July 17, 2010

Supporting Information

Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA

Qi Zhang[a], Nak-Kyoon Kim[a], Robert D. Peterson[a,b], Zhonghua Wang[a] and Juli Feigon[a,b,1]

[a]Department of Chemistry and Biochemistry, and [b]the Molecular Biology Institute,

PO Box 951569, University of California, Los Angeles, CA 90095-1569

[1]To whom correspondence should be addressed.

Department of Chemistry and Biochemistry

PO Box 951569, UCLA

Los Angeles, CA 90095-1569

Tel: 310-206-6922, Fax: 310-825-0982

E-mail: feigon@mbi.ucla.edu

SI Text

NMR Sample Preparation. Unlabeled, uniformly ^{13}C,^{15}N-labeled, and base-specifically (G, C, A, or U) ^{13}C,^{15}N-labeled P2ab and P2a1a samples, base-specifically GC-^{13}C,^{15}N-labeled E-AU-P2ab and AU-^{13}C,^{15}N-labeled E-GC-P2ab samples, and uniformly ^{13}C,^{15}N-labeled P2ab-G5 sample were prepared by *in vitro* transcription using T7 polymerase (P266L mutant) (1) with synthetic DNA templates as described previously (2). After *in vitro* transcription, the RNA samples were ethanol precipitated, purified using 15% denaturing polyacrylamide gels, electroeluted with Elutrap system (Whatman), and purified by anion-exchange with 5 ml Hi-Trap Q column (GE Healthcare). All RNA samples were then desalted and exchanged into water using an Amicon filtration system with 3K MW cut-off membranes (Millipore). The RNA samples were diluted with water to concentrations of ~10 μM, heated at 95 $^{\circ}$C for 3 mins, and snap-cooled on ice for 30 min. All the P2ab, E-P2ab, P2ab-G5 samples were then exchanged into 10 mM potassium phosphate (pH 6.4), 100 mM KCl, and 50 μM EDTA and concentrated to ~1 mM using an Amicon filtration system. The P2a1a samples were exchanged into 10 mM potassium phosphate (pH 6.4), 100 mM NaCl, and 50 μM EDTA and concentrated to ~1 mM as above. For H_2O samples, 10% D_2O was added. For D_2O samples, the RNA solutions were repeatedly lyophilized and re-dissolved in the same volume of 99.996% D_2O (Sigma).

NMR Spectroscopy. All NMR experiments were carried out on Bruker DRX 500, 600, and Avance 800 MHz spectrometers equipped with 5 mm triple-resonance cryogenic probes or QXI HCP probe (500 MHz). Exchangeable proton spectra were recorded using H_2O samples at 283K, and non-exchangeable proton spectra were recorded on D_2O samples at 293K. NMR spectra were processed and analyzed with XWINNMR 3.5 (Bruker), NMRPipe (3), and Sparky 3.110. (University of California, San Francisco, CA). The assignments were obtained using 2D NOESY, 2D TOCSY, ^{1}H-^{15}N HSQC, ^{1}H-^{13}C

HSQC, 2D HCCH-COSY, 3D HCCH-TOCSY, and a suite of 2D-filtered/edited NOESY (F2f, F1fF2e, F1fF2f, and F1eF2e) experiments on the unlabeled, uniformly labeled and base-specifically ^{13}C,^{15}N-labeled RNA samples (2, 4). The ^{31}P spin-echo difference CT-HSQC and spin-echo difference CH-HCCH correlation experiments were used to determine the ε dihedral angles of the backbone (5). One-bond C-H and N-H RDCs were measured on uniformly ^{13}C,^{15}N-labeled P2ab, P2ab-G5 and P2a.1 samples in ~15 mg/ml Pf1 phage (ASLA Biotech, Ltd) at 293K on 800 MHz spectrometer using 2D ^{1}H-^{13}C S^3CT-HSQC (6) and standard ^{1}H-^{15}N HSQC experiments as described previously (7) (Tables S1 and S4). NMR spectra for RDCs were processed and analyzed using NMRPipe/NMRDraw (3).

NMR dynamic characterizations of P2ab were performed on E-AU-P2ab and E-GC-P2ab sample at 293K on a 600 MHz spectrometer as described previously (8). The ^{1}H-^{13}C resonance intensities for C5H5, C6H6, C8H8, and C1′H1′, were measured using 2D ^{1}H-^{13}C non-constant time TROSY–HSQC experiments (8). The resonance intensities that belong to a given type of C-H spin (C1′H1’, C5H5, C6H6, and C8H8) were normalized against the lowest intensity from the elongated P2a helix to the reference value of 0.1, and was carried out independently for E-AU-P2ab and E-GC-P2ab resonances. Due to chemical shift overlap, none of C2H2 resonances from the elongated P2a helix could be sufficiently resolved to provide a reference value (Fig. S1). Therefore C2H2s were excluded from this analysis. Resonances of C5H5 from C89 and C90 display weaker intensities than the reference (Fig. 3A). This phenomenon might be due to the intrinsically faster relaxation rates of these spin pairs, as the C5H5 bonds of both C89 and C90 are oriented almost parallel to the long diffusion axis, ~10° and ~25° away from Z axis, respectively. Imino ^{15}N longitudinal (R_1) and transverse ($R_{2(CPMG)}$) relaxation rates and $\{^{1}H\}$-^{15}N heteronuclear NOEs were measured as described previously (8). The relaxation delays for R_1 and $R_{2(CPMG)}$ experiments were 0.06, 0.24 (x2), 0.48, 0.80 (x2)

and 1.2 s, and 5.2, 10.4 (x2), 20.8, 31.2 (x2), and 52.0 ms, respectively, where duplicate measurements are indicated as 'x2' (Table S2). Data were analyzed using the extended model-free formalism (9, 10) as implemented in Modelfree (Version 4.16 for Linux) (11), as described previously (8) (Table S3). The lowest-energy P2ab structure where the P2a helix was extended with 22 base pairs was used as an input structure for hydrodynamic calculation using program HYDRONMR (12) and for model-free analysis using software Modelfree Version 4.16 (11). NMR spectra for dynamic characterization were processed using NMRPipe/NMRDraw (3), and analyzed using Sparky 3.110. (University of California, San Francisco, CA) and NMRView (13).

Structure Calculations of P2ab. For structure calculation of P2ab, NOE restraints for inter-proton distances were obtained from the 2D-filtered/edited NOESY, 2D NOESY and 2D Watergate NOESY spectra acquired in D_2O and H_2O samples. NOE distances were classified as very strong (2.5 Å), strong (3.5 Å), medium (4.5 Å), weak (5.5 Å), and very weak (6.5 Å) with a lower bound of 1.8Å for all. Dihedral angle restraints (A-form values of -62.1 ± 30° for α, 180.1 ± 30° for β, 47.4 ± 30° for γ, 37.3 ± 30° for ν_2, and -74.7 ± 30° for ζ for helical residues and experimentally determined values for ε, χ, and δ), hydrogen bond distance and weak planarity restraints for the 13 Watson-Crick base pairs were incorporated in the structure calculations as previously described (14). Initially 200 structures of P2ab were calculated starting from an extended and unfolded RNA using all restraints except RDCs following standard XPLOR protocols (XPLOR-NIH-2.9.8) (15). Structures with no experimental restraint violation (distances >0.5 Å and dihedral angles >5°) were further refined with 76 RDCs. All C-H and N-H RDCs were normalized to a C-H bond length of 1.0 Å. The optimal values for the magnitude and asymmetry of the alignment tensor are Da = -20.7 Hz and -29.8 Hz and R = 0.19 and 0.25, for P2a and P2b, respectively. The force constants for RDCs were gradually increased from 0.001 to 1.0 and 0.5 $kcal \cdot mol^{-1} \cdot Hz^{-2}$ for P2a and P2b to account for the

difference in the Da values. The 20 lowest-energy structures from the final refinement are reported. Restraints and structural statistics are summarized in Table 1 in the main text. Structures were viewed and analyzed with MOLMOL (16) and PYMOL (DeLano Scientific LLC).

Preparation of Wild Type and Mutant Full Length hTRs. Full length hTR was prepared from the pBShTR plasmid (a gift from Prof. Kathleen Collins). All J2ab substitutions except for J2ab-FP (see Fig. 4A) were carried out directly on the pBShTR plasmid using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). For J2ab-FP substitution, a second mutagenesis was performed on the hTR J2ab-Δ5 plasmid. The dsDNA templates for the hTRs with J2ab substitutions were PCR amplified using a 5' primer that contained the T7 promoter added to the 5' end of the full-length hTR sequence and a 3' primer that is complementary to the 3' end of the full-length hTR. All dsDNA templates were purified using 1% Agarose gel, extracted using QIAquick Gel Extraction Kit (Qiagen), sequenced, and used directly in *in vitro* transcription of RNAs using T7 polymerase (P266L mutant) (1). All RNA samples were then purified as described above for NMR samples.

***In vitro* Reconstitution of Telomerase and Direct Telomerase Activity Assays.** All telomerases were reconstituted in Rabbit Reticulocyte Lysate (RRL) using the TNT quick coupled transcription/translation system (Promega) (17, 18). The hTERT was first expressed in a 10 μl reaction mixture, which contains 0.4 μl PCR enhancer, 0.2 μl 1 mM Methionine, 8 μl RRL, and 150 ng NFLAG-hTERT plasmid. The reaction was incubated at 30°C for 1 hour. The telomerase was then reconstituted by adding 1 μl of 10 μM *in vitro* transcribed hTR to the reaction mixture and incubating at 30 °C for an additional 30 mins. For characterizing the telomerase enzymatic activity, a 10 μl direct primer extension reaction was carried out, where 3 μl of the above reconstituted telomerase was added into the reaction mixture with 1X PE buffer (50 mM Tris-HCl, pH 8.3, 50 mM

KCl, 2 mM DTT, 3 mM $MgCl_2$, and 1 mM spermidine), 1 mM dTTP, 1 mM dATP, 2 μM dGTP, 0.25 μl ^{32}P-α-dGTP (3000Ci/mmol), and 1 μM of $(T_2AG_3)_3$ [DNA primer]. After incubating the reaction at 30 °C for 1 hour, 90 μl of a ^{32}P-labeled 15 nt DNA loading control (LC) was added to the reaction mixture. The products were extracted using phenol/chloroform/isoamyl alcohol (pH 7.9) and precipitated using ethanol. The pellets were air dried for ~15 min, resuspended in 1X formamide loading buffer, and loaded onto a 10% polyacrylamide denaturing gel. After electrophoresis, the gel was dried and the DNA products were detected using Bio-Rad FX Pro plus PhosphorImager and analyzed using Quantity One software (Bio-Rad Laboratories). Telomerase activity was characterized as the total intensity of the extended DNA products after correcting the background and normalizing against the loading control. The activity of telomerase with wild type hTR was set to be 1. In order to quantify the telomerase processivity, the intensities of each telomere repeat were measured, corrected for the number of dGTP incorporated, and normalized to the second telomere repeat whose intensity was set to 100. The normalized intensities were then plotted in logarithmic scale against the repeat number. The processivity was determined as $R_{1/2}=2\ln2/(2.303k)$, where k is the slope (17, 19).

Computational Search for Structures of 5 nt Bulge RNA Flanked by Helices. A database search for RNA structures containing 5 nt bulges was carried out using the program RNA FRABASE 2.0 at http://rnafrabase.cs.put.poznan.pl (20). Up to June 21, 2010 the database associated with RNA FRABASE 2.0 has a total of 1591 RNA structures and their complexes deposited in RCSB Protein Data Bank (PDB). The RNA sequence and secondary structure inputted for the search of our target RNA fragment is NNNNNNNNNN (((…..(((for strand 1, and NNNNNN)))))) for strand 2, where N denotes any type of nucleic acid, '(' and ')' represent Watson-Crick base paired residues, and '.' represents an unpaired residue. The search resulted in 10 fragments from 8

deposited RNA structures whose PDB IDs are 1P5M (21), 1P5O (21), 1P5P (21), 2KTZ (22), 2KU0 (22), 2NOK (23), 2PN3 (24) and 2PN4 (24). A search for structural fragments of a 5 nt bulge flanked by either 2 or 4 base pairs resulted in the same 10 fragments.

RDC-MC-Sym Approach for Structural Modeling of P2a1a. The RDC-MC-Sym approach uses a combination of NMR RDCs and computational modeling with MC-Sym (25). Here, NMR RDCs provide long-range angular restraints for the inter-helical orientation between P2a.1 and P2a, and the MC-Sym derived models provide short-range distance restraints for the J2a.1 and neighboring residues. These complimentary angular and distance restraints, together with ideal A-form restraints for P2a.1 and P2a helices, were combined together for structural calculations of P2a1a using NIH-XPLOR program following standard protocols as described above for the structure calculation of P2ab.

An initial set of 1000 P2a1a structures were generated using the online server of MC-Sym program at http://www.major.iric.ca/MC-Pipeline/, closely following the protocol in the MC-Sym's user's guide (25). The input sequence and the dot-bracket structure of P2a1a is,

GGCGUAGGCGCCGUGCUUCUUCGGAAGCCUCGGCCUGCGCC

(((((((((…..(((((….)))))….)))))))))

In addition, the following parameters were used to generate the MC-Sym script. For the MC-Sym library, the Fragment RMSD was set to be 0.1Å, and the high-resolution Nucleotide Cyclic Motifs (NCMs) were used; for the Backtrack option, both the Merge RMSD and the Clash Threshold were set to be 1.5Å; for Ribose Restraints, the Bond Threshold was set to be 2.0Å, and the Cyclic Coordinate Minimization (CCM) was used for the Construction Method option; and the Exploration Method was set to be Probabilistic. After the initial 1000 structures were generated, all structures were

subsequently minimized using Relieve, Refine, and Anneal procedures as implemented in MC-Sym.

In order to choose representative structures from the MC-Sym generated ensemble, the 1000 Anneal-minimized structures were analyzed using the tools in the online MC-Sym program in terms of base entropy, global geometry, and internal energy, as suggested in the MC-Sym's user's manual. The base entropy was characterized by bipolarity, whose mean value is 0.60 and ranges between 0.03 and 0.84 for all structures. The global geometry was measured by the molecule's radius of gyration (R_gyr), which ranges from 14.6 to 22.0 with a mean value of 18.3. The internal energy was evaluated by two parameters, a Score value and a P-Score value: the Score uses the non-bonded energy of the Amber 99 force field and ranges from -40.6 to -28.3 kcal/mol with a mean value of -35.3 kcal/mol; the P-Score is a knowledge-based measurement derived from structures in PDB and ranges from -48.1 to -12.3 with a mean value of -31.8. Based on the above information, a single structural selection criterion was applied (Bipolarity >0.60, R-gyr > 18.0, Score < -35.0, and P-Score < -30.0) to select structures for further cluster analysis. A total of 263 structures were chosen, and they were grouped into 5 clusters according to structural similarity using the "K-Means Clustering" tool in the online MC-Sym program. Clusters 1-5 have 69, 64, 52, 47, and 31 structures, respectively. The lowest energy structure (based on the Score value) from each cluster is shown in Figure S4B. The H-H distance restraints for J2a.1 and its neighboring residues were then derived from these lowest energy structures, with a H-H cut-off distance set to be 5.0 Å. All the distance restraints were categorized into 1.8-3.0 Å, 2.5-4.0 Å, 3.5-5.0 Å, and 4.5-6.0 Å ranges. A total of 355, 340, 339, 349, and 365 distance restraints were obtained for the lowest energy structure in clusters 1, 2, 3, 4, and 5, respectively. Each set of distances was then used independently for structural calculations as described below.

The structural calculation of P2a1a is similar to the procedure used in the P2ab calculations, as described previously (14). For the J2a.1 region, only MC-Sym derived distance restraints were used. For P2a.1 and P2a helices, idealized A-form restraints were used, including distance, dihedral angle, and planarity restraints. For UUCG tetraloop, restraints were adapted from the UUCG tetraloop restraints used in the structure calculations of P2ab. For each set of J2a.1 distance restraints, 200 initial structures of P2a1a were calculated using XPLOR-NIH 2.9.8 starting from an extended and unfolded RNA using all restraints except for RDCs (15), which were subsequently refined with 50 RDCs, 25 each for P2a.1 and P2a (Table S4). All C-H and N-H RDCs were normalized to C-H bond with bond length of 1.0 Å. The optimal values for the magnitude and asymmetry of the alignment tensor are Da = -37.9Hz and -45.0Hz and R = 0.14 and 0.09, for P2a.1 and P2a, respectively. The force constants for RDCs were gradually increased from 0.001 to 0.6 and 0.5 $kcal \cdot mol^{-1} \cdot Hz^{-2}$ for P2a.1 and P2a to account for the difference in the Da values.

Fig. S4C shows the 10 lowest-energy structures from the RDC refinement with J2a.1 distance restraints derived from the lowest-energy structure in MC-Sym cluster #1. The representative structures from MC-Sym clusters are well defined in terms of overall helical geometries while keeping most of the local structural features in the J2a.1 region (Fig. S4B,D). The bend and twist angles between P2a.1 and P2a for clusters 1 to 5 are 6±3° and 135±14°, 10±5°and 119±10°, 7±3° and 110±17°, 6±3° and 132±15°, and 4±1° and 129±5°, respectively. The values given are the average and standard deviation across the 10 lowest-energy structures in each cluster.

Structural Calculation of the P2/P3 Pseudoknot with Amber Force Field. An initial set of structures was generated by linking the overlapping ends of P2a1a (RDC-MC-Sym cluster 1), P2ab, and P2b-P3 pseudoknot (26) structures (Fig. S7). The top 5 lowest-energy structures from each structural ensemble were superimposed across each other to

generate a total of 125 initial structures of P2/P3 pseudoknot domain. The P2a1a structures were linked to P2ab structures by all-heavy-atom superposition of residues 79-81 and 128-130, with an average rmsd of 0.97±0.07Å. Within these residues, residues 79 and 130 from P2a1a structures and residues 80, 81, 128 and 129 from P2ab structures were included in the final structures (Fig. S7). To avoid structural errors due to terminal and loop closing base pairs in linking P2b-P3 pseudoknot and P2ab structures, an idealized A-form helix of P2b was constructed using Insight II (Molecular Simulations, Inc) to serve as a bridge between P2b-P3 and P2ab (Fig. S7). Residues 90-93 and 121-124 were used in all-heavy-atom superposition of P2ab structures onto the bridging P2b helix, with an average rmsd of 0.99±0.07Å. Residues 95-98 and 116-119 were used in all-heavy-atom superposition of P2b-P3 structures onto the bridging P2b helix, with an average rmsd of 1.29±0.01Å. For the P2ab, P2b and P2b-P3 superposition, residues 90-92 and 122-124 from P2ab structures, 93-95 and 119-121 from bridging P2b helix, and 96-98 and 116-118 from P2b-P3 structures were used in the final structures (Fig S7).

Prior to the structural refinement of the P2/P3 pseudoknot with the Amber force field, the 125 structures generated by linking overlapping residues were subjected to 20,000 steps of minimization using XPLOR-NIH 2.9.8 in order to remove possible bad local geometry at the interface between the linking residues from the sub-domain structures. The distance and dihedral angle restraints are summarized in Table S6. Experimentally derived distance restraints based on NOE measurements were applied for P2ab and P2b-P3, the MC-Sym structure derived distance restraints were applied for J2a.1 and neighboring residues, and A-form helical geometries were applied for all other helical regions. The Amber calculations were performed using AMBER version 7.0 software (27) as described previously (28), with the application of the ff99 force field (29). The standard pairwise generalized Born model (igb = 1) and the generalized Born/surface area simulation (gbsa = 1) were applied, and the salt concentration was set

to be 200 mM. The refinement was carried out over 60,000 steps with 0.2 fs/step. During the refinement, the system was heated from 0K to 500K in 5000 steps, maintained at 500K over 15,000 steps, slowly cooled down to 100K over 30,000 steps, and finally cooled down and maintained at 0K for 10,000 steps. The 20 lowest-energy structures (ranked by total Amber energy) were reported (Fig. 5), and the restraints and structural statistics for these 20 structures are summarized in Table S6.

SI References

1. Guillerez J, Lopez PJ, Proux F, Launay H, Dreyfus M (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc Natl Acad Sci USA* 102**:**5958-5963.
2. Dieckmann T, Feigon J (1997) Assignment methodology for larger RNA oligonucleotides: application to an ATP-binding RNA aptamer. *J Biomol NMR* 9**:**259-272.
3. Delaglio F, et al. (1995) Nmrpipe - a Multidimensional Spectral Processing System Based On Unix Pipes. *J Biomol NMR* 6**:**277-293.
4. Peterson RD, Theimer CA, Wu H, 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.
5. Legault P, Jucker FM, Pardi A (1995) Improved measurement of 13C, 31P J coupling constants in isotopically labeled RNA. *FEBS Lett* 362**:**156-160.
6. Meissner A, Sorensen OW (1999) The role of coherence transfer efficiency in design of TROSY- type multidimensional NMR experiments. *J Magn Reson* 139**:**439-442.
7. Zhang Q, Stelzer AC, Fisher CK, Al-Hashimi HM (2007) Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature* 450**:**1263-1267.
8. Zhang Q, Sun X, Watt ED, Al-Hashimi HM (2006) Resolving the motional modes that code for RNA adaptation. *Science* 311**:**653-656.
9. Lipari G, Szabo A (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J Am Chem Soc* 104**:**4546-4559.
10. Clore GM, et al. (1990) Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J Am Chem Soc* 112**:**4989-4991.
11. Mandel AM, Akke M, Palmer AG, 3rd (1995) Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. *J Mol Biol* 246**:**144-163.
12. Garcia de la Torre J, Huertas ML, Carrasco B (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J Magn Reson* 147**:**138-146.
13. Johnson BA, Blevins RA (1994) Nmr View - A Computer-Program For The Visualization And Analysis Of Nmr Data. *J Biomol NMR* 4**:**603-614.
14. Wu H, et al. (2001) A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. *EMBO J* 20**:**7240-7249.
15. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160**:**65-73.
16. Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14**:**51-55, 29-32.
17. Chen JL, Greider CW (2003) Determinants in mammalian telomerase RNA that mediate enzyme processivity and cross-species incompatibility. *EMBO J* 22**:**304-314.

18. Huard S, Moriarty TJ, Autexier C (2003) The C terminus of the human telomerase reverse transcriptase is a determinant of enzyme processivity. *Nucleic Acids Res* 31**:**4059-4070.
19. Wang F, et al. (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* 445**:**506-510.
20. Popenda M, et al. (2010) RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinformatics* 11**:**231.
21. Lukavsky PJ, Kim I, Otto GA, Puglisi JD (2003) Structure of HCV IRES domain II determined by NMR. *Nat Struct Biol* 10**:**1033-1038.
22. Paulsen RB, et al. (2010) Inhibitor-induced structural change in the HCV IRES domain IIa RNA. *Proc Natl Acad Sci USA* 107**:**7263-7268.
23. Dibrov SM, Johnston-Cox H, Weng YH, Hermann T (2007) Functional architecture of HCV IRES domain II stabilized by divalent metal ions in the crystal and in solution. *Angew Chem Int Ed Engl* 46**:**226-229.
24. Zhao Q, Han Q, Kissinger CR, Hermann T, Thompson PA (2008) Structure of hepatitis C virus IRES subdomain IIa. *Acta Crystallogr D Biol Crystallogr* 64**:**436-443.
25. Parisien M, Major F (2008) The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* 452**:**51-55.
26. Kim NK, et al. (2008) Solution Structure and Dynamics of the Wild-type Pseudoknot of Human Telomerase RNA. *J Mol Biol*.
27. Case DA, et al. (2002) *AMBER 7, University of California, San Francisco.*
28. Stefl R, Wu H, Ravindranathan S, Sklenar V, Feigon J (2004) DNA A-tract bending in three dimensions: solving the dA4T4 vs. dT4A4 conundrum. *Proc Natl Acad Sci USA* 101**:**1177-1182.
29. Wang JM, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21**:**1049-1074.
30. Musselman C, et al. (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.
31. Gavory G, Symmons MF, Krishnan Ghosh Y, Klenerman D, Balasubramanian S (2006) Structural analysis of the catalytic core of human telomerase RNA by FRET and molecular modeling. *Biochemistry* 45**:**13304-13311.
32. Mitchell M, Gillis A, Futahashi M, Fujiwara H, Skordalakes E (2010) Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat Struct Mol Biol* 17**:**513-518.

Table S1. RDCs measured for P2ab and P2ab-G5.

Residue		P2ab	P2ab-G5	Residue		P2ab	P2ab-G5
G77	C8H8	-20.2	-28.2	U95	C5H5	13.5	14.5
G78	C8H8	-20.0	-18.1	U95	C6H6	20.5	20.7
G78	C1'H1'	-27.1	-24.2	U95	C1'H1'	10.9	13.0
G78	N1H1	2.7	-0.9	U96	C5H5	-10.6	-11.5
U80	C5H5	-12.3	-9.1	U96	C6H6	22.5	22.9
U80	C6H6	21.0	21.1	U96	C1'H1'	-7.1	-8.6
U80	N3H3	-9.2	-9.5	C118	C5H5	5.5	0.4
U81	C5H5	7.9	14.2	C118	C6H6	-11.7	-15.6
U81	C6H6	17.7	8.6	C118	C1'H1'	-13.0	-20.1
U81	N3H3	-8.1	-4.6	G119	C8H8	1.8	-1.8
U82	C5H5	21.1	21.4	G119	C1'H1'	-3.1	-7.7
U82	C6H6	4.4	-5.9	G119	N1H1	-2.6	-2.6
U82	C1'H1'	11.9	17.6	G120	C8H8	13.0	15.0
U82	N3H3	-2.4	3.8	G120	C1'H1'	-20.5	-27.4
U83	C5H5	15.0	6.5	G120	N1H1	-4.7	-7.0
U83	C6H6	0.3	-4.3	C121	C1'H1'	-3.2	-6.2
U83	C1'H1'	17.0	17.8	A122	C2H2	9.8	10.0
G84	C8H8	2.3	-0.3	A122	C8H8	19.7	18.7
G84	C1'H1'	16.6	14.1	A122	C1'H1'	5.6	5.5
C85/G85	C5H5	-5.6	NA	C123	C5H5	12.3	14.1
C85/G85	C6H6	15.9	NA	C123	C6H6	32.2	33.0
C85/G85	C1'H1'	11.7	NA	C123	C1'H1'	12.3	14.0
U86/G86	C5H5	-8.4	NA	G124	C8H8	30.5	32.6
U86/G86	C6H6	12.5	NA	G124	C1'H1'	NA	16.3
U86/G86	C1'H1'	0.7	NA	G124	N1H1	-9.8	-11.7
C87/G87	C5H5	5.7	NA	G125	C8H8	14.9	NA
C87/G87	C6H6	8.8	NA	G125	C1'H1'	21.8	25.0
C87/G87	C1'H1'	1.3	NA	A126	C2H2	10.8	14.9
C88/G88	C5H5	15.6	NA	A126	C8H8	-4.2	-4.5
C88/G88	C6H6	-14.9	NA	A126	C1'H1'	5.7	8.0
C89	C5H5	-14.4	-16.0	A127	C8H8	-3.8	NA
C89	C6H6	-2.9	5.4	A127	C1'H1'	12.3	NA
C90	C5H5	-17.7	-14.9	A128	C8H8	NA	-5.2
C90	C6H6	21.8	27.0	A128	C1'H1'	16.1	15.6
G91	C8H8	30.5	33.0	A129	C2H2	7.0	-4.7
G91	C1'H1'	-32.4	-27.0	A129	C8H8	NA	17.7
G91	N1H1	-13.4	-13.1	G130	C8H8	18.4	21.5
U92	C5H5	23.0	27.4	G130	C1'H1'	18.1	19.6
U92	C6H6	24.4	23.2	G130	N1H1	-3.4	-5.5
U92	C1'H1'	-4.1	2.3	C131	C5H5	21.1	18.0
U92	N3H3	-11.2	-11.0	C131	C6H6	-11.2	-3.5
G93	C8H8	19.7	20.7	C131	C1'H1'	10.9	17.9
G93	N1H1	-4.6	-4.6	C132	C5H5	14.9	17.9
C94	C5H5	25.0	23.5	C132	C6H6	-25.1	-26.9
C94	C6H6	13.7	15.0	C132	C1'H1'	0.9	9.7
C94	C1'H1'	12.1	16.5				

Table S2. ^{15}N relaxation rates measured in E-P2ab.

	R_1 (Hz)	R_2 (Hz)	NOE
G77	0.716±0.010	26.78±0.35	0.644±0.021
G78	0.729±0.009	27.58±0.32	0.670±0.018
U80	0.646±0.017	26.69±0.69	0.591±0.034
U81	0.659±0.021	26.56±0.89	0.625±0.043
U82	0.570±0.064	27.60±2.88	0.739±0.098
G124	1.043±0.011	18.52±0.25	0.638±0.020
G91	1.041±0.009	18.72±0.21	0.590±0.018
U92	0.902±0.011	21.40±0.31	0.581±0.021
G93	0.888±0.007	25.03±0.22	0.495±0.016
G119	0.976±0.008	23.12±0.22	0.462±0.016

Table S3. Modelfree analysis of E-P2ab ^{15}N relaxation data.

	Model	S_f^2	S_s^2	τ_f (ps)	τ_s (ns)
G77	2	0.953±0.018		121.2±67.5	
G78	2	0.974±0.014		213.9±266.2	
U80	2	0.944±0.034		104.6±87.0	
U81	2	0.949±0.043		94.8±103.1	
U82	1	0.895±0.137			
G124	5	0.855±0.019	0.766±0.015		1.965±0.348
G91	5	0.859±0.016	0.721±0.012		1.852±0.232
U92	5	0.831±0.018	0.696±0.012		2.077±0.249
G93	5	0.827±0.012	0.693±0.009		1.625±0.118
G119	5	0.817±0.013	0.628±0.009		1.608±0.098

Table S4. RDCs measured for P2a1a.

Residue		RDC	Residue		RDC
G2	C8H8	19.2	G24	C8H8	30.5
G2	N1H1	-11.2	G24	N1H1	-15.4
G4	C8H8	26.9	A25	C2H2	32.3
G4	N1H1	-15.7	A25	C8H8	30.9
U5	C5H5	18.6	A25	C1'H1'	-23.4
U5	N3H3	-16.1	A26	C2H2	33.3
A6	C2H2	33.8	A26	C8H8	28.6
A6	C8H8	27.6	A26	C1'H1'	-17.8
A6	C1'H1'	-11.7	G27	C8H8	26.6
G7	C8H8	29.7	G27	C1'H1'	-23.2
G7	N1H1	-13.0	G27	N1H1	-13.5
G8	C8H8	25.5	C28	C6H6	20.3
G8	N1H1	-14.1	C28	C1'H1'	-25.5
G15	C8H8	28.2	G33	C8H8	28.4
G15	N1H1	-13.3	C35	C5H5	21.5
C16	C6H6	23.3	U36	C5H5	18.8
U17	C6H6	25.7	U36	C6H6	24.1
U17	N3H3	-14.8	U36	C1'H1'	-23.1
U18	C5H5	23.3	U36	N3H3	-15.2
U18	C6H6	22.8	G37	C8H8	31.9
U18	C1'H1'	-21.7	G37	N1H1	-17.5
U18	N3H3	-15.4	C38	C5H5	28.4
C19	C5H5	29.3	G39	C8H8	23.8
C19	C6H6	21.4	G39	N1H1	-11.8
C19	C1'H1'	-21.4	C40	C6H6	30.5

Table S5. Order tensor analysis of RDCs measured in helical regions of P2ab, P2ab-G5, and P2a1 using idealized A-form helices as input helix coordinates.

	Helix	N	RMSD (Hz)	Q (%)	R	CN	η	ϑ (x10^{-3})	ϑ_{int}
P2ab	P2a	28	2.4	17.2	0.99	2.4	0.28±0.06	0.69±0.04	0.69±0.04
	P2b	29	1.9	9.3	0.99	2.2	0.37±0.02	1.01±0.01	
P2ab G5	P2a	30	2.4	16.5	0.99	2.5	0.22±0.07	0.73±0.02	0.69±0.03
	P2b	28	2.2	10.5	0.99	2.2	0.35±0.02	1.06±0.04	
P2a1a	P2a.1	25	3.5	14.7	0.99	3.5	0.20±0.03	1.26±0.03	0.85±0.03
	P2a	25	3.1	10.5	0.99	2.6	0.13±0.03	1.49±0.04	

Table headings: Number of RDCs (N), RMSD between measured and back-calculated RDCs, quality factor (Q), correlation coefficient (R), condition number (CN), asymmetry parameter (η), generalized degree of order (ϑ), and internal generalized degree of order (ϑ_{int}). Errors in η, ϑ, and ϑ_{int} were calculated using the program AFORM-RDC, which takes into account of both measurement errors in RDCs and structural noise in A-form helices (30).

Table S6. Structural Statistics for P2/P3 Pseudoknot with Amber Force Field.

NOE distance and dihedral constraints	
Distance restraints	
Experimental derived from NOEs	
P2ab construct	490
P2b-P3 pseudoknot	629
MC-Sym derived	
J2a1 and neighboring residues	355
Idealized A-form helix	
P2a.1 and linking residues in P2a and P2b	2057
Hydrogen bond	192
Total dihedral angle restraints	582
Statistics of 20 lowest-energy structures	
Amber Energy (kCal/mol)	-15685.35±17.95
Violations	
Distance constraints (Å)	0.020±0.0002
Dihedral angle constraints (°)	0.311±0.139
Deviation from idealized geometry	
Bond lengths (Å)	0.007±0.00004
Bond angles (°)	1.867±0.007
Impropers (°)	1.160±0.045
Average rmsd (Å) to the mean with all heavy atoms	1.13±0.22
Number of NOE violations > 0.2 (Å)	3.50±1.12
Number of NOE violations > 0.5 (Å)	1.00±0.22
Number of Dihedral violations > 5 (°)	1.25±1.25

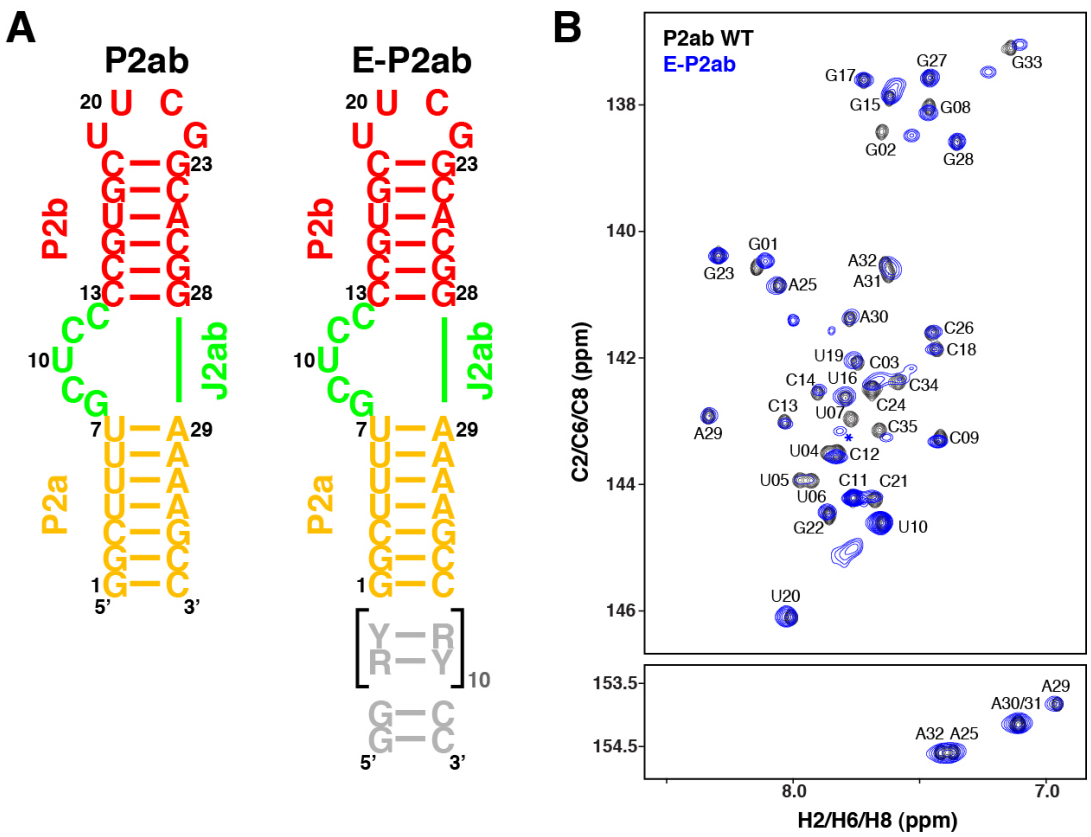

Figure S1. Domain elongation does not affect the structure and dynamical integrity of P2ab. (A) Sequence and secondary structures of P2ab and E-P2ab constructs. (B) Overlap of 2D ^{1}H-^{13}C HSQC spectra of the aromatic region (C2-H2, C6-H6 and C8-H8 resonances) from P2ab (black) and E-P2ab (blue). Note that the resonance next to C6H6 of P2ab U7 residue (depicted as the asterisk) is not C6H6 of U7 in E-GC-P2ab, it belongs to C6H6 from natural abundance U residues in E-AU-P2ab construct.

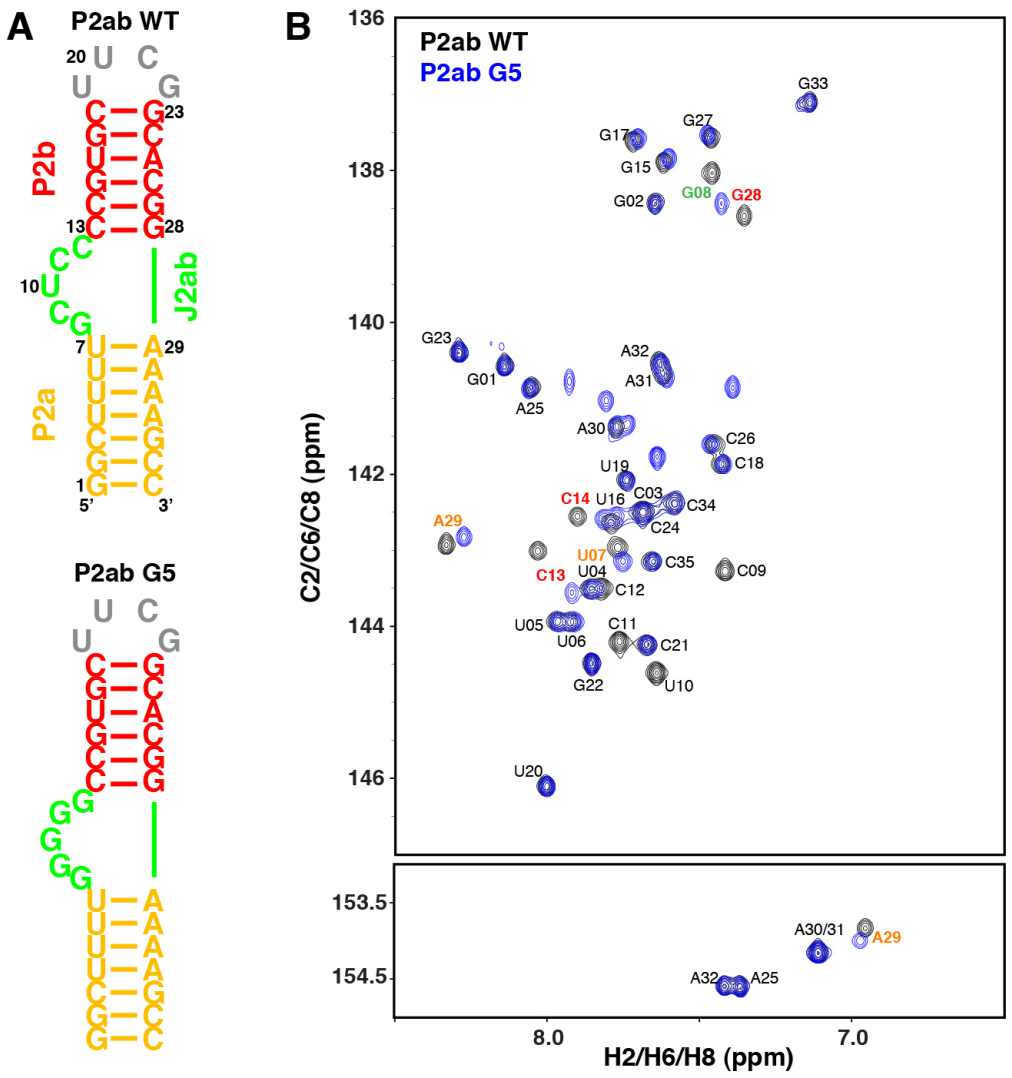

Figure S2. G5 mutation induces local structural changes around the J2ab region. (A) Sequences and secondary structures of P2ab wild-type (WT) and P2ab G5 mutant studied by NMR. (B) Overlap of 2D ^{1}H-^{13}C HSQC spectra of the aromatic region (C2-H2, C6-H6 and C8-H8 resonances) from P2ab WT and P2ab G5. The P2ab WT spectrum is shown in black and P2ab G5 spectrum is shown in blue. Only neighboring residues of J2ab are highlighted in color and display observable chemical shift changes compare to residues from helical region.

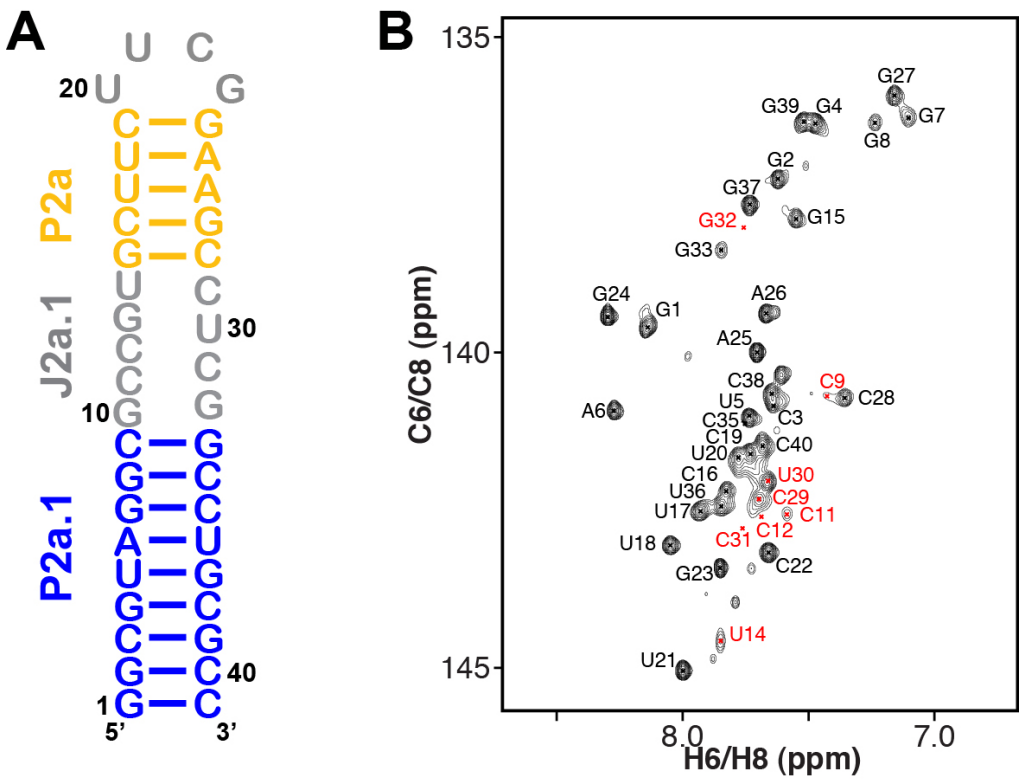

Figure S3. Conformational exchange within J2a.1 region. Shown is 2D ^{1}H-^{13}C HSQC spectrum of non-exchangeable base C6-H6 and C8-H8 resonances. Severe line broadenings are occurring for resonances belonging to J2a.1 region as highlighted with red X and labels. C8H8 resonances from G10 and G13 are too broad to be observed in the spectrum.

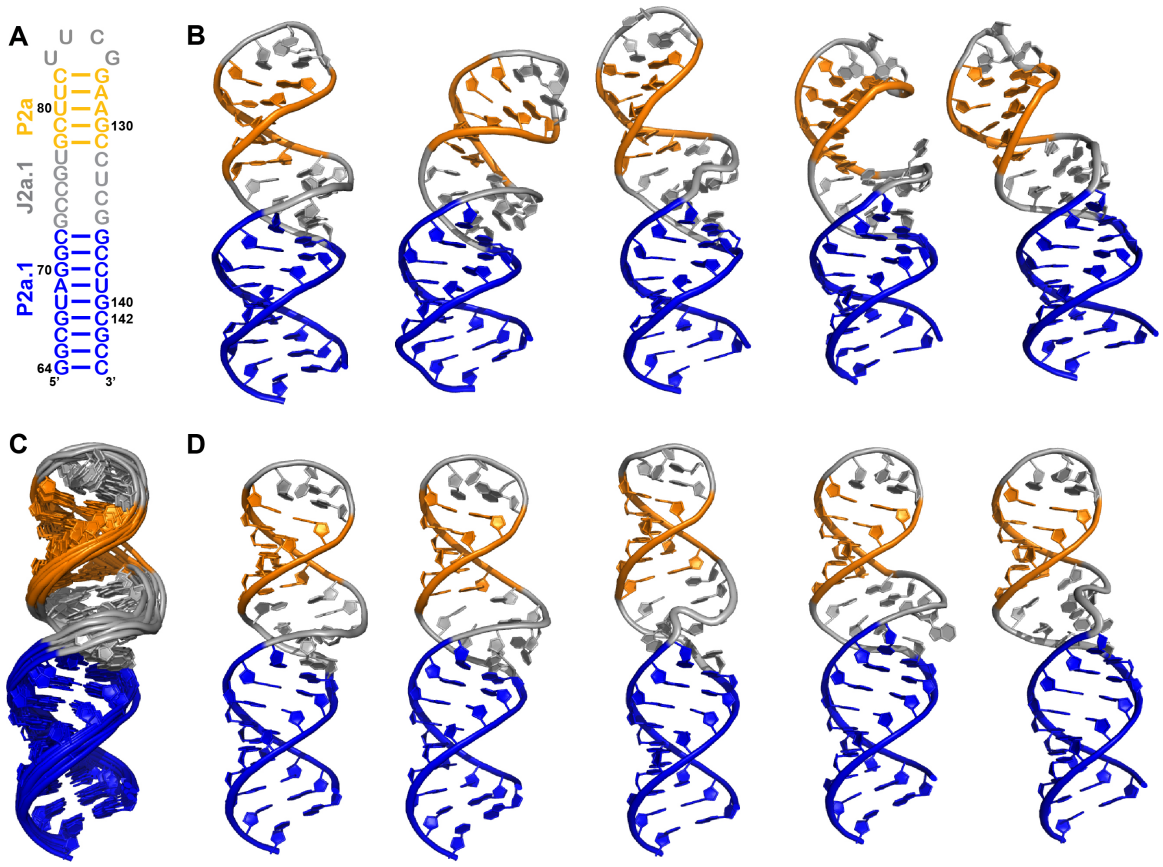

Figure S4. Structural models of P2a1a calculated by RDC-MC-Sym approach. (A) Sequence and secondary structure of P2a1a construct. (B) The lowest-energy structures from five structural clusters generated by MC-Sym. Shown from left to right are structures from clusters 1 to 5. (C) Superposition of 10 lowest-energy structures after refinement with RDCs and J2a.1 distance restraints derived from the lowest-energy structure from MC-Sym cluster 1. (D) The lowest energy structures after refinement with RDCs and J2a.1 distance restraints derived from structures shown in Figure S4B. Shown from left to right are corresponding structures from clusters 1 to 5. All structures shown in panels B and C are superimposed over P2a.1 helix.

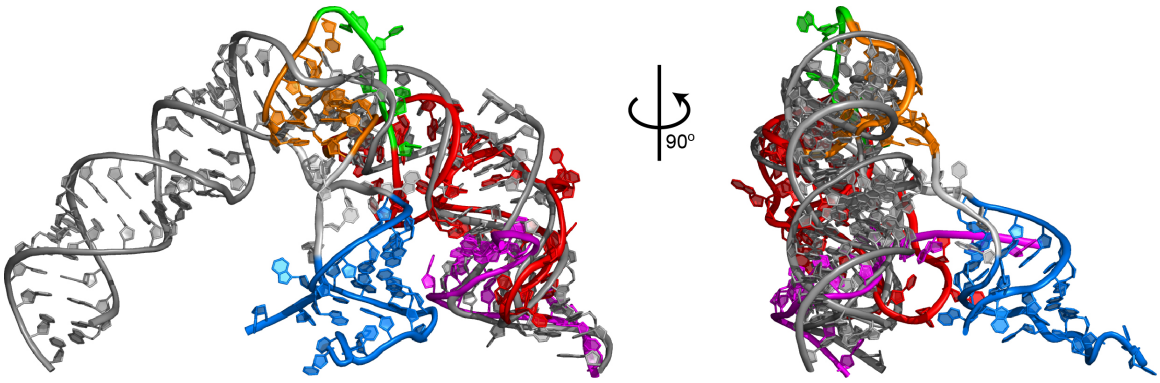

Figure S5. Structural comparison between low-resolution and high-resolution structural models of hTR core domain. Shown in color is the low-resolution model derived from FRET analysis (PDB ID 2INA) (31). For clarity, only the P2/P3 pseudoknot and the template RNA and hybrid PNA are shown. Color scheme is same as in Figure 1 in the main text. The high-resolution model from Amber refinement of sub-domains is shown in gray.

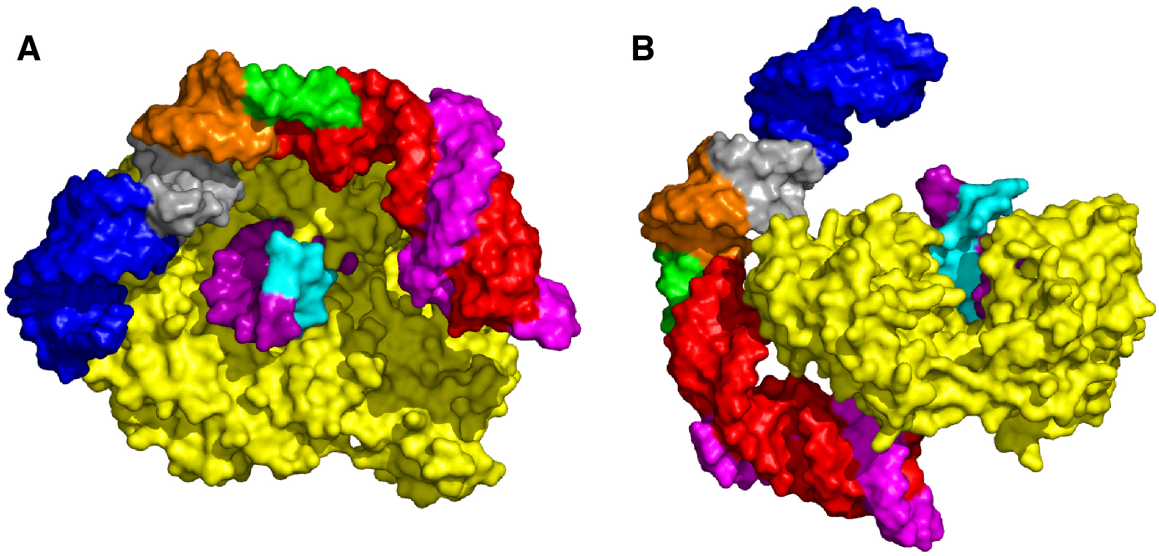

Figure S6. Surface representation of two possible orientations between P2/P3 pseudoknot domain and TERT. (A) The P2/P3 pseudoknot lies parallel to TERT, which orients the RNA/DNA hybrid hairpin perpendicular to the pseudoknot domain plane. (B) The P2/P3 pseudoknot is perpendicular to the TERT, which makes the RNA/DNA hybrid hairpin co-plane with the pseudoknot domain. Shown are the *Tribolium castaneum* TERT in complex with telomeric RNA/DNA hybrid hairpin (PDB ID 3KYL) (32) where TERT is colored in yellow, RNA template in purple and DNA in cyan. Color scheme for P2/P3 pseudoknot domain is same as in Figure 1 in main text.

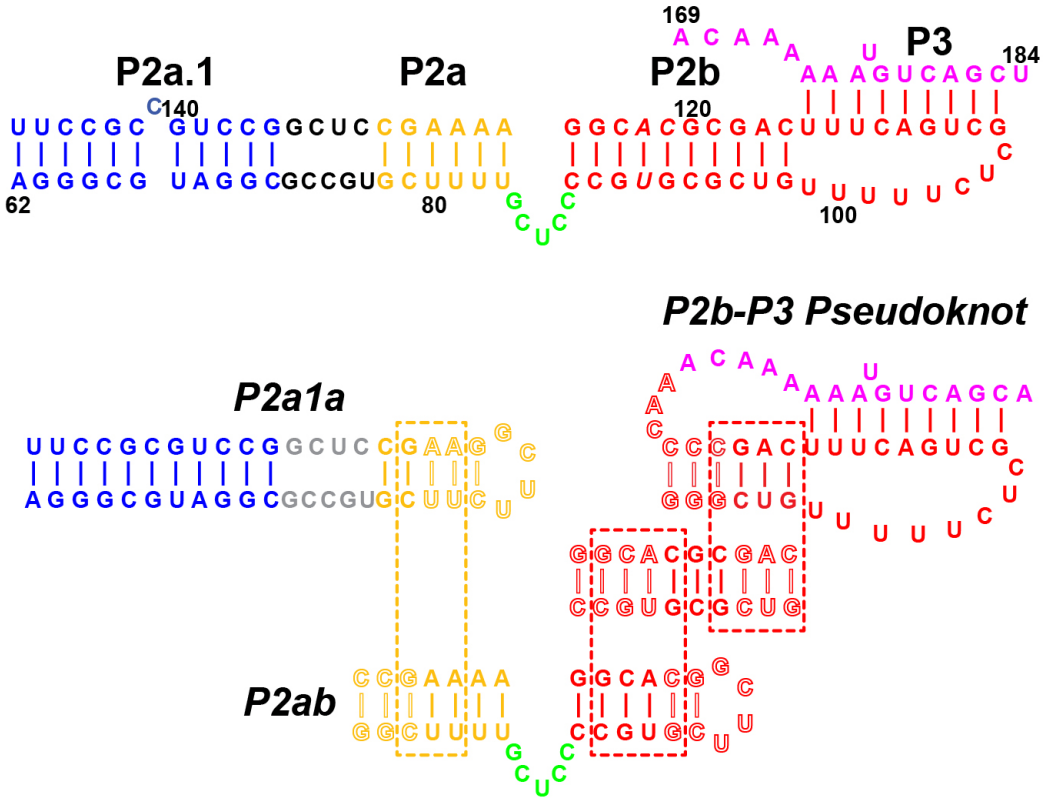

Figure S7. Sequence alignment for generating starting structures for Amber refinement from sub-domain structures. Residues shown in dash boxes are used for superimposing sub-domain structures, and residues with solid letters are used in generating the structures.