Dissecting telomerase RNA structural heterogeneity in living human cells with DMS-MaPseq

Nicholas M. Forino[1#], Jia Zheng Woo[2#], Arthur J. Zaug[3,4], Arcelia Gonzalez Jimenez[5], Thomas R. Cech[3,4*], Silvi Rouskin[2*], Michael D. Stone[5*]

[1]Department of Molecular, Cell, and Developmental Biology, University of California, Santa Cruz, CA, USA
[2]Department of Microbiology, Harvard Medical School, Boston, MA, USA
[3]Department of Biochemistry, University of Colorado, Boulder, CO, USA
[4]Howard Hughes Medical Institute, University of Colorado, Boulder, CO, USA
[5]Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA, USA

[#] authors contributed equally
[*] corresponding authors

Abstract

Telomerase is a specialized reverse transcriptase that uses an intrinsic RNA subunit as the template for telomeric DNA synthesis. Biogenesis of human telomerase requires its RNA subunit (hTR) to fold into a multi-domain architecture that includes the template-containing pseudoknot (t/PK) and the three-way junction (CR4/5). These two hTR domains bind the telomerase reverse transcriptase (hTERT) protein and are thus essential for telomerase catalytic activity. Here, we probe the structure of hTR in living cells using dimethyl sulfate mutational profiling with sequencing (DMS-MaPseq) and ensemble deconvolution analysis. Unexpectedly, approximately 15% of the steady state population of hTR has a CR4/5 conformation lacking features required for hTERT binding. Mutagenesis demonstrates that stabilization of the alternative CR4/5 conformation is detrimental to telomerase assembly and activity. We propose that this misfolded portion of the cellular hTR pool is either slowly refolded or degraded. Thus, kinetic traps for RNA folding that have been so well-studied in vitro may also present barriers for assembly of ribonucleoprotein complexes in vivo.

INTRODUCTION

Telomeres are repetitive regions of DNA found at chromosome ends that are sheathed with a coat of protective proteins[1, 2]. Telomerase is a ribonucleoprotein (RNP) enzyme that catalyzes telomere DNA repeat synthesis using a short segment of its integral RNA subunit as a template [3]. The catalytic action of telomerase counteracts the gradual erosion of telomeric DNA arising from incomplete replication of the lagging strand by the DNA replication machinery [4, 5]. In the absence of telomerase, telomeres eventually shorten to critical lengths that can trigger cellular senescence or programmed cell death [6, 7]. Telomere length maintenance by telomerase is thus an essential process for ensuring the genomic stability and replicative capacity of rapidly dividing or continually regenerating tissues[8, 9].

Underpinning the action of telomerase is the multistep co-assembly of hTR, hTERT, and several additional proteins into the functional RNP complex (**Fig. 1a**). Human telomerase is structurally organized by the hTR subunit, which adopts a phylogenetically conserved multi-domain architecture and serves as the binding scaffold for telomerase proteins [10, 11] (**Fig. 1b**). Prior chemical probing studies of hTR [12, 13] were instrumental in providing direct biochemical evidence of the proposed phylogenetically conserved secondary structure [10]. However, these studies came to differing conclusions on whether hTERT binds a pre-organized hTR or induces hTR remodeling. Central to this discrepancy is that these studies were technically limited to gathering structural information on the population average of what has been shown to be a heterogeneous pool of hTR reservoirs [14, 15]. Overexpression of telomerase has also historically been used to mitigate problems arising from its scarcity but has been shown to bypass the endogenous RNP assembly pathway[16].

Recent cryogenic electron microscopy (cryo-EM) discoveries have illuminated the structure and molecular composition of the *Tetrahymena* and human telomerase enzymes[17, 18, 19, 20, 21, 22]. While both enzymes share important structural features in the catalytic domain, human telomerase has a uniquely bilobed architecture[23, 24, 25]. The catalytic lobe contains the hTERT protein, the conserved RNA t/PK and CR4/5 domains , and an hTR-bound histone dimer (**Fig. 1c and 1d**). The biogenesis lobe includes two complements of a core tetramer of H/ACA proteins bound to a small Cajal-body RNA domain (ScaRNA) near the 3' end of hTR [26, 27]. Assembly of telomerase involves a multistep cascade of protein-binding events and nuclear trafficking of the immature telomerase RNP before finally arriving at telomeres as an active complex[28, 29]. After transcription, assembly of the H/ACA RNP licenses hTR for maturation and prevents its degradation by the RNA exosome[30, 31]. hTR transits through the nucleolar and Cajal body subnuclear compartments and accumulates in the nucleoplasm. Association with hTERT excludes hTR from the nucleolus while association with the TCAB1 permits transient association with Cajal bodies [32, 33, 34]. While the precise spatio-temporal details of hTERT association with immature telomerase RNPs remains unclear, it is well-established that mutations in telomerase components that disrupt RNP biogenesis can cause a wide range of inherited diseases [35, 36]. Moreover, mutagenesis studies on hTR have reinforced the importance of the t/PK and CR4/5 adopting precise conformations to achieve the assembly of a catalytically active telomerase RNP [37, 38, 39].

Here, we identify and characterize alternative conformations of endogenous hTR in living cells through a targeted DMS-MaPseq approach coupled with ensemble deconvolution [40, 41]. Across the entire population of hTR molecules, DMS-reactive nucleotides are encoded as mutations in sequencing reads. Thus, each read represents a single-molecule measurement of hTR structure, permitting the use of bioinformatic clustering to identify groups of reads that are defined by distinct mutation patterns. We used DMS constraints derived from the clustered MaPseq results to guide RNA folding predictions, revealing that the t/PK and CR4/5 domains each exist in a structural ensemble that includes their canonically described conformation and an alternative conformation. Based on current cryo EM models of the assembled telomerase RNP [19, 22], these alternative t/PK and CR4/5 conformations are not representative of hTR in its hTERT-bound form. CR4/5 mutants designed to destabilize the canonical three-way junction conformation while stabilizing the alternative conformation to varying degrees were misfolded in cells and exhibited catalytic defects. The decrease in activity is explained by impaired RNP

assembly. Taken together, our results suggest that the in vivo hTR folding pathway is complex and includes alternative conformations that may represent dead-end misfolded states or possible intermediates along the telomerase RNP biogenesis pathway. In either case, our results show that assembly of catalytically active telomerase RNPs may be limited by a non-productive conformation of the CR4/5 domain.

RESULTS

DMS-MaPseq reveals discrepancies with the established hTR structure model

To test whether our MaPseq approach could capture structural features of endogenous hTR in living cells, we first generated and analyzed DMS reactivities of the ensemble average in HeLa cells. Our DMS reactivities of the t/PK domain (**Fig. 2a and 2b**) are in close accord with prior chemical probing studies [12, 13] and recent cryo-EM models of assembled telomerase [19] (**Fig. 2c**). We observed nucleotides with medium to strong degrees of DMS reactivity (0.3-0.8 and >0.8, respectively) in the template region (C46-C56), the internal loop joining the P2a.1 and P2a stems, the bulge joining P2a and P2b (J2a/b), and hinge region linking P2a.1 to P3 (J2a.1/3), indicative of single-stranded RNA. High DMS reactivity of template nucleotides in vivo has also been observed for *Tetrahymena* telomerase[42]. Notably, we detected numerous nucleotides with low (0.1-0.3) and weakly reactive (<0.1) in the pseudoknot-forming P3 helix (C108, A11, C112, A174-176, C180, A181, C183).

DMS reactivity of the CR4/5 domain was consistent with a three-way junction architecture, with most of the reactive nucleotides residing within the internal loops that join the P4.2/P5 stems and the P6a/P6b stems, as well as the apical loop of P6 (C277-278) and the junction regions of the three-way junction motif (C247, C248, A252, C255, A301, C317, A318) (**Fig. 3a, b, and c**). We detected DMS reactivity in the nucleotides comprising the P6.1 stem (A302, A304, C311, C313), a finding consistent with previous chemical probing studies [12, 13, 43]. In addition, we detected unexpected DMS reactivity at a trio of cytosines (C286-288) that contribute to the P6b helix. Previously, we reported that uncharacteristic chemical reactivity of CR4/5 in vitro originates from its heterogeneous folding behavior [43]. Thus, we sought to determine whether alternative folding conformations of CR4/5 exist in the cellular environment and could explain our detected DMS reactivities.

Ensemble deconvolution reveals alternative conformations of t/PK and CR4/5 domains

Because MaPseq exploits the read-through behavior of multiple DMS modifications per RNA, it permits the use of bioinformatic deconvolution algorithms to identify distinct clusters of DMS reactivities that distinguish alternative RNA conformations [40, 44]. We applied the Deconvolution of RNA Ensembles by Expectation Maximization (DREEM) algorithm [40] to our DMS reactivities of endogenous hTR. We then used DMS reactivities derived from the clusters to guide thermodynamic secondary structure predictions.

We found that the hTR t/PK domain forms distinct major and minor clusters across multiple biological replicates with relative abundances of ~80% and ~20%, respectively. The results of the deconvolution analysis were reproducible across multiple biological replicates (**Fig. S1**). Nucleotides responsible for forming the pseudoknot P3 helix (C108, A111, C112 and A174-A176, C180, A181, C183) are overall weakly DMS reactive in the major cluster (**Fig. 4a**), while in the minor cluster they partition into strongly reactive (C108, A111, C112, C180, A181, C183) and mostly unreactive (A174-176) groups. Cluster-guided thermodynamic structure prediction of the major cluster yielded a conformation similar to the canonical t/PK domain with a slightly truncated pseudoknot P3 helix, while the predicted model of the minor cluster lacks the pseudoknot and is arranged as an extended stem-loop architecture (**Fig. 4b and 4c**). We note that the reactivity patterns for minor clusters were not as well-correlated between replicate experiments as those for the major clusters (**Fig. S1b**); however, the deconvolution results were nevertheless reproducible.

The CR4/5 domain also forms major and minor clusters with relative abundances of ~85% and ~15%, respectively. As before, the results of the deconvolution analysis were reproducible across multiple replicates (**Fig. S2**), and reactivity patterns were very well correlated for both the major and minor populations of the CR4/5 domain (**Fig. S2b**). Two cytosines that form the internal loop bisecting P6a and P6b (C266-267) are strongly reactive in the major cluster but are moderately reactive in the minor cluster (**Fig. 5a**). An additional three cytosines that form P6b (C286-288) are weakly reactive in the major cluster but strongly reactive in the minor cluster. Structure prediction of the major cluster yields the canonical three-way junction CR4/5 architecture (**Fig. 5b**). The minor cluster is predicted to form a register-shifted stem loop, with strongly reactive P6a nucleotides (A285-C290) forming the apical loop (**Fig. 5c**). Notably, this alternative conformation of CR4/5 lacks the P6.1 stem loop, which is essential for telomerase assembly and catalytic activity. This predicted alternative CR4/5 conformation is unlikely to arise from an artifact of DREEM clustering, as it is a predicted structure with similar free energy to the canonical three-way junction form even in constraint-free structure prediction runs (**Fig. S3**). We note that while major and minor clusters are observed for both the t/PK and CR4/5 domains, our sequencing library preparation method splits the two domains prior to DREEM deconvolution, and therefore does not provide information regarding whether the two minor conformations reside in the same molecule.

To quantitatively evaluate the quality of models produced by our DREEM deconvolution analysis we employed the area under the receiver operating characteristic curve (AUROC)[45] (**Fig. S4**). Previous work has established that AUROC values of ~0.95 or higher indicate a high degree of concordance between the predicted model and DMS modification pattern, whereas AUROC values approaching ~0.5 arise in the case of little to no agreement between the structural model and probing data. We find the CR4/5 canonical (AUROC = 0.97) and alternative (AUROC = 0.90) conformations are highly consistent with the deconvoluted chemical probing data. In the case of the t/PK domain, the canonical (AUROC = 0.87) and alternative (AUROC = 0.94) conformations are also congruent with the probing data. However, we note that the slightly reduced AUROC value for the t/PK canonical conformation indicates a higher level of structural heterogeneity for this domain and suggests the presence of slightly more conformations in the structural ensemble than can be resolved by our deconvolution analysis. Given the high-quality modeling (**Fig. S4**) and the excellent reproducibility of the reactivity patterns for the alternative

conformation of the CR4/5 domain (**Fig. S2b**), we elected to focus on the functional impacts of folding heterogeneity within this domain.

The alternative CR4/5 conformation does not support efficient telomerase assembly and activity

To test the biological effect of the alternative CR4/5 on telomerase biogenesis, we aimed to bias the hTR folding landscape toward this conformation in vivo and query its effects on telomerase RNP assembly and catalytic activity. We identified three different nucleotide positions within the CR4/5 P6b stem that, when mutated, are predicted to disrupt base pairing interactions in the canonical conformation and/or introduce stabilizing base pairs in the alternative conformation without directly disrupting interactions with hTERT (**Fig. 6a and 6b**). Our designed hTR mutant constructs contain either one (M1: G270C), two (M2: C277G, A281U), or three (M3: G270C, C277G, A281U) nucleotide substitutions in the CR4/5 domain.

Each of the hTR variants was sub-cloned into a plasmid under the control of the U1 promoter and transfected into HeLa cells together with a separate plasmid expressing hTERT with an N-terminal FLAG tag. Following each transfection, we performed the live cell DMS MaPseq procedure to analyze the folding properties of each transiently overexpressed hTR variant. The resultant DMS profiles of all mutants displayed elevated DMS reactivity at nucleotides C286-288, consistent with an increased relative abundance of the alternative CR4/5 conformation compared to the canonical conformation (**Fig. 6c**). Corroborating this finding, DREEM deconvolution followed by structure prediction revealed an increased representation of the alternative CR4/5 conformation in the ensemble. The CR4/5 structural ensemble of M1 and M2 is ~45% folded in the alternative conformation, while M3 adopts the alternative conformation with ~90% abundance, representing a drastic redistribution of the CR4/5 structural ensemble away from the canonical conformation (**Fig. 6c**). We note that the redistribution of molecules in the two conformations in response to these mutations provides additional evidence supporting the structural models.

As a control, we overexpressed wild type hTR and analyzed its structure. Interestingly, overexpressed wild type hTR had an increased proportion of alternatively folded CR4/5 compared to endogenous wild type hTR (~30% vs ~13%, respectively) (**Fig. S5**). This is potentially an effect of increased hTR concentration in the nucleus overwhelming some component (such as a protein chaperone) that aids hTR folding.

Having measured the increased abundance of the alternative CR4/5 conformation in cells transfected with the mutant hTR variants, we next set out to investigate the impact of the observed change in CR4/5 folding on telomerase RNP assembly and catalytic activity. For these experiments we transiently transfected either HeLa or HEK293T cells with the WT and mutant hTR constructs together with FLAG-tagged hTERT. Following immunopurification of the telomerase RNPs from each transfection experiment via the FLAG-tagged hTERT subunit, we performed direct primer extension activity assays on the IP fraction. M1 and M2 each exhibited moderately reduced total activity compared to WT (~70-80% activity) while M3 was severely deficient (~5% activity) (**Fig. 7a, 7b**). We note that the residual catalytic

activity observed for the M3 experiment may result from assembly of endogenously expressed hTR with FLAG-tagged hTERT. Importantly, Western blots for FLAG-hTERT show that perturbation of the CR4/5 folding landscape does not alter the steady-state amount of hTERT (**Fig. 7c, top panel**). Thus, the reduction in telomerase catalytic activity cannot be explained by reduced expression of the hTERT subunit.

We next asked if the reduced telomerase activity reflected less active assembled RNP enzymes or, rather, a failure to assemble RNP complexes. When telomerase was immunopurified by the FLAG epitope tag on the hTERT protein, Northern blotting for hTR in the IP and extract fractions revealed a significant enrichment of WT, M1 and M2 hTR variants (**Fig. 7d**). In contrast, the M3 hTR variant was not enriched in the IP, suggesting it has an impaired capacity to bind and assemble with hTERT (**Fig. 7d,** compare lanes 3 and 6). As a further test of whether each hTR variant was competent to assemble into a functional telomerase RNP complex, we performed Western blots for the dyskerin (DKC1) and Gar1 subunits of the H/ACA RNP complex (**Fig. 7c,** bottom two panels). Our results show that overexpression of the M1 and M2 RNAs results in modest reduction in both DKC1 and Gar1 levels in the IP fraction, while M3 knocks down DKC1 and Gar1 levels much more drastically. Since DKC1 and Gar1 bind to the ScaRNA domain of hTR, they would be expected to copurify only with FLAG-hTERT that was bound to hTR; thus, these results provide further evidence for the inability of the M3 hTR variant to co-assemble with TERT. Taken together, these results support the conclusion that the M3 hTR variant, which forms the alternative CR4/5 fold in vivo, has substantially decreased telomerase RNP assembly efficiency.

Discussion

Unlike the functional multi-state RNA folding properties of riboswitches [46], hTR structure is known to adopt a phylogenetically conserved secondary structure [10] and therefore would not be predicted to adopt long-lived alternative conformations in vivo. Thus, the observation that 12-18% of cellular hTR exists as a misfolded conformer was unexpected. The propensity of RNA to fall into kinetic traps during folding is well established in vitro; however, a reasonable expectation would be that such misfolded RNAs would be subject either to refolding or to degradation in vivo [47, 48, 49, 50]. Instead, we found that alternative conformations of the t/PK and CR4/5 – conformations that are inconsistent with the cryo EM models of telomerase – persist in the steady state at endogenous levels of expression. If the alternative hTR conformer is being refolded or degraded, the rates of such processes must be slow enough to give a substantial misfolded population in the steady state. On the other hand, it also remains possible that the alternative hTR conformer could exist in a different RNP with a non-telomerase function, though this notion is speculative and would require future exploration.

We tested whether hTR that adopted the alternative CR4/5 structure was in fact catalytically incompetent in vivo, as expected from previous work on the function of the three-way junction. Using structure-guided RNA mutagenesis, we were able to bias the hTR folding landscape away from the canonical fold and toward the alternative conformation. The hTR variants that modestly altered the hTR folding landscape to favor the alternative conformation (M1 and M2) supported substantial levels of telomerase assembly and ~75% catalytic activity. In contrast, the M3 hTR variant resulted in the majority

of cellular hTR adopting the alternative conformation, largely eliminating RNP assembly and telomerase activity.

In vitro biochemical and biophysical analyses of the hTR CR4/5 domain show that the RNA primarily folds into a heterogeneous ensemble of states[43]. Moreover, reconstitution of hTR with hTERT in rabbit reticulocyte lysates chaperones the RNA into a fold that is consistent with the recent cryo EM structures of the endogenously assembled RNP. In HeLa cells, the RNA subunit of telomerase is in stoichiometric excess of hTERT [14]. Thus, our observation that the majority of hTR in the cell, which is likely not in complex with hTERT, adopts a canonical CR4/5 conformation suggests that proteins other than hTERT promote functional hTR folding. The discovery of a histone H2A/H2B dimer directly bound to CR4/5 in recent cryo EM models [19, 21, 22] provides a strong candidate for such an RNA chaperoning activity. In fact, H2A/B could potentially perform a function analogous to that of the Tetrahymena telomerase holoenzyme protein p65, which induces an RNA conformational change to promote RNP assembly [51, 52, 53, 54, 55]. Modest stabilization of the alternative CR4/5 fold potentially allows the M1 and M2 mutants to be refolded by nuclear chaperones, explaining why they have only small effects on telomerase assembly and catalytic activity. In contrast, the M3 hTR variant, which almost exclusively adopts the alternative CR4/5 conformation, presumably has a higher energetic barrier for RNA refolding and thus is unable to be rescued.

Our characterization of the alternative pseudoknot-lacking conformer of the t/PK domain builds upon previous studies of its heterogeneous folding in human[38, 56]and ciliate telomerase RNA [57, 58, 59, 60]. We observed high and moderate levels of DMS reactivity in template sequence (nt 46-57) of both of our predicted t/PK conformations. This result suggests that assembly with hTERT is primarily associated with the formation of the pseudoknot-forming P3 helix and does not significantly reorganize the template region as was previously suggested[38]. Future experiments designed to manipulate the t/PK RNA folding ensemble will yield a richer understanding of how conformational changes of this domain promote telomerase biogenesis.

More generally, we demonstrate the efficacy of using in cell chemical probing in combination with bioinformatic deconvolution to parse individual RNA conformations of a noncoding RNA. Our findings lead us to speculate that heterogeneous hTR folding presents a barrier in the telomerase assembly pathway. This barrier may be overcome via the action of RNA folding chaperones that rescue improperly folded hTR molecules, providing a possible mechanism for regulating telomerase biogenesis pathway and catalytic activity (**Fig. 8**). These same techniques could be used to explore the populations of other functional noncoding RNAs in cells. It will be interesting to determine if the substantial subpopulation of misfolded RNA seen for hTR is specific to telomerase or perhaps occurs more generally with long noncoding RNAs.

Methods

Cell culture and DMS modification

HeLa and HEK293T/17 cells (ATCC CRL-11268, lot:70040949) were cultured in DMEM containing 10% fetal bovine serum, 100 units/mL of penicillin, 100 mg/m of streptomycin, and 1X Gibco GlutaMAX at 37°C, 5% CO2. Cells were seeded into 6-well plates at 0.3 x 10^6 cells per well and allowed to grow to 90% confluency prior to use. For DMS modification, plates of HeLa cells were divided into DMS and control groups (3 wells each). Media from the DMS group was aspirated and replaced with 2 mL of a 2% DMS (Sigma) solution in cell media (prepared by brief vortexing to incorporate DMS). Cells were incubated at 37°C for 4 min, media was aspirated from each well and replaced with 10 mL of ice cold DMS quench solution (30% beta-mercaptoethanol in phosphate-buffered saline). Plates were placed on ice and the cells were collected by scraping and brief trituration, then pelleted by centrifugation at 200 x g for 5 min. Cells were washed once with cold PBS and total RNA harvested with 1 mL Trizol (Invitrogen) following the manufacturer's protocol.

Transient transfection and purification of human telomerase

Plasmids containing hTERT (pvan107-3xFLAG) and WT or mutant hTR (pBSU1-hTR) were transfected at a 1:3 molar ratio using lipofectamine 2000 (11668019, Thermo Fisher Scientific). The cells were further expanded 3-fold 24 h after transfection and then 24 h later either harvested or used in DMS probing reactions. To purify telomerase, the cell pellet was lysed with CHAPS lysis buffer (10 mM Tris–HCl pH 7.5, 1 mM $MgCl_2$, 1 mM EGTA, 0.5% CHAPS, 10% glycerol, 5 mM beta-mercaptoethanol,) for 45 min at $4^{O}C$ on a rotator. The lysate was then clarified by centrifugation at 13,000 x g at $4^{O}C$ for 30 min. Anti-FLAG resin (A2220, Sigma-Aldrich) was added to the clarified supernatant and the samples incubated on a rotator for 4 h (or overnight) at $4^{O}C$. The anti-FLAG resin was washed 3x with wash buffer (20 mM HEPES–NaOH pH 8.0, 2 mM $MgCl_2$, 0.2 mM EGTA, 0.1% NP-40, 10% glycerol, 1 mM DTT) before elution using wash buffer supplemented with 0.25 mg/ml 3xFLAG peptide (F4799, Sigma-Aldrich). Purified telomerase complex was verified by western blotting.

Preparation of mutant hTR plasmids

Starting with the pBS-U1-hTR plasmid [61] three hTR mutants (M1: G270C, M2: C277G & A281U, M3: G270C & C277G & A281U) were generated using the mega primer method with gel extraction of intermediate products after the first PCR step[62].

Preparation of hTR-targeted MaPseq libraries

10 µg of HeLa cell derived total RNA was DNAse treated in 1X TURBO DNAse buffer with 1 µL TURBO DNAse enzyme (Thermo Fisher Scientific), followed by column-purification with DNA Clean and Concentrator-5 columns (Zymo Research) following the manufacturer protocol. 5 µg of DNAse-treated RNA was mixed with 5 pmol hTR targeting primer containing a 4-nucleotide unique molecular identifier (UMI) in a volume of 11 µL, heated to 75°C for 3 min and annealed at 35°C for 15 min. Next, 4 µL of 5X

M-MLV buffer (Promega), 1 µL 0.1M DTT, 1 µL RNAsin Plus (Promega), and 1 µL TGIRT III (Ingex) were added and the mixture was incubated at room temperature for 30 min. Then, 2 µL of 10mM dNTPs were added, and the reaction was mixed and incubated at 60 °C for 2.5 h. After RT, 1 µL of 5M NaOH was added directly to the cDNA and incubated at 95°C, followed by 2.5 µL of 2M HCl and purification by DNA Clean and Concentrator-5 columns (Zymo Research) using an 8:1 ratio of DNA Binding Buffer to cDNA volume. The cDNA was mixed into a 50 µL second strand synthesis reaction containing 10 µL 5X GC Phusion Buffer (Thermo Fisher Scientific), 1 µL 10 mM dNTP, 25 pmol of a second hTR targeting primer containing a 4 nucleotide UMI, and 1 µL Phusion polymerase. The second strand synthesis reaction was incubated in a thermocycler with the following program: 98°C 2 min, 60°C 2 min, 72°C 10 min. Second strand product was cleaned up with Ampure XP beads (Beckman Coulter) at a 0.8:1 bead to sample volume ratio and eluted in 20 µL nuclease-free H_2O. Illumina sequencing libraries of two partially overlapping hTR regions (Pseudoknot: nucleotides 1-286 and CR4/5: nucleotides 193-451) were generated by two successive 50 µL PCR reactions. For the first PCR, 4 µL second strand product was added to a 50 µL PCR reaction containing primers for either the pseudoknot or CR4/5 amplicon and run with the following program: 10 cycles at 66°C annealing temperature, followed by 10 cycles at a 60°C annealing temperature. Amplicons were bead cleaned with a 0.8:1 ratio and went into a second PCR of 10 cycles at an annealing temperature of 65°C to complete the sequencing adapters and multiplexing barcodes. Libraries were bead cleaned with a 0.8:1 ratio and quantified by Qubit (Thermo Fisher Scientific) and Tapestation (Agilent) for quality metrics before sequencing by 150 bp paired end reads on the iSeq100 (Illumina).

Western blots

The presence of dyskerin and GAR1 in the WT and mutant telomerase complexes immunopurified from HEK293T cells was analyzed by western blotting. The primary antibodies were anti-hTERT (600-401-252, Rockland), anti-DKC1 (NBP3-16405, Novus) and anti-GAR1 (NBP2-31742). Secondary antibody was anti-rabbit (711-035-152, Jackson ImmunoResearch, West Grove, PA). All primary antibodies were diluted 1:1000 for blotting. The dilution for the secondary antibody was 1:5000. Serial dilutions of the FLAG-immunopurified telomerase were run on denaturing polyacrylamide gels, transferred to Protran 0.45 mm nitrocellulose (10600002, Amersham), and probed with the primary antibodies followed by the secondary antibody. The western membrane was developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (34578, Thermo Scientific) and visualized using a FluorChem R imaging camera (Protein Simple). Quantification was done using AlphaView software version 3.5.0.927 (Protein Simple).

Telomerase activity assays

Activity of the immunopurified human telomerase complex over-expressed in either HEK293T/17 (ATCC CRL-11268, lot:70040949) or HeLa cells was determined by a direct assay modified from a published protocol (Wang et al. 2007). The reaction mixture (20 mL) contained 1x human telomerase assay buffer (50 mM Tris-HCl at pH 8.0, 50 mM KCl, 75 mM NaCl, 1 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 1 mM spermidine), 0.05 mM telomeric DNA primer, 0.5 mM dTTP, 3.3 mM dGTP, 10 mM dATP, and 0.33 mM ^{32}P-dATP (3000 Ci/mmol, 1 Ci = 37GBq, PerkinElmer). Following a 1 h incubation at 30°C, reactions were stopped with the addition of 100 mL of 3.6 M NH_4OAc containing 20 mg of glycogen. Ethanol (500 mL)

was added for precipitation. After incubating for 1 h at 80°C, samples were centrifuged for 15 min at 4°C. Pellets were washed with 70% ethanol and resuspended in 10 mL of H_2O followed by 10 mL of 2x loading buffer (94% formamide, 0.1x TBE, 0.1% bromophenol blue, 0.1% xylene cyanol). The heat-denatured samples were loaded onto a 10% polyacrylamide/7 M urea/1x TBE gel for electrophoresis (until bromophenol blue runs to the bottom of the gel). After electrophoresis, the gel was dried and quantified by using a PhosphorImager (Cytiva).

hTR Northern blots

Deproteinized CHAPs extract or deproteinized immunopurified telomerase was run on a 6% polyacrylamide/7 M urea/1xTBE gel. Nucleic acid was transferred to Hybond-N+ membrane at 2 amps in 0.5x TBE at 4°C for 1.5 h. Nucleic acid was crosslinked to the membrane using a CL-1000 ultraviolet crosslinker using 254 nm lamps with an energy level of 1200 x 100 mJ/cm^2. Membrane was blocked using rapid-hyb buffer for 30 min at 50°C. 10^7 counts of a DNA probe labeled with γ-^{32}P-ATP and polynucleotide kinase was added and incubated overnight at 50°C. Membrane was washed 3x with 2x SSC/0.1% SDS for 15 min at 50°C, followed by 1x for 15 min with 2x SSC/0.1% SDS. Membrane was wrapped in saran wrap and exposed to a phosphor screen (Cytiva).

DREEM analysis

Trimming of sequencing adapters and removal of reads smaller than 120 nucleotides was performed using TrimGalore with the following command: trim_galore –paired –nextera R1.fastq R2.fastq –length 120. FASTQ files were aligned using bowtie2[63] to the hTR reference sequence using the following parameters: --local --no-unal --no-discordant --phred33. Alignment files were analyzed with a local build of DREEM (Deconvolution of RNA Ensembles using Expectation Maximization)[40]. The start and end parameters for the Pseudoknot/Template and CR4/5 domains were 22-204, and 213-368, respectively. Folding predictions were performed using default parameters of the RNAStructure package (Reuter, 2010). Pseudoknot prediction was performed using the ProbKnots [64] function of RNAstructure with default folding parameters.

Generating receiver operating characteristic curves and computing AUROC

The AUROC quantifies how well DMS/SHAPE reactivities support the predicted RNA structure, under the assumption that paired bases should be less reactive than unpaired bases. Based on the secondary structure, each position was labeled as paired or unpaired, and the DMS reactivities were partitioned into paired and unpaired groups based on these labels. The ROC curves and AUROC values were computed using SciPy[65]. Here, “true” and “false” positives represent, respectively paired and unpaired nucleotides with DMS reactivities less than the sliding threshold.

Data visualization

Bar and scatter plots were generated with Graphpad Prism. RNA secondary structures were generated with VARNA [66] and modified with Adobe Illustrator. Arc diagrams were generated using R-chie [67].

Figure 1

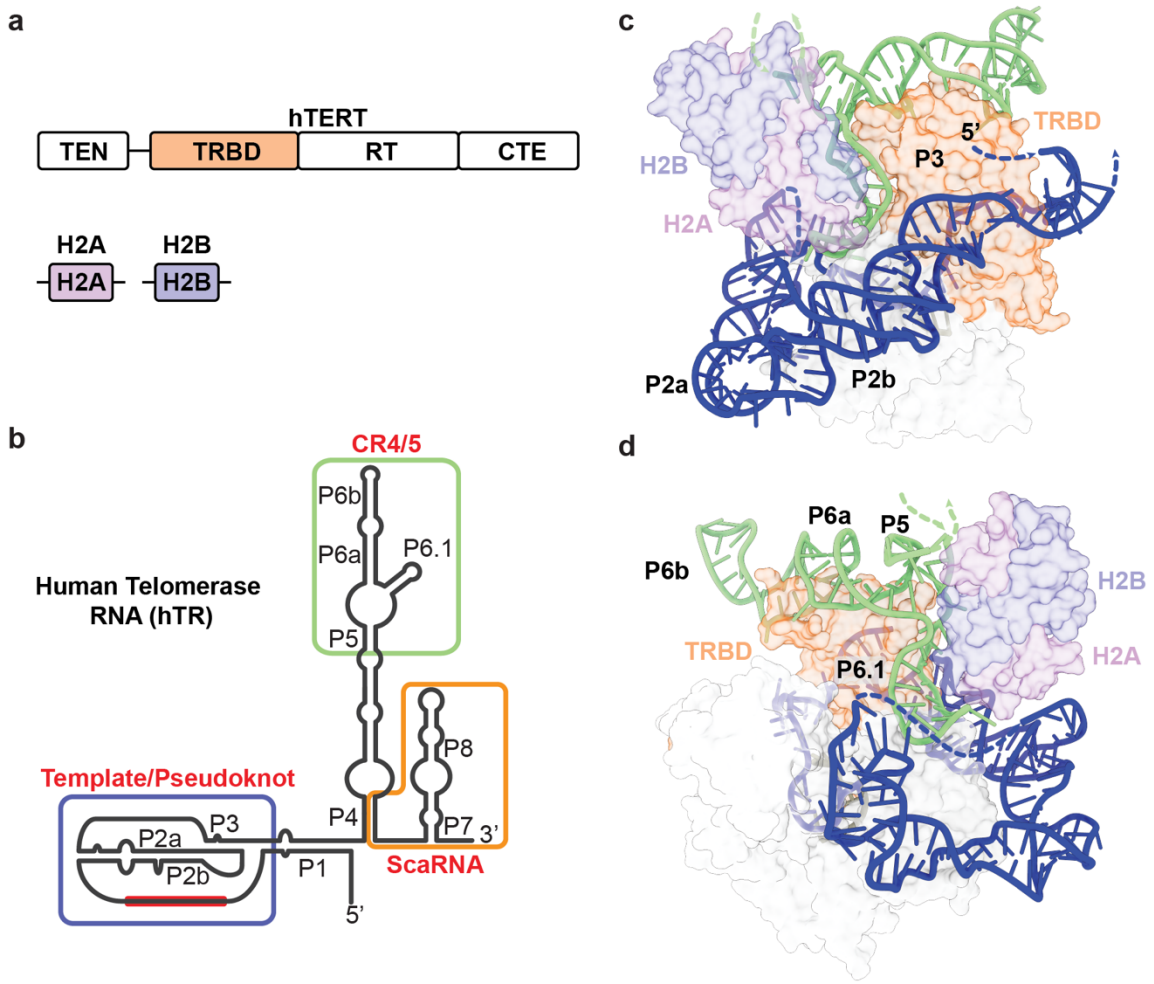

Fig 1. Overview of human telomerase components and hTR structure. (**a**) The conserved domain organization of TERT includes the telomerase essential amino-terminal domain (TEN), RNA binding domain (TRBD), reverse transcriptase domain (RT), and carboxy-terminal extension (CTE). A H2A/B dimer comprises one H2A and one H2B protein. (**b**) The vertebrate-conserved architecture of hTR includes the template/pseudoknot (t/PK, template sequence shown as red bar), CR4/5, and ScaRNA domains. P, paired region. (**c**) The hTR template/pseudoknot domain (in blue) wraps around TERT and forms the pseudoknot helix P3. TRBD, telomerase RNA binding domain. (**d**) CR4/5 (in green) adopts an 'L' shaped three-way junction and sandwiches TERT TRBD between the P6.1 and P6a/b stems. A dimer of H2A/B binds P5 and P6.1. GIVE EITHER PDB NUMBERS OR CITE REF FOR c AND d.

Figure 2

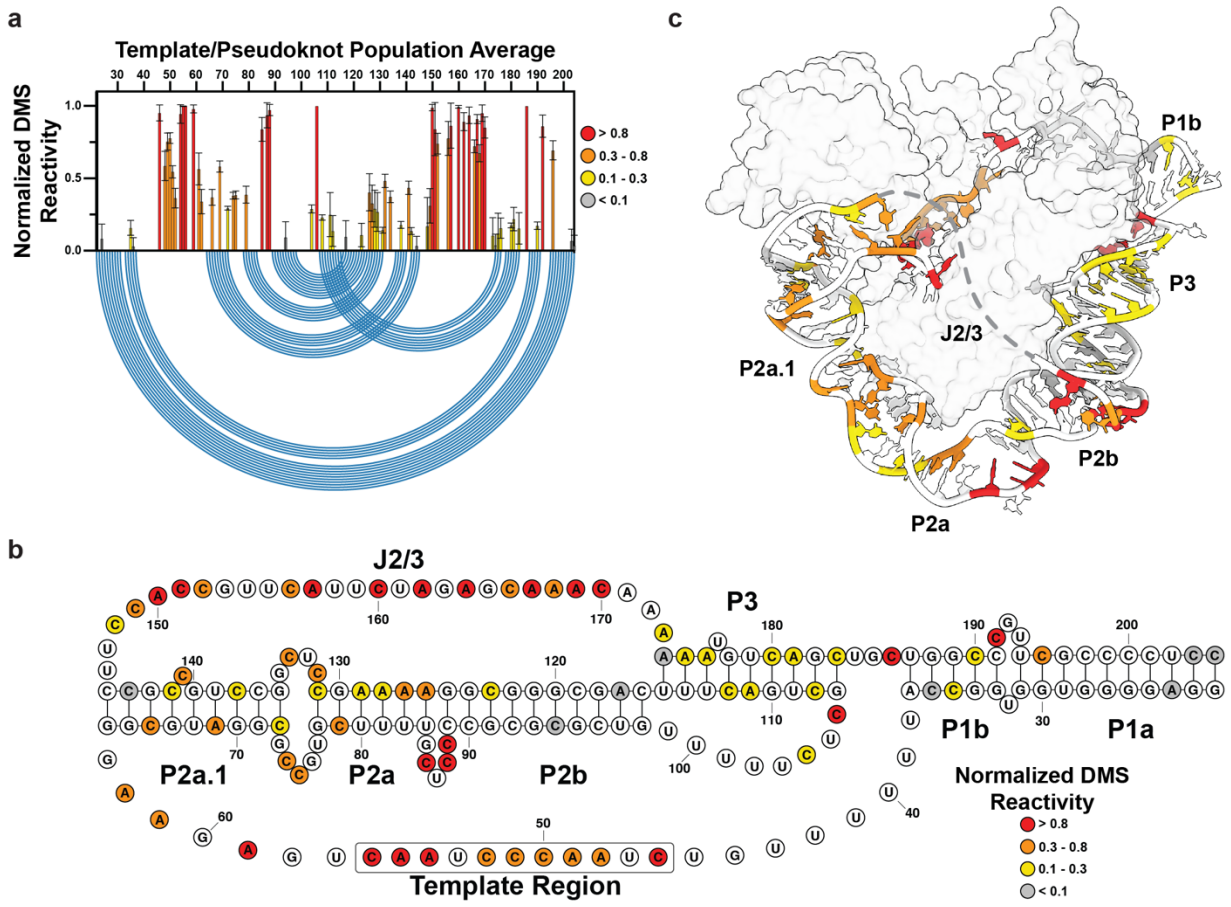

Fig 2. Population average DMS reactivity of the hTR t/PK domain. (**a**) Normalized DMS reactivity of hTR t/PK domain (nt. 22-204). Intensity of DMS reactivity colored according to the provided legend. Blue arcs designate the base pairing pattern of the canonical t/PK conformation seen by cryo-EM of assembled telomerase. (**b**) Secondary structure of the canonical t/PK conformation with DMS reactivity of the hTR population average overlaid onto the nucleotides. Stem elements (P), joining regions (J) and the template sequence are labeled. (**c**) DMS reactivity overlaid onto the cryo-EM model of assembled telomerase.

Figure 3

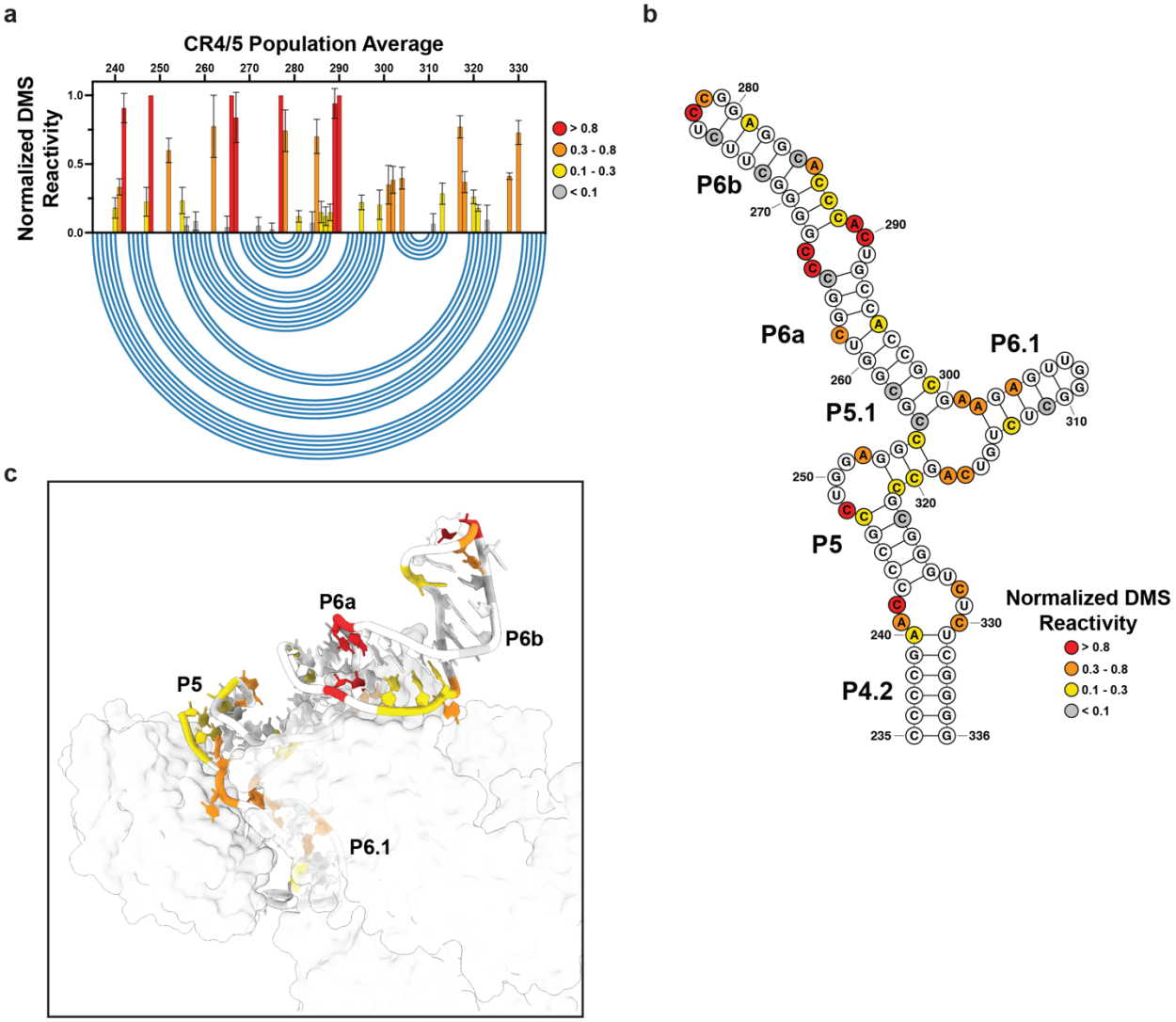

Fig 3. Population average DMS reactivity of the hTR CR4/5 domain. (**a**) Normalized DMS reactivity of hTR CR4/5 domain (nt. 235-336). Intensity of DMS reactivity colored according to the provided legend. Blue arcs designate the base pairing pattern of the canonical t/PK conformation seen by cryo-EM of assembled telomerase. **(b)** Secondary structure of the canonical CR4/5 conformation with DMS reactivity of the hTR population average overlaid onto the nucleotides. **(c)** DMS reactivity overlaid onto the cryo-EM model of assembled telomerase.

Figure 4

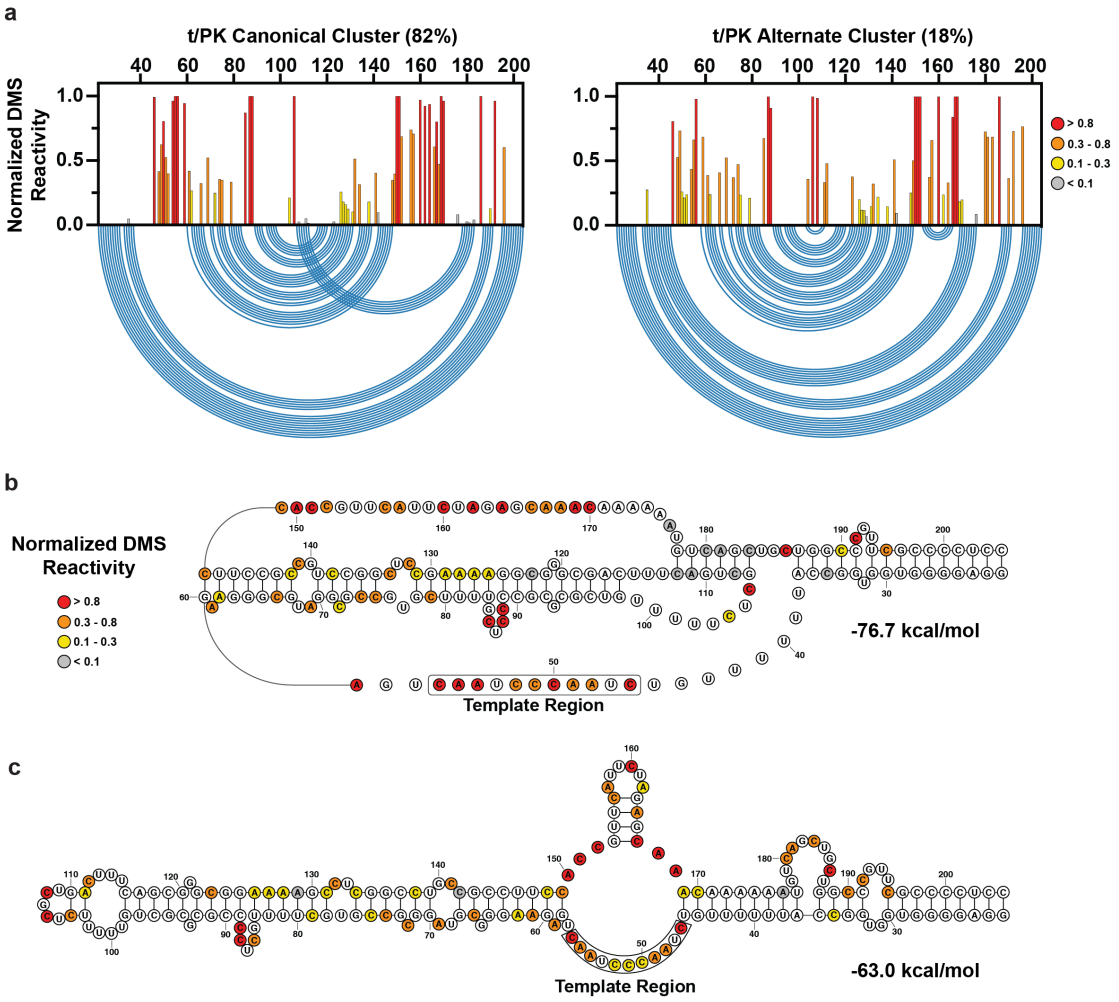

Fig 4. DREEM-deconvoluted DMS profiles and structure predictions of hTR t/PK. (**a**) Normalized DMS reactivity profiles of the two clusters predicted by DREEM. Intensity of DMS reactivity colored according to the provided legend. Blue arcs designate the base pairing pattern of the data-guided predicted t/PK secondary structure. (**b**) Data-guided secondary structure prediction of the t/PK domain from the canonical cluster of DMS reactivities. (**c**) Data-guided secondary structure prediction of the t/PK domain from the alternative cluster of DMS reactivities.

Figure 5

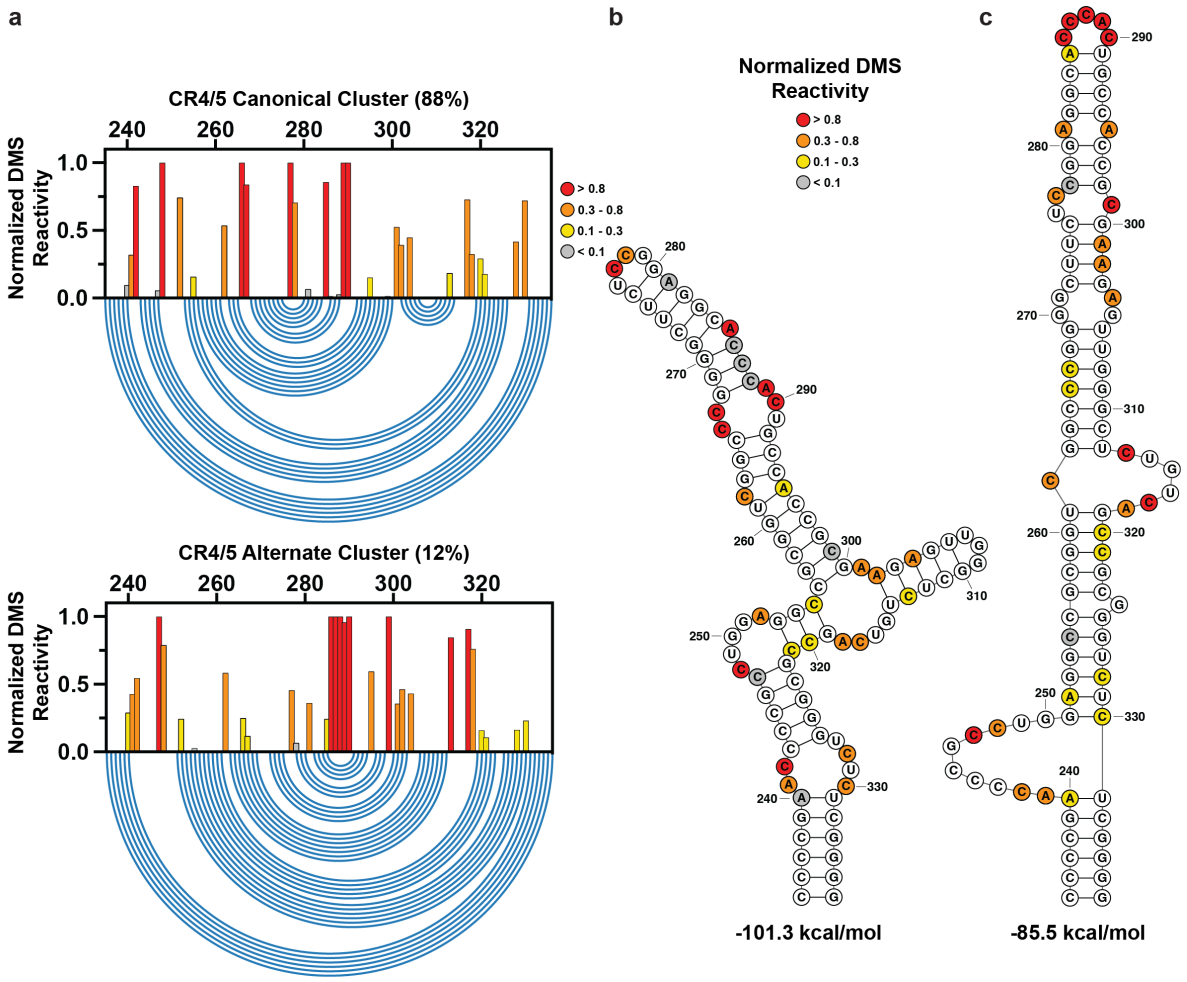

Fig 5. DREEM-deconvoluted DMS profiles and structure predictions of hTR CR4/5. (**a**) Normalized DMS reactivity profiles of the two clusters predicted by DREEM. Intensity of DMS reactivity colored according to the provided legend. Blue arcs designate the base pairing pattern of the data-guided predicted CR4/5 secondary structure. (**b**) Data-guided secondary structure prediction of the CR4/5 domain from the canonical cluster of DMS reactivities. (**c**) Data-guided secondary structure prediction of the CR4/5 domain from the alternative cluster of DMS reactivities.

Figure 6

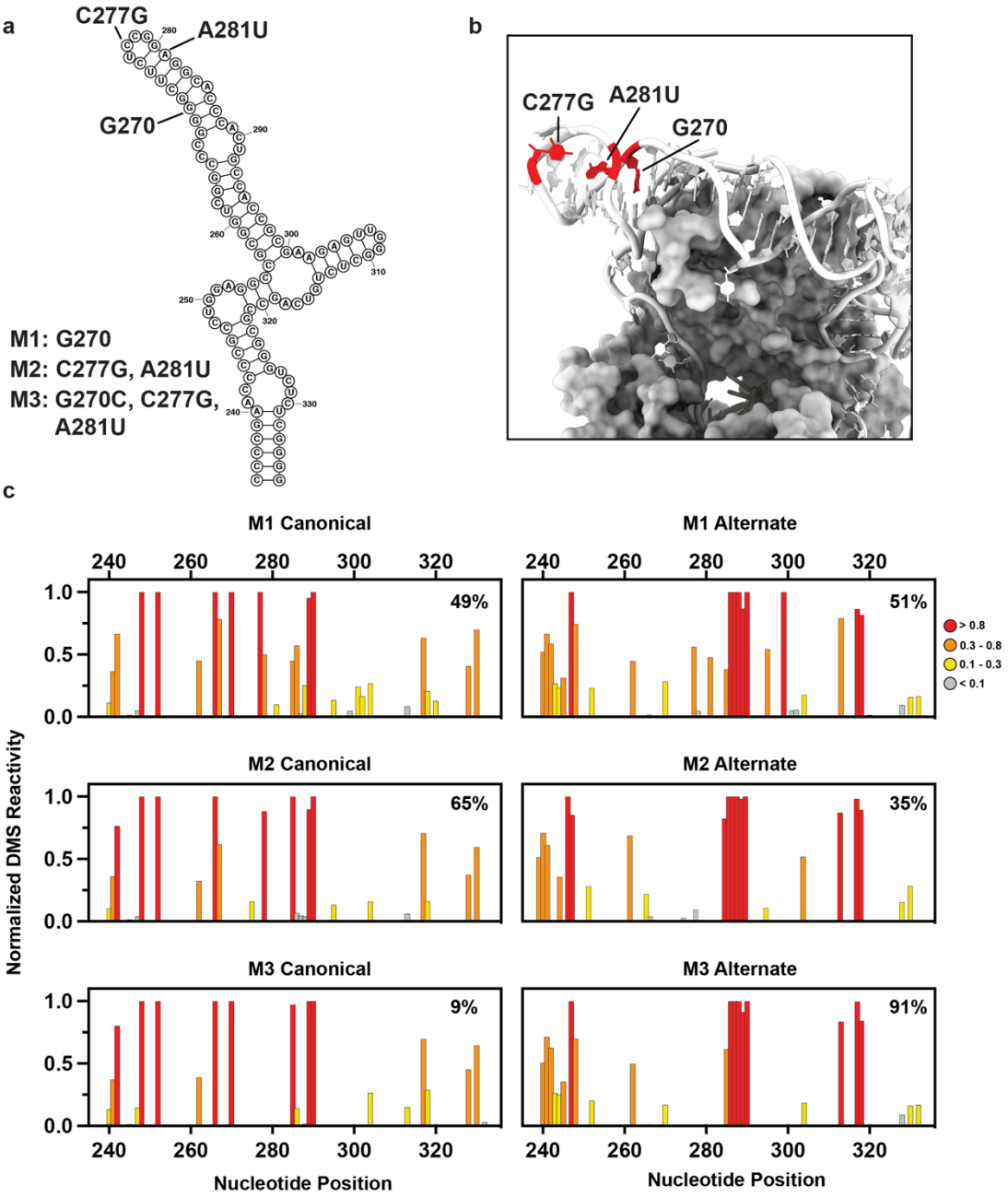

Fig. 6. Design and MaPseq validation of CR4/5 mutants. (**a**) Canonical CR4/5 secondary structure with mutated nucleotides labeled. (**b**) Cryo-EM structure of CR4/5 with mutated nucleotides highlighted. (**c**) Normalized DMS reactivity profiles of clusters predicted by DREEM for each CR4/5 mutant.

Figure 7

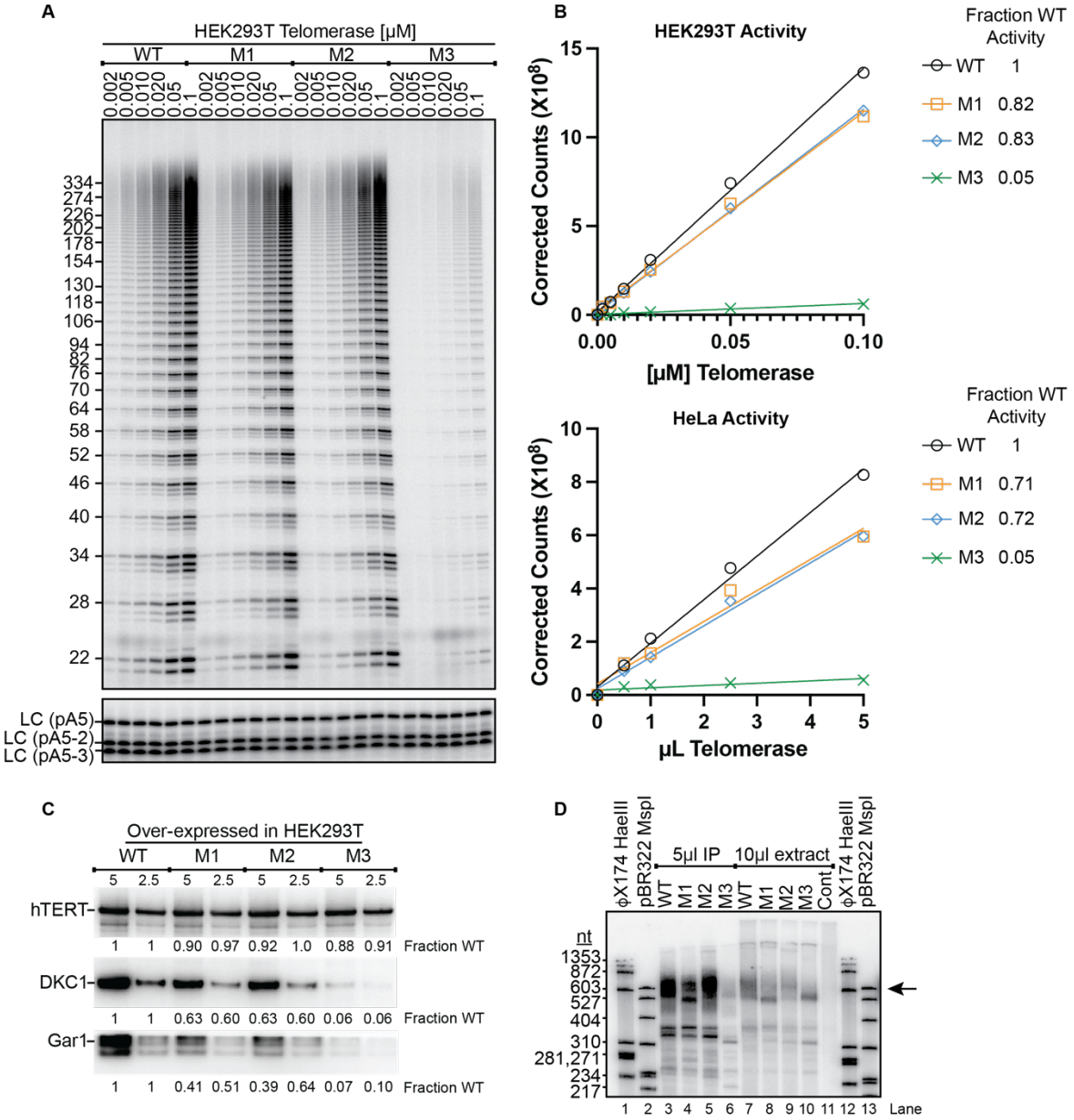

Fig. 7. Telomerase activity assay and RNP assembly for hTR mutants. (**a**) Activity assay of FLAG-purified HEK293T-derived telomerase measuring incorporation of ^{32}P-dGTP into a telomeric DNA primer for the WT, M1, M2, and M3 hTR variant constructs. LC, labeled oligonucleotide loading controls. (**b**) Quantification of telomerase activity. HEK293T- or HeLa-derived telomerase activity calculated by summing the total lane signal, correcting for differences in intensity of the loading controls, and normalizing to WT. Fractional activity relative to WT was calculated from the slopes of the linear regression fits. (**c**) Western blots for hTERT (top), DKC1 (middle), and Gar1 (bottom) all performed after IP of FLAG-hTERT. Relative amount of signal compared to wild type for two different loading amounts (normalized to hTERT signal) is shown below each lane. (Gar1 values are approximate due to accidental

truncation of tops of bands.) (**d**) Northern blot for hTR in both the input cell extract and IP fractions for HEK293T cell transfection experiments with WT and variant hTR constructs. The arrow on right side of gel indicates migration of full-length hTR which is known to migrate as multiple bands on a denaturing gel.

Figure 8

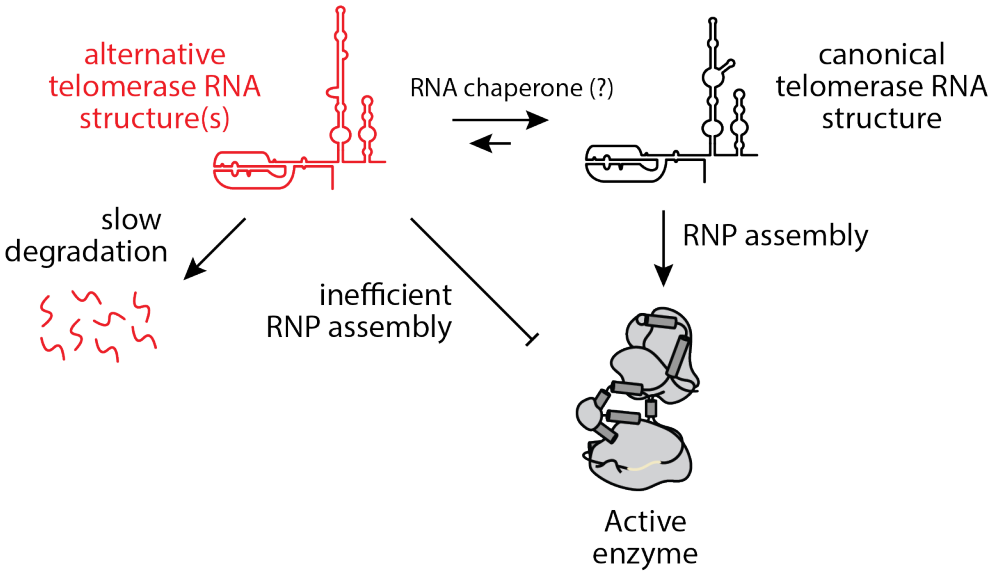

Fig. 8. Model of telomerase RNP assembly with respect to hTR folding. hTR exists as a structural ensemble in cells, with a minority population adopting alternative conformations in the t/PK and CR4/5 domains. The canonical fold of hTR may be favored by the presence of RNA folding chaperones. hTR molecules adopting the canonical fold are efficiently co-assembled into a functional RNP, whereas hTR molecules in the alternative conformation(s) are not efficiently assembled in the RNP complex or are slowly degraded.

Author Information

Contributions

N.M.F. and J.Z.W. designed, conducted, and analyzed MaPseq experiments of endogenous hTR. N.M.F. and M.D.S designed, conducted, and analyzed MaPseq experiments of transfected hTR. A.J.G. contributed to sequencing library preparation. A.J.Z. designed, conducted, and analyzed all immunopurifications, telomerase activity assays, Western blots, and Northern blots. M.D.S., S.R., and T.R.C contributed to overall experimental design, data analysis, and interpretation. N.M.F. and M.D.S. designed all manuscript figures. All authors participated in the writing of the manuscript.

Acknowledgments

This work was supported by the National Institutes of Health R01GM095850 to M.D.S. T.R.C. is an investigator of the Howard Hughes Medical Institute. S.R. is supported by the Burroughs Wellcome Fund (#1018246.01). N.M.F was supported by NIH T32 GM133391.

Ethics declarations

Competing Interests

T.R.C. is a scientific advisor for Eikon Therapeutics, Storm Therapeutics, and Somalogic, Inc.

REFERENCES

1. Blackburn EH, Collins K. Telomerase: an RNP enzyme synthesizes DNA. *Cold Spring Harb Perspect Biol* **3**, (2011).

2. de Lange T. Shelterin-Mediated Telomere Protection. *Annu Rev Genet* **52**, 223-247 (2018).

3. Greider CW, Blackburn EH. Tracking telomerase. *Cell* **116**, S83-86, 81 p following S86 (2004).

4. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature* **345**, 458-460 (1990).

5. Olovnikov AM. Telomeres, telomerase, and aging: origin of the theory. *Exp Gerontol* **31**, 443-448 (1996).

6. Hayflick L. The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp Cell Res* **37**, 614-636 (1965).

7. Shay JW, Wright WE. Hayflick, his limit, and cellular ageing. *Nat Rev Mol Cell Biol* **1**, 72-76 (2000).

8. Allsopp RC*, et al.* Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci U S A* **89**, 10114-10118 (1992).

9. Hastie ND, Dempster M, Dunlop MG, Thompson AM, Green DK, Allshire RC. Telomere reduction in human colorectal carcinoma and with ageing. *Nature* **346**, 866-868 (1990).

10. Chen JL, Blasco MA, Greider CW. Secondary structure of vertebrate telomerase RNA. *Cell* **100**, 503-514 (2000).

11. Zappulla DC, Cech TR. Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proc Natl Acad Sci U S A* **101**, 10024-10029 (2004).

12. Antal M, Boros E, Solymosy F, Kiss T. Analysis of the structure of human telomerase RNA in vivo. *Nucleic Acids Res* **30**, 912-920 (2002).

13. Zemora G, Handl S, Waldsich C. Human telomerase reverse transcriptase binds to a pre-organized hTR in vivo exposing its template. *Nucleic Acids Res* **44**, 413-425 (2016).

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

15. Vogan JM, Collins K. Dynamics of Human Telomerase Holoenzyme Assembly and Subunit Exchange across the Cell Cycle. *J Biol Chem* **290**, 21320-21335 (2015).

16. Vogan JM*, et al.* Minimized human telomerase maintains telomeres and resolves endogenous roles of H/ACA proteins, TCAB1, and Cajal bodies. *Elife* **5**, (2016).

17. Jiang J*, et al.* Structure of Telomerase with Telomeric DNA. *Cell* **173**, 1179-1190 e1113 (2018).

18. Nguyen THD*, et al.* Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* **557**, 190-195 (2018).

19. Ghanim GE*, et al.* Structure of human telomerase holoenzyme with bound telomeric DNA. *Nature* **593**, 449-453 (2021).

20. He Y*, et al.* Structure of Tetrahymena telomerase-bound CST with polymerase alpha-primase. *Nature* **608**, 813-818 (2022).

21. Liu B, He Y, Wang Y, Song H, Zhou ZH, Feigon J. Structure of active human telomerase with telomere shelterin protein TPP1. *Nature* **604**, 578-583 (2022).

22. Sekne Z, Ghanim GE, van Roon AM, Nguyen THD. Structural basis of human telomerase recruitment by TPP1-POT1. *Science* **375**, 1173-1176 (2022).

23. Forino NM, Hentschel J, Stone MD. Cryo-EM structures tell a tale of two telomerases. *Nat Struct Mol Biol* **28**, 457-459 (2021).

24. Nguyen THD. Structural biology of human telomerase: progress and prospects. *Biochem Soc Trans* **49**, 1927-1939 (2021).

25. He Y, Feigon J. Telomerase structural biology comes of age. *Curr Opin Struct Biol* **76**, 102446 (2022).

26. Egan ED, Collins K. An enhanced H/ACA RNP assembly mechanism for human telomerase RNA. *Mol Cell Biol* **32**, 2428-2439 (2012).

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

28. Collins K. Physiological assembly and activity of human telomerase complexes. *Mech Ageing Dev* **129**, 91-98 (2008).

29. Schmidt JC, Cech TR. Human telomerase: biogenesis, trafficking, recruitment, and activation. *Genes Dev* **29**, 1095-1105 (2015).

30. Roake CM, Chen L, Chakravarthy AL, Ferrell JE, Jr., Raffa GD, Artandi SE. Disruption of Telomerase RNA Maturation Kinetics Precipitates Disease. *Mol Cell* **74**, 688-700 e683 (2019).

31. Tseng CK, Wang HF, Schroeder MR, Baumann P. The H/ACA complex disrupts triplex in hTR precursor to permit processing by RRP6 and PARN. *Nat Commun* **9**, 5430 (2018).

32. Klump BM*, et al.* TCAB1 prevents nucleolar accumulation of the telomerase RNA to facilitate telomerase assembly. *Cell Rep* **42**, 112577 (2023).

33. Laprade H*, et al.* Single-Molecule Imaging of Telomerase RNA Reveals a Recruitment-Retention Model for Telomere Elongation. *Mol Cell* **79**, 115-126 e116 (2020).

34. Venteicher AS*, et al.* A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* **323**, 644-648 (2009).

35. Armanios M, Blackburn EH. The telomere syndromes. *Nat Rev Genet* **13**, 693-704 (2012).

36. Vulliamy T*, et al.* The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* **413**, 432-435 (2001).

37. Chen JL, Opperman KK, Greider CW. A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic Acids Res* **30**, 592-597 (2002).

38. Deshpande AP, Collins K. Mechanisms of template handling and pseudoknot folding in human telomerase and their manipulation to expand the sequence repertoire of processive repeat synthesis. *Nucleic Acids Res* **46**, 7886-7901 (2018).

39. Robart AR, Collins K. Investigation of human telomerase holoenzyme assembly, activity, and processivity using disease-linked subunit variants. *J Biol Chem* **285**, 4375-4386 (2010).

40. Tomezsko PJ*, et al.* Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **582**, 438-442 (2020).

41. Zubradt M, Gupta P, Persad S, Lambowitz AM, Weissman JS, Rouskin S. DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat Methods* **14**, 75-82 (2017).

42. Zaug AJ, Cech TR. Analysis of the structure of Tetrahymena nuclear RNAs in vivo: telomerase RNA, the self-splicing rRNA intron, and U2 snRNA. *RNA* **1**, 363-374 (1995).

43. Palka C, Forino NM, Hentschel J, Das R, Stone MD. Folding heterogeneity in the essential human telomerase RNA three-way junction. *RNA* **26**, 1787-1800 (2020).

44. Morandi E*, et al.* Genome-scale deconvolution of RNA structure ensembles. *Nat Methods* **18**, 249-252 (2021).

45. Lan TCT*, et al.* Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nat Commun* **13**, 1128 (2022).

46. Breaker RR. Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* **4**, (2012).

47. Cruz JA, Westhof E. The dynamic landscapes of RNA architecture. *Cell* **136**, 604-609 (2009).

48. Herschlag D. RNA chaperones and the RNA folding problem. *J Biol Chem* **270**, 20871-20874 (1995).

49. Huang H, Karbstein K. Assembly factors chaperone ribosomal RNA folding by isolating helical junctions that are prone to misfolding. *Proc Natl Acad Sci U S A* **118**, (2021).

50. Woodson SA. Taming free energy landscapes with RNA chaperones. *RNA Biol* **7**, 677-686 (2010).

51. Berman AJ, Gooding AR, Cech TR. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol Cell Biol* **30**, 4965-4976 (2010).

52. Prathapam R, Witkin KL, O'Connor CM, Collins K. A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nat Struct Mol Biol* **12**, 252-257 (2005).

53. Singh M*, et al.* Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell* **47**, 16-26 (2012).

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

55. Witkin KL, Collins K. Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes Dev* **18**, 1107-1118 (2004).

56. Hengesbach M, Kim NK, Feigon J, Stone MD. Single-molecule FRET reveals the folding dynamics of the human telomerase RNA pseudoknot domain. *Angew Chem Int Ed Engl* **51**, 5876-5879 (2012).

57. Cash DD, Feigon J. Structure and folding of the Tetrahymena telomerase RNA pseudoknot. *Nucleic Acids Res* **45**, 482-495 (2017).

58. Cole DI*, et al.* New models of Tetrahymena telomerase RNA from experimentally derived constraints and modeling. *J Am Chem Soc* **134**, 20070-20080 (2012).

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

60. Theimer CA, Finger LD, Trantirek L, Feigon J. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci U S A* **100**, 449-454 (2003).

61. Cristofari G, Lingner J. Telomere length homeostasis requires that telomerase levels are limiting. *EMBO J* **25**, 565-574 (2006).

62. Tyagi R, Lai R, Duggleby RG. A new approach to 'megaprimer' polymerase chain reaction mutagenesis without an intermediate gel purification step. *BMC Biotechnol* **4**, 2 (2004).

63. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).

64. Bellaousov S, Mathews DH. ProbKnot: fast prediction of RNA secondary structure including pseudoknots. *RNA* **16**, 1870-1880 (2010).

65. Virtanen P*, et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261-272 (2020).

66. Darty K, Denise A, Ponty Y. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25**, 1974-1975 (2009).

67. Tsybulskyi V, Mounir M, Meyer IM. R-chie: a web server and R package for visualizing cis and trans RNA-RNA, RNA-DNA and DNA-DNA interactions. *Nucleic Acids Res* **48**, e105 (2020).