Published online 11 December 2013 *Nucleic Acids Research, 2014, Vol. 42, No. 5* ***3395–3408***
doi:10.1093/nar/gkt1276

Structure and sequence elements of the CR4/5 domain of medaka telomerase RNA important for telomerase function

Nak-Kyoon Kim[1,2], Qi Zhang[1,3] and Juli Feigon[1,*]

[1]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095-1569, USA, [2]Advanced Analysis Center, Korea Institute of Science and Technology, Seoul, 136-791, Republic of Korea and [3]Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Received August 11, 2013; Revised November 13, 2013; Accepted November 16, 2013

ABSTRACT

Telomerase is a unique reverse transcriptase that maintains the 3′ ends of eukaryotic chromosomes by adding tandem telomeric repeats. The RNA subunit (TR) of vertebrate telomerase provides a template for reverse transcription, contained within the conserved template/pseudoknot domain, and a conserved regions 4 and 5 (CR4/5) domain, all essential for catalytic activity. We report the nuclear magnetic resonance (NMR) solution structure of the full-length CR4/5 domain from the teleost fish medaka (*Oryzias latipes*). Three helices emanate from a structured internal loop, forming a Y-shaped structure, where helix P6 stacks on P5 and helix P6.1 points away from P6. The relative orientations of the three helices are Mg^{2+} dependent and dynamic. Although the three-way junction is structured and has unexpected base pairs, telomerase activity assays with nucleotide substitutions and deletions in CR4/5 indicate that none of these are essential for activity. The results suggest that the junction is likely to change conformation in complex with telomerase reverse transcriptase and that it provides a flexible scaffold that allows P6 and P6.1 to correctly fold and interact with telomerase reverse transcriptase.

INTRODUCTION

Telomeres are the 3′ ends of linear chromosomes and comprise short tandem double-stranded DNA repeats, ending in a single-stranded 3′ overhang, and telomere-specific DNA binding proteins (1,2). By capping the 3′ end single-stranded overhang, the associated telomere end-binding proteins stabilize genomic DNA and prevent end-to-end fusion and enzymatic degradation. In eukaryotes, incomplete replication of the 3′ ends during each cell cycle and chromosome end processing result in progressive attrition of telomeric DNA (3). Eventually, critically shortened telomeres lead to chromosome instability and replicative senescence (4,5). In most eukaryotes, to compensate for the telomeric DNA loss, telomerase, a specialized reverse transcriptase (RT), maintains the telomere by synthesizing tandem DNA repeats. In most somatic cells, telomerase activity is low or undetectable. Telomerase is active in proliferative cells such as epithelial, germline and hematopoietic cells and in adult stem cells, where it partially compensates for telomere loss associated with increasing age (6), and it is highly active in ~90% of cancer cells, likely enabling their immortal phenotype (7).

Telomerase is a ribonucleoprotein (RNP) enzyme complex with two catalytically essential elements responsible for telomeric DNA synthesis, the telomerase reverse transcriptase (TERT) and the telomerase RNA (TR) (8,9). The telomerase holoenzyme also contains species-specific proteins important for *in vivo* telomerase activity, RNP assembly and localization (1,8). TERT, unlike other RTs, requires an internal RNA template, contained within the TR, for synthesis of the telomere repeat. Telomere repeat addition processivity requires a translocation step. TERT possesses four evolutionarily conserved domains: telomerase essential N-terminal (TEN), telomerase RNA binding domain (TRBD), RT and C-terminal extension (CTE) (10–13). The TEN domain contains an anchor site that binds single-stranded DNA product and is important for repeat addition processivity (12,14). The TRBD is required for specific high affinity interactions with TR for telomerase catalytic function (14,15). The RT domain includes the enzyme active site where the TR template is located. The CTE domain may promote telomerase processivity and enhance nucleic acid association (16,17).

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

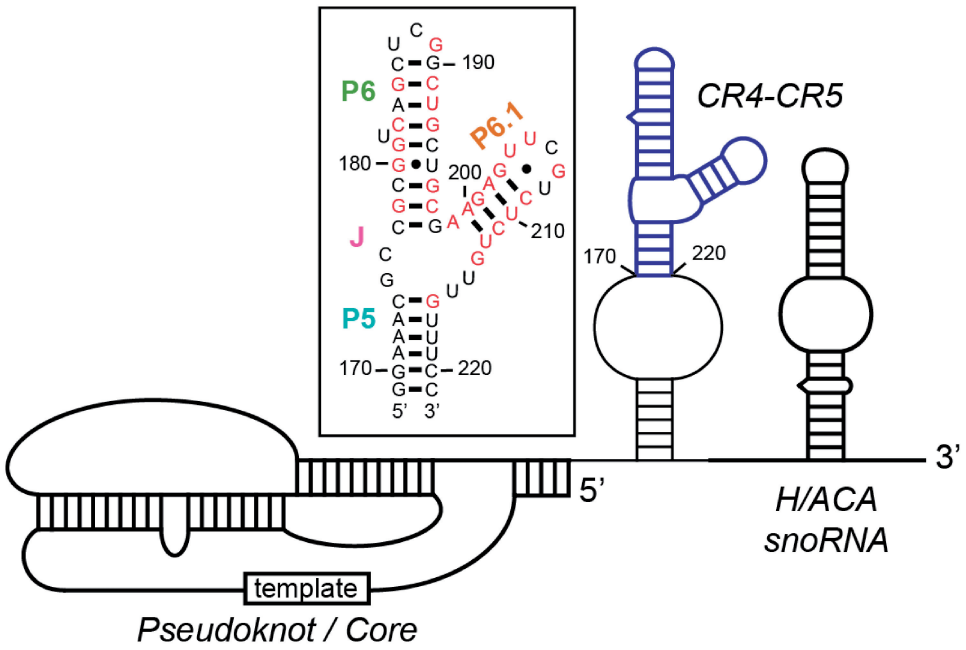

Figure 1. Secondary structure of medaka TR. Schematic of secondary structure of medaka TR with CR4/5 domain colored in blue. mdCR4/5 sequence used in the NMR study and base pairs predicted by phylogenetic comparative analysis are shown in the box. The 100% conserved nucleotides in the five identified teleost fish TR are highlighted in red (30).

In contrast to the conserved structure of TERT, TR is highly variable in sequence and size ranging from ~150 nt in ciliates to >2000 nt in *Candida* and *Aspergilla* species (18–20). However, phylogenetic comparisons of TR secondary structures and functional studies of wild-type and mutant TRs have revealed two catalytically essential domains almost universally shared among known species (1,18). These are the template/pseudo-knot (t/PK) (also called core) domain and stem terminus element (STE) (also called activation domain) (1,2), which is CR4/5 (conserved regions 4 and 5) in vertebrates. The t/PK domain includes the RNA template and a template-enclosing pseudoknot, with a conserved pyrimidine motif triple helix (21–24). The other essential domain, the STE, stimulates telomerase activity and binds to the TRBD (1,25). The STE generally contains a stem (in yeasts) or stem-loop (in filamentous fungi, vertebrates and ciliates) required for catalysis. This stem is connected to a three-way junction, except in ciliates, with a proposed conserved secondary structure (20,26). These two domains of TR can reconstitute activity *in vitro* when added to TERT expressed in rabbit reticulocyte lysate (RRL), even when added as separate domains (25).

In vertebrates, CR4/5 forms a three-way junction structure with flanking P5, P6 and P6.1 stem-loops. The human TR (hTR) P6 is extended by an additional helix P6b and forms a linear P6a-P6b stem-loop structure (Supplementary Figure S1A) (27). The P6b stem-loop is dispensable for activity, but the P6a helix and part of the internal loop sequence that connects P6a and P6b are required for telomerase activity (25). In the highly conserved P6.1 stem and loop, the P6.1 helix is required for TERT binding as well as for telomerase activity, whereas the absolutely conserved U307 and G309 in the P6.1 loop are essential for telomerase activity but not for TERT binding (28,29). Based on secondary structure prediction, the junction region of CR4/5 in hTR consists of 16 single-stranded nucleotides, including several conserved residues near P6a and P6.1, which form a large internal loop (18). There are several sets of complementary sequences in the internal loop that may form base pairs, including those predicted by comparison among different species (20,26), but the effect of mutations that would disrupt and compensatory mutations that would restore these bases pairs as well as other nucleotide substitutions on TERT binding and activity provided little evidence for formation of base pairs within the internal loop in the holoenzyme context (29). The structural role of the three-way junction on telomerase assembly or catalytic activity remains elusive.

Medaka (*Oryzias latipes*), a teleost fish, TR (mdTR) is the smallest identified TR in vertebrates (Figure 1) (30). Although mdTR has only 312 nt, it includes the t/PK, CR4/5 and H/ACA scaRNA domains found in all vertebrate TRs. In general, the sizes of these domains are smaller than those in hTR. The medaka CR4/5 (mdCR4/5) differs from human CR4/5 (hCR4/5) primarily in having a shorter P6, lacking the P6b of hTR and possibly a smaller internal loop at the three-way junction. The highly conserved mdCR4/5 P6.1 is identical to the hCR4/5 P6.1 domain except for two non-conserved nucleotides in the apical loop.

Here, we have investigated the structural and sequence requirements of medaka TR for telomerase catalytic activity. We report the first structure of a complete CR4/5 domain, by solution-state nuclear magnetic resonance (NMR) spectroscopy. The mdCR4/5 forms a Y-shaped structure in which the positions of all three

stems are dynamic relative to each other, even in the presence of Mg^{2+}. The structure along with the results of activity assays on medaka telomerase reconstituted with TRs containing nucleotide substitutions suggests that the three-way junction provides a flexible scaffold that allows P6 and P6.1 to correctly fold for interaction with TERT and that it likely rearranges when mdCR4/5 binds TERT. The mdCR4/5 structure can be modeled onto TERT consistent with cross-linking data, which places P6.1 at the TRBD-CTE interface.

MATERIALS AND METHODS

NMR sample preparation

For NMR studies of mdCR4/5, unlabeled, selectively deuterated (deuterated at the ribose 3′, 4′, 5′, 5′ of all nucleotides and at the five positions of pyrimidines; Cambridge Isotope Laboratories, Inc.) and uniformly and base-specifically (A, U, G or C) ^{13}C,^{15}N-labeled RNAs were transcribed *in vitro* using T7 RNA polymerase (P266L mutant) (31) with synthetic DNA templates. All transcribed RNAs were purified as previously described (32). All purified RNAs were desalted and exchanged extensively into an NMR buffer using Amicon ultrafiltration (Millipore). For structure determination, the NMR buffer was 10 mM sodium phosphate (pH 6.3) and 100 mM KCl. A variety of different cation conditions were initially tested, and this buffer gave spectra with the best spectral dispersion and narrowest linewidths. NMR samples were prepared by heating the RNA in a dilute (1–10 μM RNA) condition to 95°C for 5 min, fast-cooling on ice for 30 min and then concentrating to 1–1.5 mM.

To investigate the effects of Mg^{2+} on the mdCR4/5 spectra, $MgCl_2$ (1 M) was added in stepwise increments to a ~0.5 mM mdCR4/5 sample in the NMR buffer, to final concentrations of 1–30 mM, and chemical shift changes in non-exchangeable protons were monitored in ^{1}H-^{13}C HSQC spectra. To identify the sites of divalent metal ion binding on mdCR4/5, $MnCl_2$ was added in stepwise increments to a ~1 mM mdCR4/5 sample in the NMR buffer to final concentrations of 100 μM $MnCl_2$, and line broadening of residues near bound Mn^{2+} were monitored in ^{1}H-^{13}C HSQC spectra (33).

For preparation of NMR samples of mdCR4/5 with Mg^{2+} present, the final concentration of 20 mM $MgCl_2$ was added to the dilute RNA (1–10 μM) in NMR buffer. The RNA was then concentrated and exchanged to NMR buffer with 0.5 mM Mg^{2+}. This method of sample preparation ensures that all Mg^{2+} are saturated while giving spectra with narrower linewidths than if higher concentration of Mg^{2+} is used in the final solution. Although the linewidths are larger, the chemical shifts observed in ^{1}H-^{13}C HSQC NMR spectra under these conditions are identical to the ones in 20 mM Mg^{2+}.

NMR spectroscopy and structure calculations

NMR spectra were recorded on Bruker DRX 500, 600 and Avance 800 MHz spectrometers equipped with HCN cryoprobes or QXI probe at 10°C for exchangeable and at 20°C for non-exchangeable proton spectra. NMR spectra were processed and analyzed using XWINNMR 3.5 (Bruker), TOPSPIN (Bruker) and Sparky 3.110. (University of California, San Francisco, CA, USA).

The imino and amino exchangeable protons were unambiguously assigned from 2D NOESY, 2D ^{15}N-correlated CPMG NOESY, 2D ^{1}H-^{15}N HMQC and 2D H5(C5C4N)H spectra (34). RNA base pairs were confirmed with 2D HNN-COSY experiments (35,36). The 2′-hydroxyl protons were identified as described (32). The standard H1'-base proton sequential assignments were initially obtained from analysis of 2D NOESY and 2D TOCSY experiments with unlabeled RNA samples. The assignments for all non-exchangeable protons were achieved from analysis of 2D NOESY, 2D ^{1}H-^{13}C HSQC, 2D TOCSY, 2D HCCH-COSY and 3D HCCH-TOCSY (37–39). A suite of 2D-filtered/edited proton NOESY (F2f, F1fF2e, F1fF2f and F1eF2e) experiments on base-specific ^{13}C,^{15}N-labeled RNAs was used to resolve ambiguous assignments in overlapped regions and to obtain nuclear Overhauser effect (NOE) restraints, as previously described (40). The sugar pucker was determined as C2′*endo* ($\delta = 145 \pm 30°$) for residues 187–189 and 205–208 that showed strong H1′-H2′ cross-peaks in 2D DQF-COSY (41). $^{3}J_{H2'P}$ and $^{3}J_{CP}$ were measured using ^{31}P spin echo difference CT-HSQCs to determine the ε dihedral angles for the loop residues (42). The base conformation of G207 was determined to be *syn* ($\chi = 60 \pm 30°$), based on the presence strong intra-nucleotide H1′-H8 and weak H2′-H8 NOE cross-peaks in a 2D NOESY. All other bases were restrained to *anti* ($\chi = -160 \pm 30°$). For structural refinement and dynamic characterization, one-bond C-H (C1′H1′, C2H2, C5H5, C6H6 and C8H8) and N-H (N1H1 and N3H3) residual dipolar couplings (RDCs) were measured on uniformly ^{13}C,^{15}N-labeled samples in the absence and presence of ~15 mg/ml Pf1 phage (ASLA Biotech, Ltd.) at 20°C on the 800 MHz spectrometer. C-H RDCs were measured from the splittings of ^{1}H-^{13}C doublets along the ^{1}H dimension using 2D ^{1}H-^{13}C S^3CT-HSQC experiments (43), and N-H RDCs were measured and averaged from the splittings of ^{1}H-^{15}N doublets along both the ^{1}H and ^{15}N dimensions using standard 2D ^{1}H-^{15}N HSQC experiments. A total of 77 one-bond C-H and N-H RDCs were obtained for mdCR4/5 in the absence of Mg^{2+}. For structural refinement, all C-H and N-H RDCs were input as normalized C-H RDCs with a fixed bond length of 1.0 Å as previously described (44).

Inter-proton distances were measured from normalized cross-peak volumes in 2D NOESY as well as 2D-filtered/edited NOESY experiments acquired with various mixing times (40). NOE distance restraints were classified as very strong ($2.5 + 1.0/-0.7$ Å), strong (3.5 ± 1 Å), medium (4.5 ± 1 Å), weak (5.5 ± 1 Å) and very weak (6.5 ± 1 Å). The dihedral angle restraints (α, β, γ, δ, ε, χ, ν_2 and ζ) were used in the structure calculations, where the δ, ε and χ were determined experimentally as described earlier in the text. The α, β, γ, ν_2 and ζ dihedral angles for the nucleotides involved in base pairs in the helix were constrained to the A-form values ($-62.1 \pm 30°$, $179.9 \pm 30°$, $47.4 \pm 30°$, $37.3 \pm 30°$ and $-73.6 \pm 30°$, respectively).

Final structure calculations included hydrogen bond distance restraints and weak base pair planarity restraints for the 16 Watson–Crick (WC) base pairs and 3 G•U base pair as previously described (32). Two hundred initial structures of mdCR4/5 were calculated using XPLOR-NIH 2.9.8 starting from an extended unfolded RNA using the NOE distance and dihedral angle restraints following standard XPLOR protocols. The 50 lowest energy structures were further refined by inclusion of RDCs for the final structures.

To account for different degrees of alignment due to inter-helical motions, the values for the magnitude (Da) and asymmetry (R) of the alignment tensor for each stem (P5, P6 and P6.1) of mdCR4/5 were first evaluated by analyzing their RDCs using the program RAMAH (45). During this initial order tensor analysis (46), idealized 'A'-form helices corresponding to P5, P6 and P6.1 were generated using Insight II software (Accelrys) and used as input coordinates (47). These initial values were then used as the starting values and were further optimized using the structure calculated without RDCs. The optimized values for P5, P6 and P6.1 are Da = −17.1 Hz, 29.2 Hz and 10.4 Hz, respectively, and R = 0.63, 0.62 and 0.38, respectively. During the final structure calculations with RDCs, the force constants for the RDCs were gradually increased from 0.001 to 0.85, 0.50 and 1.40 kcal mol^{-1} Hz^{-2} for P5, P6, and P6.1, respectively, to account for the differences in the Da values. Experimental restraints and structural statistics for the 20 lowest energy structures are shown in Table 1. All structures were viewed and analyzed using MOLMOL (48) and PyMOL (DeLano Scientific LLC).

The inter-helical dynamics were characterized by the resultant generalized degree of orders (GDO, $\vartheta = \sqrt{2(S_{XX}^2+S_{yy}^2+S_{ZZ}^2)/3}$ where S_{xx}, S_{yy}, and S_{zz} are the order tensor values in the principle axis system) obtained from order tensor analysis of each stem using RAMAH program with the RDC refined structure. The relative motions between each helix were then measured by the relative ratios of ϑ values determined for P5, P6, and P6.1, where $\vartheta_{int,P5} = \vartheta_{P5}/\vartheta_{P6}$ and $\vartheta_{int,P6.1} = \vartheta_{P6.1}/\vartheta_{P6}$, and ϑ_{int} is the inter-helical generalized degree of order (also called GDO_{int}). Assuming a cone motional model, the amplitude of inter-helical motion can be described by the cone radius angle, Ψ_{cone}, which can be calculated from ϑ_{int} values based on the equation $\vartheta_{int} = 1/2\ \cos\Psi_{cone}(1+\cos\Psi_{cone})$ (49).

For structure determination of mdCR4/5 in the presence of Mg^{2+}, the line broadening observed in the presence of Mg^{2+} precluded obtaining accurate NOE-based distance constraints. Therefore, an NMR-based model structure was determined by generating a starting structure using the same NOE and dihedral angle restraints that were used for mdCR4/5 without Mg^{2+}, and refining this structure with 60 RDCs acquired on the mdCR4/5 sample with Mg^{+2}, prepared as described earlier in the text. Owing to differences in chemical shifts, NMR line broadening and spectral overlap in the two samples, there are some differences in the sets of RDCs collected (see Supplementary Table S1). The RDCs report on structural differences between the samples with and without Mg^{2+}.

Table 1. Restraints and structure statistics for the 20 lowest energy structures of mdCR4/5 without Mg^{2+}

Distance restraints	
Total NOE	1085
Intraresidue	462
Interesidue	623
Hydrogen bond	104
Total dihedral angle restraints	313
Total RDCs ($^1D_{CH}$ and $^1D_{NH}$)	77
Violations	
Distance constraints (Å)	0.026 ± 0.001
Dihedral angle constraints (°)	0.23 ± 0.05
Dipolar couplings (Hz)	0.51 ± 0.03
Deviation from idealized geometry	
Bond lengths (Å)	0.005 ± 0.0001
Bond angles (°)	1.09 ± 0.01
Impropers (°)	0.61 ± 0.01
Average root-mean-square deviation (Å) from the mean	
All heavy atoms	0.94 ± 0.19
Number of NOE violations > 0.2 Å	0.30 ± 0.56
Number of NOE violations >0.5 Å	0
Number of dihedral violations >5°	0.15 ± 0.32
Number of RDC violations >2 Hz	0

Modeling of mdCR4/5–mdTERT (medaka TERT) complex

The medaka TRBD (mdTRBD) domain was generated using the SWISS MODEL workplace server with the TRBD sequence of mdTERT (50) and the crystal structure of TRBD from *Takifugu rubripes* (PDB ID: 4LMO) (51) as a structural template. The modeled mdTRBD domain was then superimposed on the TRBD domain in the structure of the *Tribolium castaneum* TERT (PDB ID: 3KYL) (52), and the best fit was used to replace the *Tribolium castaneum* TRBD with the modeled mdTRBD. The model structure of mdCR4/5 in the presence of Mg^{2+} was used and positioned on the mdTRBD in a way that satisfied the constraints (within 10 Å) from the cross-linking study (53). The resulting position of the P6.1 loop at the interface between the CTE and TRBD results in some steric clash, suggesting that some conformational rearrangement would need to occur.

Direct telomerase activity assays

For telomerase activity studies, the DNA (a gift from Prof Julian J.L. Chen) for a full-length (312 nt) wild-type medaka TR was cloned into pUC19 (pUC19mdTR). The mdCR4/5 mutants were made by site-directed mutagenesis (Stratagene). Both wild-type and mutant RNAs were transcribed *in vitro* using linearized DNA templates obtained from polymerase chain reaction (PCR) amplification of the pUC19mdTR. RNA products were purified by 5% polyacrylamide gel electrophoresis and electroelution as described earlier in the text and concentrated.

N-terminally FLAG-tagged mdTERT in pCITE4a was expressed using the TNT quick-coupled transcription/translation RRL system (Promega) according to the manufacturer's instruction. Briefly, mdTERT protein

was translated in 10 µl of RRL for 1 h at 30°C. *In vitro* synthesized mdTR was added to a final concentration of 1 µM, and telomerase RNP was reconstituted by incubating the mixture for 30 min at 30°C. *In vitro* telomerase activity was measured by using the direct telomere extension assay. For telomere extension reactions, 3 µl of the reconstituted telomerase RNP was added to 10 µl of reaction mix, which includes 1× TA buffer (50 mM Tris–HCl (pH 8.3), 2 mM DTT, 0.5 mM $MgCl_2$, and 1 mM spermidine), 1 mM dATP, 1 mM dTTP, 50 µM dGTP, 0.25 µl of ^{32}P-α-dGTP (3000 Ci/mmol) and 1 µM DNA primer $(T_2AG_3)_3$. The reaction mixture was incubated for 1 h at 30°C in a heat block. A ^{32}P 5′-end labeled 15 nt DNA recovery control was added, the telomeric DNA products were extracted by using phenol-chloroform-isoamyl alcohol (pH 7.9) and ethanol-precipitated for 1 h at −80°C. The products were resolved by gel electrophoresis on a 10% denaturing gel (19:1 cross-linking ratio) in 1× TBE buffer. Gels were dried, and products were detected and analyzed using a phosphorimager (BioRad FX Pro). The relative activity was determined by measuring the total intensity of extended telomere primers after background correction and normalizing against the recovery control (30). Each activity assay was repeated a minimum of two times.

RESULTS

Secondary structure of mdCR4/5

The mdCR4/5 construct for NMR study includes the full sequence (nts 170–220) (Figure 1, boxed region). The secondary structure of mdCR4/5 was determined from the NMR data by identifying the base pairs in P5, P6, P6.1 and the junction. The WC and G•U base pairs were identified from the NOEs between imino and imino (Figure 2A), imino and aromatic and imino and amino protons in 2D ^{1}H-^{1}H NOESY and 2D ^{15}N-correlated NOESY spectra, and confirmed by 2D HNN-COSY spectrum (Figure 2B). Observed base pairs in stems P5, P6 and P6.1 agree with those predicted by phylogenetic analysis (30). The P5 stem shows sequential imino to imino cross-peaks from the G170-C220 base pair to the C174-G216 base pair (Figure 2B, cyan). P6 forms a helix containing one G•U base pair and a single bulge nucleotide U182 and is capped by a 3-nt terminal loop. The G180•U195 base pair is identified by the typical NOE pattern between the imino resonances in the base pair and to the neighboring base pairs in 2D NOESY spectra (Figure 2A, green). The sequential connectivity between the imino cross-peaks of the P6 stem in the 2D NOESY spectrum are disrupted by the bulge U182, and the remaining imino proton connectivities for base pairs above the bulge were assigned using the NOEs observed in other chemical shift regions. In P6.1, four sequential cross-peaks between the imino protons from G201 to G207 were identified (Figure 1A, orange). Interestingly, base pairs A200-U212 and U204-U208 that were predicted by phylogenetic analysis are not formed in P6.1 of mdCR4/5. Instead, we observed an unexpected NOE cross-peak between G203H1 and G207H1, suggesting that G207

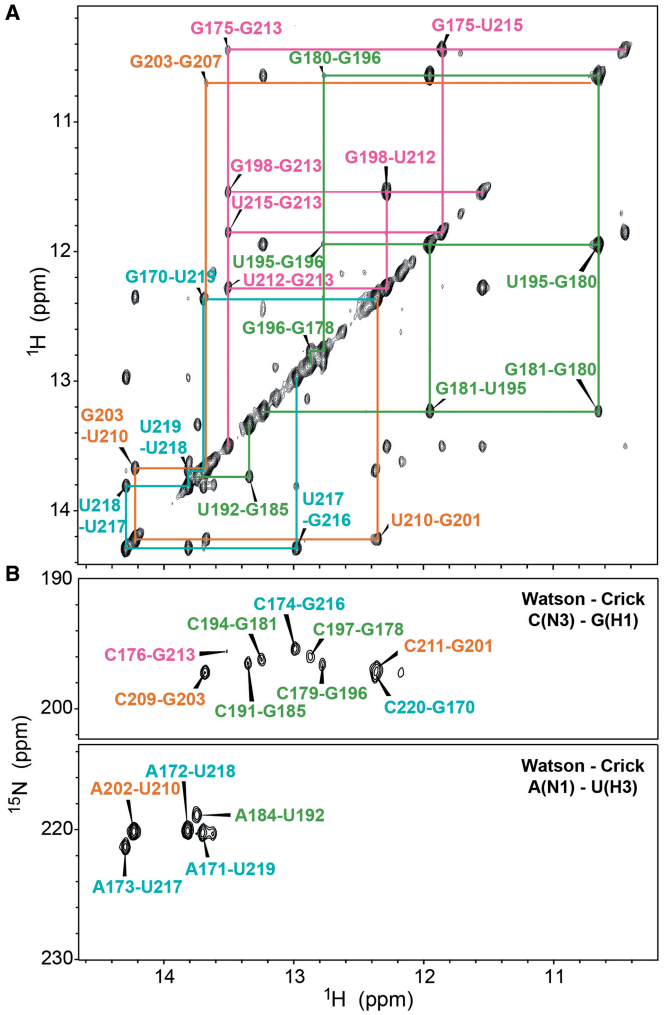

Figure 2. Identification of stem and junction base pairs in mdCR4/5. (**A**) Imino proton region of 2D ^{1}H-^{1}H NOESY spectrum reveals the base pairs of mdCR4/5. NOE cross-peaks between the neighboring imino protons are indicated for P5 (cyan), P6 (green), P6.1 (orange) and the junction (magenta). Three distinct canonical G•U wobble pairs are identified; U195•G180 is in the P6 helix, and G175•U215 and G198•U212 are in the junction. The 2D HNN COSY spectra in (**B**) show a direct correlation between imino protons and their hydrogen-bonded nitrogen atoms in the canonical base pairs (upper box: G-C, lower box: A-U). Cross-peak labels are colored as in (A).

base pairs to U204 (see later in the text). The junction flanked by P5, P6 and P6.1 has a unique base pairing pattern, which is also different from the prediction from phylogenetic analysis. Except for U214, all the imino protons in the junction are involved in hydrogen bonding for base pairing. In the imino region of the 2D NOESY spectrum, a G•U ↔ C-G ↔ G•U sequential cross-peak pattern was observed (Figure 2A, magenta). These base pairs were assigned as G175•U215, C176-G213 and G198•U212, respectively. The remaining junction residues C177, A199, A200 and U214 are unpaired.

Solution structure of mdCR4/5

The structure of mdCR4/5 in 10 mM sodium phosphate (pH 6.3), 100 mM KCl was calculated using 1085

NOE-based distance restraints with 623 inter-residue NOEs. In addition to the NOE restraints, the relative helical orientations were determined by incorporating 77 RDCs in the structure calculations (Table 1 and Supplementary Table S1). The 20 lowest energy structures (Figure 3A) have a root-mean-square deviation of 0.93 Å for all heavy atoms.

The solution structure of mdCR4/5 reveals that the P5, P6 and P6.1 helices are splayed in a Y-shape around a well-structured junction. P6, the connecting junction nucleotides ($J_{5/6}$), and P5 stack sequentially on each other and P6.1 sticks out from the junction to form a V with P6 (Figure 3). Helical angle analysis showed that the helical axis of P5 is inclined ~60° and ~85° relative to P6 and P6.1, respectively, and the P6.1 helix is inclined from the P6 helical axis by ~35° [analyzed by Curves 5.3 (54)]. P5 forms a six WC base-paired A-RNA helix with an average major groove width of ~11 Å, P6 has eight base pairs including one G•U wobble pair and is capped by a 3-nt loop, and P6.1 has three WC base pairs capped by a UUCGU pentaloop. The nucleotides in the three-way junction (G175–C177, G198–A200 and U212–U215) form a unique tertiary structure with unpredicted base pairings (Figure 4). Nucleotides G175 and C176 at the 5/6 junction form canonical wobble and WC base pairs with U215 and G213 in the 5/6.1 junction, respectively. There is a slight twist (~20°) and roll toward the major groove between the C174-G216 base pair of P5 and the bottom junction G175•U215 base pair. C177 and U214 are not involved in base pairing, but stabilize the junction by providing additional base-stacking interactions—the C177 base is stacked on the C176 base, and the U214 base stacks between the bases of G213 and U215 at the top of P5. The phosphate backbone unwinds after C177, and the P6 stem starts at the G178-C197 base pair. Nucleotide U212, which has been predicted to form a terminal A-U base pair in the P6.1 stem of vertebrate TR (18), instead forms a wobble base pair with G198. The bases of G198•U212 are almost perpendicular to those of C176-G213 and G178-C197, resulting in a sharp turn in the phosphate backbone between C197 and G198 (Figure 4A). The conserved A199 and A200 are stacked in between G198•U212 and G201-C211 at the bottom of P6.1.

The mdCR4/5 P6 has a single bulge in the stem and a 3-nt loop (Figure 3B, green, Supplementary Figure S1A). In the P6 helix, U182 is bulged out of the helix and lies perpendicular to the major groove. This universally conserved pyrimidine bulge induces an S-shape kink in the helix backbone. The P6 loop is well structured, with U187 pointing into the minor groove and C188 and G189 into the major groove of the P6 helix.

In the P6.1 pentaloop, the first four nucleotides form a structure similar to the well-characterized UUCG tetraloop, and the last nucleotide, U208 is flipped out (Figure 4C). Unlike the secondary structure prediction, U204 does not form a base pair with U208; no imino protons of U204 and U208, and no inter-residue NOEs between U204 and U208 are observed. Instead the conserved G207 forms a sheared pair with U204 as seen in the UUCG tetraloop (55). Because the U204

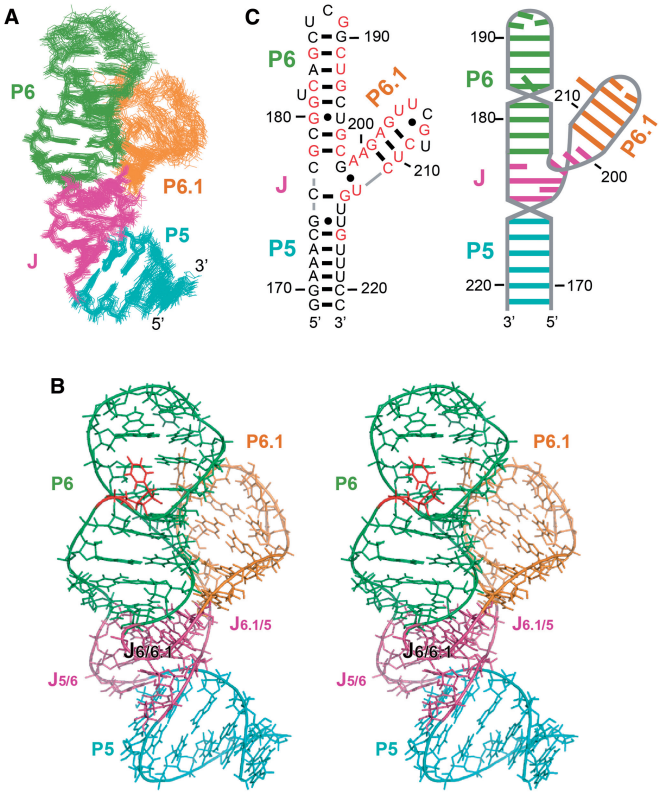

Figure 3. NMR solution structure of mdCR4/5. (**A**) Superposition of the 20 lowest energy structures. (**B**) Stereo view of the lowest energy structure. The junction (J) is classified as $J_{5/6}$, $J_{6/6.1}$ and $J_{6.1/5}$ based on their flanking subdomain names. U182 is highlighted in red. (**C**) Schematic of the tertiary fold of mdCR4/5. The phosphate backbone is indicated by gray line, base pairs are shown as bars and the unpaired nucleotides are shown as half bars. P6, P6.1, P5 and J are colored as in Figures 1 and 2.

2′-hydroxyl proton resonance and the G207 imino proton resonance both had broad linewidths, the expected NOE between the 2′-hydroxyl proton of U204 and G207H1 that is usually observed in the U•G base pair in a UUCG tetraloop were not seen. However, there are inter-residue NOEs between G203 imino and G207 amino and imino protons, and between U204 sugar protons and G207 imino protons, indicative of the sheared U204•G207 base pair seen in the structure. Also as seen in standard UUCG tetraloops, the bases of U204, C206 and G207 stack on U203, U204 and C209, respectively, which increases the stability of P6.1, and the universally conserved U205 is flexibly positioned in the minor groove.

In the mdCR4/5 structure, the P6.1 loop is positioned near the minor groove of P6 near base pair C183-G193, and the flipped out loop nucleotide U208 points toward this groove. This U208 base is dynamic, as evidenced by the intensities of the C6H6 and C5H5 cross-peaks in the ^{1}H-^{13}C HSQC spectra (Figure 5A and Supplementary Figure S2A). There is a sharp turn in the backbone between C197 at the bottom of P6 and G198 (where it forms the G198•U212 base pair), which results in close approach of the phosphodiester backbone from P6 C194 through C197 to P6.1. Only a few long distance (>4.5 Å)

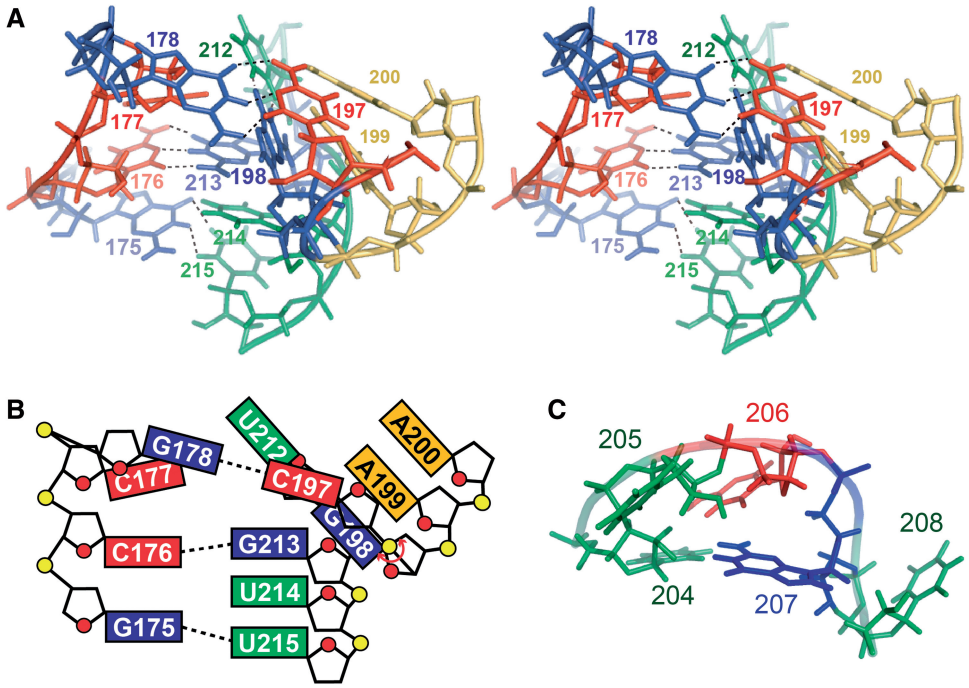

Figure 4. Structure of the three-way junction and P6.1 loop of mdCR4/5. (**A**) Stereo view of the junction in the lowest energy structure. Hydrogen bonds in base pairs are shown by dotted lines. (**B**) Schematic showing the base interactions in the three-way junction. (**C**) Structure of the P6.1 loop. Nucleotides A, U, G and C are colored yellow, green, blue and red, respectively.

NOEs are observed between P6.1 and P6, specifically between C209 base and sugar at the top of P6.1 and the G193 sugar (2′H and 2′OH) in the middle of P6, indicative of close approach of these two helices at least some of the time. However, no NOEs indicative of close-packing of the P6.1 loop or stem against P6 are observed.

Inter-helical dynamics

To investigate how fixed the position of the three stems are relative to each other, we carried out order tensor analysis of the RDC data measured for each stem, which provides information on inter-helical dynamics faster than millisecond time scale. The obtained stem-specific order parameters indicate that both stems P5 and P6.1 have a large amplitude of inter-helical motions relative to P6, as predicted by $\vartheta_{int,P5} = \vartheta_{P5}/\vartheta_{P6} = 0.55 \pm 0.03$ and $\vartheta_{int,P6.1} = \vartheta_{P6.1}/\vartheta_{P6} = 0.30 \pm 0.02$, where ϑ_{int} is the inter-helical GDO as described in 'Materials and Methods' section (46). Using the cone-motional model to estimate the inter-helical dynamics, the determined amplitudes of the motion for P5 and P6.1 are ~49° and ~60°, respectively, relative to P6. These results are consistent with the paucity of inter-helical NOEs observed between P6 and P6.1. We conclude that the determined solution structure is the averaged structure of a dynamic ensemble.

Divalent metal ion binding and Mg^{2+} induced global change of mdCR4/5 structure

Given the dynamic nature of mdCR4/5 in the absence of divalent cations, we investigated whether divalent cations had any effect on its folding. To determine the effect of Mg^{2+} on the mdCR4/5 structure, $MgCl_2$ was added in increments up to 30 mM to a ~0.5 mM mdCR4/5 sample and chemical shift changes were monitored by ^{1}H-^{13}C HSQC experiments (Figure 5 and Supplementary Figure S2A). Most of the residues show small chemical shift changes and/or line broadening upon addition of Mg^{2+} ($\Delta\delta_{max}$ <0.12, Supplementary Figure S2C), indicating fast or intermediate exchange between Mg^{2+}-free and Mg^{2+}-bound conformations. The NOESY cross-peak patterns of the imino proton resonances that identify the base pairs are the same (Supplementary Figure S3) and the NOESY cross-peak patterns for the non-exchangeable spectra are also similar (data not shown), indicating that addition of Mg^{2+} does not cause significant local structural changes or changes in base pairing.

The largest chemical shift perturbations (~0.1 ppm) are observed for resonances from G180 and G189 in the P6 stem and loop, respectively, C177, G213 and U214 in the junction and G201 in P6.1. Except for G189, all of these residues are either directly involved in or near a G•U wobble pair, which is known to be a divalent metal-binding site (56). Mg^{2+} binding could stabilize the flexible loop nucleotide G189. Negligible changes in chemical shifts and signal intensities were observed for ^{1}H-^{13}C resonances from the non–base-paired loop or bulge residues: U187 and C188 in the P6 loop, U205, C206 and U208 in the P6.1 loop and U182. Based on their large relative intensity in ^{1}H-^{13}C HSQC spectra, these residues are more dynamic than base-paired residues. In contrast, relatively large chemical shift changes are observed for non base-paired residues in the junction (C177, U214, A199 and A200).

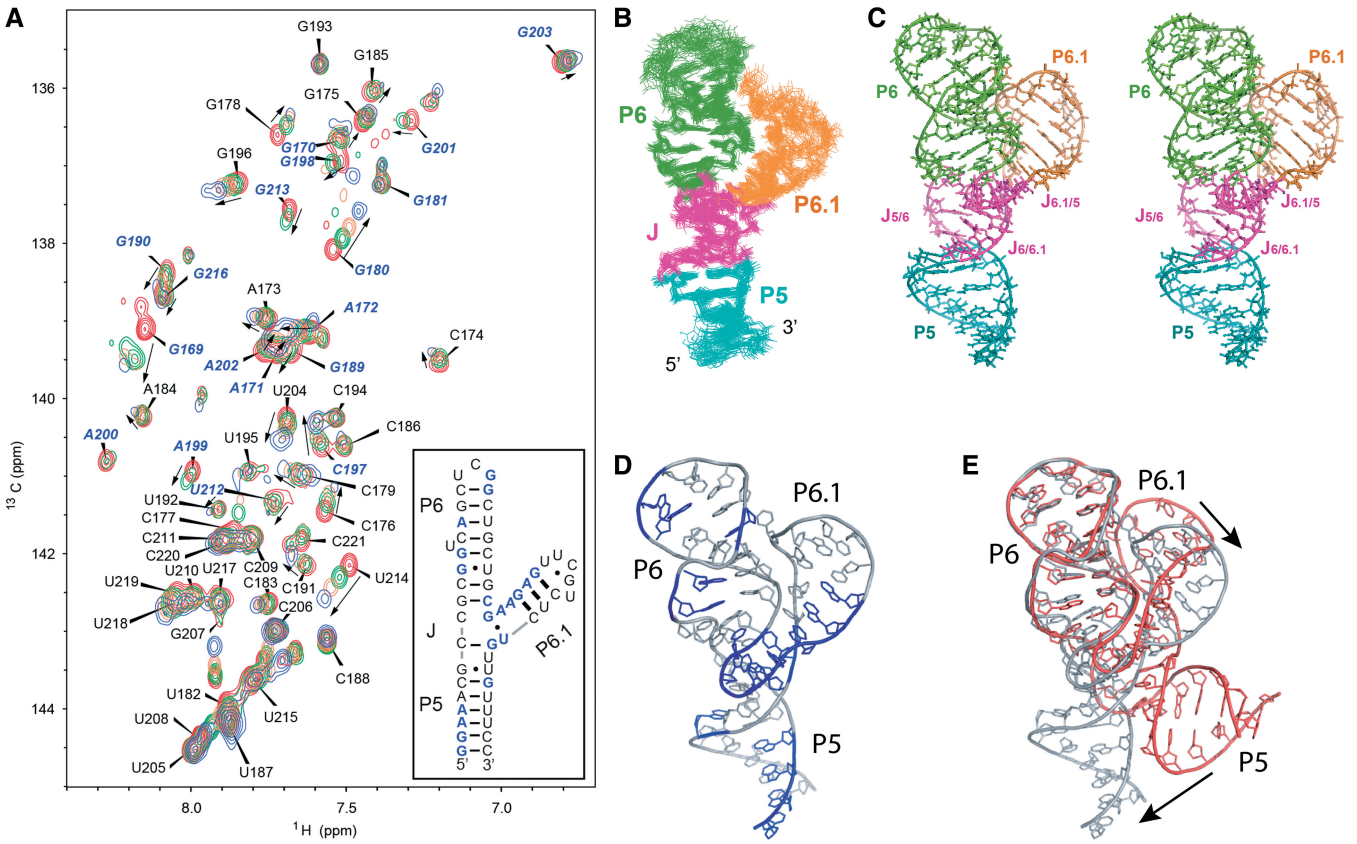

Figure 5. Mg^{2+} ion induces a change in the relative orientations of the P5, P6 and P6.1 helices (**A**) ^{1}H-^{13}C HSQC spectra (C8H8 and C6H6 regions) of mdCR4/5 as a function of added Mg^{2+}. $MgCl_2$ was added to a final concentration (mM) of 0 (red), 5 (green), 10 (yellow) and 20 (blue) with 0.5 mM mdCR4/5. (Inset) Sites of divalent cation localization as determined from paramagnetic line broadening in the presence of Mn^{2+} are colored in blue (see also Supplementary Figure S2) (**B**) Superposition of the 20 lowest energy structures of mdCR4/5 structure refined using RDCs measured for the sample with Mg^{2+} (mdCR4/5-Mg^{2+}). (**C**) Stereo-view of the lowest energy structure of mdCR4/5-Mg^{2+}. For B and C, subdomains are colored as in Figure 2. (**D**) Lowest energy structure of mdCR4/5-Mg^{2+} with regions of divalent cation localization shown in blue. (**E**) Superposition of lowest energy structures of mdCR4/5 (red) and mdCR4/5-Mg^{2+} (gray). Superposition is on P6 residues 178–197, and arrows indicate changes in the positions of P6.1 and P5.

To further characterize the divalent metal ion binding sites in mdCR4/5, we used paramagnetic line broadening experiments with Mn^{2+}. Mn^{2+} is a sensitive and direct probe for identifying metal binding residues in RNA using NMR (33). The paramagnetic Mn^{2+} induces a distance dependent (r^{-6}) line broadening in resonances of residues proximal to bound Mn^{2+} that results in reduced NMR signals. Addition of 40 μM Mn^{2+} results in line broadening of resonances, observed in ^{1}H-^{13}C HSQC spectra, of residues near the G•U base pairs in the P6 stem (G180 and G181) and in the junction (C197, G198, U212 and G213), consistent with the results with Mg^{2+} (Supplementary Figure S2B). Line broadening was observed around tandem purine bases throughout mdCR4/5: G169 through A172 in P5, A184, G189 and G190 in P6, and A199 through G203 in P6.1. The latter region is also where there is a sharp turn in the backbone between P6 and $J_{6/6.1}$, and divalent metal ion binding could help reduce the electrostatic repulsion between the phosphate backbones at this turn.

We next investigated the effect of Mg^{2+} on the overall folding of mdCR4/5. Titration of Mg^{2+} up to 30 mM did not saturate the chemical shift perturbations, suggesting weak binding affinity of Mg^{2+} to RNA. Significant NMR line broadening was also observed at higher Mg^{2+} concentrations, which may be largely attributed to non-specific Mg^{2+} binding to RNA. Consequently, further analysis at higher saturating Mg^{2+} concentration was not possible due to difficulties in resonance assignments and NOE-based distance measurements in NMR spectra. Therefore, we calculated a model structure of mdCR4/5 in the presence of Mg^{2+} starting with the NMR restraints previously used for mdCR4/5 without Mg^{2+} and a set of RDC data obtained for the sample in the presence of Mg^{2+}. As described earlier in the text, as the RNA retains the same base pairs and NOESY cross-peak patterns in the presence of Mg^{2+}, we conclude that the local RNA structures such as the helices, P6 and P6.1 loops, the U182 bulge and the junction are not significantly changed. Based on this assumption, the structure was calculated using the initial structure calculated without RDCs for the sample without Mg^{2+} as a starting structure and then refining with a set of 60 RDCs measured on the sample with Mg^{2+}, prepared as described in the 'Materials and Methods' section (see Supplementary Table S2 for structural statistics).

The calculated model structure of mdCR4/5 with Mg^{2+} showed a change in the relative orientations of P5, P6 and P6.1 (Figure 5E). In the presence of Mg^{2+}, P5 and P6 are nearly coaxial, with only ~2.5° helical angle deviation between each other. The angle between the helical axes of P6 and P6.1 increases from ~35° to ~65° (Figure 5B and C). The internal GDO values for P5 and P6.1 are $\vartheta_{int,P5} = \vartheta_{P5}/\vartheta_{P6} = 0.57 \pm 0.02$ and $\vartheta_{int,P6.1} = \vartheta_{P6.1}/\vartheta_{P6} = 0.49 \pm 0.02$, respectively. In the cone-helical motion model as described earlier, the amplitude of the inter-helical dynamics are ~47° for P5 and ~50° for P6.1 relative to P6, respectively. This indicates that both P5 and P6.1 remain mobile relative to P6 even in the presence of Mg^{2+} ions, albeit with a lower amplitude of motions.

Effect of mdCR4/5 junction nucleotide substitutions on telomerase activity

To identify if any structural elements of the mdCR4/5 are correlated with telomerase function, we incorporated nucleotide substitutions into the full-length mdTR and performed direct telomerase activity assays *in vitro* using mdTERT translated in RRL (Figure 6). We first investigated the effect of base pairing in the junction using a compensatory mutagenesis strategy (Figure 6B, lane 10–12). In the NMR solution structure of mdCR4/5, three base pairs, G175-U215, C176-G213 and G198-U212, are formed in or next to the junction (Figure 6A, boxes). Simultaneous disruption of all three base pairs by the mutations G175U/C176U/G198U or U212A/G213A/U215A (Figure 6B, lane 10–11) surprisingly showed no effect or slightly increased telomerase activity. Compensatory mutations that potentially restored base pairing also had little effect on activity *in vitro* (Figure 6B, lane 12). Hence, formation of the base pairs in the junction of mdCR4/5 does not appear to be required for catalytic activity.

The conserved P6.1 terminal A-U base pair in vertebrate TR (A200 and U212 in mdCR4/5) (18,20) predicted by phylogenetic comparison does not form in the solution structure of mdCR4/5. Instead, A200 is unpaired and U212 forms a U•G base pair with G198. Thus, we next examined whether an A200-U212 base pair might be important for mdCR4/5 folding and telomerase catalytic activity even though they are not paired in the solution structure. The compensatory substitutions A200U/U212A have no effect on telomerase activity (Figure 6, lane 7). The A200C substitution that would disrupt the previously predicted A200-U212 base pair decreased activity somewhat to 79% (Figure 6B, lane 5). The U212G that would disrupt A200-U212 would also disrupt the G198-U212 base pair seen in the solution structure and decreases telomerase activity to 60% (Figure 6B, lane 9). However, the U212C substitution that would disrupt A200-U212 and at the same time replace the G198-U212 base pair with a G-C base pair decreases activity to 20% (Figure 6B, lane 8). Finally, G198U, which would change the G198-U212 base pair to U-U but leave the A200-U212 base pair intact decreases activity only slightly to 87% (Figure 6B, lane3), consistent with the lack of effect of this mutation in the context of disruption

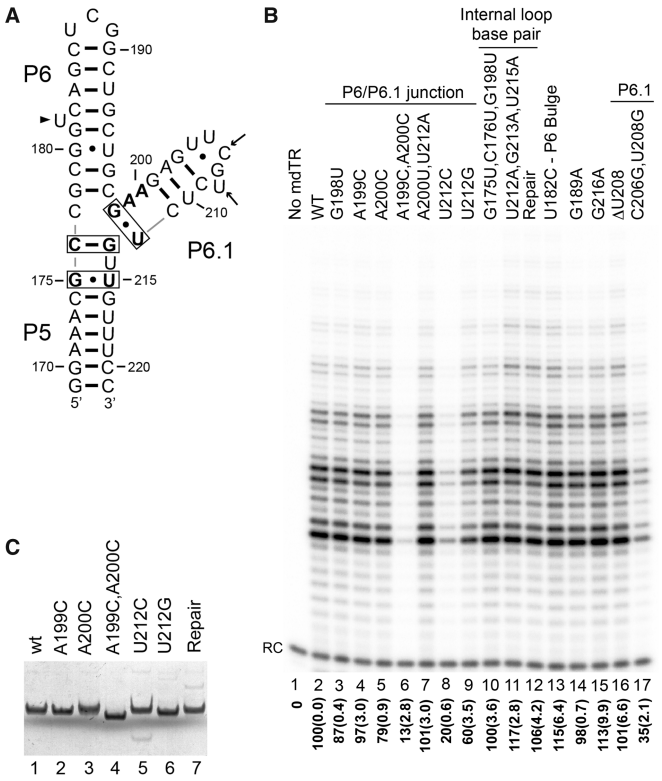

Figure 6. Telomerase activity assays of mdCR4/5 mutants. (**A**) Secondary structure of mdCR4/5 with junction nucleotides that were changed highlighted in bold, base pairs that were changed shown boxed, substitutions in the P6.1 loop indicated by arrows and the U182C bulge mutation in P6 indicated by a wedge. (**B**) Effect of mdCR4/5 mutations on telomerase activity. Except for lane 1, full-length mdTERTs synthesized in RRL were assembled with full-length mdTR mutants. Relative activities compared with wild-type (WT) mdTR are shown in bold below the lane numbers and errors in the parentheses. (**C**) Electrophoretic mobility shift assay of the mdCR4/5 domain mutants. 'Repair' in (B) and (C) stands for simultaneous substitutions of G175U-C176U-G198U and U212A-G213A-U215A.

of all three junction nucleotides discussed earlier in the text. Thus, it seems likely that the G198-U212 base pair seen in the solution structure is not present in or important for the complex formation with mdTERT TRBD, whereas the A200-U212 base pair predicted by phylogeny may be. This would be the terminal base pair of the P6.1 helix and if present would most likely play a role in stabilizing the helix.

A199 is a universally conserved nucleotide linking P6 and P6.1. Nevertheless, the substitution A199C had no significant effect on telomerase activity (Figure 6B, lane 5). In previous studies on hTR, the same result was shown for the mutation on the corresponding (A301) residue, where A301U (see Supplementary Figure S1C, FL CR4/5 for the sequence identity) substitution had only a minor impact on hTR-hTERT (human TERT) interaction and catalytic activity (29). In contrast to the negligible and small effects of individual A199C and A200C substitutions, respectively, simultaneous substitution of A199C/A200C significantly reduced catalytic activity (Figure 6B lane 6). However, this decrease can be explained by the misfolding of mdCR4/5 on A199C/A200C double mutation. Although

secondary structure prediction by Mfold (57) (Supplementary Figure S4) does not calculate the internal loop base pairing, it does predict the formation of the catalytically essential P6.1. Mfold generates only one conformation for A199C with the P6.1 stem-loop intact (Supplementary Figure S4B), but in the case of the A200C substitution two mdCR4/5 conformers were generated, where the less thermodynamically stable structure does not form the catalytically essential P6.1 hairpin (Supplementary Figure S4C, right). The presence of an inactive conformer in A200C could explain the decrease in telomerase activity compared with the A199C mutant. For the A199C/A200C double mutant, mdCR4/5 is predicted to fold into a stem-loop-stem hairpin structure, where the substituted C199 and C200 are base-paired with G190 and G189, respectively. In this case, the P6.1 stem-loop is completely lost, which could explain the almost complete loss of telomerase activity (Figure 6B, lane 6 and Supplementary Figure S4D). Consistent with the predicted change in overall fold, the A199C/A200C substituted TR migrates faster than the other wild-type and mutant mdCR4/5 RNAs in a non-denaturing gel (Figure 6C).

In summary, none of the mutations in junction residues including disruption of the three base pairs seen in the solution structure and changes in the conserved nucleotides decrease telomerase activity <75%, with the exception of U212G (60%) and U212C (20%). These mutations change both the observed G198-U212 and the predicted U212-A200 base pairs, so it is difficult to rationalize their effect. Previous studies of mdCR4/5 show that the minimal domain for >75% TRBD binding comprises nucleotides 176–214 (53). This minimal domain retains P6, P6.1 and the junction nucleotides but deletes P5. Similarly, for hCR4/5, residues 253–318, which comprise P6ab and P6.1 with 3 and 4 single-stranded nucleotides at the 5′ and 3′ ends, respectively, are sufficient to confer 50–70% activity relative to full-length hCR4/5 when combined *in trans* with the core domain and reconstituted with hTERT (32). Therefore, the role of the internal loop residues and P5 residues is likely to help position P6, P6.1 and flanking residues for interaction with TERT in the context of the full-length TR. The sequence of the internal loop residues including the conserved nucleotides may also be important for proper RNA folding to maintain the P6.1 and P6 stem-loops.

Effect of mdCR4/5 loop and bulge substitutions on telomerase activity

We next investigated the importance of the conserved P6 bulge (U182), found in all vertebrate TRs, and the loop G189 residues for telomerase activity. In the previous UV cross-linking studies of mdCR4/5 with the TRBD of mdTERT, the bulge U182 and P6 loop residue U187 were identified as being intimately involved in TRBD binding (53). Deletion of U182 or substitution to U182C eliminated or reduced by ~3-fold, respectively, the binding of mdCR4/5 to mdTERT TRBD (53). However, our activity assays using full-length TR and TERT showed that U182C substitution slightly increased telomerase activity (Figure 6B, lane 13). The single bulge nucleotide in the P6 stem is not sequence conserved, but is present in all vertebrate TRs except for one murine species (50). Because the bulge U182 base is looped out into the major groove of the P6 helix and has no tertiary interactions with other parts of mdCR4/5, the U182C substitution is unlikely to change the P6 stem structure. Thus, although previous cross-linking and binding assays of CR4/5 with the isolated TRBD from TERT indicate that U182 may form specific contacts with TRBD, changing it to C does not appear to reduce the interaction with full-length TERT sufficiently to affect activity.

G189, which is 100% conserved in teleost fish and conserved as U or G for all identified vertebrate TR (42 species), is near the P6 loop residue U187 that cross-links to mdTRBD. Our assays show that G189A substitution has no effect on activity (Figure 6B lane 14). Substitution of all three loop-residues (U187/C188/G189) with a GAAA tetraloop has been shown to nearly abolish binding of mdCR4/5 to the isolated TRBD (53). As with U182, this suggests that G189 may be involved in TRBD binding, but is not directly involved in telomerase activity. The differences in binding and telomerase activity can be explained by the fact that telomerase activity was assayed in the context of full-length TR and TERT, whereas the binding assays were performed with mdCR4/5 and TRBD only. Modeling based on the cross-linking studies predicts that the P6.1 loop interacts with the TERT CTE domain that is absent in the TRBD binding assay, which would provide additional binding affinity (53).

Finally, we tested the importance of the P6.1 loop residues for telomerase activity. P6.1 is an essential and conserved stem-loop in the CR4/5 domain of vertebrate TR (28,29,32,53,58). Base pairing of P6.1 is critical for TRBD binding, which is required for telomerase activity, whereas the loop is important for catalysis but not for TERT binding (28,29). A previous study showed that simultaneous mutation of the two absolutely conserved P6.1 loop nucleotides, U205A/G207C in mdCR4/5, did not affect TERT binding affinity, but decreased activity to 15% *in vitro* (53). The same mutation in murine TR eliminated telomerase activity (28). In the mdCR4/5 P6.1 loop structure, 204-UUCG-207 forms a tetraloop, and U208 is flipped out toward the minor groove of P6 (Figure 4C). Our results show that deletion of U208 has no effect on activity (Figure 6B, lane16). Substitution of U205A/G207C would be expected to result in structural changes to the loop, whereas deletion of the bulge U208 should have little effect. These results show that the non-conserved U208 is not required for telomerase activity.

Interestingly, the double mutation of the non-conserved nucleotides C206G/U208G that changes the sequence of the P6.1 loop to that of hTR P6.1 (UUGGG) decreased telomerase activity to ~35% of wild-type mdTR. Because hTR P6.1 loop folds into a structure different from mdTR (Supplementary Figure S1B) (58), we propose that telomerase activation by P6.1 is loop sequence specific.

DISCUSSION

In this work, we present the first structure of a complete CR4/5 domain, determined by NMR. Medaka TR has the

simplest and smallest CR4/5 domain and contains all the conserved elements, i.e. a three-way junction flanked by three helices P5, P6 and P6.1. The solution structure of mdCR4/5 revealed that P6 and P6.1 fold into a V-shape conformation, and P6 is stacked on P5 through the three-way junction. Although the TERT-free CR4/5 forms a three-way junction structure that is stabilized by well-ordered base pairs and stacking interactions at the junction region, none of these structural elements appear to be essential for telomerase activity. These results, together with the observed intrinsic dynamics of the three stems relative to each other, suggest that the three-way junction of mdCR4/5 could undergo a conformational rearrangement that transitions the TERT-free three-way junction and relative positions of P6 and P6.1 into the telomerase-active conformation.

Based on comparison of TR sequences from fungi, yeasts and vertebrates, a consensus three-way junction secondary structure was proposed, with a single conserved A nucleotide (A199 in mdCR4/5) connecting P6.1 with P6 and conserved UG connecting P6.1 to P5 (U212, G213 in mdCR4/5; in hCR4/5 there are two additional non-conserved nucleotides), and a conserved A-U base pair at the bottom of P6.1 (A200-U212 in mdCR4/5) (20). The predicted three-way junction secondary structure is different from what we observe in solution. In particular, the A200-U212 base pair does not form, rather U212 pairs with G198. Formation of this base pair results in a sharp turn in the backbone between P6 and P6.1. G•U base pairs seem to be important for stabilizing RNA backbone turns (56), e.g. in many junctions and at the sharp turns in $tRNA^{Phe}$ (59). The conserved purines AA GAG (199–203) may also contribute to stabilization of the close contact between phosphates at the tight turn between P6 and $J_{6/6.1}$, as these consecutive purines were identified as divalent cation binding sites. Given that disruption of the three-way junction base pairs observed in the mdCR4/5 solution structure had little effect on telomerase activity *in vitro* and the phylogenetic predictions for junction base pairing, it is likely that the free mdCR4/5 structure represents a stable inactive state. It can be envisioned that shifting G198-U212 to A200-U212 could release the tight turn between P6 and P6.1 helices and the compact base stacking from G198 to A200, which would induce the repositioning of P6.1 and P6 and/or open the junction region that might be important for interaction with TERT. We note that other rearrangements of the junction pairing, including tertiary interactions, might also occur in the complex of mdCR4/5 with TERT.

Previous telomerase activity assays with mutations on the conserved residues in the internal loop of hCR4/5 have suggested that some of the conserved three-way junction residues are involved in direct tertiary contact with TERT, and others may be required for organizing the internal loop conformation for TERT binding (29). Because the most important structural element in CR4/5 for TERT binding as well as telomerase activity is P6.1 formation (25), it is probable that the three other conserved nucleotides (A199, G213 and U214 in mdCR4/5) observed in the internal loop of CR4/5 might function to help ensure the formation of an active three-way junction structure and proper formation of P6.1.

Outside the junction, there are three regions of CR4/5 that are important in terms of structure, TRBD binding and activity: a conserved pyrimidine bulge in P6, the P6 loop and the P6.1 loop. The structure of mdCR4/5 P6.1 loop is well defined but is different from that of hCR4/5 (Supplementary Figure S1B). The hCR4/5 P6.1 loop contains pseudo-uridines (Ψ) at two potential positions (306 and 307), and structures for it have been determined both with and without these Ψs (32,58). In the absence of Ψ, the first three bases (U306, U307 and G308) in P6.1 loop shows some structural similarity to mdCR4/5 P6.1 loop, where U306 and G308 are positioned in the major groove with G308 stacking on U306, and the conserved U307 points out on the minor groove side. However, unlike mdCR4/5 P6.1 loop, the first U306 forms a wobble pair with G310 and the conserved G309 is flipped out to the minor groove. The structure of Ψ-modified hCR4/5 P6.1 loop shows different orientations of loop residues than the unmodified P6.1 loop, where Ψ306-G310 base pair is intact but G308 and G309 are all in the minor groove side and Ψ307 is stacked above Ψ308. Based on these structural differences, we propose that telomerase activation by P6.1 is loop sequence specific.

In the middle of P6, the conserved pyrimidine bulge nucleotide U182 sticks out from the major groove and results in the S-shape twist of the phosphate backbone. A similar kink is also observed at the conserved bulge (C262) in hCR4/5 P6 structure (Supplementary Figure S1A) (27). This unique structural feature, along with the deleterious effect on TRBD binding of deletion of mdCR4/5 U182 (53), and the minimal effect of U182→C substitution on activity, suggests that the conserved single pyrimidine bulge in vertebrate CR4/5 may affect proper positioning of the P6 loop for effective TERT binding or it may be important for TERT binding.

To gain insights into the potential interaction between the CR4/5 and TERT, we made a model structure of the mdCR4/5–mdTRBD complex based on the UV cross-linking data shown in a previous study, where three uridines, U182 in P6, U187 in P6 loop and U205 in P6.1 loop, were found to cross-link to mdTRBD residues Y503, F355 and W477, respectively (see 'Materials and Methods' section) (Supplementary Figure S5, left) (53). We tested modeling the complex using both Mg^{2+}-free and -bound mdCR4/5 structures by manually placing each of these uridine bases to be within ~10 Å from the corresponding cross-linked amino acid. Interestingly, the Mg^{2+}-bound CR4/5 structure, which has an inter-helical angle between P6 and P6.1 of 65° compared with 35° to for the Mg^{2+}-free CR4/5, fits better than the Mg^{2+}-free one, in a way that all three cross-linking pairs can be better placed within the ~10 Å distance constraints. Given the mutational data and the intrinsic positional flexibility of the helices, it is plausible that the inter-helical angle between P6 and P6.1 can increase further in the mdCR4/5-TERT complex, which would bring mdCR4/5 closer to TRBD and allow some tertiary interactions between the conserved nucleotides in the junction and TRBD. In the

mdCR4/5-mdTRBD model, the P6.1 loop is close to the CTE domain of TERT, as predicted from the cross-linking study, where it could function to help bring the TRBD and CTE together to close the TERT ring (53). Consistent with an interaction at the CTE, substitution of the two conserved P6.1 loop nucleotides, U205A and G207C in mdCR4/5, has little effect on TRBD binding *in vitro*, but a large effect on activity (53). It has been shown in hCR4/5 that formation of the P6.1 stem but not its sequence is important for TERT binding and telomerase activity, whereas the P6.1 loop sequence is critical to telomerase activity but not for assembly of CR4/5 with TERT (29), indicating that the P6.1 stem plays a role for positioning the P6.1 loop close to TERT in a manner required for telomerase activity. Notably, in the EM structure of *Tetrahymena* telomerase holoenzyme, the model of TERT and TR fit into the EM maps also placed TR loop 4, the functional equivalent of P6.1, in contact with the CTE at the CTE/TRBD interface (Supplementary Figure S5, right) (60). The human telomerase P6.1 loop differs in sequence from that of medaka, and changing the sequence of the medaka telomerase P6.1 loop to that of hTR, which changes non-conserved nucleotides but not the conserved U205 and G207 (U307 and G309 in hTR), has a large effect on activity. Thus it appears that TERT and TR STE loop (P6.1 loop in vertebrates) have co-evolved for optimization of this interaction for activity.

Three-way junctions with stable and locked structures are important for many functional RNAs, such as some riboswitches that recognize specific small metabolites or the hammerhead ribozyme that performs precise phosphodiester bond cleavage (61,62). In these three-way junctions, the positions of the helices are often fixed by distal inter-stem or stem-loop interactions through base stacking or base pairing interactions (63,64). In contrast, in mdCR4/5 no tertiary interactions indicative of close packing of P6 and P6.1 were identified. The isolated individual P6 and P6.1 hairpins showed similar NOE patterns in the loop region compared with the ones from mdCR4/5 (data not shown), indicating that the stem-loops have the same structure in the isolated forms and in the three-way junction. The absence of inter-helical tertiary interactions results in the P6, P6.1 and P5 stems being relative dynamic with respect to each other, even in the presence of Mg^{2+}. The dynamic nature of the structure of mdCR4/5 and the generally small effects of nucleotide substitutions and compensatory mutations in the junction on telomerase activity taken together indicate that the three-way junction and P5 function as a scaffold that allows P6 and P6.1 positional flexibility for interaction with TERT and correct positioning of the P6.1 loop for contribution to catalysis. In ciliates, the three-way junction is replaced by a stem-loop with a conserved bulge. A large (105°) bend in stem-loop 4, induced in the holoenzyme by binding of the LARP7 protein p65, is required for proper assembly of TERT with TR (65,66), where the conformational change positions loop 4 to interact with TERT (60). It is possible that the three-way junction in vertebrate CR4/5 serves the same role, with tight interactions between P6 and TRBD providing initial stabilization of CR4/5-TERT interactions and the flexible junction rearranging to allow interaction between P6.1 and TERT at the TRBD-CTE interface.

COORDINATE DEPOSITION

Coordinates for the 20 lowest energy structures of mdCR4/5 have been deposited in the Protein Data Bank under accession code 2MHI, and chemical shifts have been deposited in the BioMagResBank under accession code 19634.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank E.J. Miracco for helpful comments.

FUNDING

National Institutes of Health [GM48123 to J.F.]; Korea Institute of Science and Technology [2V02900, 2E23810 (in part) to N.-K.K.]. Funding for open access charge: National Institute of Health [GM48123].

Conflict of interest statement. None declared.

REFERENCES

1. Blackburn,E.H. and Collins,K. (2011) Telomerase: an RNP enzyme synthesizes DNA. *Cold Spring Harb. Perspect. Biol.*, **3**, a003558.
2. Nandakumar,J. and Cech,T.R. (2013) Finding the end: recruitment of telomerase to telomeres. *Nat. Rev. Mol. Cell Biol.*, **14**, 69–82.
3. Harley,C.B., Futcher,A.B. and Greider,C.W. (1990) Telomeres shorten during ageing of human fibroblasts. *Nature*, **345**, 458–460.
4. Palm,W. and de Lange,T. (2008) How shelterin protects mammalian telomeres. *Annu. Rev. Genet.*, **42**, 301–334.
5. Bernardes de Jesus,B. and Blasco,M.A. (2013) Telomerase at the intersection of cancer and aging. *Trends Genet.*, **29**, 513–520.
6. Armanios,M. (2013) Telomeres and age-related disease: how telomere biology informs clinical paradigms. *J. Clin. Invest.*, **123**, 996–1002.
7. 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.
8. Collins,K. (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat. Rev. Mol. Cell Biol.*, **7**, 484–494.
9. Theimer,C.A. and Feigon,J. (2006) Structure and function of telomerase RNA. *Curr. Opin. Struct. Biol.*, **16**, 307–318.
10. Lingner,J. and Cech,T.R. (1998) Telomerase and chromosome end maintenance. *Curr. Opin. Genet. Dev.*, **8**, 226–232.
11. Mason,M., Schuller,A. and Skordalakes,E. (2011) Telomerase structure function. *Curr. Opin. Struct. Biol.*, **21**, 92–100.
12. Egan,E.D. and Collins,K. (2012) Biogenesis of telomerase ribonucleoproteins. *RNA*, **18**, 1747–1759.
13. Podlevsky,J.D. and Chen,J.J. (2012) It all comes together at the ends: telomerase structure, function, and biogenesis. *Mutat. Res.*, **730**, 3–11.
14. Robart,A.R. and Collins,K. (2011) Human telomerase domain interactions capture DNA for TEN domain-dependent processive elongation. *Mol. Cell*, **42**, 308–318.
15. Lai,C.K., Mitchell,J.R. and Collins,K. (2001) RNA binding domain of telomerase reverse transcriptase. *Mol. Cell. Biol.*, **21**, 990–1000.

16. Wyatt,H.D., West,S.C. and Beattie,T.L. (2010) InTERTpreting telomerase structure and function. *Nucleic Acids Res.*, **38**, 5609–5622.
17. Autexier,C. and Lue,N.F. (2006) The structure and function of telomerase reverse transcriptase. *Annu. Rev. Biochem.*, **75**, 493–517.
18. Chen,J.L., Blasco,M.A. and Greider,C.W. (2000) Secondary structure of vertebrate telomerase RNA. *Cell*, **100**, 503–514.
19. 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.
20. Kuprys,P.V., Davis,S.M., Hauer,T.M., Meltser,M., Tzfati,Y. and Kirk,K.E. (2013) Identification of telomerase RNAs from filamentous fungi reveals conservation with vertebrates and yeasts. *PLoS One*, **8**, e58661.
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. 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.
23. 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.
24. 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. USA*, **110**, 10970–10975.
25. Mitchell,J.R. and Collins,K. (2000) Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. *Mol. Cell*, **6**, 361–371.
26. Brown,Y., Abraham,M., Pearl,S., Kabaha,M.M., Elboher,E. and Tzfati,Y. (2007) A critical three-way junction is conserved in budding yeast and vertebrate telomerase RNAs. *Nucleic Acids Res.*, **35**, 6280–6289.
27. Leeper,T.C. and Varani,G. (2005) The structure of an enzyme-activating fragment of human telomerase RNA. *RNA*, **11**, 394–403.
28. Chen,J.L., Opperman,K.K. and Greider,C.W. (2002) A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic Acids Res.*, **30**, 592–597.
29. Robart,A.R. and Collins,K. (2010) Investigation of human telomerase holoenzyme assembly, activity, and processivity using disease-linked subunit variants. *J. Biol. Chem.*, **285**, 4375–4386.
30. Xie,M., Mosig,A., Qi,X., Li,Y., Stadler,P.F. and Chen,J.J. (2008) Structure and function of the smallest vertebrate telomerase RNA from teleost fish. *J. Biol. Chem.*, **283**, 2049–2059.
31. 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. USA*, **102**, 5958–5963.
32. Kim,N.K., Theimer,C.A., Mitchell,J.R., Collins,K. and Feigon,J. (2010) Effect of pseudouridylation on the structure and activity of the catalytically essential P6.1 hairpin in human telomerase RNA. *Nucleic Acids Res.*, **38**, 6746–6756.
33. Feigon,J., Butcher,S.E., Finger,L.D. and Hud,N.V. (2001) Solution nuclear magnetic resonance probing of cation binding sites on nucleic acids. *Method Enzymol.*, **338**, 400–420.
34. Wohnert,J., Ramachandran,R., Gorlach,M. and Brown,L.R. (1999) Triple-resonance experiments for correlation of H5 and exchangeable pyrimidine base hydrogens in (13)C,(15)N-labeled RNA. *J. Magn. Reson.*, **139**, 430–433.
35. Dingley,A.J., Masse,J.E., Peterson,R.D., Barfield,M., Feigon,J. and Grzesiek,S. (1999) Internucleotide scalar couplings across hydrogen bonds in Watson-Crickand Hoogsteen base pairs of a DNA triplex. *J. Am. Chem. Soc.*, **121**, 6019–6027.
36. 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.
37. 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.
38. Cromsigt,J., van Buuren,B., Schleucher,J. and Wijmenga,S. (2001) Resonance assignment and structure determination for RNA. *Method Enzymol.*, **338**, 371–399.
39. Cavanagh,J. and Rance,M. (1990) Sensitivity improvement in isotropic mixing (TOCSY) experiments. *J. Magn. Reson.*, **88**, 72–85.
40. 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.
41. Varani,G., Aboulela,F. and Allain,F.H.T. (1996) NMR investigation of RNA structure. *Prog. Nucl. Magn. Reson. Spectrosc.*, **29**, 51–127.
42. Legault,P., Jucker,F.M. and Pardi,A. (1995) Improved measurement of 13C, 31P J coupling constants in isotopically labeled RNA. *FEBS Lett.*, **362**, 156–160.
43. Meissner,A. and Sorensen,O.W. (1999) The role of coherence transfer efficiency in design of TROSY-type multidimensional NMR experiments. *J. Magn. Reson.*, **139**, 439–442.
44. Zhang,Q., Kim,N.K., Peterson,R.D., Wang,Z. and Feigon,J. (2010) Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. *Proc. Natl Acad. Sci. USA*, **107**, 18761–18768.
45. Hansen,A.L. and Al-Hashimi,H.M. (2006) Insight into the CSA tensors of nucleobase carbons in RNA polynucleotides from solution measurements of residual CSA: towards new long-range orientational constraints. *J. Magn. Reson.*, **179**, 299–307.
46. Prestegard,J.H., al-Hashimi,H.M. and Tolman,J.R. (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Q. Rev. Biophys.*, **33**, 371–424.
47. Musselman,C., Pitt,S.W., Gulati,K., Foster,L.L., Andricioaei,I. and Al-Hashimi,H.M. (2006) Impact of static and dynamic A-form heterogeneity on the determination of RNA global structural dynamics using NMR residual dipolar couplings. *J. Biomol. NMR*, **36**, 235–249.
48. Koradi,R., Billeter,M. and Wuthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph. Model.*, **14**, 51–55, 29–32.
49. Tolman,J.R., Flanagan,J.M., Kennedy,M.A. and Prestegard,J.H. (1997) NMR evidence for slow collective motions in cyanometmyoglobin. *Nat. Struct. Biol.*, **4**, 292–297.
50. Podlevsky,J.D., Bley,C.J., Omana,R.V., Qi,X.D. and Chen,J.J.L. (2008) The telomerase database. *Nucleic Acids Res.*, **36**, D339–D343.
51. Harkisheimer,M., Mason,M., Shuvaeva,E. and Skordalakes,E. (2013) A motif in the vertebrate telomerase N-terminal ;linker of TERT contributes to RNA binding and telomerase activity and processivity. *Structure*, **21**, 1870–1878.
52. Mitchell,M., Gillis,A., Futahashi,M., Fujiwara,H. and Skordalakes,E. (2010) Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat. Struct. Mol. Biol.*, **17**, 513–518.
53. 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. USA*, **108**, 20333–20338.
54. Lavery,R. and Sklenar,H. (1989) Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.*, **6**, 655–667.
55. Allain,F.H. and Varani,G. (1995) Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.*, **250**, 333–353.
56. Varani,G. and McClain,W.H. (2000) The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep.*, **1**, 18–23.
57. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
58. 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.

59. Clark,B.F. and Klug,A. (1975) Structure and function of tRNA with special reference to the three dimensional structure of yeast phenylalanine tRNA. *Proc. Tenth FEBS Meet.*, **39**, 183–206.
60. 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.
61. Chi,Y.I., Martick,M., Lares,M., Kim,R., Scott,W.G. and Kim,S.H. (2008) Capturing hammerhead ribozyme structures in action by modulating general base catalysis. *PLoS Biol.*, **6**, e234.
62. Serganov,A. and Nudler,E. (2013) A decade of riboswitches. *Cell*, **152**, 17–24.
63. Lescoute,A. and Westhof,E. (2006) Topology of three-way junctions in folded RNAs. *RNA*, **12**, 83–93.
64. Serganov,A. and Patel,D.J. (2012) Molecular recognition and function of riboswitches. *Curr. Opin. Struct. Biol.*, **22**, 279–286.
65. 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.
66. Stone,M.D., Mihalusova,M., O'Connor,C.M., Prathapam,R., Collins,K. and Zhuang,X. (2007) Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature*, **446**, 458–461.