Cryo-EM structure of substrate-bound human telomerase holoenzyme

Thi Hoang Duong Nguyen1,2,3,4, Jane Tam1, Robert A. Wu1,6, Basil J. Greber2,3, Daniel Tos02, Eva Nogales1,2,3,5*, & Kathleen Collins1,2,5*

The enzyme telomerase adds telomeric repeats to chromosome ends to balance the loss of telomeres during genome replication. Telomerase regulation has been implicated in cancer, other human diseases, and ageing, but progress towards clinical manipulation of telomerase has been hampered by the lack of structural data. Here we present the cryo-electron microscopy structure of the substrate-bound human telomerase holoenzyme at subnanometre resolution, showing two flexibly RNA-tethered lobes: the catalytic core with telomerase reverse transcriptase (TERT) and conserved motifs of telomerase RNA (hTR), and an H/ACA ribonucleoprotein (RNP). In the catalytic core, RNA encircles TERT, adopting a well-ordered tertiary structure with surprisingly limited protein–RNA interactions. The H/ACA RNP lobe comprises two sets of heterotetrameric H/ACA proteins and one Cajal body protein, TCAB1, representing a pioneering structure of a large eukaryotic family of ribosome and spliceosome biogenesis factors. Our findings provide a structural framework for understanding human telomerase disease mutations and represent an important step towards telomerase-related clinical therapeutics.

**Abstract**

Linear eukaryotic chromosomes are capped by telomeric repeat tracts that recruit the telomere-binding proteins that are essential to distinguish telomeres from DNA breaks and thus to avoid telomere end-resection and interchromosomal fusions. Owing to inherently incomplete genome replication, telomeres are progressively shortened in each cell cycle. Critically short telomeres result in genome instability and cell death. To compensate for this sequence loss, a specialized reverse transcriptase, telomerase, adds telomeric repeats to the chromosome end, using TERT and an integral telomerase RNA subunit (hTR) with an internal template for repeat synthesis. Human telomerase activation in embryogenesis and its repression in somatic tissues govern cellular renewal capacity, with telomerase deficiency imposing haematopoietic and epithelial failures and aberrant telomerase activation enabling tumorigenesis.

Despite its medical relevance, efforts to design drugs that target telomerase have been hampered by limited structural information. Endogenous human telomerase holoenzyme is scarce. Most studies of human telomerase use cellular overexpression of limiting TER (hTR in humans) and TERT. Two hTR domains are critical for activity: a domain with the template and adjacent pseudoknot (t/PK) and the conserved regions 4 and 5 (CR4/5), consisting of a branched junction of stems P5 and P6, and the activity-critical stem-loop P6.1 (Fig. 1a). Together with TERT, these domains are sufficient to reconstitute telomerase activity with the repeat addition processivity characteristic of the catalytic core. TERT is composed of four domains: the N-terminal (TEN) domain, the high-affinity RNA-binding domain (TRBD), the reverse transcriptase domain (RT), and a C-terminal extension (CTE) analogous to a polymerase thumb (Fig. 1a). The high-resolution structure of a streamlined TERT from the flour bee telomerase revealed its overall architecture, including a monomeric TERT–TER catalytic core. However, the subunit composition of the Tetrahymena telomerase holoenzyme is distinct from its human counterpart, and available structural information on human telomerase holoenzyme is limited to a 30 Å negative-stain electron microscopy reconstruction. Therefore, both the structure and the composition of the human telomerase holoenzyme remain poorly characterized.

Telomerase assembly initiates in human cells with nascent hTR binding to H/ACA proteins and their chaperones. Human H/ACA proteins are shared by telomerase and the small nucleolar (sno) and small Cajal body (sca) RNP families, which catalyse ribosomal and spliceosomal RNA modifications, respectively. Eukaryotic H/ACA snoRNPs contain the pseudouridine synthase dyskerin (Cbf5), NOP10, NHP2, and GAR1, assembled on each of two hairpin stems; and scaRNPs also contain TCAB1/WDR79 bound to loop CAB-box motifs. The hTR/HACA 3’ hairpin loop has a CAB box but the 5’ hairpin instead extends into CR4/5 (Fig. 1a). All eukaryotic H/ACA RNP families, including telomerase, require both hairpins of the H/ACA motif for RNP biogenesis in cells, unlike their single-hairpin archaeal counterparts, but it remains unclear whether the active human telomerase holoenzyme has one or two sets of H/ACA proteins, only dyskerin, or only NOP10 and dyskerin. Syndromes of human telomerase deficiency have been traced to mutations in dyskerin, NOP10, NHP2, and their chaperone NAF1. While the structure of the related single-hairpin archaeal RNP is known, there is no high-resolution structure of any complete eukaryotic H/ACA RNP.

Here, to address the structure and composition of human telomerase holoenzyme, we optimized active telomerase purification and used cryo-EM to determine its structure at subnanometre resolution, which
Telomerase holoenzyme reconstitution and characterization. a, Domain architecture of TERT (top) and secondary structure of hTR (bottom). Paired stems (P) are shown compacted in length. b, Silver-stained SDS–PAGE gel showing fractionation of purified, heterogeneously active TERT RNP. Proteins were detected by mass spectrometry and immunoblotting (Extended Data Fig. 1a, Extended Data Table 1). c, In vitro telomerase assays performed on the three elution fractions shown in b, the beads after elution, and fivefold diluted E1. A 12-nucleotide (nt) oligonucleotide was used as recovery control (RC). Processive repeat additions are numbered or indicated by dots. d, e, Representative negative-stain 2D class averages for fractions E1 and E2, respectively. We also observed a small fraction of E2-like particles in E1 (data not shown), while E2 contained predominantly the particles shown in e. Experiments in b and c were performed three times and twice, respectively, with similar results. For gel source data, see Supplementary Fig. 1.

Activities of distinct structural states
We reconstituted human telomerase holoenzyme by transient transfection of HEK 293T cells with vectors expressing TERT and hTR. After affinity purification steps to enrich for hTR and TERT, the predominant additional proteins in the sample were TCAB1 and the H/ACA proteins (dyskerin, NOP10, NHP2 and GAR1), each of which was assigned by immunoblotting and, with the exception of glycine/arginine-rich GAR1, also confirmed by mass spectrometry (Fig. 1b, Extended Data Fig. 1a, Extended Data Table 1). Furthermore, protein identities were confirmed by subunit drop-out purifications (see below). Direct primer-extension assays of eluted fractions showed that the amount of TERT did not fully correlate with the amount of telomerase activity; also, dilution of the less-active fraction E1 increased rather than decreased activity (Fig. 1c). Visualization by negative-stain electron microscopy of E1 and E2, which we refer to as inactive and active fractions, respectively, revealed morphologically distinct populations of particles. Particles from the inactive fraction were more elongated (typically 400 Å or more in length) and more heterogeneous than those in the active fraction (Fig. 1d, e). Particles from the active fraction consisted of two asymmetric lobes, nearly 300 Å in length (Fig. 1e). Previous studies have shown that cellular human telomerase reconstitution yields complexes both monomeric and dimeric for TERT15,17, explaining the observed particle heterogeneity. For cryo-EM studies, we modified our purification protocol to recover only active fractions (Extended Data Fig. 1b, c), added a DNA substrate (T$_{14}$A$_{14}$G$_{3}$) during purification, and confirmed the presence of the bound substrate in the purified RNP (Extended Data Fig. 1c). We also demonstrated that our tagged TERT was active in vivo (Extended Data Fig. 2a–e).

Architecture and flexibility of telomerase
We prepared ΔTCAB1 telomerase with an hTR CAB box mutant that disrupted TCAB1 binding25,26 and separately prepared TERT–hTRMin27,28, in which hTRMin contains only the activity-essential t/PK and CR4/5 domains of hTR. By comparing their negative-stain 2D class averages to those of the intact holoenzyme, we established the particle regions corresponding to TCAB1, the H/ACA RNA motif and proteins, and TERT with the two activity-essential domains of hTR (Extended Data Fig. 1d–h). This revealed that one lobe of the holoenzyme contained TERT and its associated hTR motifs (the catalytic core) and the other contained the H/ACA RNP components and TCAB1 (the H/ACA lobe). We next collected cryo-EM data to determine the structure of the substrate-bound human telomerase holoenzyme (Extended Data Fig. 3a, b, Extended Data Table 2a). Initial image processing revealed extensive conformational flexibility between the two lobes, as well as within each lobe, which limited the quality of the overall reconstruction (Extended Data Figs. 3c, 4a). We performed focused classification with signal subtraction29 to obtain more homogeneous subsets of particles for each lobe separately. This procedure yielded reconstructions at 7.7 Å and 8.2 Å resolution for the catalytic core and H/ACA lobes, respectively (Extended Data Figs. 3c, 4b–e). At these resolutions, densities for protein α-helices, β-barrels, and double-stranded RNA are discernible, allowing us to unambiguously fit available protein and RNA structures or homology models, and ultimately to account for the majority of the density, except for flexible regions of single-stranded RNA and protein extensions (Fig. 2a, b, Extended Data Fig. 4f). We also obtained a reconstruction of the entire holoenzyme at 10.2 Å resolution (Extended Data Fig. 3c).
Subunit distribution across the two lobes

The bilobal human telomerase includes one TERT molecule and one hTR (Fig. 2a, b), in agreement with suggestions of an evolutionarily conserved single-TERT catalytic core for telomerases from ciliate and yeast model systems to human. The complex also contains one TCAB1 and two sets of the four mature H/ACA RNP proteins. The previous human telomerase negative-stain electron microscopy reconstruction also had a bilobal architecture, interpreted as dimeric TERT, although gold-labeling experiments indicated the presence of a mixture of RNPs containing both one (50%) and two (36.5%) TERT molecules. Incomplete separation of active from inactive TERT assemblies may account for the mixture of complexes in the previous study. Our finding of separate lobes for the catalytic core and the H/ACA RNP gives a structural explanation for why telomerase holoenzyme RNP biogenesis strategies have such plasticity across eukaryotes.

The extended architecture of hTR scaffolds the overall subunit organization across both lobes (Fig. 3a, b). The catalytic core and H/ACA lobes are bridged by stems P1a and P4.2 (Fig. 2a, b, 3b). The crystal structure of *Tribolium* TERT in complex with an RNA–DNA duplex could be readily fitted into the catalytic core density (Fig. 4a, Extended Data Figs. 5c, d, 6b). As the *Oryzias latipes* (medaka) TRBD structure includes vertebrate TERT motifs not present in the streamlined *Tribolium* TERT (Extended Data Fig. 7e), we removed extensions from the medaka TRBD structure that did not fit our density and used this model to replace the *Tribolium* TRBD (Extended Data Fig. 5b). Adjacent to the TERT ring, there is globular density with approximately the dimensions of the *Tetrahymena* TEN domain crystal structure (Fig. 4b, Extended Data Fig. 5a). The *Tetrahymena* and human TEN domains share only 20% sequence identity and have differences in predicted secondary structure; therefore, we removed regions of the former that are likely to be less structurally congruent with the human structure for tentative fitting into our map (Extended Data Fig. 6a). The NMR structure for the full human PK and the crystal structure of the medaka CR4/5 in the conformation bound to TRBD were fitted in one side of the catalytic core (Figs. 2b, 3a, b, 4a, Extended Data Fig. 5h). The PK is connected to the template region located in the TERT active site by a bent helical RNA density that we attributed to the sequential junction and stem elements of P2 (P2b–J2a/b–P2a–J2a/a.1–P2a.1)10,37,38 (Figs. 3a, 4a).
We docked two copies of the crystal structure of an archaeal H/ACA RNP\textsuperscript{28}, each consisting of Cbf5 (orthologous to dyskerin), GAR1, NOP10 and L7Ac (structurally related to NHP2) and a single-hairpin RNA, into the H/ACA lobe (Figs. 2a–c, 5a, b, Extended Data Figs. 5e, f, 7f, 8a–d). The single-hairpin RNA was truncated to a 9-bp stem flanked by the ACA motif to mimic P4 and P7 stems flanked by box H and ACA, respectively, in hTR. The hTR segments bound by each archaeal RNP equivalent can be unambiguously assigned on the basis of their connectivity with the rest of the RNA densities. The P4–P4.1 stem is connected to the CR4/5 domain in the catalytic core via the P4.2 stem, and P7 is connected to the P8 stem-loop (Fig. 3a, b, 5a, b, Extended Data Fig. 5i). Density in the regions of P1, P4.1 and P4.2 was fitted with A-form RNA duplexes. TCAB1 harbours a WD40 domain, and its homology model fit uniquely at the top of the H/ACA lobe (Figs. 2b, 5a, b, Extended Data Figs. 5g, 7g), consistent with the missing density from the hTR CAB-box mutation (Extended Data Fig. 1h). This assignment of TCAB1 places it adjacent to an RNA-hairpin-like density that was fitted with the NMR structure of the P8 stem-loop\textsuperscript{19}.\

The catalytic core\
The t/PK and CR4/5 domains of hTR wrap around TERT to form the catalytic core (Fig. 4a). In turn, the human TERT ring encircles the fitted Tribolium TERT RNA–DNA bound to the TERT active site (Fig. 4a, c), which shifts the template position compared to that in the substrate-free Tetrahymena telomerase holoenzyme (Extended Data Fig. 7i). The majority of the TERT RNA binding surface is contributed by the TRBD and CTE, which are sandwiched between the PK and CR4/5 (Fig. 4d). The full human PK (P2b–P3) is connected to P2a by the J2a/b bulge, and P2a connects to the less conserved helical extension P2a.1 before the transition to the single-stranded 3′-template-flanking region (Figs. 1a, 3a, 4a). The t/PK forms a rigid, arc-like structure (Figs. 1a, 4a), as predicted by modelling using NMR structures of its subdomains\textsuperscript{30,36,38}. The least conserved end of the PK P3 stem approaches the TRBD close to its interface with the CTE and co-axially stacks with P2b (Fig. 4d). The unexpectedly limited interaction between the hTR PK and TERT favours a model for indirect PK function in the co-folding of TERT and hTR\textsuperscript{30,46}. The location of P2a.1 positions the immediately 5′ single-stranded 3′-template-flanking RNA such that it could thread between the TEN domain and the TERT ring (Fig. 4b, Extended Data Fig. 7c), similar to the path of the 3′-template-flanking TERT in the Tetrahymena telomerase holoenzyme\textsuperscript{3,10,19} (Fig. 4e), despite major changes in the remainder of the t/PK secondary and tertiary structure (Figs. 1a, 4f, Extended Data Fig. 7j). The proximity of the TEN domain to the template 3′ end further reinforces its proposed role in stabilization of RNA–DNA duplex at the template 3′ end\textsuperscript{14,22}. The position of the Tetrahymena 5′ template-boundary element, stem-loop 2 (SL2), relative to the TRBD, is roughly similar to that of human P1b (Fig. 4d, g), consistent with their common function in 5′ template boundary determination\textsuperscript{41}.

The CR4/5 domain, which consists of P5, P6 and P6.1 stems connected as a three-way junction, adopts a similar conformation to that in the medaka CR4/5–TRBD crystal structure\textsuperscript{22} (Fig. 4d, Extended Data Fig. 7d), except that human P6 is longer and composed of two stems, P6a and P6b. The P6a stem binds along TRBD (Fig. 4d), while the P6.1 hairpin emanates from the three-way junction almost perpendicularly to P6a, inserting into the interface between the TRBD and CTE. Clear helical density is observed for P6b, which makes no RNA–RNA or RNA–protein interactions. P5 co-axially stacks with P6a, but like P6b, it does not interact with TERT (Fig. 4d, Extended Data Fig. 7d). The P6.1-TERT interaction is critical for TERT association with hTR and telomerase catalytic activity\textsuperscript{42,43}. The P6.1 apical loop was disordered in the TRBD–CR4/5 crystal structure\textsuperscript{33} and, although our resolution is insufficient for de novo building, its density is visible in our cryo-EM map, appearing packed against the CTE (Extended Data Fig. 7d). Our structure suggests that interaction with the CTE stabilizes this loop, consistent with previous crosslinking studies\textsuperscript{44}. In the Tetrahymena telomerase cryo-EM structure, the stem-terminus element (SL4) is bound at the interface between the TRBD and CTE of TERT\textsuperscript{19}, similar to P6.1 in our human telomerase structure (Fig. 4d, 4g). Whether or not they are related evolutionarily, P6.1 and SL4 serve a similar structural role in telomerase catalytic core architecture. Finally, the catalytic core has extra density that is likely to correspond to TERT and hTR regions that were not possible to model (Extended Data Fig. 4f), such as the ‘insertion in fingers’ (IFD) subdomain of the RT domain (residues 734–802) that is missing in Tribolium TERT (Extended Data Figs. 6b, 7a, b), and the non-conserved flexible linker between the TEN domain and the TERT ring (Fig. 4c).
The H/ACA RNP and human disease mutations
Our structure offers the first glimpse of the subunit architecture of an intact eukaryotic H/ACA RNP, with two sets of H/ACA heterotetramer proteins that bind two RNA hairpins (Fig. 5a, b). As expected, the 3′ hairpin loop associates with the single holoenzyme TCAB1 subunit (Fig. 5a–c). We observe an interaction between the two H/ACA protein heterotetramers, predominantly mediated by the two dyskerin molecules (Fig. 5d, Extended Data Fig. 7h). The first set of H/ACA proteins contacts the 5′ hairpin P4 stem exclusively via dyskerin (Fig. 5a, b, Extended Data Fig. 5e). By contrast, the second set forms more extensive interactions with the 3′ hairpin P7 stem and P8 stem-loop. The P7 stem is bound by dyskerin, similarly to the 5′ hairpin P4, and continues to P8, which interacts with NOP10, NHP2 and TCAB1 (Fig. 5a, b, Extended Data Fig. 5f, i). The fact that binding of dyskerin at the base of the 5′ hairpin is sufficient to anchor the entire core heterotetramer accounts for the evolutionary divergence of the vertebrate TER 5′/H/ACA hairpin and for the hTR 5′ hairpin tolerance of changes in stem and pocket structure. Loss of hTR 5′ hairpin contacts to H/ACA proteins is compensated by an hTR-specific increase in protein binding to the 3′hairpin. The position of dyskerin defined by our structure explains why H/ACA RNP subunit exchange occurs in cell extracts for GAR1, NOP10 and NHP2, but not dyskerin. The surprisingly limited protein–RNA interaction of the hTR 5′ hairpin implies that a wider breadth of RNA structures can assemble H/ACA proteins than has been predicted from the snoRNAs and scaRNAs, which conserve not only H/ACA protein binding but also requirements for dyskerin catalytic function.

The single subunit of TCAB1 in human telomerase contacts both the 3′ hairpin CAB box and dyskerin in the second set of H/ACA proteins (Fig. 5c, Extended Data Fig. 7g), giving a unifying explanation for its simultaneous identification as an RNA-independent dyskerin-binding protein and a sequence-specific, direct RNA binding protein. Unexpectedly, TCAB1 also contacts GAR1 (Fig. 5c, Extended Data Fig. 7g). Protein–protein interactions by TCAB1 explain why such a limited amount of CAB-box sequence can stabilize its association with H/ACA RNPs. Structure fitting of the archaeal H/ACA RNP, combined with sequence homology between the archaeal and human proteins, allowed mapping of human dyskeratosis congenita and Hoyeraal–Hreidarsson disease mutations in our structure (Extended Data Fig. 8a–d). Notably, a large number of dyskerin mutations cluster at the dyskerin–dyskerin interface, which is also close to its RNA binding surface (Fig. 5d, Extended Data Fig. 7f, h). In addition, dyskerin R158 is near TCAB1, and NOP10 R34 and NHP2 V126 are near the P8 stem (Fig. 5c). Mutations at any of these newly evident interfaces would be expected to have more severe impact on hTR than snoRNAs or scaRNAs, which conserve not only H/ACA protein–RNA interaction of the hTR 5′ hairpin makes 5′ hairpin assembly more dependent on cross-hairpin dyskerin–dyskerin interactions; (b) hTR is uniquely reliant on enhanced H/ACA–protein affinity for the 3′–hairpin P8 stem-loop; and (c) hTR has one rather than two CAB box locations for TCAB1 association. Our ability to map disease mutations to the H/ACA RNP structure expands the scope of telomerase deficiency mechanisms beyond decreased assembly of the core heterotrimer (dyskerin, NOP10, NHP2) to include decreased dyskerin interaction with itself and with TCAB1 as well as with hTR.

Conclusions
The structure presented here provides, to our knowledge, the first architectural characterization of a telomerase holoenzyme from a multicellular species, addressing the longstanding question of vertebrate telomerase holoenzyme subunit content and arrangement. Our structure opens new opportunities for the design and screening of inhibitors targeting hTR folding and TERT interaction that will hold promise as general anti-cancer therapeutic agents.

Online content
Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0062-x.

Received: 10 January 2018; Accepted: 28 March 2018; Published online 25 April 2018.
35. Petrova, O. A. et al. Structure and function of the N-terminal domain of the yeast telomerase reverse transcriptase. *Nucleic Acids Res.* **46**, 1525–1540 (2018).

36. Kim, N.-K. et al. Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA. *J. Mol. Biol.* **384**, 1249–1261 (2008).

37. Zhang, Q., Kim, N.-K. & Feigon, J. Architecture of human telomerase RNA. *Proc. Natl Acad. Sci. USA* **108**, 20325–20332 (2011).

38. Wang, Y., Yesselman, J. D., Zhang, Q., Kang, M. & Feigon, J. Structural conservation in the template/pseudoknot domain of vertebrate telomerase RNA from teleost fish to human. *Proc. Natl Acad. Sci. USA* **113**, E5125–E5134 (2016).

39. Theimer, C. A. et al. Structural and functional characterization of human telomerase RNA processing and Cajal body localization signals. *Mol. Cell* **27**, 869–881 (2007).

40. Cash, D. D. & Feigon, J. Structure and folding of the *Tetrahymena* telomerase RNA pseudoknot. *Nucleic Acids Res.* **45**, 482–495 (2017).

41. Jansson, L. I. et al. Structural basis of template-boundary definition in *Tetrahymena* telomerase. *Nat. Struct. Mol. Biol.* **22**, 883–888 (2015).

42. Mitchell, J. R. & Collins, K. Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. *Mol. Cell* **6**, 361–371 (2000).

43. Chen, J.-L., Opperman, K. K. & Greider, C. W. A critical stem-loop structure in the CR4-CRS domain of mammalian telomerase RNA. *Nucleic Acids Res.* **30**, 592–597 (2002).

44. Bley, C. J. et al. RNA–protein binding interface in the telomerase ribonucleoprotein. *Proc. Natl Acad. Sci. USA* **108**, 20333–20338 (2011).

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

46. Egan, E. D. & Collins, K. An enhanced H/ACA RNP assembly mechanism for human telomerase RNA. *Mol. Cell. Biol.* **32**, 2428–2439 (2012).

47. Kittur, N., Darzacq, X., Roy, S., Singer, R. H. & Meier, U. T. Dynamic association and localization of human H/ACA RNP proteins. *RNA* **12**, 2057–2062 (2006).

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

49. Tycowski, K. T., Shu, M. D., Kukoyi, A. & Steitz, J. A. Conserved WD40 protein binds the Cajal body localization signal of scaRNP particles. *Mol. Cell* **34**, 47–57 (2009).

50. Sarek, G., Marzec, P., Margalef, P. & Boulton, S. J. Molecular basis of telomere dysfunction in human genetic diseases. *Nat. Struct. Mol. Biol.* **22**, 867–874 (2015).

Acknowledgements We thank P. Grob, J. Fang, A. Chintangali, P. Tobias and H. Upton for technical support; X. Zhang and J. Vogan for establishing the suspension cell transfection protocol used in early stages of this work and sharing cell lines for in vivo studies and CAB-box mutant hTR; L. Kohlstaedt and QB3 mass spectrometry facility for analysis of purified telomerase; Y. Wang, J. Jiang and J. Feigon for sharing the coordinates of the *Tetrahymena* catalytic core and the modelled human t/PK; E. Rodina for sharing the *Hansenula polymorpha* TEN coordinates before publication; and H. Upton and A. Deshpande for comments on the manuscript. We thank the National Energy Research Scientific Computing Center supported by the Office of Science of the US Department of Energy for providing computing resources under contract number DE-AC02-05CH11231. This work was funded by N.I.H. grant GM054198 to K.C. T.H.D.N. is a Fellow of the University of California, Berkeley Miller Institute for Basic Research in Science. B.J.G. was supported by fellowships from the Swiss National Science Foundation (projects P300PA_166983, P300PA_174395). E.N. is a Howard Hughes Medical Investigator.

Reviewer information Nature thanks S. Scheres and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions K.C. and E.N. directed the study. T.H.D.N. developed the telomerase purification procedure and performed in vitro biochemical assays, negative-stain and cryo-EM specimen preparation, data collection, data processing and model fitting. J.T. performed in vivo experiments and assisted T.H.D.N with cell culture and biochemistry. R.A.W guided T.H.D.N with telomerase biochemistry and purification at the initial stage of the project. B.J.G. and D.T. helped T.H.D.N with data collection. T.H.D.N, K.C. and E.N. wrote the paper with contributions from all other authors.

Competing interests The authors declare no competing interests.

Additional information Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0062-x.

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0062-x.

Reprints and permissions information is available at http://www.nature.com/reprints.

Correspondence and requests for materials should be addressed to E.N. or K.C.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.


METHODS

Telomerase reconstitution. Human telomerase was reconstituted in cells with the TERT expression plasmid pcDNA3.1-3Z-Tev-TwinStrep-TERT expressing ZZ-SS-TERT, which has a Tobacco Etch Virus protease cleavage site between the tandem Protein A and TwinStrep tag modules, and the hTERT expression plasmid pcDNA3.1-U3-hTERT-HD. HDV is the hepatitis delta virus ribozyme, which was previously shown to increase hTERT accumulation26,46. Initially we used suspension HEK 293 cells lacking N-acetylglucosaminyl transferase (GlcNAcI) and transfected the cells using polyethyleneimine (PEI)32. These cells were cultured in FreeStyle 293 Expression medium (ThermoFisher). After 48 h of transfection, whole cell extract was prepared by CHAPS lysis33. We obtained populations of inactive and active telomerase RNP purified from these extracts. To enrich for the active telomerase RNP, we used adherent HEK 293T cells cultured in DMEM with GlutaMAX (ThermoFisher) supplemented with 10% FBS, 100 μM MgCl₂, 0.2 mM EGTA, 10% glycerol, 0.1% NP40, 1 mM DTT and 0.1 mM PMSF. Whole-cell extract was prepared by three freeze–thaw cycles and adjusted to 300 mM before being cleared by centrifugation. The clarified extract was adjusted to 150 mM NaCl. For ΔTCA1 telomerase and TERT–hTrm1n, we used a previously characterized CAB-box mutant hTIR (G414C)35 and LiliTrm35. HEK 293 cell lines are long-term laboratory stocks and had the expected morphology and growth. They have not been authenticated or tested for mycoplasma recently.

Telomerase purification. The prepared whole-cell extract (from 120 plates of adherent HEK 293T cells with suspension HEK 293 cells incubated in the adherent HEK 293T cell extract) was incubated with an immobilized α-Strep (Sigma–Aldrich) pre-bound to a 5′-biotinylated oligonucleotide at room temperature for 3 h. The resin was washed with wash buffer (20 mM HEPES NaOH pH 8.0, 150 mM NaCl, 2 mM MgCl₂, 0.2 mM EGTA, 10% glycerol, 0.1% Igepal CA-630, 1 mM DTT) and eluted with a competitor oligonucleotide35. The eluate was subsequently incubated with MagStrep XT resin (IBA LifeSciences) for 2 h at 4°C. The resin was washed with wash buffer and eluted in batch with bioron elution buffer (100 mM Tris HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 5–10 mM bion, 2 mM MgCl₂, 0.1% Igepal CA-630, 10% glycerol, 1 mM DTT). We collected three sequential batch elution fractions for which the elution buffer was incubated with the resin for 30 min each time. For mass spectrometry, the eluate was combined and precipitated with trichloroacetic acid (TCA). The pellet was washed extensively with acetone and resuspended in water. The sample was analysed by MudPIT to identify the protein components present in the sample34. For cryo-EM studies, the final eluate was concentrated and buffer-exchanged with cryo-EM buffer (20 mM HEPES NaOH pH 8.0, 150 mM NaCl, 2 mM MgCl₂, 0.05% Igepal CA-630, 1% trehalose, 1 mM DTT). To prepare DNA-substrate bound telomerase, oligonucleotide T₄ΔAG₃ (1 μM) was added in the MagStrep binding step, and excess DNA substrate was washed away with TBS. The solution was collected and subsequently incubated with the enzyme for 30 min at 30°C for 40 min and stopped with TES buffer (50 mM Tris HCl pH 7.5, 20 mM EDTA, 0.2% SDS). The products were extracted and precipitated together with a 12-nM- end-labelled oligonucleotide as a recovery control and resolved on a denaturing polyacrylamide gel. The dried gels were exposed to phosphorimager screen and imaged on a Typhoon Trio system (GE Healthcare).

HCT116 cell culture and genome editing. HCT116 cells were cultured in suspension in DMEM with GlutaMAX (ThermoFisher), supplemented with 10% FBS and 100 μg/ml Primocin (InvivoGen). Clonal cell lines with homozygous TERT disruption were generated previously35. Transgene integration at the AAVS1 safe harbour locus was conducted as previously described35 using CAGGS promoter for TERT expression and a neomycin resistance cassette for selection. Genome engineering in HCT116 cells was performed using the manufacturer’s transfection protocol (Lipofectamine 3000, ThermoFisher) with media exchange at 24 h into selection with 200 μg/ml G418 and 100 μg/ml hygromycin. Cells were passaged continuously with selection medium for more than 3 months after selection for cells with the integrated TERT transgenes. HCT116 cells were obtained from Molecular and Cell Biology Department Tissue Culture Facility (UC Berkeley); this cell line was validated by short tandem repeat (STR) profiling with GenePrint and tested for mycoplasma by the Tissue Culture Facility.

Telomeric repeat amplification assays. Cell extract was prepared using the HLB method as described above and normalized to 2 mg/ml, which was quantified using the Bio-Rad Bradford protein assay. Dilutions were made of the whole cell extract, which were used for the PCR-based telomeric repeat amplification protocol (TRAP) as previously described45. Two microtitretes with 400 ng total protein was used per 20 μl QTRAP reaction, consisting of 0.1 μg/ml TS primer (5′-ATCCGTGGAAGACAGTTT-3′), 0.1 μg/ml ACX primer (5′-CCGCGCCGTCACCCCTACCCCTACCCCTAACC-3′), and Tαq Universal Green Supermix (BioRad). Samples were incubated on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad) at 30°C for 30 min, followed by a hot start of 95°C for 2 min. The samples were then cycled 35 times between 95°C for 20 s and 61°C for 90 s. Relative telomerase activity was calculated using ΔCₙ normalized to the parental sample. Serial dilutions of 10, 30, and 100 ng were used for the hotTRAP protocol. The TS and ACX primers were used to amplify telomeric repeats labelled by incorporation of α-32P-dGTP. TSNT (5′-ATCCGTGGAAGACAGTTT-3′) was the dedicated TSNT reverse primer. Samples were resolved on a native gel. The gel was dried, exposed on a phosphorimager screen and imaged on a Typhoon scanner (GE Healthcare).

Southern blotting. Genomic DNA preparation was conducted as previously described43. Four to eight micrograms of DNA was digested for between 8 h and overnight with AluI and MboI at 37°C before electrophoresis on a 0.7% agarose gel (Seakem ME agarose). The gel was dried and denatured in 1.5 M NaCl and 0.5 M NaOH for 1 h at 50°C. The gel was then washed twice with 4× SSC + 0.1% SDS between washes and prehybridized with Church’s buffer (1% BSA, 1 mM EDTA, 0.5 M NaPO₄ pH 7.2, 0.2% SDS) for 30 min at 50°C. 32P-end-labelled probes were made of the telomeric repeat (T₄ΔG₃) as well as the O’Generuler 1 kb plus ladder (ThermoFisher). These probes were then added and hybridized overnight at 50°C. The gel was washed with 4× SSC + 0.1% SDS at 40°C before exposing on a phosphorimager screen and imaging on a Typhoon scanner (GE Healthcare).

Negative-stain electron microscopy and data processing. For negative-stain preparation, 4 μl of telomerase sample was incubated on glow-discharged 400-mesh copper grids, which had been previously coated with a homemade carbon film on nickel grids2. Grids were incubated for 90 s with 2% w/v uranyl acetate, washed briefly with distilled water, blot dry and negatively stained1. The inactive and active fractions of telomerase were collected on an FEI Tecnai 12 transmission electron microscope operated at 120 kV and equipped with a TVIPS TemCam F416. Data was collected using the Leginon package56 and data processing was done in RELION 1.45 unless otherwise indicated. Initial particle picking of ~1,500 particles was performed manually, particles were extracted with a 220 pixel box at a pixel size of 2.18 Å per pixel and subjected to reference-free 2D classification to obtain 2D initial references for auto-picking within RELION34. A total of 51,423 and 26,350 particles were picked for the inactive and active fractions, respectively. We performed two rounds of reference-free 2D classification for each dataset.

Micrographs for the enriched active fraction of intact holoenzyme, ΔTCA1 telomerase, and TERT–hTrm1n samples were collected on an FEI Tecnai F20 transmission electron microscope operated at 120 kV and equipped with a Gatan US4000 CCD camera. Data collection was done using the Leginon software package56. Initial particle picking for intact telomerase holoenzyme was done using DOG PICKER36 within the APION pipeline36. This yielded approximately 16,000 particles, which were extracted using a 400 pixel box at a pixel size of 1.5 Å per pixel and subjected to reference-free 2D classification to obtain initial 2D class averages. These 2D class averages were subsequently used as references for automatic picking within RELION36 to yield 15,293 and 23,922 particles for intact telomerase holoenzyme and ΔTCA1 mutant, respectively. Reference-free 2D classification was again performed on these particles. For the TERT–hTrm1n complex, which is a much smaller RNP than telomerase holoenzyme, DOG PICKER was initially used to pick approximately 48,000 particles, which were extracted using a 180 pixel box at a pixel size of 1.5 Å per pixel. 2D class averages resulting from reference-free 2D classification of these particles were used as references for the next round of particle picking. A total of 50,720 particles, which were subjected to another round of reference-free 2D classification. A comparison of the 2D class averages of intact telomerase holoenzyme, ΔTCA1 mutant and TERT–hTrm1n is shown in Extended Data Fig. 1e–h.

© 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
Cryo-electron microscopy. For cryo-EM analysis, 3yl of the purified telomerase was applied to C-flat CF 4/2 grids (Protosigs) that had been previously coated with a 6-nm-thick layer of homemade continuous carbon film and glow-discharged in amylamine to overcome the preferred orientation of the telomerase holoenzyme on untreated carbon film. The grids were blotted for 4–5 s at 100% humidity and 4°C and plunged into a mixture of liquid ethane and propane cooled by liquid nitrogen using an FEI Vitrobot MK IV. The grids were loaded onto an FEI Titan Krios transmission electron microscope operated at 300kV and equipped with a Gatan K2 Summit direct electron detector camera mounted behind a G2 Quantum energy filter. Images were collected automatically using the Serial-EM package in super-resolution counting mode at a calibrated magnification of 43,477 × (physical pixel size of 1.15 Å per pixel), with a total electron dose of 42 electrons per Å2 during a total exposure time of 7 s, dose-fractionated into 28 movie frames. We used a slit width of 25 eV on the energy filter and a defocus range of 1.5–3.5 μm. A total of 11,654 micrographs were collected from two separate data collection sessions.

Image processing. Most steps of data processing were done in RELION 2.04 unless otherwise stated (Extended Data Fig. 3c). The movie stacks for each micrograph were corrected for drift and dose-weighted using MOTIONCOR2 running within FOCUS during data collection. Contrast transfer function (CTF) parameters were estimated for the drift-corrected micrographs using GCTF26. Initial particle picking was performed on a subset of 350 micrographs using Gautomatch (from K. Zhang, MRC-LMB, Cambridge) and templates from a previous dataset collected on substrate-free telomerase at a different magnification (data not used and not shown in this study), resulting in approximately 32,000 particles. These particles were extracted with a 400 Å pixel box for reference-free 2D classification to obtain 2D references for automatic picking of all micrographs by Gautomatch. A total of 1,091,177 particles from 11,450 micrographs were picked, binned by 2 and extracted with a box size of 200 Å2. The data were initially processed in two batches, one for each data collection session. 3D classification into five classes was run for 30 iterations with an angular sampling of 7.5°, a regularization parameter T of 4 and a 60 Å low-pass filtered initial model obtained from the previous dataset of substrate-free telomerase (data not used and not shown in this study). The initial model used for the substrate-free telomerase was that of the published 30 Å negative stain reconstruction of human telomerase17. For dataset 1, a subset of 121,116 particles from one class with the best features was selected and un-binned. For dataset 2, we selected two 3D classes and subjected them to one round of reference-free 2D classification to remove more particle images showing contaminants or damaged telomerase complexes. A subset of 214,157 particles from this dataset was unbinned and combined with the subset of particles selected from dataset 1. The combined subset of 335,283 particles was refined and yielded a reconstruction that was hard to interpret due to the conformational heterogeneity between the two lobes and within each lobe. The angular assignments resulting from this refinement were used to produce 3D classification results with signal substation of either lobe to resolve the conformational heterogeneity within each lobe28. The two lobes are henceforth referred to as the H/ACA lobe and the catalytic core, according to our subunit assignments using negative-stain EM (Extended Data Fig. 1e–h). The best class consisting of 103,363 particles for the catalytic lobe was refined to 7.8 Å resolution while that for the H/ACA lobe was refined to 8.8 Å resolution from 104,330 particles. These reconstructions show clear secondary structure features as expected for this resolution range. To further reduce conformational heterogeneity, we performed another round of alignment-free 3D classification, using the alignment parameters from the refinement run for each lobe, and refined the resulting best class. We obtained a 7.7 Å resolution reconstruction from 39,860 particles for the catalytic core and 8.2 Å resolution reconstruction from 36,210 particles for the H/ACA lobe. These reconstructions were used for model fitting (see below).

We also performed alignment-free 2D classification on the original (unsubtracted) particles using the orientations determined by the 3D refinement of each lobe to assess the relative orientation of the other lobe (Extended Data Fig. 4a). For both cases, the other lobe adopts a wide range of conformations relative to the lobe that is aligned, clearly showing that the quality of any reconstruction for the whole structure is limited by both the continuous movements between the two lobes relative to each other and the conformational heterogeneity within each lobe. To obtain a better reconstruction encompassing both lobes, we performed global 3D classification for the subset of 335,283 particles before signal subtraction (25 iterations, T = 8, angular sampling of 7.5°) (Extended Data Fig. 3c). As expected for a particle population exhibiting continuous motion, it resulted in four classes with nearly equal distribution and the two lobes adopting different conformations relative to each other. We refined each of these classes and obtained 10.2 Å resolution for the best class.

All refinements were performed using fully independent data half-sets (gold-standard refinement) and resolutions are reported based on the Fourier shell correlation (FSC) = 0.143 criterion26,40. (Extended Data Fig. 4b). FSC curves were calculated with a soft mask. For visualization, maps were corrected for the modulation transfer function of the detector and sharpened by applying a negative B factor as determined by the post-processing function of RELION. Local resolution (Extended Data Fig. 4d, e) was calculated within RELION 2.0. For model versus map FSC curves calculations, we converted the catalytic core and H/ACA RNP models into density using EMAN2. Owing to small regions of unmodelled density in our maps (Extended Data Fig. 4f), we prepared a soft mask (6-pixel soft edge) from the model and extracted the cryo-EM maps of each model using this mask. FSC model versus map was calculated using Xnippy31 and EX3 (Extended Data Fig. 4c).

Model fitting. Initial fitting of all X-ray structures, NMR ensembles and homology models was performed manually by visual inspection of the best maps of the individual two lobes, followed by fit optimization in Chimera22 (Extended Data Table 2b, Supplementary Data 1 and Supplementary Video 1). The fitted structure for each lobe was subsequently rigid-body fitted into the overall map resulting from refinement of the best 3D class from global 3D classification. A homology model for TCA1B1 WD40 domain was obtained using Phyre23 based on the structure of coatomer subunit alpha in the COPI coat linkage structure (PDB 2A15) and fitted into the density located by comparisons of the 2D class averages for wild-type and ATCA1B1 telomerase (Extended Data Fig. 1e–h). The Tetrahymena TEN domain (PDB 2B2A) and the TRBD, RT and CTE domains of the Tribolium TERT (PDB 3KYL) were fitted by rigid-body fitting into the map separately. We compared the fittings of the Tribolium TRBD and the medaka TRBD (PDB 4O26) (Extended Data Fig. 7e). The former has a truncated TRBD, while the latter is closely related to the human TRBD in sequence. Some of the medaka TRBD extensions close to the RNA binding surface fit our map, while those in the periphery of the domain (including residues 403–425 and residues 545–572) did not. We also fitted a model of TRBD that did not fit from the medaka TRBD and replace the Tribolium TRBD with this truncated version. Our substrate can form a 6-bp duplex with hTR template region. Thus we truncated the DNA/RNA duplex model in the duplex-bound Tribolium TERT (PDB 3KYL) to 6 bp. Overall, the relative orientations of the domains and the DNA–RNA duplex are very similar to that of the crystal structure. For the H/ACA RNP, the RNA used in the single-hairpin archaean crystal structure (PDB 2HIV) was truncated to a 9-bp hairpin flanked by the 3′ A motif. We used this model to fit both H/ACA copies because no structure of H/ACA proteins bound to the H box is available. Fitted structures of available hTR domains include the human P2b–P3 pseudoknot26 (PDB 2B2A), P2a–J2a–b/p2–b–P2b (PDB 2L3E), P2a–b, P8 stem (or CR7)30 (PDB 2QH2), and the medaka TRBD-bound CR435 (PDB 4O26). The P4 and P7 stems bound by the H/ACA proteins were adapted from the truncated archaean hairpin RNA model, as mentioned above. We used COOT22 to place idealized A-form RNA helices in the map regions assigned to the P1, P4.2, P4.2 and P6b stems.

Map and model visualization. Maps were visualized in Chimera22 and all model illustrations were prepared using either Pymold (https://www.pymol.org) or Chimera.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. Figure source data are provided in Supplementary Fig. 1. Cryo-EM maps of the catalytic core, H/ACA lobe and overall human telomerase holoenzyme have been deposited with the Electron Microscopy Data Bank under accession numbers EMD-7518, EMD-7519 and EMD-7521, respectively. The PyMol session used for fitting of homology models into the EM density is provided as Supplementary Information.

51. Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proc. Natl Acad. Sci. USA 99, 13419–13424 (2002).
52. Kim, N. W. et al. Specific association of human telomerase activity with immortal cells and cancer. Science 266, 2011–2015 (1994).
53. Schimer, C. E., Yates, J. R., Ili & Gerace, L. MudPIt: A powerful proteomics tool for discovery. Discov. Med. 3, 38–39 (2003).
54. Sexton, A. et al. Gene structure and molecular identification of three human TTP1 functions in telomerase action: recruitment, activation, and homeostasis set point regulation. Genes Dev. 28, 1885–1899 (2014).
55. Suloway, C. et al. Automated molecular microscopy: the new Legion system. J. Struct. Biol. 151, 41–62 (2005).
56. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530 (2012).
57. Scheres, S. H. W. Semi-automated selection of cryo-EM particles in RELION-1.3. J. Struct. Biol. 189, 114–125 (2015).
58. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S. & Carragher, B. DOG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. J. Struct. Biol. 166, 205–213 (2009).

© 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
61. Nguyen, T. H. D. et al. The architecture of the spliceosomal U4/U6/U5 tri-snRNP. Nature 523, 47–52 (2015).
62. Tivol, W. F., Briegel, A. & Jensen, G. J. An improved cryogen for plunge freezing. Microsc. Microanal. 14, 375–379 (2008).
63. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movement. J. Struct. Biol. 152, 36–51 (2005).
64. Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. eLife 5, 18722 (2016).
65. Zheng, S. Q., Palovcak, E., Armache, J.-P., Cheng, Y. & Agard, D. A. MotionCor2: anisotropic correction of beam-induced motion for improved single particle electron cryo-microscopy. Nat. Methods 14, 331–332 (2017).
66. Biyani, N. et al. Focus: The interface between data collection and data processing in cryo-EM. J. Struct. Biol. 198, 124–133 (2017).
67. Zhang, K. G. T. Real-time CTF determination and correction. J. Struct. Biol. 193, 1–12 (2016).
68. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745 (2003).
69. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. Nat. Methods 9, 853–854 (2012).
70. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 35–46 (2007).
71. Scheres, S. H. W., Núñez-Ramírez, R., Sorzano, C. O. S., Carazo, J. M. & Marabini, R. Image processing for electron microscopy single-particle analysis using XMIPP. Nat. Protocols 3, 977–990 (2008).
72. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. J. Struct. Biol. 157, 281–287 (2007).
73. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protocols 10, 845–888 (2015).
74. Dodonova, S. O. et al. Vesicular transport. A structure of the COPI coat and the role of coat proteins in membrane vesicle assembly. Science 349, 195–198 (2015).
75. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010).
76. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324 (2014).
Extended Data Fig. 1 | Protein identification by immunoblotting, enriching active telomerase, substrate pre-binding, and comparison of intact, ΔTCAB1, and TERT–hTRmin RNPs. a, Immunoblotting of TERT, TCAB1, dyskerin, GAR1, NHP2 and NOP10 in telomerase purified after CHAPS lysis protocol as shown in Fig. 1b. We used primary antibodies against each protein, except ZZ-SS-TERT, for which we used rabbit IgG. Owing to the wide range of molecular weights of the proteins in our sample, TERT, TCAB1, dyskerin and GAR1 were detected in one blot, while NHP2 and NOP10 were detected in a separate blot. The use of the same sample to probe all proteins was performed only once, but TERT, dyskerin and TCAB1 were also probed individually twice. b, Silver-stained SDS–PAGE gel of purified telomerase fractions obtained from adherent cells lysed using the hypotonic lysis method, which enriches active telomerase. This experiment was repeated more than five times with similar results. c, Direct primer-extension assays of the purified telomerase fractions shown in b, confirming that E1 is no longer inactive (left), and of the substrate-bound purified telomerase fractions with additional DNA substrate omitted from the assays (right). The activity observed confirmed that purified telomerase contains the DNA substrate. The activity assays with substrate added were repeated over five times and the activity assays with substrate pre-bound were repeated twice. All repeats showed similar results. d, Silver-stained SDS–PAGE gel of purified intact and ΔTCAB1 telomerase and TERT–hTRmin telomerase prepared for subunit assignments. This experiment was done only once to provide a direct comparison between these different purified telomerase complexes. e–g, Negative-stained 2D class averages of intact and ΔTCAB1 telomerase and TERT–hTRmin, respectively. h, Comparison of representative 2D class averages of intact and ΔTCAB1 telomerase and TERT–hTRmin showing the inferred localization of TCAB1 and TERT. For gel source data, see Supplementary Fig. 1.
Extended Data Fig. 2 | Cellular function of the tagged TERT used for structural analysis. a, Western blot detection of ZZ-SS-TERT in TERT knockout (KO) cells rescued by untagged TERT or ZZ-SS-TERT expression. Whole-cell extracts were probed using Strep antibody. HCT116 is the parental cell line. Lysate prepared from HEK 293T cells transiently transfected with ZZ-SS-TERT and hTR was used as a positive control (Ctrl). Tubulin was detected as a loading control. This experiment was performed only once to confirm the success of ZZ-SS-TERT incorporation into the HCT116 TERT KO cells. b, Telomeric restriction fragment analysis of HCT116 parental cells, TERT KO cells (before senescence), and TERT KO cells rescued with untagged or ZZ-SS-TERT transgene. Transgene-expressing cells were sampled at 31, 62 and 98 days post-transfection with transgene vectors. This experiment was performed twice with similar results. c, TRAP assay detection of telomerase activity in HCT116 parental cells, TERT KO cells, and TERT KO cells rescued with untagged TERT or ZZ-SS-TERT transgene. Whole-cell extracts were normalized by total protein concentration and assayed at 100, 30 or 10 ng of total protein per reaction. IC, internal control. This experiment was repeated four times with similar results. d, Quantification of Q-TRAP assay detection of telomerase activity in HCT parental cells, TERT KO cells and TERT KO cells rescued with untagged TERT or ZZ-SS-TERT transgene. Error bars were calculated by taking the s.d. of the average ΔCt from four time points. Data points were shown as overlays. e, Direct primer-extension assay of telomerase after template-complementary oligonucleotide purification from extracts of TERT KO cells rescued by untagged TERT or ZZ-SS-TERT transgene. Assays were performed on clarified cell lysate (crude), flow-through (O-FT) and elution (OE) using equivalent amounts of cell extract. This experiment was performed only once to re-confirm the results in c. For gel source data, see Supplementary Fig. 1.
Extended Data Fig. 3 | Image processing procedures. a, Representative raw micrograph. We collected a total of 11,654 micrographs for this study. b, Representative 2D class averages obtained from reference-free 2D classification. c, Data processing strategy used in this study.
Extended Data Fig. 4 | Resolution estimation and analysis of the flexibility of the complex. **a**, Representative 2D class averages obtained from 2D classification without alignment of particles that were aligned on either the catalytic core or the H/ACA lobe. For both cases, the other lobe adopts a wide range of conformations, as illustrated by the blurriness of the density. **b**, FSC curves for the overall map and the maps of the catalytic core and H/ACA lobe resulting from focused classification with signal subtraction and gold-standard refinement. **c**, Model versus map FSC curves for the catalytic core and the H/ACA RNP. We fitted only homology models as rigid bodies into the map and did not perform model coordinate refinement owing to the limited resolutions of the maps. Therefore, we used a lower FSC threshold of 0.25 for resolution estimates. **d**, Local resolution for the catalytic lobe (d) and the H/ACA lobe (e) estimated by RELION 2.06.4. Most of the catalytic core is resolved at 6–8 Å while most of the H/ACA lobe is resolved at 7–9 Å. **f**, Front (left) and back (right) views of the reconstruction showing modelled (grey) and unmodelled (gold) density. Most of the unmodelled density corresponds to single-stranded RNA regions or RNA bulges, and human protein extensions that cannot be built de novo at this resolution.
Extended Data Fig. 5 | Fittings of proteins and RNA into the cryo-EM map. 

a–d, Domains of TERT. a, The TEN domain from *Tetrahymena* (PDB 2B2A). b, The truncated medaka TRBD domain (PDB 4O26). c, d, The RT and CTE domains from *Tribolium* (PDB 3KYL). e, Front (top) and back (bottom) views of the 5′ hairpin set of H/ACA proteins (dyskerin, red; GAR1, cyan; NOP10, wheat; NHP2, pink) bound to P4 stem (dark blue) fit by the archaeal H/ACA RNP (PDB 2HVY). f, Front (top) and back (bottom) views of the 3′ hairpin set of H/ACA proteins using the same model and colour scheme as e. g, Homology model of TCAB1 WD40 domain. h, Front (top) and bottom (bottom) views of hTR in the catalytic core. i, hTR in the H/ACA lobe.
Extended Data Fig. 6 | Sequence alignment of TERT with secondary structure assignments based on known structures. a, Sequence alignment of *Tetrahymena* and human TEN domains. The secondary structure assignments of the *Tetrahymena* TEN domain (PDB 2B2A) are shown above the aligned sequences. Regions removed before fitting are indicated with dashed lines below the sequences. b, Sequence alignment of the *Tribolium*, human, and *Tetrahymena* TERT, with the latter two N-terminally truncated to match *Tribolium*. Secondary structure assignments of the *Tribolium* TERT are shown on top, with conserved motifs labelled in blue. Throughout the figure, the η symbol refers to a 310-helix. Strict β-turns and strict α-turns are displayed as TT and TTT. The three catalytic aspartic acids are indicated with black arrowheads. ESPript was used to generate this figure. © 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
Extended Data Fig. 7 | See next page for caption.
Extended Data Fig. 7 | Selected protein–protein and protein–RNA interactions in telomerase holoenzyme and comparisons between human and *Tetrahymena* TERT. a, Interactions between the RT and CTE domains of TERT and the substrate–template duplex. The RT domain is divided into two subdomains, the palm (green) and fingers (orange), that are commonly observed in retroviral reverse transcriptases. The CTE (cyan) is the putative thumb. The IFD insertion that is missing in the *Tribolium* TERT is indicated. b, Region of the cryo-EM reconstruction shown in a. Unassigned density close to the IFD insertion is highlighted in magenta. c, Cryo-EM density of the TEN domain in the same view as that in Fig. 4b. Connecting density is observed between the template region and the P2a.1 stem. d, Map of the CR4/5 three-way junction (wheat) and the nearby TERT domains highlighting the position of the P6.1 loop near the interface of the CTE (cyan) and TRBD (blue) domains of TERT. This loop was not ordered in medaka CR4/5 bound to the TRBD alone33. e, Comparison of the *Tribolium* (left) and medaka (right) TRBD with the medaka CR4/5 domain of hTR13,35. Extensions of the medaka TRBD that did not fit the map were truncated for visualization. f, Cryo-EM map with H/ACA components fitted. g, h, Detailed views of regions boxed in f show TCAB1 interactions with dyskerin, GAR1 and the P8 stem-loop (g), and interactions between the two dyskerin molecules (h), where a cluster of DC mutations are found (Fig. 5d). i, Comparison of the human and *Tetrahymena* TERT superposed on the RT domain. Domains of human TERT are coloured as in Fig. 1a, while *Tetrahymena* TERT is coloured grey. The bound human and *Tetrahymena* templates are coloured dark and light red, respectively. j, Comparison of human and *Tetrahymena*19 catalytic cores fitted into the corresponding cryo-EM maps. Domains of TERT were coloured as in Fig. 1a and TER is coloured yellow. We used the catalytic core and H/ACA lobe densities resulting from our focused classification/refinement for the human telomerase and the overall 9.4 Å *Tetrahymena* telomerase map (EMD-6442).
Extended Data Fig. 8 | Sequence alignments of H/ACA proteins with secondary structure assignments based on known structures.

a–d, Sequence alignments of Pyrococcus furiosus (archaeal) and human Cbf5/dyskerin (a), GAR1 (b), NOP10 (c), and L7Ae/NHP2 (d). Secondary structure assignments displayed on the top are from the archaeal H/ACA RNP structure (PDB 2HVY). The η symbol refers to a 3_10-helix. Strict β-turns and strict α-turns are displayed as TT and TTT, respectively. Known human dyskeratosis congenita and Hoyeraal–Hreidarsson disease mutations50 in H/ACA proteins are indicated with arrowheads. Blue arrowheads indicate residues that can be mapped onto the archaeal structure and black arrowheads indicate residues that were not mapped. ESPript was used to generate this figure76.

© 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
Extended Data Table 1 | Mass spectrometry analysis of the purified telomerase sample

| Protein name                              | Sequence count | Sequence coverage (%) | Length (aa) | MW (KDa) |
|-------------------------------------------|----------------|-----------------------|-------------|----------|
| NHP2 (H/ACA subunit 2)                    | 2              | 19                    | 153         | 17.2     |
| NOP10 (H/ACA subunit 3)                   | 1              | 17.2                  | 64          | 7.7      |
| Nuclease-sensitive element binding protein| 4              | 13.6                  | 324         | 35.9     |
| Methylosome subunit pICln                 | 2              | 13.5                  | 237         | 26.2     |
| TERT                                      | 11             | 10.6                  | 1132        | 127      |
| Insulin-like growth factor 2              | 3              | 8.6                   | 548         | 60.1     |
| TCAB1/WDR79/WRAP53beta                    | 3              | 8.6                   | 548         | 59.3     |
| Squamous cell carcinoma antigen            | 6              | 8                     | 963         | 110      |
| recognised by T-cells 3 (SART3)           |                |                       |             |          |
| Importin subunit α-1                      | 3              | 7.8                   | 529         | 57.9     |
| Dyskerin (H/ACA subunit 4)                | 5              | 7.2                   | 514         | 57.7     |
| hnRNP A2/B1                               | 2              | 7.7                   | 313         | 33.9     |
| hnRNP A3                                  | 1              | 5.8                   | 378         | 39.6     |
| Protein arginine N-methyltransferase 5    | 1              | 5.6                   | 531         | 61.1     |
| ATP-dependent RNA helicase DHX36          | 2              | 2.7                   | 1008        | 114.8    |

Proteins are ranked according to their percentage of sequence coverage eliminating keratins, histones, and abundant cytoskeletal proteins. Only proteins with sequence coverages of above 5% or known telomerase-associated proteins (red background) with any level of detection are shown.
Extended Data Table 2 | Summary of cryo-EM data collection and modelling of protein and RNA subunits

### a. Cryo-EM data collection

| Data collection and processing | Catalytic Core (EMDB-7518) | H/ACA lobe (EMDB-7519) | Overall (EMDB-7521) |
|-------------------------------|---------------------------|------------------------|----------------------|
| Magnification                 | 300                       | 300                    | 300                  |
| Voltage (kV)                  |                           |                        |                      |
| Electron exposure (e-/Å²)     | 42                        | 42                     | 42                   |
| Defocus range (µm)            | 1.5-3.5                   | 1.5-3.5                | 1.5-3.5              |
| Pixel size (Å)                | 1.15                      | 1.15                   | 1.15                 |
| Symmetry imposed              | C1                        | C1                     | C1                   |
| Initial particle images (no.) | 1,091,177                 | 1,091,177              | 1,091,177            |
| Final particle images (no.)   | 39,860                    | 36,210                 | 88,200               |
| Map resolution (Å)            | 7.7                       | 8.2                    | 10.2                 |
| FSC threshold                 | 0.143                     | 0.143                  | 0.143                |
| Map resolution range (Å)      | 6-9                       | 7-9                    |                      |

### b. Modelling of protein and RNA subunits

| Protein/RNA | Total residues | M.W. (KDa) | Domain                | PDB code or modelling         | Reference |
|-------------|----------------|------------|-----------------------|-------------------------------|-----------|
| TERT        | 1132           | 127        | TEN                   | 2B2A (Tetrahymena model)      | 34        |
|             |                |            | TRBD                  | 4O26 (medaka model)           | 33        |
|             |                |            | RT                    | 3KYL (Tribolium model)        | 13        |
| TCAB1       | 548            | 59         | β-propeller domain    | homology model by Phyre       |           |
| Dyskerin    | 514            | 57         |                       |                                |           |
| GAR1        | 217            | 21         |                       | 2HVY (archaeal model)         | 24        |
| NOP10       | 64             | 8          |                       |                                |           |
| NHP2        | 153            | 17         |                       |                                |           |
| hTR         | 451            | 138        |                       |                                |           |
| P1          |                |            | A-form double helix   | model from Chan et al, 2017   | 10, 38    |
| P2a.1       |                |            |                       |                                |           |
| P2a-J2a/b-P2b|               |            | 2L3E                  |                                | 37        |
| template    |                |            | 3KYL                  |                                | 13        |
| P3          |                |            | 2K95                  |                                | 36        |
| P4          |                |            | 2HVY as part of the H/ACA RNP | 24        |
| P4.1        |                |            | A-form double helix   |                                |           |
| P4.2        |                |            | A-form double helix   |                                |           |
| P5-P6.1-P6a |                |            | 4O26 (medaka model)   |                                | 33        |
| P6b         |                |            | A-form double helix   |                                |           |
| P7          |                |            | 2HVY as part of the H/ACA RNP | 24        |
| P8          |                |            | 2QH2                  |                                | 39        |
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
- Clearly defined error bars
- State explicitly what error bars represent (e.g. SD, SE, CI)

Software and code

Policy information about availability of computer code

**Data collection**

EM data collection: Serial-EM and Leginon. Image Studio v5.2 on the Li-COR-Odyssey. Typhoon Scanner Control v5.0 on the Typhoon Trio.

**Data analysis**

The EM softwares were used: Focus, Leginon, APPION, Relion v1.4, Relion v2.0, EMAN2, Motioncorr v2, GCTF v1.06, Xmipp v2.4, Gautomatch_v0.53. Structures were built and visualized using Coot v0.8.8, Chimera v1.11.2, and Pymol v1.7.4 and v2.0.6. Gel images were analyzed using ImageJ v1.51. EScript v3.0 was used for secondary structure assignment.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data
Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Cryo-EM maps of the catalytic core, H/ACA lobe and overall human telomerase holoenzyme have been deposited with the Electron Microscopy Database under accession number EMD-7518, EMD-7519 and EMD-7521, respectively. PDB coordinates are included in a Pymol session as Supplementary Information.

Field-specific reporting
Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑️ Life sciences
☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | Cryo-EM data were collected over approximately 7 days, yielding the number of particles required for extensive classification as described to achieve interpretable maps presented here. |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | Through 2D and 3D classification procedures, we discarded "bad particles" or classes of particles that did not result in good 3D reconstruction. This procedure is standard in the EM field. |
| Replication | Our biochemical purification and activity assays are all replicated successfully. The Western blots used for protein identification were not replicated because they were only used to confirm the identities of the proteins on the SDS page. |
| Randomization | Randomization was not relevant to this study. |
| Blinding | Blinding was not relevant to this study. |

Materials & experimental systems

Policy information about availability of materials

n/a Involved in the study

☑☐ Unique materials
☑☐ Antibodies
☐☐ Eukaryotic cell lines
☐☐ Research animals
☐☐ Human research participants

Antibodies

Antibodies used

Rabbit anti-dyskerin (Santa Cruz Biotechnology, sc48794, Lot E0214), mouse anti-TCAB1 (Sigma, SAB1400602, Lot 12178-1F12), rabbit anti-GAR1 (Novus, NBP2-31742, Lot R82648), rabbit anti-NHP2 (Proteintech, 15128-1-AP), rabbit anti-NOP10 (Abcam, ab133726, Lot YJ021509CS), mouse anti-Strep (Abcam, ab184224, Lot GR165043-30), mouse anti-tubulin (Abcam, ab44928, Lot GR293438-2 ) and rabbit IgG (Sigma, 15006). Secondary Antibodies (Invitrogen): Anti-rabbit (A21109), Anti-mouse (A21057, Lot 1792100), Anti-mouse (A32720, Lot SC243837)

Validation

Rabbit anti-dyskerin (Santa Cruz Biotechnology): Western blots with different cell lysates
Mouse anti-TCAB1 (Sigma): validated by Western blots of transfected and non-transfected cell lysates
Rabbit anti-GAR1 (Novus): Specificity of GAR1 antibody verified on a Protein Array containing target protein plus 383 other non-specific proteins.
Rabbit anti-NHP2 (Proteintech): Western blots with different cell lysates
Rabbit anti-NOP10 (Abcam): Western blots with different cell lysates
**Eukaryotic cell lines**

| Policy information about cell lines | HEK 293 cells were long-term Collins lab stock and HCT116 cells were obtained from UC Berkeley Tissue Culture facility. |
|------------------------------------|--------------------------------------------------------------------------------------------------|
| Cell line source(s)                | None of the cell lines used were authenticated.                                                  |
| Authentication                     | HEK 293 cells were not tested for mycoplasma contamination. HCT116 cells were tested for mycoplasma. |
| Mycoplasma contamination            | No commonly misidentified lines were used                                                         |

Commonly misidentified lines (See ICLAC register)

**Method-specific reporting**

| Method                  | Involved in the study |
|-------------------------|-----------------------|
| ChIP-seq                | ✔️                    |
| Flow cytometry          | ✔️                    |
| Magnetic resonance imaging | ✔️                  |