the rest of the helix lacks the extensive bulges predicted by these two methods. Comparison of chemical shifts and NOE crosspeak patterns in the imino NOESY of TRE-stem with t/PK confirmed that the secondary structure of TRE-stem also forms in t/PK. However, there are a few additional broad imino resonances in t/PK that suggest that other alternate conformations may be present.

For the NMR studies described above, the RNA was heated and snap cooled to achieve a homogeneous well folded sample. However, it is possible that TER could fold into a different native structure during transcription. To test for this possibility, we also purified t/PK under nondenaturing conditions, using ion exchange and size exclu-

sion column chromatography. NMR spectra of t/PK purified under denaturing and non-denaturing conditions were essentially identical, indicating that co-transcriptional folding does not trap a different structure than the one obtained after heating and snap-cooling (Supplementary Figure S5). We also investigated a construct that spans the template, TRE and pseudoknot regions (nts 42–100; Stem 3alt), which comprises the residues that form the alternative stem–loop, but opens the t/PK circle. Interestingly, the NMR data shows that for this construct the pseudoknot is present in ~1:1 equilibrium with the alternate conformation (Supplementary Figure S7). Apparently the topological constraints of the closed TER t/PK circle also influence the structural equilibrium, disfavoring the pseudoknot. The identification of an alternate stem–loop structure, in which the template and the TRE are paired with residues from the pseudoknot sequence, explains why the pseudoknot does not form in the context of the full length TER in the absence of TERT. Phylogenetic analysis of the t/PK from 12 *Tetrahymena* species shows that each species can form comparable template- and TRE-stems (Figure 5). Taken together, these results suggest that not only the folded pseudoknot but also the Stem 3alt structure in the free TER may be biologically relevant.

DISCUSSION

Comparison of ciliate, human, and yeast pseudoknots

The *Tetrahymena* TER pseudoknot has a number of similarities and differences to those of human (hPK) and yeast
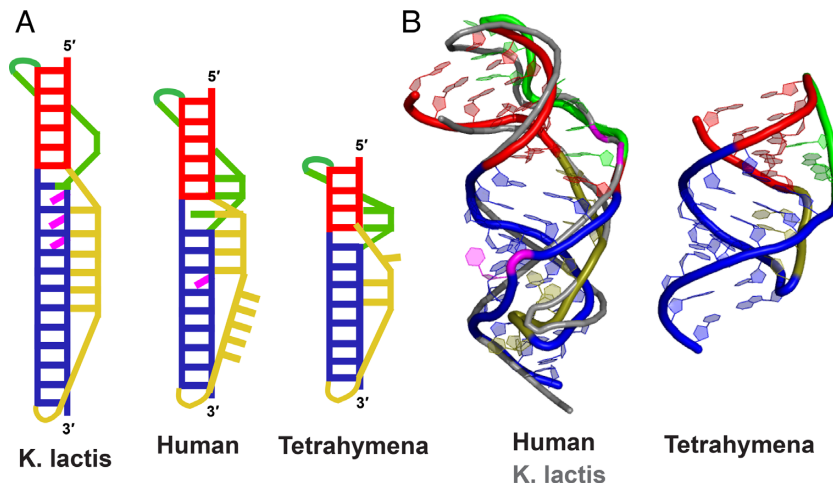
Figure 6. (**A**) Schematic of TER pseudoknot secondary structures from *K. lactis*, human and *Tetrahymena*. The Stem B:Loop A (Stem 2:Loop 1) nt numbers are given for each pseudoknot. (**B**) Solution NMR structures of TER pseudoknots from human (colored) and *K. lactis* (backbone only, gray), which are overlayed (left), and from *Tetrahymena* (right). The structures are aligned through the base triple region. Stems and loops are colored as in Figure 1.

K. lactis (kPK), which were previously solved by NMR (Figure 6) (15,21,29). All three pseudoknots have continuous helical stacking interactions through the stems and junction. Only hPK has an intervening loop-loop interaction at the junction, while tetPK and kPK have stem-stem stacking. All three pseudoknots have a stabilizing triple helix that includes major groove tertiary interactions, and both tetPK and hPK also have minor groove base triples. kPK cannot form the typical minor groove interactions due to the lack of adenines in its second loop, but has a much longer major groove triplex. The base triples were shown to be important for proper pseudoknot folding for all three pseudoknots. In both tetPK and kPK, mutations in the loop residues that disrupt major groove base triples abolished formation of Stem A (*Tetrahymena*) and the equivalent stem (Stem 1) in yeast (29). In hPK, mutations in loop residues that disrupt major groove base triples destabilized the base pairs in both stems near the junction (77). The stems of TER pseudoknots have varying degrees of stability depending on sequence, bulges, G-C content and length, but all require the base triple interactions to ensure proper folding.

TetPK has a significantly different tertiary structure when compared to hPK and kPK (Figure 6B). The backbone of hPK and kPK overlay very well and the structures appear remarkably similar despite differences in sequence and tertiary interactions. TetPK is more compact. This may be partly due to the different ratios in the number of nucleotides in Loop A (human J2b/3) to Stem B (human P3). hPK has a ratio of ~1:1, kPK is 1:2, and tetPK is closer to ~1:3 (Loop A: Stem B nt ratio) (Figure 6A). TetPK has eight base pairs in Stem B and only three Loop A nts, while hPK has nine stem base pairs and eight loop nts. This means that tetPK Loop A must span a greater distance per nucleotide and is consistent with a more compact molecule. These pseudoknots belong to different organisms with significantly different TER size, sequence and structure, and to a lesser extent different TERT sequence and structure, and therefore their similarities may reflect a conserved function while their differences reflect their diverse environments.

One study investigated a chimeric human TER, where hPK was replaced with tetPK, which could potentially compensate for function since the secondary structures of the two pseudoknots are similar (Figure 6A) (78). However TERT assembled with the chimeric TER was only 'weakly active' when assayed *in vitro*, supporting the conclusion that the pseudoknot differences reflect species-specific requirements.

Correlation between pseudoknot structure, p65, and telomerase activity

Previous studies of ciliate, yeast, and human telomerase RNA have shown that formation of the pseudoknot and its tertiary interactions are essential for activity in the context of the minimal telomerase RNP (i.e. TERT + TER only) (14,15,33,65,79–81). It has also been shown that for *Tetrahymena* p65 can rescue mutations that affect assembly, if they are not too severe, but not mutations that affect catalysis (42,65,82). In the cryo-EM model of the Tetrahymena telomerase holoenzyme, the pseudoknot is on the opposite side of the TERT ring from the template (active site), too far away to be directly involved in catalysis. Thus, it has been proposed that the pseudoknot has a role in assembly, with the tertiary interactions stabilizing the correct fold on TERT (12). Consistent with this, all of the pseudoknot substitutions reported here, which disrupt the U•A-U and C•G-A$^+$ base triples, were rescued by p65, in the context of assembly of TER with TERT in RRL.

Some early studies seemed to indicate that the pseudoknot was not essential for telomerase activity in *Tetrahymena,* since mutations which were presumed to abolish pseudoknot structure, such as deleting nt 86–89 or adding a 4nt (CAAU) bulge at position 81–82 did not significantly affect *in vitro* or *in vivo* activity (assayed in the presence of p65) (83). However, the structural analysis presented here and in yeast TER pseudoknots (29) suggest that the pseudoknot might still form with these nucleotide changes albeit less stably. TER-Δ86-89 would disrupt the terminal A-U base pairs of Stem A, but would leave the G–C base pairs at
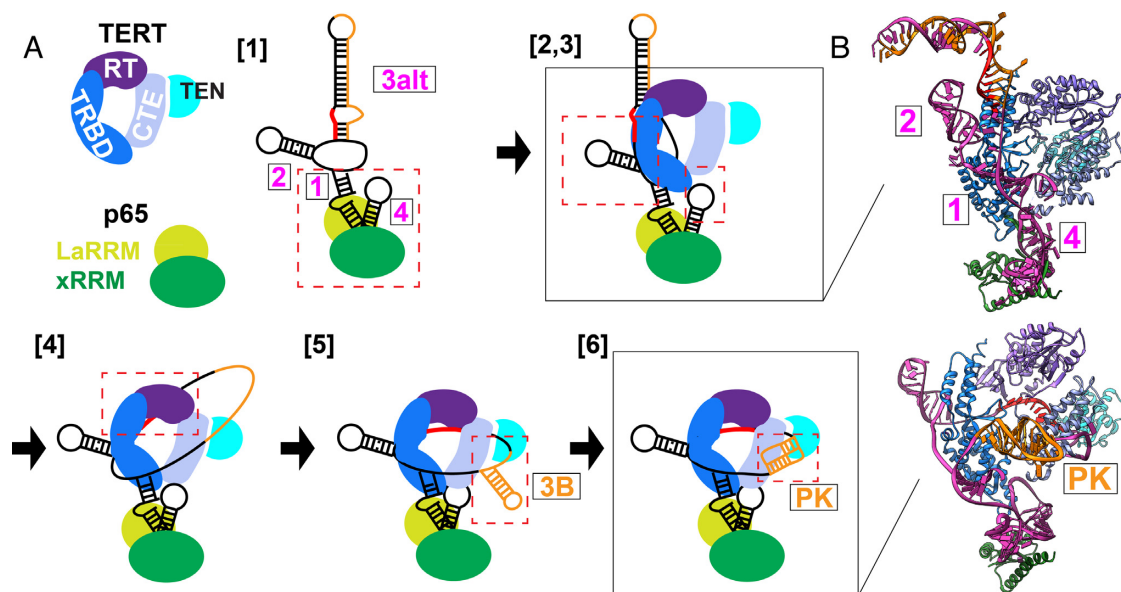
Figure 7. (**A**) (left) Schematics of TERT and p65. The TEN domain is behind the TERT ring and is connected to TRBD by a long linker. (1–6) Illustration of the proposed steps in p65-TER-TERT assembly: (1) p65 binds and bends Stem 4; (2, 3) Entering from the bottom of the TER circle (with respect to this view of TER, with SL4 on the bottom), the TERT-TRBD binds Stem 2, TRBD-CTE binds SL4, and the template-paired stem opens; (4) The template interacts with RT, and Stem 3alt fully opens as TERT enters the t/PK circle; (5) the pseudoknot Stem B forms; (6) the pseudoknot fully folds, acting like a watchband ratchet clasp. Domains are colored as follows: TERT TEN (cyan), TRBD (blue), RT (purple), CTE (silver); p65 LaRRM (lime), xRRM (green); TER (black), with template (red), pseudoknot residues (orange), and numbered TER stems boxed (magenta). Note that the position of the TEN domain is probably not fixed until after the TER circle goes over the TERT ring, since TER is between the TERT ring and the TEN domain. (**B**) (Top) Hypothetical model of the interaction of TERT with TER as Stem 3alt begins to open. Model was built as described in Materials and Methods. (Bottom) Pseudoatomic model of the catalytic core, from reference (12). Colors are as in A, except TER is magenta.

the junction and the tertiary interactions which stabilize the pseudoknot intact (14,21). Adding a bulge in Stem B would not necessarily abolish tetPK formation, since all secondary and tertiary interactions remain intact. A similar bulge is predicted in the *S. cerevisiae* pseudoknot (29,84). Interestingly, when pseudoknot formation was indeed abolished, by deleting Stem B, telomerase activity only decreased to 63% (85), in assays where p65 was presumably present. This deletion would prevent formation of Stem 3alt and would result in a smaller t/PK circle. Activity in this mutant might be explained by the assembly pathway postulated below, since deletion of Stem B would decrease the size of the t/PK circle, similar to that of the folded PK and allow assembly to proceed up to the last step.

The Stem 3alt and pseudoknot structures explain FRET and chemical probing data

The structure of Stem 3alt is completely consistent with the identification of flexible residues at the apical loop and near the bulge A80 in the free TER by SHAPE chemical probing (Supplementary Figure S8A) (47). The SHAPE data for the pseudoknot in the presence of TERT indicates that a significant portion of Stem B is flexible (Supplementary Figure S8A). The SHAPE data was acquired at pH 8, which we have shown destabilizes the pseudoknot due to the pH dependence of the tertiary $C \bullet G$–A^+ base triple, and in the absence of p65, which assembles more active complexes. These conditions could explain the observed 'flexibility' of Stem B, since some proportion of the complexes might have partially or misfolded pseudoknots. Telomerase activity assays

showed lower activity at pH 8 versus pH 7 and two-fold lower activity in the absence (versus presence) of p65.

Previous FRET studies showed that the pseudoknot does not form in the protein-free *Tetrahymena* TER and suggested that instead of a Stem 3B hairpin, an alternative structure forms (46). In the Stem 3alt structure the FRET dyes would be farther apart than they would be if a Stem B hairpin forms, consistent with the FRET results (Supplementary Figure S8A). FRET studies of the much larger human t/PK also indicated that the pseudoknot was not fully folded in the free TER (77). Examination of human TER sequence suggests that it can form a similar alternate structure with P2 flipping out to mimic the TRE-stem and the template base pairing with pseudoknot residues (Supplementary Figure S8B). This alternate structure, in which the FRET dye-labeled pseudoknot residues (C92 and U85) would be far from each other (Supplementary Figure S8B), would explain the FRET data on human TER and TERT (Supplementary Figure S8B) (77).

A model for assembly of the catalytic core

Although the *Tetrahymena* pseudoknot does not form in TERT-free TER, it is fully folded in the pseudoatomic model of *Tetrahymena* telomerase from the cryo-EM map (12). We note in this context that, in the absence of TERT, it was only by studying the isolated tetPK that the pseudoknot structure could be determined. The cryo-EM structure is of the apo enzyme, without telomeric DNA. Single-molecule FRET experiments where the conformation of the pseudoknot was monitored during catalysis showed that only RNP

molecules containing a properly folded pseudoknot were catalytically active and that the pseudoknot does not unfold during catalysis (46).

Human cells contain pools of free hTER and hTERT which assemble into the telomerase RNP as necessary (86). Telomerase RNP assembly begins with the co-transcriptional binding of species-specific accessory proteins, i.e. H/ACA proteins in vertebrates and p65 in ciliates (1,87), which bind TER regions outside the t/PK, and would not be expected to affect the folding of the (TERT-free) t/PK. The structure of free TER potentially plays an important role in telomerase biogenesis. The Stem 3alt structure in *Tetrahymena* t/PK has more base pairs and is more stable than the 'open circle' structure with a single-stranded template-TRE and PK. The double-helical RNA would better protect TER from degradation since single-stranded RNA is more likely to self-cleave due to the more accessible 2′OH groups. Base-paired RNA is also more protected from mutations, since single-stranded cytosines are deaminated (converted to U) at a much higher rate (88). This is especially critical in the template region, where mutations would be propagated to the telomere repeat (47).

Aside from stabilizing free TER, the alternate structure may play a role in assembly, given that it appears to be conserved to base pair. In the assembled catalytic core, the t/PK encircles the TERT ring approximately perpendicular to the plane of the ring. The free t/PK circle, with the Stem 3alt, is too small to allow entry of the TERT ring and also sequesters the template. However, the TBE is still exposed. Based on the secondary structure of the free t/PK, we can speculate on a step-wise pathway for p65-TER-TERT pathway (depicted in Figure 7): (Step 1) First, p65 binds to two sites in SL4: the p65 La-RNA recognition motif (LaRRM, or La module) binds the 3′ poly U tail (40,42) and p65 C-terminal xRRM binds Stem 4, bending it so that Loop 4 is closer to SL2 (40–43). (Steps 2, 3) Next, the TRBD binds the high affinity binding site TBE, which is still accessible in the Stem 3alt structure. Binding of the TBE to TERT also positions Loop 4 (if bound to p65) close to the TRBD. Loop 4 interacts at the interface of the TRBD and CTE, where it is proposed to stabilize the closed TERT ring (12,89,90). Complete binding to the TRBD requires opening of the unstable template-stem. (Step 4) The template and RT interact due to their affinity and proper positioning. Further entry of RT-TRBD into the t/PK circle facilitates the opening of the rest of Stem 3alt (TRE-stem). (Step 5) The residues of the open circle re-fold to form stem–loop 3B of the pseudoknot. (Step 6) Stem 3A and the pseudoknot tertiary interactions form (stabilized by TERT). The pseudoknot acts like a 'watchband ratchet clasp,' decreasing the size of the TER t/PK circle around the TERT ring, and locking the RNP complex into place. We note that the TER circle passes between the TERT ring and the TEN domain, which is connected to the TRBD by a long linker. Telomerase activity can be reconstituted with TEN added in *trans* to the TERT ring and TER (91). It is therefore likely that the TEN domain stacks over the CTE only after step 6, or even after association of TERT with the holoenzyme protein p50 (39). The TEN domain interacts directly with p50, where p50 serves as a central hub connecting the RNP catalytic core to other accessory proteins to promote processivity (12,39,92).

In this speculative model of the catalytic core assembly, TERT enters the TER circle from the 'bottom', i.e. on the same side as SL4 (as viewed in Figure 7). TERT could possibly also enter from the top, but entry from the bottom would be facilitated by p65, supporting its role in assembly, since by bringing SL4 closer to SL2, it would help bring the TERT ring to the TER circle. Either way, the determined secondary structure of the t/PK in the absence of TERT and, from the cryo-EM model, the t/PK structure in the presence of p65 and TERT, provides a working model for thinking about assembly of the catalytic core. Furthermore, this model could potentially explain why p65 can 'rescue' (or mask) pseudoknot destabilizing substitutions *in vitro*, since both components work in tandem during TERT-TER assembly. In the absence of p65, assembly of the TER circle around TERT and insertion of Loop 4 at the TRBD-CTE interface would occur independently of each other, since stem Loop 4 would point away from TERT without bound p65. Thus, the chance of properly assembling TER with TERT would be decreased. In the presence of p65, once the TER circle assembles around the TERT ring, even a destabilized pseudoknot might be able to fold on TERT.

In summary, this work highlights the important conserved features of TER pseudoknots, in particular base triples that stabilize the fold. Examination of the sequences of *Tetrahymena* and human TER t/PK suggests that an alternative structure, in which the pseudoknot nucleotides are paired with other regions of the t/PK in the TERT-free TER, may also be a common feature of TERs and play a role in assembly.

COORDINATE DEPOSITION

Coordinates and restraints for the 10 lowest energy structures of the recalculated tetPK have been deposited in the Protein Data Bank with accession code 5KMZ (supercedes 2N6Q).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Collins,K. (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat. Rev. Mol. Cell. Biol.*, **7**, 484–494.

2. Harrington,L. (2003) Biochemical aspects of telomerase function. *Cancer Lett.*, **194**, 139–154.

3. Blackburn,E.H. (2000) The end of the (DNA) line. *Nat. Struct. Biol.*, **7**, 847–850.

4. Nandakumar,J. and Cech,T.R. (2013) Finding the end: recruitment of telomerase to telomeres. *Nat. Rev. Mol. Cell. Biol.*, **14**, 69–82.

5. Blackburn,E.H. (2001) Switching and signaling at the telomere. *Cell*, **106**, 661–673.

6. Blackburn,E.H. and Collins,K. (2011) Telomerase: an RNP enzyme synthesizes DNA. *Cold Spring Harb. Perspect. Biol.*, **3**, a003558.

7. O'Sullivan,R.J. and Karlseder,J. (1994) Telomeres: protecting chromosomes against genome instability. *Nat. Rev. Mol. Cell. Biol.*, **11**, 171–181.

8. Mason,M., Schuller,A. and Skordalakes,E. (2010) Telomerase structure function. *Curr. Opin. Struct. Biol.*, **21**, 92–100.

9. Kachouri-Lafond,R., Dujon,B., Gilson,E., Westhof,E., Fairhead,C. and Teixeira,M.T. (2009) Large telomerase RNA, telomere length heterogeneity and escape from senescence in Candida glabrata. *FEBS Lett.*, **583**, 3605–3610.

10. Egan,E.D. and Collins,K. (2012) Biogenesis of telomerase ribonucleoproteins. *RNA*, **18**, 1747–1759.

11. Chen,J.L. and Greider,C.W. (2005) Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 8080–8085.

12. Jiang,J., Chan,H., Cash,D.D., Miracco,E.J., Ogorzalek Loo,R.R., Upton,H.E., Cascio,D., O'Brien Johnson,R., Collins,K., Loo,J.A. *et al.* (2015) Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science*, **350**, aab4070.

13. Ly,H., Blackburn,E.H. and Parslow,T.G. (2003) Comprehensive structure-function analysis of the core domain of human telomerase RNA. *Mol. Cell. Biol.*, **23**, 6849–6856.

14. Qiao,F. and Cech,T.R. (2008) Triple-helix structure in telomerase RNA contributes to catalysis. *Nat. Struct. Mol. Biol.*, **15**, 634–640.

15. Theimer,C.A., Blois,C.A. and Feigon,J. (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol. Cell*, **17**, 671–682.

16. Chen,J.L. and Greider,C.W. (2003) Template boundary definition in mammalian telomerase. *Cancer Lett.*, **17**, 2747–2752.

17. Lai,C.K., Miller,M.C. and Collins,K. (2002) Template boundary definition in Tetrahymena telomerase. *Genes Dev.*, **16**, 415–420.

18. Miller,M.C. and Collins,K. (2002) Telomerase recognizes its template by using an adjacent RNA motif. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 6585–6590.

19. Autexier,C., Pruzan,R., Funk,W.D. and Greider,C.W. (1996) Reconstitution of human telomerase activity and identification of a minimal functional region of the human telomerase RNA. *EMBO J.*, **15**, 5928–5935.

20. Kim,N.K., Zhang,Q. and Feigon,J. (2014) Structure and sequence elements of the CR4/5 domain of medaka telomerase RNA important for telomerase function. *Nucleic Acids Res.*, **42**, 3395–3408.

21. Kim,N.K., Zhang,Q., Zhou,J., Theimer,C.A., Peterson,R.D. and Feigon,J. (2008) Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J. Mol. Biol.*, **384**, 1249–1261.

22. Chen,Y., Fender,J., Legassie,J.D., Jarstfer,M.B., Bryan,T.M. and Varani,G. (2006) Structure of stem–loop IV of Tetrahymena telomerase RNA. *EMBO J.*, **25**, 3156–3166.

23. Leeper,T., Leulliot,N. and Varani,G. (2003) The solution structure of an essential stem–loop of human telomerase RNA. *Nucleic Acids Res.*, **31**, 2614–2621.

24. Leeper,T.C. and Varani,G. (2005) The structure of an enzyme-activating fragment of human telomerase RNA. *RNA*, **11**, 394–403.

25. Richards,R.J., Theimer,C.A., Finger,L.D. and Feigon,J. (2006) Structure of the Tetrahymena thermophila telomerase RNA helix II template boundary element. *Nucleic Acids Res.* **34**, 816–825.

26. Richards,R.J., Wu,H., Trantirek,L., O'Connor,C.M., Collins,K. and Feigon,J. (2006) Structural study of elements of Tetrahymena telomerase RNA stem–loop IV domain important for function. *RNA*, **12**, 1475–1485.

27. Theimer,C.A., Finger,L.D., Trantirek,L. and Feigon,J. (2003) Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 449–454.

28. Theimer,C.A., Jady,B.E., Chim,N., Richard,P., Breece,K.E., Kiss,T. and Feigon,J. (2007) Structural and functional characterization of human telomerase RNA processing and Cajal body localization signals. *Mol. Cell*, **27**, 869–881.

29. Cash,D.D., Cohen-Zontag,O., Kim,N.K., Shefer,K., Brown,Y., Ulyanov,N.B., Tzfati,Y. and Feigon,J. (2013) Pyrimidine motif triple helix in the Kluyveromyces lactis telomerase RNA pseudoknot is essential for function in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 10970–10975.

30. Shefer,K., Brown,Y., Gorkovoy,V., Nussbaum,T., Ulyanov,N.B. and Tzfati,Y. (2007) A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. *Mol. Cell. Biol.*, **27**, 2130–2143.

31. Tzfati,Y., Knight,Z., Roy,J. and Blackburn,E.H. (2003) A novel pseudoknot element is essential for the action of a yeast telomerase. *Genes Dev.*, **17**, 1779–1788.

32. Gilley,D. and Blackburn,E.H. (1999) The telomerase RNA pseudoknot is critical for the stable assembly of a catalytically active ribonucleoprotein. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 6621–6625.

33. Lai,C.K., Miller,M.C. and Collins,K. (2003) Roles for RNA in telomerase nucleotide and repeat addition processivity. *Mol. Cell*, **11**, 1673–1683.

34. ten Dam,E., van Belkum,A. and Pleij,K. (1991) A conserved pseudoknot in telomerase RNA. *Nucleic Acids Res.*, **19**, 6951.

35. Blackburn,E.H., Greider,C.W. and Szostak,J.W. (2006) Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nat. Med.*, **12**, 1133–1138.

36. Greider,C.W. and Blackburn,E.H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell*, **43**, 405–413.

37. Romero,D.P. and Blackburn,E.H. (1991) A conserved secondary structure for telomerase RNA. *Cell*, **67**, 343–353.

38. Ye,A.J. and Romero,D.P. (2002) Phylogenetic relationships amongst tetrahymenine ciliates inferred by a comparison of telomerase RNAs. *Int. J. Syst. Evol. Microbiol.*, **52**, 2297–2302.

39. Jiang,J., Miracco,E.J., Hong,K., Eckert,B., Chan,H., Cash,D.D., Min,B., Zhou,Z.H., Collins,K. and Feigon,J. (2013) The architecture of Tetrahymena telomerase holoenzyme. *Nature*, **496**, 187–192.

40. Singh,M., Wang,Z., Koo,B.K., Patel,A., Cascio,D., Collins,K. and Feigon,J. (2012) Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol. Cell*, **47**, 16–26.

41. Akiyama,B.M., Loper,J., Najarro,K. and Stone,M.D. (2012) The C-terminal domain of Tetrahymena thermophila telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA. *RNA*, **18**, 653–660.

42. O'Connor,C.M. and Collins,K. (2006) A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. *Mol. Cell. Biol.*, **26**, 2029–2036.

43. Stone,M.D., Mihalusova,M., O'Connor,C.M., Prathapam,R., Collins,K. and Zhuang,X.W. (2007) Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature*, **446**, 458–461.

44. Bhattacharyya,A. and Blackburn,E.H. (1994) Architecture of telomerase RNA. *EMBO J.*, **13**, 5721–5731.

45. Ulyanov,N.B., Shefer,K., James,T.L. and Tzfati,Y. (2007) Pseudoknot structures with conserved base triples in telomerase RNAs of ciliates. *Nucleic Acids Res.*, **35**, 6150–6160.

46. Mihalusova,M., Wu,J.Y. and Zhuang,X. (2011) Functional importance of telomerase pseudoknot revealed by single-molecule analysis. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 20339–20344.

47. Cole,D.I., Legassie,J.D., Bonifacio,L.N., Sekaran,V., Ding,F., Dokholyan,N.V. and Jarstfer,M.B. (2012) New models of Tetrahymena telomerase RNA from experimentally derived constraints and modeling. *J. Am. Chem. Soc.*, **134**, 20070–20080.

48. Batey,R.T., Inada,M., Kujawinski,E., Puglisi,J.D. and Williamson,J.R. (1992) Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res.*, **20**, 4515–4523.

49. Dieckmann,T. and Feigon,J. (1997) Assignment methodology for larger RNA oligonucleotides: application to an ATP-binding RNA aptamer. *J. Biomol. NMR*, **9**, 259–272.

50. Pley,H.W., Flaherty,K.M. and McKay,D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372**, 68–74.

51. Scott,W.G., Finch,J.T. and Klug,A. (1995) The crystal structure of an all-RNA hammerhead ribozyme. *Nucleic Acids Symp. Ser.*, 214–216.

52. Gullerez,J., Lopez,P.J., Proux,F., Launay,H. and Dreyfus,M. (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 5958–5963.

53. Petrov,A., Wu,T., Puglisi,E.V. and Puglisi,J.D. (2013) RNA purification by preparative polyacrylamide gel electrophoresis. *Methods Enzymol.*, **530**, 315–330.

54. Cromsigt,J.A., Hilbers,C.W. and Wijmenga,S.S. (2001) Prediction of proton chemical shifts in RNA. Their use in structure refinement and validation. *J. Biomol. NMR*, **21**, 11–29.

55. Peterson,R.D., Theimer,C.A., Wu,H. and Feigon,J. (2004) New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA-protein complexes. *J. Biomol. NMR*, **28**, 59–67.

56. Wu,H., Yang,P.K., Butcher,S.E., Kang,S., Chanfreau,G. and Feigon,J. (2001) A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. *EMBO J.*, **20**, 7240–7249.

57. Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide (2)J(NN) couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.

58. Dingley,A.J., Nisius,L., Cordier,F. and Grzesiek,S. (2008) Direct detection of N-H(...)N hydrogen bonds in biomolecules by NMR spectroscopy. *Nat. Protoc.*, **3**, 242–248.

59. Hansen,M.R., Mueller,L. and Pardi,A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.*, **5**, 1065–1074.

60. Schwieters,C.D., Kuszewski,J.J., Tjandra,N. and Clore,G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.*, **160**, 65–73.

61. Bermejo,G.A., Clore,G.M. and Schwieters,C.D. (2016) Improving NMR structures of RNA. *Structure*, **24**, 806–815.

62. de Alba,E. and Tjandra,N. (2002) NMR dipolar couplings for the structure determination of biopolymers in solution. *Progr. Nuclear Magn. Reson. Spectrosc.*, **40**, 175–197.

63. Davis,I.W., Leaver-Fay,A., Chen,V.B., Block,J.N., Kapral,G.J., Wang,X., Murray,L.W., Arendall,W.B. 3rd, Snoeyink,J., Richardson,J.S. *et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.*, **35**, W375–W383.

64. Bryan,T.M., Goodrich,K.J. and Cech,T.R. (2000) A mutant of Tetrahymena telomerase reverse transcriptase with increased processivity. *J. Biol. Chem.*, **275**, 24199–24207.

65. Berman,A.J., Gooding,A.R. and Cech,T.R. (2010) Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol. Cell. Biol.*, **30**, 4965–4976.

66. Popenda,M., Szachniuk,M., Antczak,M., Purzycka,K.J., Lukasiak,P., Bartol,N., Blazewicz,J. and Adamiak,R.W. (2012) Automated 3D structure composition for large RNAs. *Nucleic Acids Res.*, **40**, e112.

67. Jansson,L.I., Akiyama,B.M., Ooms,A., Lu,C., Rubin,S.M. and Stone,M.D. Structural basis of template-boundary definition in Tetrahymena telomerase. *Nat. Struct. Mol. Biol.*, **22**, 883–888.

68. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.

69. Abu Almakarem,A.S., Petrov,A.I., Stombaugh,J., Zirbel,C.L. and Leontis,N.B. (2012) Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res.*, **40**, 1407–1423.

70. Kazantsev,A.V. and Pace,N.R. (1998) Identification by modification-interference of purine N-7 and ribose 2′-OH groups critical for catalysis by bacterial ribonuclease P. *RNA*, **4**, 937–947.

71. Sklenar,V. and Feigon,J. (1990) Formation of a stable triplex from a single DNA strand. *Nature*, **345**, 836–838.

72. Legault,P. and Pardi,A. (1994) 31P chemical shift as a probe of structural motifs in RNA. *J. Magn. Reson.*, **103**, 82–86.

73. Pan,B., Mitra,S.N. and Sundaralingam,M. (1999) Crystal structure of an RNA 16-mer duplex R(GCAGAGUUAAAUCUGC)2 with nonadjacent G(syn).A+(anti) mispairs. *Biochemistry*, **38**, 2826–2831.

74. Chen,J.L., Blasco,M.A. and Greider,C.W. (2000) Secondary structure of vertebrate telomerase RNA. *Cell*, **100**, 503–514.

75. Su,L., Chen,L., Egli,M., Berger,J.M. and Rich,A. (1999) Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat. Struct. Biol.*, **6**, 285–292.

76. Zaug,A.J. and Cech,T.R. (1995) Analysis of the structure of Tetrahymena nuclear RNAs in vivo: telomerase RNA, the self-splicing rRNA intron, and U2 snRNA. *RNA*, **1**, 363–374.

77. Hengesbach,M., Kim,N.K., Feigon,J. and Stone,M.D. (2012) Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angew. Chem. Int. Ed. Engl.*, **51**, 5876–5879.

78. Marie-Egyptienne,D.T., Cerone,M.A., Londono-Vallejo,J.A. and Autexier,C. (2005) A human-Tetrahymena pseudoknot chimeric telomerase RNA reconstitutes a nonprocessive enzyme in vitro that is defective in telomere elongation. *Nucleic Acids Res.*, **33**, 5446–5457.

79. Licht,J.D. and Collins,K. (1999) Telomerase RNA function in recombinant Tetrahymena telomerase. *Genes Dev.*, **13**, 1116–1125.

80. Sperger,J.M. and Cech,T.R. (2001) A stem–loop of Tetrahymena telomerase RNA distant from the template potentiates RNA folding and telomerase activity. *Biochemistry*, **40**, 7005–7016.

81. Lin,J., Ly,H., Hussain,A., Abraham,M., Pearl,S., Tzfati,Y., Parslow,T.G. and Blackburn,E.H. (2004) A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 14713–14718.

82. Prathapam,R., Witkin,K.L., O'Connor,C.M. and Collins,K. (2005) A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nat. Struct. Mol. Biol.*, **12**, 252–257.

83. Cunningham,D.D. and Collins,K. (2005) Biological and biochemical functions of RNA in the tetrahymena telomerase holoenzyme. *Mol. Cell. Biol.*, **25**, 4442–4454.

84. Liu,F., Kim,Y., Cruickshank,C. and Theimer,C.A. (2012) Thermodynamic characterization of the Saccharomyces cerevisiae telomerase RNA pseudoknot domain *in vitro*. *RNA*, **18**, 973–991.

85. Autexier,C. and Greider,C.W. (1998) Mutational analysis of the Tetrahymena telomerase RNA: identification of residues affecting telomerase activity in vitro. *Nucleic Acids Res.*, **26**, 787–795.

86. Xi,L. and Cech,T.R. (2014) Inventory of telomerase components in human cells reveals multiple subpopulations of hTR and hTERT. *Nucleic Acids Res.*, **42**, 8565–8577.

87. Fu,D. and Collins,K. (2003) Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol. Cell*, **11**, 1361–1372.

88. Frederico,L.A., Kunkel,T.A. and Shaw,B.R. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry*, **29**, 2532–2537.

89. Bley,C.J., Qi,X., Rand,D.P., Borges,C.R., Nelson,R.W. and Chen,J.J. (2011) RNA-protein binding interface in the telomerase ribonucleoprotein. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 20333–20338.

90. Huang,J., Brown,A.F., Wu,J., Xue,J., Bley,C.J., Rand,D.P., Wu,L., Zhang,R., Chen,J.J. and Lei,M. (2014) Structural basis for protein-RNA recognition in telomerase. *Nat. Struct. Mol. Biol.*, **21**, 507–512.

91. Eckert,B. and Collins,K. (2012) Roles of telomerase reverse transcriptase N-terminal domain in assembly and activity of Tetrahymena telomerase holoenzyme. *J. Biol. Chem.*, **287**, 12805–12814.

92. Hong,K., Upton,H., Miracco,E.J., Jiang,J., Zhou,Z.H., Feigon,J. and Collins,K. (2013) Tetrahymena telomerase holoenzyme assembly, activation, and inhibition by domains of the p50 central hub. *Mol. Cell. Biol.*, **33**, 3962–3971.