Pyrimidine motif triple helix in the *Kluyveromyces lactis* telomerase RNA pseudoknot is essential for function in vivo

Darian D. Cash^a, Osnat Cohen-Zontag^b, Nak-Kyoon Kim^a, Kinneret Shefer^b, Yogev Brown^b, Nikolai B. Ulyanov^c, Yehuda Tzfati^{b,1}, and Juli Feigon^{a,1}

^aDepartment of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, CA 90095-1569; ^bDepartment of Genetics, The Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel; and ^cDepartment of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158-2517

Contributed by Juli Feigon, May 23, 2013 (sent for review April 17, 2013)

Telomerase is a ribonucleoprotein complex that extends the 3' ends of linear chromosomes. The specialized telomerase reverse transcriptase requires a multidomain RNA (telomerase RNA, TER), which includes an integral RNA template and functionally important template-adjacent pseudoknot. The structure of the human TER pseudoknot revealed that the loops interact with the stems to form a triple helix shown to be important for activity in vitro. A similar triple helix has been predicted to form in diverse fungi TER pseudoknots. The solution NMR structure of the *Kluyveromyces lactis* pseudoknot, presented here, reveals that it contains a long pyrimidine motif triple helix with unexpected features that include three individual bulge nucleotides and a C⁺•G-C triple adjacent to a stem 2-loop 2 junction. Despite significant differences in sequence and base triples, the 3D shape of the human and *K. lactis* TER pseudoknots are remarkably similar. Analysis of the effects of nucleotide substitutions on cell growth and telomere lengths provides evidence that this conserved structure forms in endogenously assembled telomerase and is essential for telomerase function in vivo.

RNA triplex | yeast | RNA structure | Hoogsteen

Telomerase is a ribonucleoprotein complex that extends the 3' ends of eukaryotic chromosomes by adding successive telomere DNA repeats using an internal RNA template and a specialized reverse transcriptase (1, 2). Telomeres are the protein-DNA complexes that form the ends of linear chromosomes and protect them from end-to-end fusion and degradation (3, 4). Telomerase is of significant medical interest owing to the correlation between telomere length and human health and the association of telomerase activity with cancer (5, 6). In the absence of telomerase activity, telomeres shorten with each round of cell division because of exonuclease digestion and the inability of conventional DNA polymerases to fully replicate linear chromosomes. Shortening past a critical length leads to cell cycle arrest and/or apoptosis (7). Telomerase activity is undetectable in most somatic cells, resulting in telomere attrition with each cell cycle (8, 9). On the other hand, telomerase is active in, and essential for the proliferation of, the germ line, some epithelial, haemopoietic, and stem cells, as well as ~90% of cancer cell lines (10, 11). A number of inherited diseases are associated with telomere shortening due to telomerase insufficiency, such as dyskeratosis congenita, aplastic anemia, and pulmonary fibrosis (12–15).

The telomerase holoenzyme consists of the telomerase reverse transcriptase (TERT) and telomerase RNA (TER), which are essential and sufficient for catalytic activity in vitro (16), and several species-specific accessory proteins. TERs are highly divergent in size and sequence between species, ranging from ~150 nt in ciliates, ~450 nt in vertebrates, to more than 2,000 nt in some fungi (17). TERs provide the template for telomeric DNA synthesis but also contain other domains that are essential for telomerase assembly and function. The most well-conserved region of TER is the template/pseudoknot (t/PK), or core domain, which contains a pseudoknot, the template, and a template boundary

element (TBE) (18, 19) (Fig. 1). The t/PK interacts with TERT and together with another conserved region of TER, the stem terminus element, is critical for telomerase activity (17, 20–22).

The solution NMR structure of minimal human TER pseudoknots revealed that the loops interact with the stems to form a triple helix surrounding the helical junction (22, 23). Loop 1 interacts with the major groove of stem 2 to form a short pyrimidine motif triple helix extended by a loop 1-loop 2 Hoogsteen base pair, whereas loop 2 interacts with stem 1 to form two minor groove triples. Both DNA and RNA polymers and oligonucleotides readily form pyrimidine motif triplexes with U•A-U Watson-Crick (WC)/Hoogsteen paired triples (24–26), but this was the first observation of a pyrimidine motif triplex in a biologically functional RNA. The triple-helical interactions were shown to be important for catalytic activity in vitro. Yeast TERs are substantially larger than vertebrate TERs and have three long “arms” extending from the catalytic core at the center (4, 27–29) that serve as a scaffold for the binding of accessory and regulatory proteins (30–32). The reverse transcriptase, termed Est2 in yeast, binds to the central core domain (33). Mutational analysis and structure modeling has provided evidence for functionally important pyrimidine motif triplexes in yeast TER pseudoknots (34, 35), similar to the human pseudoknot loop 1 and stem 2 interactions (22, 23).

Here we report the NMR solution structure of a minimal *Kluyveromyces lactis* TER pseudoknot and provide evidence that the tertiary interactions observed are present in the endogenously assembled telomerase and are essential for telomerase function in vivo. Comparison of the *K. lactis* and human TER pseudoknots shows that despite different sequences, junctions, and base triples, the 3D structures of the triple helices are remarkably similar, presumably serving a conserved role in telomerase function.

Results and Discussion

Minimal Pseudoknot Design and Secondary Structure. For structural studies, a minimal *K. lactis* pseudoknot ΔU959 construct (kPKDU) was designed according to a proposed model derived from in vivo mutational studies (Fig. 1B) (34). The predicted bulged U959 was deleted and a terminal A added, to increase the stability of the pseudoknot. Deletion of the bulged U in the context of full-length TER has little or no effect on telomerase activity in vivo, as revealed by the BclI-tagged telomeric repeats incorporated by this mutant telomerase and the normal telomere length observed by

Author contributions: D.D.C., N.B.U., Y.T., and J.F. designed research; D.D.C., O.C.-Z., N.-K.K., K.S., and Y.B. performed research; D.D.C., O.C.-Z., K.S., N.B.U., Y.T., and J.F. analyzed data; and D.D.C., Y.T., and J.F. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2M8K).

¹To whom correspondence may be addressed. E-mail: tzfati@mail.huji.ac.il or feigon@mbi.ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1309590110/-DCSupplemental.

Fig. 1. (A) tPK (or core) domain of *K. lactis* TER. Conserved sequences (CS) 3 and 4 make up the pseudoknot, which was truncated to the boxed nucleotides for structural studies. (B) Minimal *K. lactis* TER pseudoknot construct (kPKDU) for NMR studies with predicted secondary structure (34). Secondary structure elements are colored: stem 1 (red), loop 1 (gold), stem 2 (blue), loop 2 (green). (C) Southern analysis of telomere restriction fragments from *K. lactis* strains harboring WT or Δ U959 TER1 with BclI template mutation to mark telomerase action, shown in D. Genomic DNA samples were digested with EcoRI (-) or EcoRI + BclI (+) restriction endonucleases, electrophoresed in a 1% agarose gel, blotted, and hybridized first with a BclI-specific telomere probe (Upper) and then with a WT telomere probe (Lower). (D) Schematic representation showing a telomere containing WT (blue) and BclI (green) repeats.

Southern blotting and hybridization with a BclI-specific and WT telomeric probes (Fig. 1C and D and Fig. S1). 1D imino spectra of kPKDU in different salt conditions reveal that pseudoknot formation is stabilized by either MgCl_2 or KCl (Fig. S2B). Further NMR structural studies were done with added Mg^{+2} (Materials and Methods).

Fig. 2A shows HNN- correlation spectroscopy (COSY) spectra of ^{13}C -, ^{15}N -AU-, and ^{13}C -, ^{15}N -GC-labeled kPKDU aligned with the imino proton region of the 2D H_2O NOESY. The stems and loop 1 imino protons were assigned by standard methods using sequential NOE connectivities in the NOESY, along with HNN-COSY, which directly detects hydrogen bonds to differentiate A-U WC, A-U Hoogsteen, and G-C WC base pairs (36). Although not directly detected by HNN-COSY, an additional resonance is identified as a hydrogen bonded protonated C(861) imino by its distinct downfield chemical shift. Six sequential imino NOE connectivities were observed between the Hoogsteen base pairs of loop 1–stem 2 (C861 to U866), indicating a continuous triple helix (Fig. 2A). A Hoogsteen base paired triplex NOE pattern is also observed in the nonexchangeable NOESY spectra for the six nucleotides of loop 1 (22) (SI Materials and Methods).

At the helical junction between stem 1 and stem 2, a surprising difference from the predicted secondary structure is observed. There are two consecutive G-C base pairs at the top of stem 2, which could only form if an additional two nucleotides (U955 and C957) are bulged out (Fig. 2A, green). Sequential NOEs are observed through the stems and junction from U860–G954–G956–U876, which indicates that the pseudoknot is continuously base paired and stacked through the two stems. Furthermore, one of these G-C base pairs forms a base triple with the first C in loop 1 (C861•G956–C877). Thus, analysis of these spectra indicates that an extended pyrimidine motif triple helix forms between loop 1 and stem 2 that includes one $\text{C}^+\bullet\text{G-C}$ triple and five $\text{U}\bullet\text{A-U}$ triples, stem 2 extends by two G-C bases pairs, including a stem 2–loop 2 junction interaction, and there are two bulge bases at the top of stem 2 (Fig. 2B).

Solution Structure of kPKDU. The solution structure of kPKDU was solved using 1,027 NOE distance restraints and 549 dihedral restraints and refined with 98 residual dipolar couplings (RDCs)

(Table S1). Superposition of the lowest 20 energy structures shows that the pseudoknot is well defined, with an overall rmsd to the mean of 1.1 ± 0.2 Å (Fig. 3A and Fig. S3A). A continuous A-RNA helix is formed by 17 consecutive WC stacked base pairs: 5 in stem 1, 1 in the stem 2–loop 2 junction (C878–G954), and 11 in stem 2. There are no significant bends in the helical axis. Near the junction, U955 and C957 are bulged out in the major and minor grooves, respectively (Fig. 3C and Fig. S3B) to allow for base pairing in the junction and top of stem 2. Loop 1 forms six successive Hoogsteen base pairs in the major groove of stem 2, creating a pyrimidine motif triple helix, whereas loop 2 lies in the minor groove of stem 1 on the same face of the structure. Stem 1 U860 and junction G954 are at stem/loop interfaces and adopt C2' endo sugar pucker conformations. This increases the phosphate to phosphate distance, allowing placement of the adjoining residues in the grooves and stacking of U860 and G954. There is a sharp turn in the backbone at the end of the triple-helix (loop 1/stem 2) due to the consecutive residues U866 and A867 having tertiary/secondary interactions that separate them by five base pairs (Fig. 3D and Fig. S3C). These two residues are ~ 10 Å apart.

Because Mg^{+2} stabilized the folded pseudoknot, the sites of divalent metal ion localization were investigated by titrating MnCl_2 into kPKDU and observing changes in ^1H - ^{13}C heteronuclear single-quantum coherence (HSQC) spectra (Fig. S4). Resonances of residues close to the bound Mn^{+2} will broaden and disappear owing to paramagnetic relaxation (37). Two sites of divalent cation localization are observed, in the major groove near residues at the sharp loop 1/stem 2 turn and in the minor groove of stem 1 where loop 2 binds. These ions likely stabilize the close

Fig. 2. (A) (Upper) AU and GC (Inset) JNN-HNN-COSY spectra aligned with (Lower) imino proton region of 2D H_2O NOESY at 278K and 800 MHz. NOE cross-peaks of stem 1/stem 2 WC pairs are connected above the diagonal in black and loop1–stem 2 Hoogsteen pairs below the diagonal in red. Iminos from the stem junction are connected below the diagonal and colored green. (B) Determined secondary structure of kPKDU (colored as in Fig. 1) with bulge nucleotides (U955, C957) in magenta.

Fig. 3. (A) Twenty lowest energy structures from NMR solution structure determination of kPKDU superpositioned over all heavy atoms. (B) Schematic representation of the kPKDU tertiary structure. (C) Junction and (D) loop 1/stem 2 turn tertiary structures. Stereo views are presented in Fig. S3.

approach of the negatively charged phosphate backbone at the loop 1/stem 2 turn and the minor groove interactions of loop 2, respectively.

To determine whether the bulges and/or stem junction induce flexibility in the pseudoknot, we analyzed the RDC data to determine the generalized degree of order (GDO, or θ) parameter for the individual stems and assess the degree of interhelical motion. The internal GDO ($\theta_{\text{int}} = \theta_{\text{stem1}}/\theta_{\text{stem2}}$) is equal to 1 for rigidly connected helices and decreases as interhelical motions increase dependent on motional amplitude and direction (38). kPKDU has $\theta_{\text{int}} = 0.96$, indicative of a rigid structure. Thus the junction stem 2-loop 2 base pair (C878-G954) and adjacent bulge nucleotides do not induce significant flexibility in the structure. Analysis of the RDCs from the WT minimal human pseudoknot, hPKWT (23), also gave a $\theta_{\text{int}} = 0.96$. The two pseudoknots have equally rigid structures despite differences in the stem junctions and number of bulge bases.

kPKWT Has the Same Tertiary Structure as kPKDU. After determining that two additional nucleotides were bulged out near the junction in close proximity to the deleted U959, a WT pseudoknot construct was made that included U959 (kPKWT) to further confirm that deletion of U959 did not cause significant structural differences. Analysis of the imino region of the 2D H_2O NOESY spectra shows the same NOE pattern for the two pseudoknots, with the only chemical shift differences for residues near the bulge U959 in kPKWT (Fig. S5). kPKWT has a base paired junction and stacked stems, with U955 and C957 bulged out. The NOE between the base paired iminos of U876 and U875 indicates that U959 is bulged out as predicted. The triple helix NOE pattern is also present for all six nucleotides of loop 1. We conclude that there are no significant tertiary structure differences between kPKDU and kPKWT.

Extended Triple Helix Stabilizes TER Pseudoknot Formation. To investigate the contribution of the tertiary interactions to the structure and folding of the pseudoknot, RNA constructs were made with nucleotide substitutions in the triple helix and examined by NMR. First, kPKDU, which has a truncated stem 1, was extended to a full-length stem 1 (kPKFL). Stem 2-loop 2 hairpin (S2L2) constructs were made, and their 1D imino spectra were compared to identify the conformation of the pseudoknot constructs in solution (Fig. 4, green). Under low salt conditions, kPKDU is primarily in the S2L2 conformation, whereas kPKFL primarily forms the pseudoknot, indicating the kPKFL pseudoknot is more stable (Fig. 4 A and B, blue). Addition of Mg^{2+} to kPKDU and kPKFL (Materials and Methods) shifts the equilibrium

of both constructs completely to the pseudoknot conformation (Fig. 4 A and B, red).

Previous mutational data showed that loop 1 substitution 864-6UUU:CCC, which should disrupt three of the six base triples, abolished telomerase activity in vivo (33). NMR studies of this substitution as well as 861-3CUU:UCC in kPKFL show that these RNAs do not form a pseudoknot, even when Mg^{2+} is present; instead only an elongated S2L2 forms (Fig. 4 C and D). Additional imino peaks in these hairpins arise from formation of additional U-A base pairs at the top of stem 2 (diagrammed in Fig. 4 C and D). Thus, base triples are needed to stabilize the formation of stem 1 and therefore folding of the pseudoknot as it is transcribed, revealing an important role for the extended triple helix in *K. lactis* TER. In the human pseudoknot, changing loop 1 residues (99-100UU:CC), which greatly decreases activity in vitro (20), abolished formation of stem 1 and stem 2 base pairs around the junction and all tertiary triplex interactions (39). This nucleotide substitution also stabilized an alternate conformation in high salt conditions, as shown by FRET analysis (39).

Mutational Studies Indicate the Determined Structure is Important for Function in Vivo.

To test whether the determined tertiary interactions of kPKDU are important for function in vivo, a series of TER mutations were made, and their effects on telomerase activity, telomere length, and colony phenotype were investigated. Our structure revealed a triple helix with five U•A-U base triples in a distinct register (register 1), starting with the first stem 2 A-U base pair (relative to the junction) (Fig. 5A). However, modeling studies indicated the possibility of an alternate register (register 2) in which only four U•A-U triples would form, starting with the second stem 2 A-U base pair (Fig. 5B). To determine which register formed in vivo, mutations were designed to affix each of the conformations separately (Fig. 5 C and D). First, alternate As in the top strand (A) of stem 2 were replaced with Gs (S2A). Second, alternate Us in the bottom strand (B) of stem 2 were replaced with Cs, and alternate Us in loop 1 were replaced with Cs starting with the first U (862) or the second U (863) (S2B+L1 and SB2+L1',

Fig. 4. One-dimensional imino proton spectra at 283K and 500 MHz of RNA constructs of the *K. lactis* pseudoknot and stem 2-loop 2 hairpin in 10 mM Tris-D (pH 6.3) (green and blue) and with added Mg^{2+} (red) (Materials and Methods). Hairpin (green, bracketed) and pseudoknot sequences are shown below the spectra. (A) kPKDU and S2L2, (B) kPKFL and corresponding S2L2-FL. kPKFL and S2L2-FL have the full-length stem 1. (C) kPKFL(861-3CUU:UCC) and S2L2-FL, and (D) kPKFL(864-6UUU:CCC) and S2L2-FL. Dashed lines on the sequences indicated observed tertiary base pairs. Nucleotide substitutions in C and D were designed to prevent formation of the first three or last three triples. Spectra are labeled with the RNA construct and primary tertiary fold, hairpin (HP), or pseudoknot (PK), under the given conditions.

Fig. 5. (A and B) Two alternative base triple registers, 1 and 2, predicted for the *K. lactis* pseudoknot. Register 1 (A) is observed in the solution structure. (C and D) Compensatory mutations designed to affix (C) register 1 and (D) register 2 in vivo. Mutations were made in the top strand (S2A), bottom strand (S2B), and loop (L1, L1'). (E) Representative Southern analysis of telomere restriction fragments prepared from WT, Δ TER1, and pseudoknot mutants (shown in red in A, C, and D) (Materials and Methods and Fig. S1).

respectively). The individual S2A, S2B+L1, and S2B+L1' mutations, which are expected to disrupt stem 2 formation and potential tertiary interactions, abolished telomerase activity in vivo, as apparent by the lack of BclI repeat incorporation, the smeared pattern of telomere restriction fragments typical of an alternative, recombination-dependent, telomere elongation mechanism (Fig. 5E, assay schematic in Fig. S1), and a rough colony phenotype indicative of impaired telomere maintenance (Fig. S6). These results demonstrate that stem 2 and triplex formation is essential for telomerase function in vivo. The combined S2A+S2B+L1 or S2A+S2B+L1' mutations were designed such that alternating U•A-U and C⁺•G-C triples would form along the helix, because adjacent C⁺•G-C triples are very unstable (owing to the requirement for protonation of the Hoogsteen paired C) and fail to support normal telomerase function (34). Strikingly, the yeasts with TER S2A+S2B+L1 (register 1) substitutions have normal telomerase activity, telomere length, and colony morphology, whereas S2A+S2B+L1' (register 2) results in barely detectable telomerase activity, severely short telomeres, and rough colonies (Fig. 5E and Fig. S5). These results indicate that the particular register is essential for telomerase function in vivo. The importance of the length of the triple helix was also investigated by making mutations that extend or shorten it by two base triples. These sets of mutations caused moderate and severe telomere shortening, respectively, but did not completely abolish telomerase activity, as indicated by the incorporation of BclI repeats (Fig. 5E). Although the shortening of the triplex may destabilize the structure, the effect of its extension indicates that the length of the triplex or the position of nucleotides at the apical part of loop 1 or adjacent stem 2 are crucial for telomerase function. The importance of the length of loop 1 was tested by deleting the apical nucleotide U866 or inserting a C downstream to it. Both of these mutations also caused moderate but significant shortening of telomeres (Fig. 5E). Deletion of U866 would abolish the last triple and likely disrupt the terminal A867-U986 base pair, and its effect is consistent with the length of the triple helix being important for telomerase activity in vivo. The extra C in the loop could affect the positioning of the third strand or could interfere with a TERT

specific interaction. In *Saccharomyces cerevisiae*, TERT is predicted to interact in this region adjacent to the triple helix (35).

An unexpected feature of the *K. lactis* WT pseudoknot was the presence of three bulge nucleotides (C955, U957, and U959) in the purine strand of the triplex adjacent to the top three triples. A bulged U177 in the human telomerase pseudoknot below the bottom triple significantly destabilizes stem 2 and therefore the loop 2–stem 2 hairpin (23, 36, 40), and deletion of U177 decreases activity in vitro (22, 40). On the basis of the kPKDU structure, we predicted that deletion of the three bulge nucleotides from the *K. lactis* pseudoknot would stabilize the pseudoknot but otherwise not affect its conformation. To test the importance of the bulge nucleotides in the *K. lactis* pseudoknot on telomerase function in vivo, we deleted them simultaneously. This triple deletion had no significant effect on telomerase activity or telomere length (Fig. 5E), indicating that the bulge residues in the *K. lactis* pseudoknot are not important for telomerase function in vivo. Finally, G954 and G956 were substituted with Cs, disrupting the junction base pair and C⁺•G-C base triple. This substitution caused significant telomere shortening, demonstrating the importance of this end of stem 2 and the triplex (Fig. 5E).

The previous mutational analysis correctly identified a loop 1–stem 2 triple helix but failed to define the junction base pairs (Fig. 1A) (34). Deleting the bulge (Δ 876-8) in the initially predicted structure or making a more definitive bulge (876-8UCC:AUG + 957-9CAU:GGA) abolished activity, whereas base pairing the bulge (876-8UCC:AUG, or 957-9CAU:GGA) maintained WT activity, indicating that the junction must base pair to function (34). The solution structure explains these data, revealing that three residues bulge out to allow for a base paired junction and formation of two additional base triples, including a C⁺•G-C triple. Taken together, the effects of the pseudoknot substitutions on telomere length in vivo confirm that the secondary and tertiary interactions in kPKDU determined by NMR are present in the context of the full-length TER in the telomerase holoenzyme and that formation of the extended triple helix is critical for telomerase activity in vivo.

***Kluyveromyces* Species TER Pseudoknots Have a Conserved Tertiary Structure.** Previous analysis of TER pseudoknot sequences from the *Kluyveromyces marxianus* cluster of species revealed conservation of a potential triple helix (34). We have cloned the TER gene from a newly identified species, *Kluyveromyces siamensis* (41). Six of the 40 nucleotides forming the *K. siamensis* pseudoknot differ from the *K. lactis* sequence, and it lacks all three of the nucleotides that are bulged out in the *K. lactis* pseudoknot. We predicted the secondary structure on the basis of the *K. lactis* structure and found that it could form an identical secondary structure except for the absence of the three bulge nucleotides (Fig. S7). We reanalyzed the five other known *Kluyveromyces* sp. pseudoknot sequences and found that by bulging out none, one, two, or three residues, as necessary, each could form a similar triple helix with five consecutive U•A-U triples and in all but one case an additional C⁺•G-C or U•G-C triple (Fig. S7). The absence of conservation of the bulge nucleotides along with the results of the mutagenesis data indicate that the bulge nucleotides are dispensable for function, in contrast to the case for the bulge U177 in the human pseudoknot.

Comparison with *S. cerevisiae* Telomerase Pseudoknot. *S. cerevisiae* TER (TLC1) was shown to form a pseudoknot with base triple interactions involving stem 2 (A803-806), which are important for telomerase activity (35). However, the precise secondary and tertiary structure of this pseudoknot has yet to be defined. A recent thermodynamics study predicts that stem 2 is extended by 6 base pairs beyond a 6-nt bulge loop and that loop 1 forms additional triples with this stem, eventually supported by sequence conservation of the U•A-U residues (Fig. S8) (42). We noticed that two additional C⁺•G-C triples could form with this part of stem 2 by protonating the Cs (741-742) in loop 1. Preliminary NMR data of an *S. cerevisiae* minimal pseudoknot shows two distinct downfield shifted iminos that correspond to protonated

Cs (Fig. 6B), supporting the hypothesis of an extended bipartite triple helix in the *S. cerevisiae* pseudoknot. Pseudoknots with a bipartite triple helical region have been predicted to form in *Candida* yeast TER as well (43). At the helical junction, the *S. cerevisiae* pseudoknot has an apparent 3-nt bulge that would preclude the formation of the continuous stem 1–stem 2 stacking interactions seen in the *K. lactis* pseudoknot. However, we noticed that by bulging out residues A733–G734 the three As in the predicted bulge could base pair to extend stem 1 (Fig. 6A). This would allow for a stacked junction proximal to an extended triple helix, resulting in a pseudoknot structurally similar to that of *K. lactis*.

Human and *K. lactis* Pseudoknot Structures Are Remarkably Similar.

Comparison of the tertiary structures of the pseudoknots from *K. lactis* (kPK) and human (hPK) (22, 23) reveals that despite the lack of significant sequence identity the 3D structures are strikingly similar (Fig. 7). Both *K. lactis* and human PKs have continuous stacking interactions between the two stems across the junction. kPK has WC base pairing in the junction allowing the stems to coaxially stack, while the hPK stems are separated by a loop–loop Hoogsteen interaction. This causes underwinding and a small bend in the helical axis at the junction of the two stems in hPK. However, a second bend at the bulge U177 results in an overall straight stem similar to kPK. Both pseudoknots have a sharp loop 1/stem 2 turn where the phosphates in the backbone come very close together, and it is likely that cation binding in this region is important for stabilizing the tertiary structure (39). Although the extended triple helix has a different arrangement of base triples around the junction for kPK vs. hPK, the triple-helical region is the same length. hPK has three Hoogsteen base triples in the major groove of stem 2, one loop–loop Hoogsteen base pair in the junction, and two base triples in the minor groove of stem 1, which stack consecutively (nucleotides 99–102, 172–173; Fig. 7B). kPK forms a triple helix with the six sequential residues of loop 1, including a C⁺•G-C triple, and has no minor groove or loop–loop base interactions (Fig. 7A).

Although both human and *K. lactis* TER pseudoknots have bulged residues in the purine-rich strand of stem 2, examination of the structures reveals they are at different locations (Fig. 7). The bulged residues in kPK are near the junction at the 5' end of the stem 2, and deletion of one or all three of these bulge residues has little effect on telomerase activity or telomere length in vivo. hPK has a bulged residue (U177) in stem 2 just below the triple helix, deletion of which decreases activity by 50% in vitro (22, 40). Backbone superposition of the two structures (Fig. 7C) reveals that the bulge residues are on opposite faces of the helix. In the model of the human TER core domain (44), the bulge U177 is near the template and facing toward the modeled TERT and putative active site. In contrast, bulges near the junction like those found in the *K. lactis* pseudoknot would apparently face away from the active site, potentially explaining their lack of effect on telomerase function.

Fig. 6. (A) Minimal *S. cerevisiae* pseudoknot with predicted secondary structure and tertiary interactions. Interactions are validated by assays (blue), sequence conservation (green), or unconfirmed (red). (B) One-dimensional imino spectra of *S. cerevisiae* pseudoknot with protonated C⁺ iminos (NH) indicated by arrows, recorded at 283K on 600-MHz NMR. Sample buffer is 10 mM Na phosphate (pH 6.3) and 50 mM KCl.

Fig. 7. Solution structures of minimal TER pseudoknots from (A) *K. lactis* (kPK) and (B) *Homo sapiens* (hPK), colored as in Fig. 1. Tertiary structure schematics (Left) and lowest energy structures (Right) are shown. (C) Superposition of the backbones of the kPK (red) and hPK (green) tertiary structures, with bulges as stick/ball for kPK (black) and hPK (magenta).

Pyrimidine Motif Triplexes Stabilize Other Noncoding RNAs. Pseudoknots are a prevalent RNA motif, and a variety of major groove triples have been found in riboswitch aptamers that fold into pseudoknots for ligand capture [e.g., SAM-II riboswitch (45)], as well as in the group II intron near the catalytic site (46). However, pyrimidine motif triplexes (i.e., three or more consecutive U•A-U triples) have only been found in pseudoknots from human and yeast telomerase. Recently, pyrimidine motif triplexes were discovered in the viral polyadenylated nuclear (PAN) RNA of Kaposi's sarcoma-associated herpesvirus (47, 48) and the long noncoding RNAs (lncRNAs) metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) and multiple endocrine neoplasia-β (MENβ), where they function to protect the 3' A-rich ends from degradation (49, 50). Five U•A-U triples form in PAN RNA, whereas a bipartite triple helix containing two regions of three to five U•A-U triples separated by a bulge (with a possible C⁺•G-C triple) form in MENβ and MALAT1, similar to *S. cerevisiae* and *Candida* pseudoknot predictions (42, 43). Although the secondary structure of the pseudoknot has two stems, it is likely that the viral RNAs, lncRNAs, and TER pseudoknots all form the triplex in the same way, with formation of an initial stem (stem 1 in pseudoknots) as the RNA is transcribed followed by insertion of a 3' A-rich strand into a U-rich hairpin (36) or internal loop.

Conclusions

The triple helix has emerged as a conserved and essential element that stabilizes the pseudoknot fold in TERs. Folding of the TER pseudoknot is likely important for positioning the template and the TBE that regulates telomere repeat synthesis, at the catalytic site of TERT (33, 35). TERT does not seem to interact strongly with the triple helical region of the pseudoknot but rather with the adjacent base pairs near the core enclosing helix or TBE (33, 35). Previous studies of the human telomerase pseudoknot established the importance of the triple helix interactions for activity of telomerase reconstituted in vitro (22, 35). However, mutations that affect catalysis in vitro often have different effects in vivo (51). Here we have directly correlated the secondary and tertiary interactions observed in a determined structure of a pseudoknot with the effects of mutations that disrupt those interactions in vivo on telomerase activity, telomere length, and cell viability.

Studies of telomere-associated proteins have revealed strong structural similarities by divergent polypeptide sequences from distant species (e.g., TPP1 in human, Est3 in yeast, and TEBP-β in ciliates) (52). Similarly, the results presented here also indicate that although TER is highly divergent and presents no apparent sequence similarity across yeast and vertebrates, human and *K. lactis* TERs form remarkably similar structures around a conserved extended triple helix. Maintaining such a high structural conservation by RNA molecules that evolve almost as rapidly as

intergenic regions reflects evolutionary constraints to conserve an essential telomerase function.

Materials and Methods

NMR Sample Preparation, Data Acquisition, and Structure Calculations. RNAs were synthesized by in vitro transcription with T7 RNA polymerase on a synthetic DNA template and purified by PAGE as previously described (22). For structure determination of kPKDU, unlabeled, (^{13}C , ^{15}N)-AU and GC labeled, and fully labeled RNA samples were made. RNA concentrations were ~ 1 mM in a sample buffer of 10 mM Tris-D (98% deuterated) (pH 6.3), 0.5 mM MgCl_2 . NMR spectra were recorded on Bruker DRX 500- and 600- and Avance 800-MHz spectrometers equipped with HCN cryoprobes. Spectra were recorded at 298K in 100% D_2O , except for imino data, which were recorded at 278K and 283K in 10% $\text{D}_2\text{O}/90\%$ (vol/vol) H_2O . NMR experiments used and methods for obtaining assignments and restraints, RDC analysis, and structure calculations were done as previously described (22, 23). Details are given in the *SI Materials and Methods*.

In Vivo Telomere Length Assay. *K. lactis ter1 Δ* strain yJR27 was used for the plasmid shuffling of WT with mutant TER genes encoded on a CEN-ARS plasmid, as described previously (53). Both WT and mutant *TER1* genes contained an additional *BclI* template mutation that is incorporated into the nascent telomeric repeats, introducing *BclI* restriction sites. Otherwise it does not affect telomerase activity or telomere length and is thus used to mark the action of the examined telomerase. Genomic DNA was prepared from the yeast strains at their sixth passage and analyzed by Southern blotting and hybridization first to a *BclI*-specific and then to a WT *K. lactis* telomeric probe, as described previously (34). A schematic of the assay is shown in Fig. S1. At least two clones were examined for each mutation.

ACKNOWLEDGMENTS. We thank Savitree Limtong, Kasetsart University, for the gift of *K. siamensis* and Jing Zhou for help to D.D.C. This work was supported by National Science Foundation Grant MCB1022379 and National Institutes of Health (NIH) Grant GM048123 (to J.F.) and United States-Israel Binational Science Foundation Grant 2009204 (to N.B.U. and Y.T.). D.D.C. is a member of the Cellular and Molecular Biology Training Grant program at University of California, Los Angeles supported by NIH Grant GM007185.

- Blackburn EH (2001) Switching and signaling at the telomere. *Cell* 106(6):661–673.
- Blackburn EH, Collins K (2011) Telomerase: An RNP enzyme synthesizes DNA. *Cold Spring Harb Perspect Biol* 3(5):3.
- Londoño-Vallejo JA, Wellinger RJ (2012) Telomeres and telomerase dance to the rhythm of the cell cycle. *Trends Biochem Sci* 37(9):391–399.
- Cech TR (2004) Beginning to understand the end of the chromosome. *Cell* 116(2):273–279.
- de Jesus BB, Blasco MA (2012) Potential of telomerase activation in extending health span and longevity. *Curr Opin Cell Biol* 24(6):739–743.
- Blasco MA (2003) Telomeres and cancer: A tale with many endings. *Curr Opin Genet Dev* 13(1):70–76.
- Collins K, Mitchell JR (2002) Telomerase in the human organism. *Oncogene* 21(4):564–579.
- Aubert G, Lansdorp PM (2008) Telomeres and aging. *Physiol Rev* 88(2):557–579.
- Blasco MA (2007) Telomere length, stem cells and aging. *Nat Chem Biol* 3(10):640–649.
- Wong JM, Collins K (2003) Telomere maintenance and disease. *Lancet* 362(9388):983–988.
- Shay JW, Wright WE (2006) Telomerase therapeutics for cancer: Challenges and new directions. *Nat Rev Drug Discov* 5(7):577–584.
- Armanios M, Blackburn EH (2012) The telomere syndromes. *Nat Rev Genet* 13(10):693–704.
- Chen JL, Greider CW (2004) Telomerase RNA structure and function: Implications for dyskeratosis congenita. *Trends Biochem Sci* 29(4):183–192.
- Kirwan M, Dokal I (2008) Dyskeratosis congenita: A genetic disorder of many faces. *Clin Genet* 73(2):103–112.
- Calado RT, Young NS (2008) Telomere maintenance and human bone marrow failure. *Blood* 111(9):4446–4455.
- Collins K (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat Rev Mol Cell Biol* 7(7):484–494.
- Egan ED, Collins K (2012) Biogenesis of telomerase ribonucleoproteins. *RNA* 18(10):1747–1759.
- Chen JL, Greider CW (2003) Template boundary definition in mammalian telomerase. *Genes Dev* 17(22):2747–2752.
- Chen JL, Greider CW (2004) An emerging consensus for telomerase RNA structure. *Proc Natl Acad Sci USA* 101(41):14683–14684.
- Chen JL, Greider CW (2005) Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc Natl Acad Sci USA* 102(23):8080–8085, discussion 8077–8079.
- Ly H, Blackburn EH, Parslow TG (2003) Comprehensive structure-function analysis of the core domain of human telomerase RNA. *Mol Cell Biol* 23(19):6849–6856.
- Theimer CA, Blois CA, Feigon J (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* 17(5):671–682.
- Kim NK, et al. (2008) Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J Mol Biol* 384(5):1249–1261.
- Felsenfeld G, Rich A (1957) Studies on the formation of two- and three-stranded polyribonucleotides. *Biochim Biophys Acta* 26(3):457–468.
- Arnott S, Bond PJ (1973) Structures for Poly(U)-poly(A)-poly(U) triple stranded polynucleotides. *Nat New Biol* 244(134):99–101.
- Rajagopal P, Feigon J (1989) Triple-strand formation in the homopurine:homopyrimidine DNA oligonucleotides d(G-A)₄ and d(T-C)₄. *Nature* 339(6226):637–640.
- Tzfati Y, Fulton TB, Roy J, Blackburn EH (2000) Template boundary in a yeast telomerase specified by RNA structure. *Science* 288(5467):863–867.
- Dandjinou AT, et al. (2004) A phylogenetically based secondary structure for the yeast telomerase RNA. *Curr Biol* 14(13):1148–1158.
- Chappell AS, Lundblad V (2004) Structural elements required for association of the *Saccharomyces cerevisiae* telomerase RNA with the Est2 reverse transcriptase. *Mol Cell Biol* 24(17):7720–7736.
- Bertuch AA, Lundblad V (2006) The maintenance and masking of chromosome termini. *Curr Opin Cell Biol* 18(3):247–253.
- Hug N, Lingner J (2006) Telomere length homeostasis. *Chromosoma* 115(6):413–425.
- Zappulla DC, Cech TR (2004) Yeast telomerase RNA: A flexible scaffold for protein subunits. *Proc Natl Acad Sci USA* 101(27):10024–10029.
- Lin J, et al. (2004) A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. *Proc Natl Acad Sci USA* 101(41):14713–14718.
- Shefer K, et al. (2007) A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. *Mol Cell Biol* 27(6):2130–2143.
- Qiao F, Cech TR (2008) Triple-helix structure in telomerase RNA contributes to catalysis. *Nat Struct Mol Biol* 15(6):634–640.
- Theimer CA, Finger LD, Trantirek L, Feigon J (2003) Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci USA* 100(2):449–454.
- Butcher SE, Allain FH, Feigon J (2000) Determination of metal ion binding sites within the hairpin ribozyme domains by NMR. *Biochemistry* 39(9):2174–2182.
- Al-Hashimi HM, et al. (2002) Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. *J Mol Biol* 315(2):95–102.
- Hengesbach M, Kim NK, Feigon J, Stone MD (2012) Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angew Chem Int Ed Engl* 51(24):5876–5879.
- Comolli LR, Smirnov I, Xu L, Blackburn EH, James TL (2002) A molecular switch underlies a human telomerase disease. *Proc Natl Acad Sci USA* 99(26):16998–17003.
- Am-In S, Yongmanitchai W, Limtong S (2008) *Kluyveromyces siamensis* sp. nov., an ascomycetous yeast isolated from water in a mangrove forest in Ranong Province, Thailand. *FEMS Yeast Res* 8(5):823–828.
- Liu F, Kim Y, Cruickshank C, Theimer CA (2012) Thermodynamic characterization of the *Saccharomyces cerevisiae* telomerase RNA pseudoknot domain in vitro. *RNA* 18(5):973–991.
- Gunisova S, et al. (2009) Identification and comparative analysis of telomerase RNAs from *Candida* species reveal conservation of functional elements. *RNA* 15(4):546–559.
- Zhang Q, Kim NK, Feigon J (2011) Architecture of human telomerase RNA. *Proc Natl Acad Sci USA* 108(51):20325–20332.
- Gilbert SD, Rambo RP, Van Tyne D, Batey RT (2008) Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat Struct Mol Biol* 15(2):177–182.
- Toor N, Keating KS, Taylor SD, Pyle AM (2008) Crystal structure of a self-spliced group II intron. *Science* 320(5872):77–82.
- Mitton-Fry RM, DeGregorio SJ, Wang J, Steitz TA, Steitz JA (2010) Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. *Science* 330(6008):1244–1247.
- Tycowski KT, Shu MD, Borah S, Shi M, Steitz JA (2012) Conservation of a triple-helix-forming RNA stability element in noncoding and genomic RNAs of diverse viruses. *Cell Rep* 2(1):26–32.
- Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA (2012) Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MEN β noncoding RNAs. *Proc Natl Acad Sci USA* 109(47):19202–19207.
- Wilusz JE, et al. (2012) A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* 26(21):2392–2407.
- Robart AR, Collins K (2010) Investigation of human telomerase holoenzyme assembly, activity, and processivity using disease-linked subunit variants. *J Biol Chem* 285(7):4375–4386.
- Lewis KA, Wuttke DS (2012) Telomerase and telomere-associated proteins: Structural insights into mechanism and evolution. *Structure* 20(1):28–39.
- Roy J, Fulton TB, Blackburn EH (1998) Specific telomerase RNA residues distant from the template are essential for telomerase function. *Genes Dev* 12(20):3286–3300.