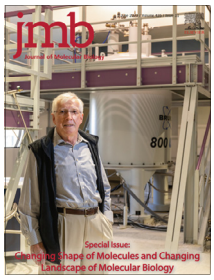

Structure of LARP7 Protein p65–telomerase RNA Complex in Telomerase Revealed by Cryo-EM and NMR

Yaqiang Wang [1†], Yao He [1,2†], Yanjiao Wang [1†], Yuan Yang [1†], Mahavir Singh [1‡], Catherine D. Eichhorn [1‡], Xinyi Cheng [1], Yi Xiao Jiang [1], Z. Hong Zhou [2,3] and Juli Feigon [1*]

*1 - **Department of Chemistry and Biochemistry,** University of California Los Angeles, Los Angeles, CA 90095-1569, USA*
*2 - **Department of Microbiology,** Immunology, and Molecular Genetics, University of California Los Angeles, Los Angeles, CA 90095, USA*
*3 - **California NanoSystems Institute,** University of California Los Angeles, Los Angeles, CA 90095, USA*

***Correspondence to Juli Feigon:** feigon@mbi.ucla.edu (J. Feigon)*
https://doi.org/10.1016/j.jmb.2023.168044
Edited by M.F. Summers

Abstract

La-related protein 7 (LARP7) are a family of RNA chaperones that protect the 3′-end of RNA and are components of specific ribonucleoprotein complexes (RNP). In *Tetrahymena thermophila* telomerase, LARP7 protein p65 together with telomerase reverse transcriptase (TERT) and telomerase RNA (TER) form the core RNP. p65 has four known domains—N-terminal domain (NTD), La motif (LaM), RNA recognition motif 1 (RRM1), and C-terminal xRRM2. To date, only the xRRM2 and LaM and their interactions with TER have been structurally characterized. Conformational dynamics leading to low resolution in cryo-EM density maps have limited our understanding of how full-length p65 specifically recognizes and remodels TER for telomerase assembly. Here, we combined focused classification of *Tetrahymena* telomerase cryo-EM maps with NMR spectroscopy to determine the structure of p65–TER. Three previously unknown helices are identified, one in the otherwise intrinsically disordered NTD that binds the La module, one that extends RRM1, and another preceding xRRM2, that stabilize p65–TER interactions. The extended La module (αN, LaM and RRM1) interacts with the four 3′ terminal U nucleotides, while LaM and αN additionally interact with TER pseudoknot, and LaM with stem 1 and 5′ end. Our results reveal the extensive p65–TER interactions that promote TER 3′-end protection, TER folding, and core RNP assembly and stabilization. The structure of full-length p65 with TER also sheds light on the biological roles of genuine La and LARP7 proteins as RNA chaperones and core RNP components.

Introduction

The eukaryotic genuine La and La-related protein (LARP) superfamily comprise a diverse class of RNA binding proteins with various functions in RNA metabolism.[1–3] Genuine La protein binds nascent RNA polymerase III transcripts at their 3′-polyuridine (polyU) tail via its La module, which comprises a La motif (LaM) and an RNA recognition motif (RRM1).[4] Most genuine La proteins also have a C-terminal atypical RRM2 that has RNA folding and chaperone activity.[5,6] Although all LARP families contain a similar La module, their C-terminal domain structures diverge significantly depending on the cognate RNA and function. LARPs bind to specific RNAs to function in folding, assembly, and biogenesis of ribonucleoproteins (RNP). Among LARPs, La-

related protein 7 (LARP7) members are the closest relatives to genuine La protein.[7]

All LARP7 identified to date are components of 7SK or telomerase RNPs. Larp7 (a specific protein in metazoa) is an essential component of 7SK RNP and required for the sequestration of positive transcription elongation factor b (P-TEFb), which regulates the transition of RNA polymerase II from promoter proximal paused to productive elongation.[8–10] Telomerase is essential for the maintenance of telomeric DNA at linear chromosome ends in most eukaryotes. Telomerase comprises a noncoding RNA (telomerase RNA, TER), telomerase reverse transcriptase (TERT), and species-specific associated proteins.[11,12] The size of TER varies among species, but it has two evolutionarily conserved domains, template/pseudoknot (t/PK) and stem-terminus element (STE), that independently bind TERT and are required for telomerase activity.[13] Ciliate TERs, like metazoan 7SK RNA, are synthesized by RNA polymerase III. LARP7 proteins p65 and p43 in *Tetrahymena thermophila* and *Euplotes aediculatus,* respectively, and Pof8 in *Schizosaccharomyces pombe* (fission yeast) are constitutive components of telomerase[14–20] that bind TER and are required for its biogenesis.

In *Tetrahymena* telomerase holoenzyme, p65 is required for TER accumulation *in vivo*[21] and is an essential component of the core RNP.[12] Biochemical, single-molecule, and structural studies showed that p65 initiates hierarchical assembly of telomerase through its interactions with TER stem 4.[22–25] *Tetrahymena* p65 has four known domains: N-terminal domain (NTD), LaM, RRM1, and a C-terminal xRRM2 (Figure 1(A)).[23,26] The LaM and RRM1 comprise the La module, which is highly conserved in sequence among LARP7s and genuine La.[1] Crystal structures of La modules of human Larp7 and human genuine La protein show conserved interactions with RNA UUU-3′-OH in a V-shaped configuration.[4,27] However, to date, most structural studies of p65 were focused on its xRRM2 interaction with TER stem-loop 4 (SL4, STE in *Tetrahymena* TER). xRRMs have an atypical mode of RNA recognition divergent from that of canonical RRMs.[25,28,29] In *Tetrahymena* telomerase, p65 xRRM2 specifically recognizes and bends SL4 to promote telomerase assembly.[25,30] However, since TER stem 1 (S1) and SL4 are connected by a single-stranded linker (Figure 1(B)), knowledge of the xRRM2–S4 structure[25] outside the context of the entire core RNP does not completely explain how p65 positions loop 4 of TER for interaction with TERT, where it is inserted between TERT RNA binding domain (TRBD) and C-terminal element (CTE), during telomerase RNP assembly.[31–33]

We previously reported a 3.3 Å overall resolution cryo-EM structure of *Tetrahymena* telomerase with DNA bound at the second step of telomere repeat synthesis.[34] While high resolution features of the TERT–TER catalytic core could be clearly discerned, the p65–TER region has a much lower resolution due to conformational dynamics and its location at the periphery of the core RNP. Globally, the structure shows that p65 and TER form a U-shape structure under TERT,[31–33] and p65 LaM and xRRM2 bind near the two ends of the U, respectively.[34] However, the structure of the rest of p65 (NTD, RRM1, and linker between RRM1 and xRRM2) and interactions with TER have not been determined. Here we report another cryo-EM reconstruction of *Tetrahymena* telomerase with improved local resolution of p65–TER obtained by

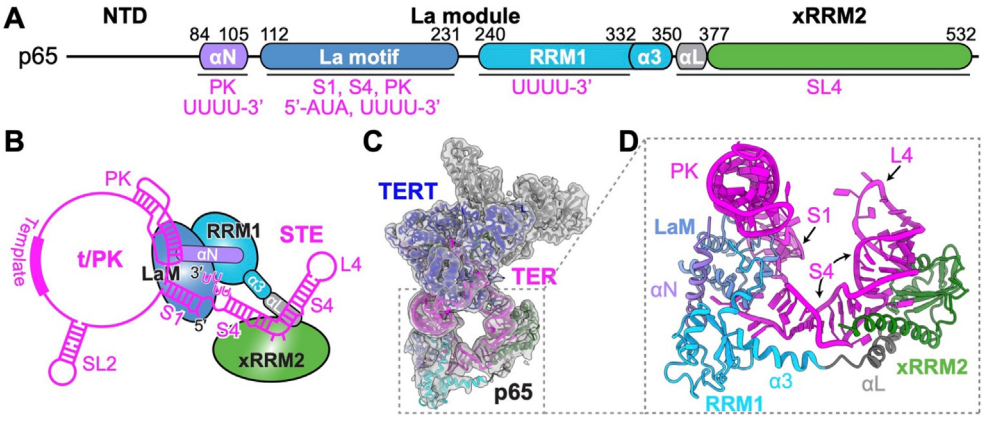

Figure 1. Structure of p65–TER complex in *Tetrahymena* telomerase. **(A)** Schematic of determined domain structure of p65, including newly identified helices αN, RRM1 α3, and αL, and corresponding TER interacting regions. **(B)** Schematic of the secondary structure of TER and binding sites for p65 NTD, LaM, RRM1, and xRRM2 (shown schematically). **(C)** Cryo-EM density map of *Tetrahymena* telomerase at 3.8 Å resolution with improved density for the p65-TER region. **(D)** Molecular model of p65–TER complex in telomerase. Newly identified helices αN, RRM1 α3, and αL are labeled. Coloring scheme for p65 is αN (violet), LaM (blue), RRM1 including α3 (cyan), αL (gray), xRRM2 (green), and TER is magenta.

focused classification on a larger dataset (Figure 1(C)). Although the resolution in parts of the cryo-EM map is not sufficient for *de novo* model building, by integrating information from NMR studies and Rosetta modeling, we were able to build a model of full-length p65 bound to TER in the telomerase holoenzyme. We find that almost all of p65, including a helix in the otherwise unstructured NTD, interacts with TER, thereby linking stem 1 and stem 4 that close the catalytically essential t/PK and STE, respectively, as well as the PK. This structure of the complete p65–TER complex also sheds light on how genuine La and other Larp7 proteins may chaperone their RNA partners.

Results

Overall structure of p65–TER in telomerase

To obtain a model of p65–TER in telomerase, we combined telomerase particles from two different cryo-EM datasets, and performed focused classification on the p65 La module (LaM and RRM1) (Figure S1(A)). An additional round of focused classification on p65 LaM was performed to further improve densities at its interfaces with TER S1, PK and 3′-polyU tail (UUUU-3′-OH). The resultant map yielded an overall 3.8 Å resolution telomerase holoenzyme with 4–5 Å local resolution for p65–TER (Figure 1(C) and Figure S1(B-D)), improved from ~6 Å.[34] The reconstruction shows well-resolved density for almost all of p65 starting from residue 83 in the NTD and of the interacting TER S1-SL4 and PK (Figure 1(D), Figure S1). S1-SL4 including the single-strand regions were *de novo* modeled assisted by NMR data on TER constructs. The crystal structure of xRRM2[25] fit the density with minor refinement. The previously reported homology model of LaM and its interface with S1, PK and the 3′-polyU tail[34] was improved by manual fitting and Rosetta refinement. p65 RRM1, whose structure had not been previously determined, was initially built using a homology model to human genuine La protein RRM1[34]. An additional helix following RRM1 was identified in the density and assigned as RRM1 α3 (Figure 1(D)). Two more protein helices were identified in the density and, based on our NMR and other studies discussed below, assigned to the NTD (αN) and the linker region between RRM1 α3 and xRRM2 (αL) (Figure 1(D)). Combining cryo-EM and NMR studies we also obtained a better model of p65-bound TER, including the 5′ and 3′ ends, S1, SL4, PK, and single-stranded regions linking PK to S1 and S1 to SL4.

p65 interacts along almost the entire S1-SL4 and additionally contacts the PK together forming a U-shape attached to the TERT ring (Figure 1(D)). The LaM and xRRM2 are at the left and right of the "U" separated by ~64 Å. In addition to the La module binding the 3'-polyU tail, LaM interacts with the ends of S1 and stem 4 (S4) and with the PK, and xRRM2 interacts with central S4 and the GA bulge (Figure 1(B,D)). Binding of p65 fixes the relative positions of S1-SL4. All three newly identified p65 helices—αN, RRM1 α3, and αL—contribute to RNA recognition and/or p65 domain stability, as discussed in detail below.

La module RRM1 has a C-terminal helix α3 required for stable folding

RRM1 is positioned on the outer edge of telomerase near the bottom of the U. In the cryo-EM density map the local resolution of RRM1 (~5 Å) is lower than that of LaM (~4 Å) (Figure S1(D)), suggesting RRM1 is positionally dynamic. Consequently, we were only able to build a homology model of this highly conserved domain refined with Rosetta. The La module RRM1 has the typical β-α-β-β-α-β fold with anti-parallel β-strands stacked on the two α-helices and conserved RNP1 sequence on β3 and RNP2 sequence on β1,[1] plus additional N-terminal α0 and C-terminal α3 helices (Figure 2(A)). In the Rosetta model, the long (~18 residues) helix α3 following RRM1 β4 appears to be positioned on the β-sheet at its N-terminal end by a stacking interaction between W336 from α3 and Y250 from RRM1 β1 (Figure 2(B)). Y250 is part of RNP2 that typically binds RNA, but here mediates the interaction with β3. Sequence alignment suggests a similar stacking between α3 and β1 residues could form in all other identified LARP7 proteins (Figure 2(C)). Consistent with this, the crystal structure of RRM1 in human genuine La protein La module shows stacking between β1 Y114 and a helix α3 Y188,[4] and the recent cryo-EM structure of human Larp7 in core 7SK RNP has a stacking interaction between RRM1 β1 Y128 and α3 W384.[35] (A previous crystal structure of human Larp7 La module was missing β4 and α3.[27]) Beyond the stacking interactions there is a small kink in the helix and the remainder of the helix protrudes out from the β-sheet of RRM1. This part of α3 parallels the backbone of proximal S4, but seems to have no direct interactions (Figure 1(D)).

Previous attempts to express or purify p65 La module or RRM1 alone failed, likely due to lack of knowledge on the domain boundaries.[25,26,30] Based on the cryo-EM model, we designed p65 La module (LaMod(108–355)) and RRM1 (RRM1(240–355)) constructs that include the newly identified α3 to test the importance of this helix for RRM1 folding and stability, and found that the protein domains could be expressed and purified. RRM1 constructs that include the entire α3 (RRM1(240–355)) or the first three turns of helix α3 (RRM1(240–342)) give ^{1}H-^{15}N heteronuclear single quantum coherence spectroscopy (HSQC) NMR spectra that indicate RRM1 is folded (Figure 2(D,E)). In contrast, ^{1}H-^{15}N HSQC spectrum of RRM1(240–339) shows line broadening or disappearance of the majority of peaks, as well as sample precipitation over time,

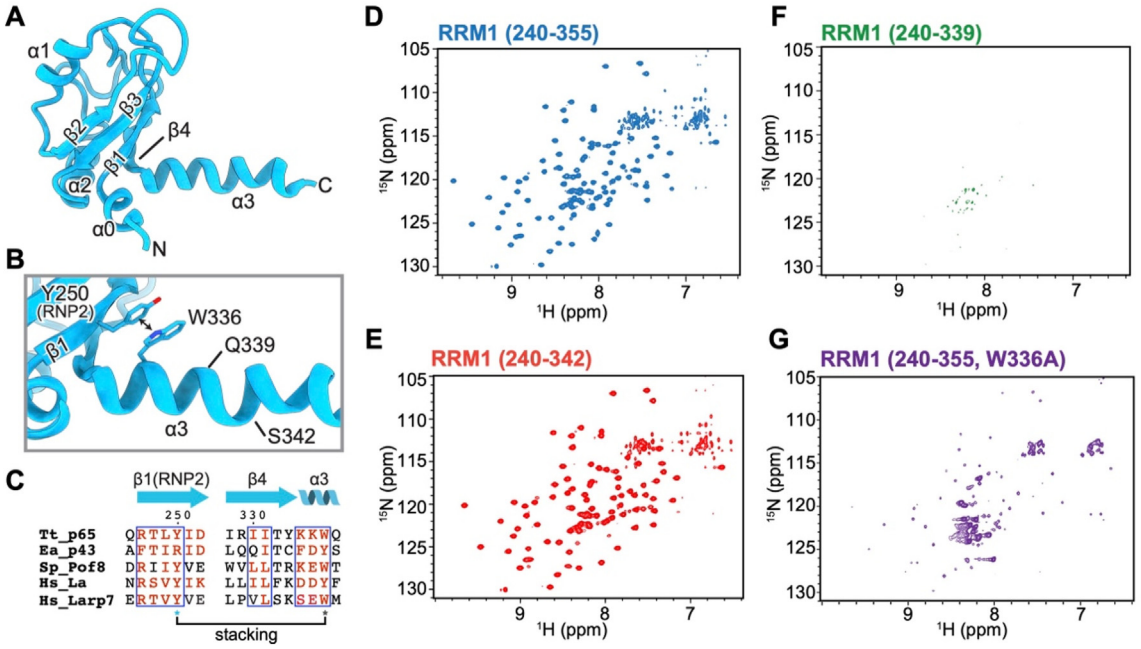

Figure 2. Structure and stability of La Module RRM1. (**A**) Structure of RRM1 including non-canonical helices α0 and α3. (**B**) Region of RRM1 showing stacking interaction between Y250 on β1 and W336 on α3. (**C**) Sequence alignment of β1 and α3 from various LARP7 proteins (Tt, *Tetrahymena*; Ea, *Euplotes aediculatus*; Sp, *Schizosaccharomyces pombe*; Hs, *Homo sapiens*). (**D-G**) ^{1}H-^{15}N HSQC spectra of RRM1 constructs: (**D**) RRM1(240–355), (**E**) RRM1(240–342), (**F**) RRM1(240–339), (**G**) RRM1(240–355, W336A).

indicating that it is not stably folded (Figure 2(F)). Although this shorter construct potentially includes 2 turns of helix α3, it is apparently not long enough to form a stable interaction between W336 and Y250. Similar results were obtained for αN-LaMod (67–355) versus αN-LaMod(67–335) (Figure S2). To confirm the importance of the stacking interaction between α3 W336 and β4 Y250 for stable RRM1 folding, we made a W336A substitution in the context of RRM1(240–355). ^{1}H-^{15}N HSQC spectrum of this construct shows decrease in intensity and chemical shift changes of many peaks, as well as sample precipitation over time, indicating significant destabilization of the domain (Figure 2 (G)). Together, the cryo-EM model, sequence conservation of these residues, observation of similar stacking interactions in other LARP7 proteins, and the NMR data indicate that α3 is part of RRM1 and is required for a stable fold. More generally, it shows that α3 is a conserved element of LARP7 La modules.

The intrinsically disordered NTD contains a helix αN that binds the La module

The La module was originally characterized for binding the 3′-UUU end of RNA Pol III transcripts and functioning to protect them from degradation.[3] It consists of LaM and RRM1 domains connected by a linker that is typically unstructured in the absence of RNA.[1,3] LaM is a winged-helix motif with an alpha–beta fold comprising a 3-stranded β-sheet and 6 α-helices (Figure 3).[1,36] p65 LaM has a non-conserved 23 residue loop between LaM α4 and α5, and α5 is unusually long (Figure S3). Deletion of the LaM α4-α5 Loop (Δ174-197), which is not visible in the cryoEM map, does not affect the overall fold of La module, as determined by ^{1}H-^{15}N HSQC spectra (Figure S2). After refining the previously reported p65 LaM structure[34] and modeling RRM1 into the cryo-EM map, there was still unassigned albeit relatively weak density between LaM and RRM1 that could be fit with an ∼18 residue helix (Figure 3(A,B)).

Since all residues in the La module had been assigned to sequence during modeling, we investigated whether this helix came from the NTD. NTD was predicted by Jpred[37] to be an intrinsically disordered protein region (IDR). We expressed and purified NTD (residues 1–109), and determined its backbone ^{1}H, ^{15}N, and ^{13}C assignments by heteronuclear NMR spectroscopy (Figure 3(C) and Figure S4). NMR ^{1}H-^{15}N HSQC (Figure 3(C)) and circular dichroism (CD) spectra (Figure S4(C)) confirm that NTD is indeed an IDR. Secondary structure propensity (SSP) predictions derived from NMR backbone chemical shifts[38] also indicate that most of the NTD is unstructured, but the C-terminus shows moderate interspaced helical propensity (Figure 3(D)). In the presence of 25 mM SDS, the helical propensity is significantly increased consistent with formation of a continuous helix for residues 87–104, which matches the size

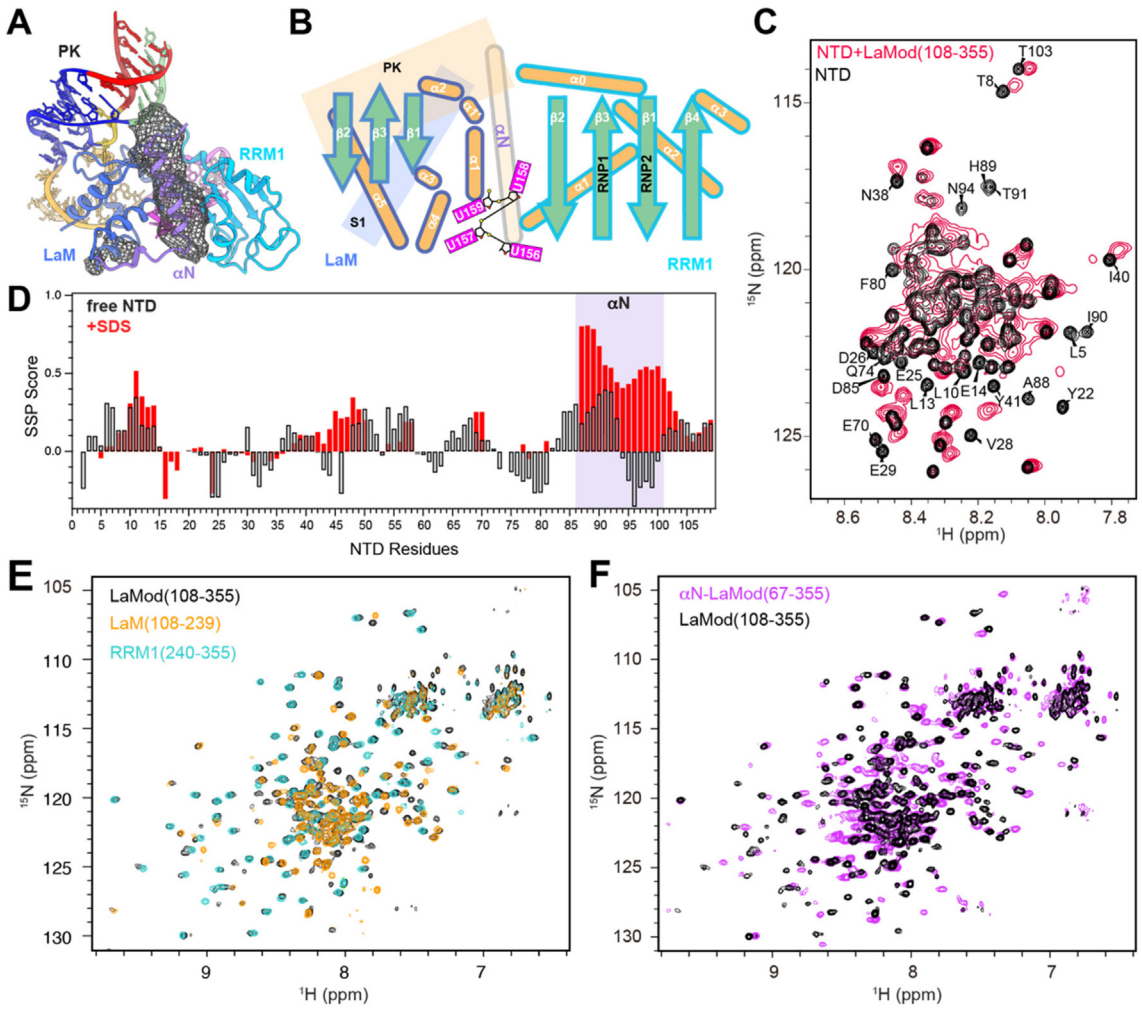

Figure 3. p65 NTD αN binds the La module. (**A**) Model of La module with TER, with cryo-EM density between LaM and RRM1 that was assigned to NTD αN shown as mesh surface. (**B**) Schematic representation of secondary structure elements of La module LaM and RRM1 and interactions with PK, S1, and UUUU-3′-OH. (**C**) ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled NTD (1–109) in the absence (black) and presence of LaMod(108–355) (red). See Supplementary Figure S3 for sequence and S4 for amide assignments. The peaks in the free NTD that are perturbed in the complex are labeled. (**D**) Secondary Structure Propensity (SSP) Score of p65 NTD in 25 mM SDS. + 1 means 100% population of α-helical conformation and −1 means 100% population of β-strand conformation. (**E**) ^{1}H-^{15}N HSQC spectra of p65 ^{15}N-labeled LaM (orange), RRM1 (cyan) and LaMod(108–355) (black). (**F**) ^{1}H-^{15}N HSQC spectra of p65 ^{15}N-labeled LaMod(108–355) (black) and αN-LaMod(67–355) (purple).

of the unassigned helical density in the cryo-EM map (Figure 3(A,D)). No other distinct density from the NTD was evident in the cryo-EM map. Based on the cryo-EM density map and the NMR data, we conclude that the NTD contains a C-terminal helix that interacts with the La module in the telomerase p65–TER structure.

We investigated whether NTD interacts with the La module in the absence of TER or telomerase in solution by NMR spectroscopy. Titration of NTD (residues 1–109) with the La Module (residues 108–355) showed that the two domains interact and the binding is in slow exchange on the NMR time scale, as assayed by ^{1}H-^{15}N HSQC (Figure 3 (C) and Figure S4(D)). The resonances that are affected the most are from αN residues. We then investigated the interactions of αN, LaM, and RRM1 by comparison of ^{1}H-^{15}N HSQC spectra of LaM, RRM1, La Module (LaMod(108–355)), and αN-La Module (αN-LaMod(67–355)). Individual spectra of LaM and RRM1 overlay well with LaMod(108–355), indicating that LaM and RRM1 fold as independent domains connected by a flexible linker, as observed for genuine La and human Larp7 (Figure 3(E)).[4,39] However, when αN is included (αN-LaMod(67–355)), there are significant chemical shift and linewidth changes in the spectrum consistent with αN interacting with LaM and RRM1 and the domains no longer tumble independently (Figure 3(F)). We conclude that αN binds

the La module in the absence of RNA as well as in the context of telomerase.

The αN-La module interacts with four uridines at the 3′-end

In the cryo-EM structure, NTD αN inserts between LaM and RRM1 (Figure 4(A)). αN has apparent interactions with LaM α1, α5-β2 loop, and linker following β3 on one helical face, and with RRM1 α1, β2 and β3-α2 loop on the opposite face. The insertion of αN results in the RRM1 and LaM being wedged apart by 50° (Figure 4(B)). Previous NMR studies of human genuine La protein indicated that upon binding with UUU-3′-OH the LaM and RRM1 domains form hydrogen-bonding interactions (Y23-N139 and R57-D125) and tumble as a single domain.[40] Similar to genuine La, human Larp7 also has individually tumbling LaM and RRM1 in the free La module, and the two domains form a hydrogen-bonding interaction (K53-E172) upon RNA binding. In contrast, p65 LaM and RRM1 bound to TER in the telomerase RNP are in a more open conformation that precludes the formation of the interdomain hydrogen bonds between the equivalent residues on LaM and RRM1 (Figure 4(A)). The four Us found at the TER 3′-end all interact with the αN-La module (Figure 4(C)). In the crystal structure of human genuine La with polyU[4] and cryo-EM structure of Larp7 with 7SK RNA,[35] three terminal nucleotides (UUU) bind in a conserved binding pocket between LaM and RRM1 in a V-shape, having alternating interactions with LaM and RRM1 (Figure 4(D)). A similar V-shaped arrangement is observed in p65, with the first and third U from the 3′-end stacking and binding the LaM (Figure 4(C)). However, instead of the penultimate U binding to RRM1, which is too far away due to αN wedging LaM and RRM1 apart, it interacts with the C-terminal end of αN (Figure 4

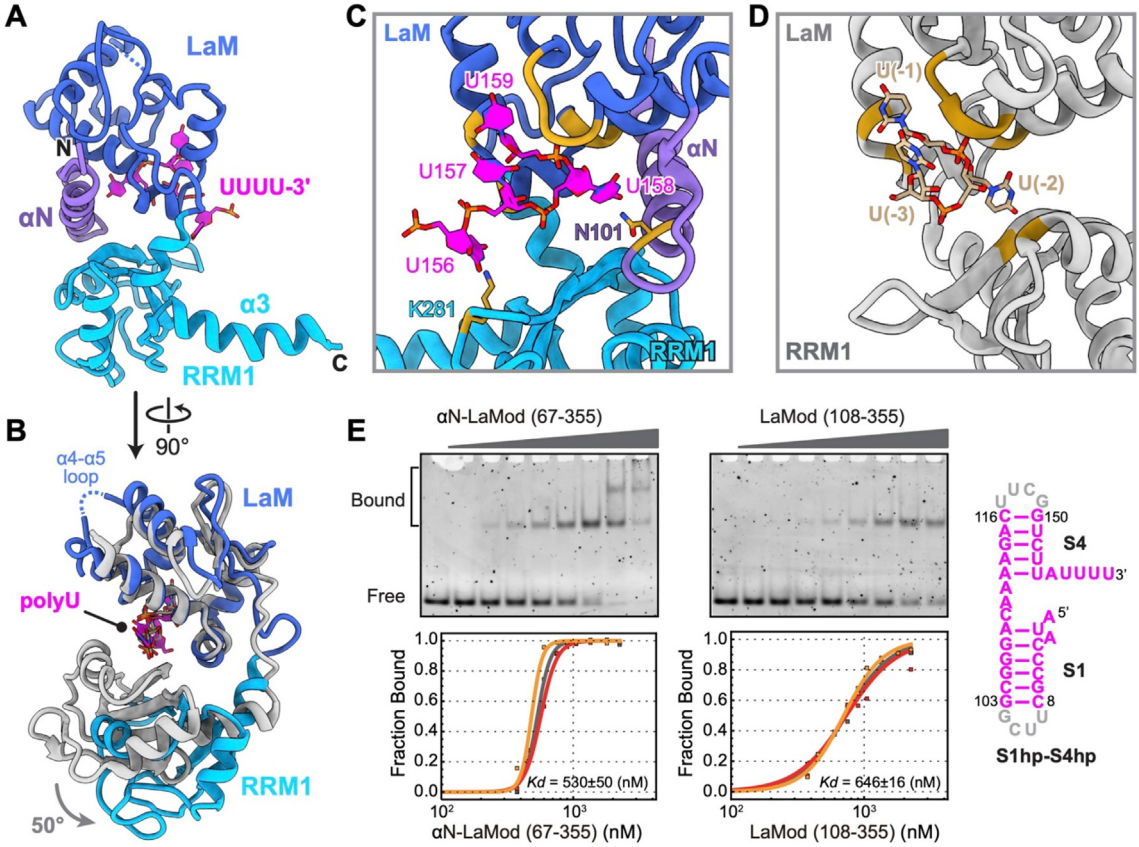

Figure 4. Interactions between p65 and the polyU tail of TER. (**A**) Model of p65 αN-La module with UUUU-3′. (**B**) Superposition of La modules of p65 (colored) and La protein (gray, PDB ID 2VOO) in complex with polyU tail highlighting the more open form of p65 due to αN interaction. p65 RRM1 α3 and αN, which are not present in the La protein structure, are omitted for clarity. (**C**) Locations of UUUU-3′ on p65 LaM, RRM1, and αN and potential interactions. (**D**) Locations of UUU-3′ on La protein La module (PDB ID 2VOO). Colors are as in Figure 1. (**E**) Representative EMSA of αN-LaMod (left) and LaMod (right) and corresponding binding curves from 3 experiments. For each sample, the concentration of S1hp-S4hp was 100 nM, and the protein concentration was increased stepwise from 375 to 2250 nM. Sequence and secondary structure of RNA (S1hp-S4hp) used in the EMSAs is shown at far right. Gray nucleotides are non-native sequence.

(C)). The fourth-to-last nucleotide (U156) interacts with RRM1, where it appears to contact K281 in the β2-β3 loop (as judged by their merged densities in the cryoEM map (Figure 4(C)). Gel shift experiments show that LaMod(108–355) binds to TER construct S1hp-S4hp, that includes the 3′-UUUU-OH, 1.2 fold more weakly than αN-La module (αN-LaMod(67–355)) (Table 1 and Figure 4(E)). Based on the cryo-EM, EMSA, and NMR data we conclude that αN-La module interacts with the last four nucleotides in the 3′-polyU tail, which bind alternately RRM1-LaM-αN-LaM.

The C-terminus of αN also contacts the PK in the minor groove near U92 (Figure 3(A)), although the density in this region is too weak to assign specific protein side-chain interactions. The αN contacts with U158 in the polyU tail and with the PK explain previous biochemistry results that deletion of NTD decreases p65 binding affinity with TER by two-fold and increases the susceptibility of S1 to RNase One.[30] In summary, the NTD αN functions to stabilize a more open form of the La module (compared to La protein and Larp7) and plays a role in both polyU and PK interactions.

LaM binding remodels the PK and fixes the positions of multiple regions of TER

The improved map and resulting model of LaM provide new details about its interaction with other regions of TER besides the 3′-UUUU-OH (Figure 5).[34] LaM inserts at a V-shaped junction between PK-S1-CAA-S4, fixing their positions relative to each other. The PK forms one side of the V while S1-CAA-S4 proximal end stack on each other to form the other side of the V (Figure 5(A–C)). Binding of LaM induces a conformational rearrangement of the PK relative to that observed for the free PK by NMR.[41] In the absence of p65, the free PK has a C75•G95$_{syn}$–A80 base triple and a U74•A94–U81 base triple formed between Stem B and Loop A (Figure 5(G)). In complex with LaM, A80 forms a non-Watson-Crick base pair with G95, while C75 flips out and inserts into a positively charged channel formed by LaM β2 and β3 (K221, K223, K228, K230, R232), which also interact with the backbone of neighboring nucleotides (U74 and A76) (Figure 5 (A,D,E,H)). In addition, K150 from β1 contacts the backbone of C75 (Figure 5(B,D)). The PK junction region is also rearranged. Although the stem-stem stacking remains the same, U73 from Loop A swaps with U92 from Stem B and forms a new base pair U73–A83 in Stem A and U92 flips out in Loop B (Figure 5(H)). This rearrangement shortens Loop A to 2 nucleotides that span the 7-base pair Stem B along the major groove, and results in a more compact PK structure. The unexpected interaction between PK and p65 LaM provides a structural basis for why and how p65 rescues the misfolding of PK caused by mutations.[41,42]

LaM also interacts with S1 to lock S1 in a fixed position relative to PK. Positively charged residues from LaM α3 (K156, K160) and α5 (K203) interact with the end of S1 on the major groove side (Figure 5(A,B)), however, individual nucleotides of S1 near the 5′-end could not be unambiguously assigned to the density. Therefore, to determine the positions of nucleotides at the end of S1, including the predicted single stranded 5′-AUA$_{1\text{-}3}$ and 3′ ACAA$_{108\text{-}111}$, we synthesized S1 hairpin constructs that contain the 5′-end AUA$_{1\text{-}3}$ with or without ACAA$_{108\text{-}111}$ linker to SL4 (S1hp and S1hp-ACAA, respectively). The NMR data indicate that in free S1hp-ACAA, U$_2$ is unexpectedly base paired with A$_{108}$ (Figure S5), extending S1 5′-end by one base pair and decreasing the single-strand S1-S4 linker from 4 to 3 nts (CAA). The U$_2$-A$_{108}$ base pair can be modeled into the cryo-EM density map, with the terminal A$_1$ base flipped out to stack on the PK and LaM α2 stacking on the terminal U$_2$-A$_{108}$ base pair (Figure 5(F)). The relative orientation of S1 and SL4 is fixed by the interaction between LaM and S1-SL4 linker CAA$_{109\text{-}111}$. Thus, LaM interacts with both the 5′ and 3′ ends of TER, where it could function to stabilize the terminal base pairs in S1 and protect both ends from nucleolytic degradation. The 5′-end AUA$_{1\text{-}3}$ and S1-SL4 ACAA$_{108\text{-}111}$ linker are highly conserved among *Tetrahymena* species so that similar ends could form,[43] while other identified ciliate TERs (*Paramecium*, *Euplotes*, *Oxytricha*, and *Stylonychia families*) lack the 5′ single-stranded residues altogether.

In contrast to the numerous interactions that fix LaM on TER, RRM1 has only limited interactions with TER, contacting a single U in the polyU tail. RRM1 α3 parallels proximal S4 but seems to have no direct or fixed interactions (Figure 1(D)). This is

Table 1 Binding affinity of p65 and TER constructs determined from EMSA experiments.

Protein	RNA	*Kd** (nM)	Relative *Kd*
αN-LaMod(67–355)	S1hp-S4hp	530 ± 50	~1.2
LaMod(108–355)		646 ± 16	
αL-xRRM2	SL4	816 ± 57	~1.6
xRRM2		1319 ± 102	
αL-xRRM2	SL4	816 ± 57	—
	SL4ΔU	848 ± 71	

*Determined Kds are average of three experiments.

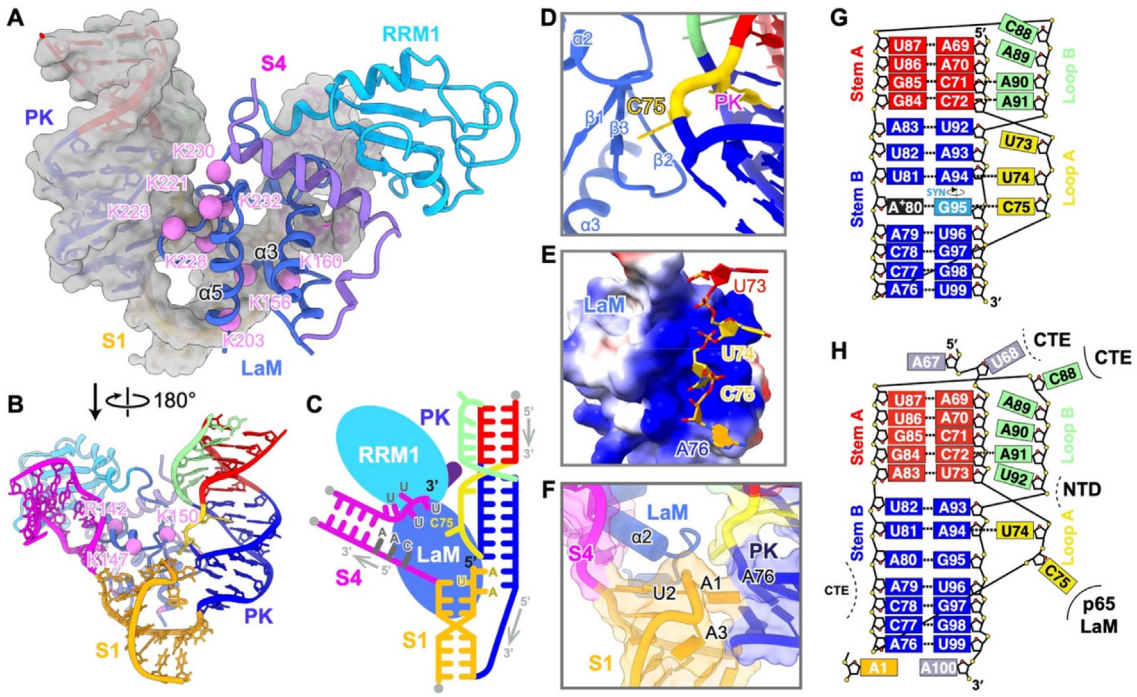

Figure 5. p65 αN-LaM binds the TER 5′-end, UUUU-3′-OH, and PK. (A) The interactions between αN-LaM (shown as ribbon) and S1, PK and S4 (shown as sticks inside surface). Positively charged residues located on LaM–PK and LaM–S1 interfaces are shown as pink spheres. **(B)** 180° rotation of (**A**) showing additional positively charged residues from LaM that contact S1, S4, and the 5′ end. **(C)** Schematic of αN-La module interactions with TER. **(D)** PK loop residue C75 inserts into a groove between LaM β2 and β3 (left). (**E**) 90° rotation of (**D**) showing PK loop residues 73–76 backbones in the basic groove on LaM. LaM is shown as grasp surface. **(F)** Region of LaM at intersection of PK, S1, and S4 shows that LaM α2 stacks on the terminal base pair of S1 and 5′-end nucleotide A1 of TER stacks underneath PK. (**G,H)** Schematics of PK secondary structures without p65 in solution determined by NMR (**F**) and in telomerase by cryo-EM (**G**). The domains of p65 and TERT contacting PK nucleotides are annotated in (**G**). Solid and dashed lines by protein domain names indicate base and backbone interactions, respectively.

consistent with the higher conformational heterogeneity of RRM1 observed during cryo-EM data processing (Figure S1(A)) and the lower resolution of RRM1 in the final cryo-EM reconstruction (Figure S1(D)) compared to LaM. The structure determined here illustrates that LaM serves multiple roles in remodeling and orienting regions of TER, including linking the 5′ and 3′ ends. La family proteins are known as RNA chaperones in addition to binding the polyU tail to protect the RNA from degradation, but the structural basis of their chaperone function has remained largely unknown. Here we see that for p65 not only the xRRM2 but also the LaM play specific roles in TER folding.

A helix αL stabilizes the interaction of xRRM2 to S4

The xRRM2 was proposed as the recognition domain of p65 specific for *Tetrahymena* TER.[44] The xRRM2 structure consists of an atypical RRM β-α-β-β-α-β'-β-α fold, where α3 lies across the surface of the β-sheet where single-stranded RNA nucleotides usually bind in canonical RRMs, and it lacks the canonical RNA-binding RNP1 and RNP2 sequences, but contains an RNP3 sequence that can recognize single-stranded RNA nucleotides.[19,44] Previous biochemical and structural studies have shown that p65 xRRM2 induces a conformational change in S4 via interactions with the conserved GA bulge and adjacent base pairs.[23,25,26,42] This conformational change bends S4 by ~60° to help position loop 4 for interaction with TERT, where it inserts between TRBD and CTE, thereby stabilizing the TERT ring. As observed in the xRRM2–S4 crystal structure,[25] xRRM2 α3 inserts from the major groove side between the base pairs on either side of the GA bulge, to bend S4, and the GA bulge residues flip out to interact in a binding pocket between β4 (RNP3) and helix α3. In the context of full-length S4 in telomerase, we find that p65 has additional interactions to S4 from xRRM2 α1 and the RRM1-xRRM2 linker, that were not present in the xRRM2 structure co-crystalized with central S4 and the GA bulge (Figure 6(A,B)). K392 from helix α1 of p65 xRRM2 contacts the minor groove of distal S4, interacting with the backbone of C142 in S4 (Figure 6(C)). The S4 construct in the crystal struc-

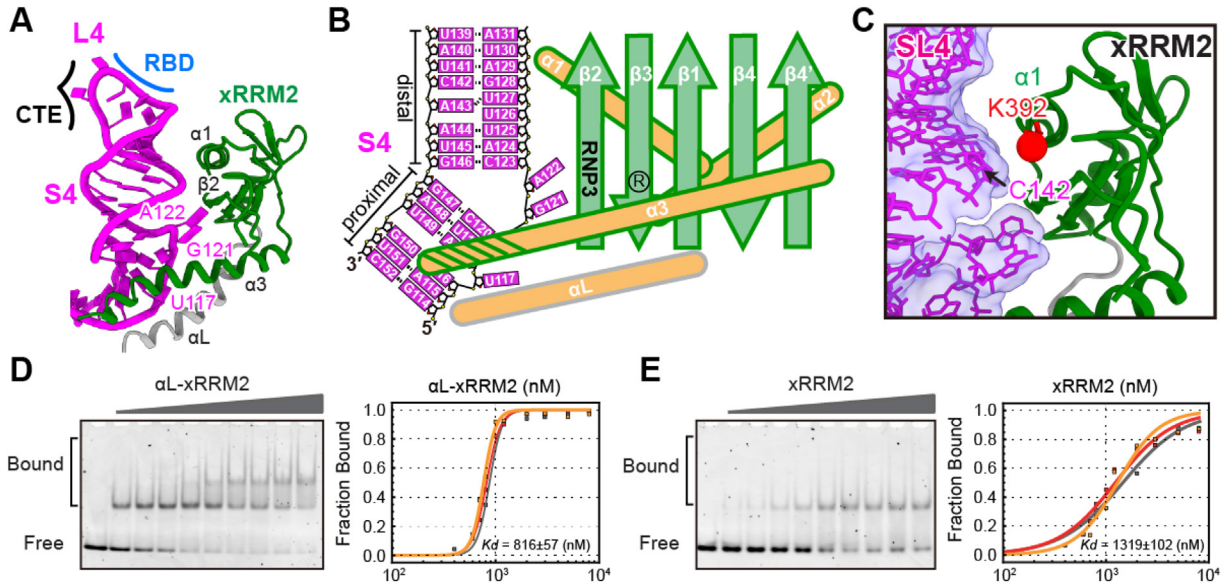

Figure 6. Interactions between p65 αL-xRRM2 and SL4. (**A**) Structure of αL-xRRM2–SL4 in telomerase. Helix αL is gray, xRRM2 is green, and SL4 is magenta. (**B**) Schematic of secondary structure and interactions of αL-xRRM2 with SL4. (**C**) Model of xRRM2 (ribbon) with distal stem 4 (stick model with transparent cryo-EM density), showing p65 residue K392, a conserved positively charged residue in LARP7 family xRRM2 α1, interaction near C142. Representative EMSA of αL-xRRM2 (**D**) and xRRM2 (**E**) with SL4 and binding curves for three experiments. For each sample, the concentration of SL4 was 150 nM and the protein concentration was increased stepwise from 600 to 8000 nM. The sequence and secondary structure of SL4 used in the EMSAs is similar to what is shown in (**B**) except the upper stem is capped by a UUCG apical loop. See Supplementary Figure S7 for the sequence and secondary structure of SL4 used in these assays.

ture had a distal stem of 4 base pairs and did not include the C_{142}-G_{128} base pair.[25] Sequence alignment of LARP7s and human La protein shows that the equivalent K392 position is always a positively charged residue (Figure S6). The equivalent residue from human Larp7 xRRM2 α1 (R468) forms a hydrogen bond with 2′OH of the RNA.[28]

A C-terminal helix α3 in RRM1 and a linker between RRM1 α3 and xRRM2 are present in all identified LARP7 proteins (Figure S6). The cryo-EM density assigned to the linker indicates that it is a helix, here named αL. Helix αL and xRRM2 α3 are positioned almost antiparallel to each other in the complex. The density for helix αL, which was not part of the xRRM2 construct used in the crystal structure[25] is relatively weak compared to helix α3, so sidechain interactions cannot be discerned. Based on the model, αL and α3 interact, most likely via charge interactions as well as a hydrophobic interaction at each end. Helix αL also appears to contact S4, most likely via the positively charged residue R362 to the RNA backbone near U117. While the non-conserved bulge U117 appears flipped out to potentially contact αL, deletion of this residue has no effect on xRRM–SL4 binding (Table 1 and Figure S7) or on activity.[45]

Although αL appears to be part of xRRM2 in the complex (Figure 6), p65 xRRM2 as well as xRRM2 from other LARP7 proteins folds as a standalone functional domain without the N-terminal helix αL, based on structures determined to date.[19,20,25,29] NMR ^{1}H-^{15}N HSQC spectra show that p65 xRRM2 domain folds the same with or without helix αL and that helix αL has a minor effect on the chemical shifts of residues from xRRM2, mainly affecting residues 376–379 preceding β1 (Figure S8). TALOS-N scores calculated from backbone chemical shift assignments showed that the αL residues do not form a helix in the absence of RNA; however, additional resonances from α3x residues (aa 527–539) were observed in αL-xRRM2 that were not present in xRRM2 alone and TALOS-N scores indicate helical conformation for residue 527 (Figure S8(C)). These α3x residues were previously indicated to be dynamic in xRRM2 and only form a helix upon binding S4.[25] Thus, these NMR observations indicate that αL residues do transiently interact with α3x residues in the absence of RNA to promote the stability of α3x, but both αL and α3x only fold stably upon binding RNA. This is consistent with binding studies using gel shifts, where xRRM2 binds to SL4 about ~1.6-fold weaker than αL-xRRM2 does (Figure 6(D,E), and Table 1). Therefore, αL interaction with α3 stabilizes the interaction of xRRM2 with S4, likely through stabilization of α3x folding and enhanced

interactions with SL4. Thus, these results suggest that the αL cooperates with xRRM2 to tightly bind and orient SL4 at a precise angle to dock loop 4 into the TRBD–CTE interface.

Discussion

Previous studies of genuine La and LARP7 proteins have shown that the La module binds and protects the 3′-polyU ends of RNA polymerase III transcripts[46] while a C-terminal xRRM2, separated by a linker of variable length, determines specificity for specific (or non-specific for genuine La) RNA substrates.[1,47] Earlier studies established that p65 xRRM binds and bends SL4 to position loop 4 for interaction with TERT at the interface between TRBD and CTE. However, it has been unclear how p65–TER interactions can collectively orient the SL4 at the correct location during assembly, because the free TER has several intrinsically flexible regions (i.e., single-stranded linkers and bulges) between and within the rigid structural units (i.e., S1, SL4, and PK), and the free p65 has domains (La motif, RRM1, xRRM2) connected by flexible linkers. Here, using a combination of cryo-EM, NMR spectroscopy and Rosetta modeling, we determined the structure of *Tetrahymena* telomerase LARP7 protein p65 bound to TER in the context of its native RNP telomerase. We find that virtually the entire 'linker' between the canonical end of RRM1 and the start of xRRM2 forms helices—RRM1 α3 and αL—contribute with the La module and xRRM2 in orienting TER for interaction with TERT, and that the intrinsically disordered NTD contains an incipient helix that orders LaM and RRM1 in the La module.

The hybrid approach used here should be generally applicable to other complex biological systems where flexibility, conformational dynamics, or structural heterogeneity limit the overall resolution of cryoEM maps. In this study, NMR was used to characterize the intrinsically disordered NTD and identify NTD helix αN; show that RRM1 helix α3, in particular stacking interactions between α3 W336 and β1 Y250, is required for domain stability; show that αN binds LaM and RRM1 in the absence of TER in solution as well as with TER in telomerase; show that αL and xRRM2 α3x only fold stably when xRRM2 binds TER; and identify RNA base pairs at the 5′ and 3′ ends of TER. These results not only made it possible to model the p65–TER complex in the cryoEM density map, but also provided insights into assembly of p65 with TER.

Telomerase activity requires two conserved domains of TER, t/PK that is closed by a helix, stem 1 in *Tetrahymena*, and STE, stem-loop 4 in *Tetrahymena*, that binds the TRBD and CTE interface to help close the TERT ring. The structure of p65–TER reveals that p65 links these two elements, binding the PK, stem 1 that includes the 5′-end, the polyU tail at the 3′-end of stem 4, and stem 4. A helix αN in the otherwise unstructured NTD cooperates with the La module to bind the 4 Us at the 3′-end of TER in a binding pocket similar to but divergent from genuine La protein. While RRM1 has only one apparent contact with TER, with the U nucleotide fourth from the 3′-end, the LaM interacts not only with the polyU tail, but, as previously observed at lower resolution,[34] also with S1, proximal S4, PK, and the 5′-end. These contacts are buttressed by the unanticipated interaction of an NTD helix αN with the La module, with direct contacts from αN to both the PK and the penultimate U. Together, the interactions of p65 along the entire length of the RNA from PK to S1 to S4 reveal how it can precisely position loop 4 for interaction with the TERT ring. These interactions also link the positions of the PK and the STE, two essential elements of TER for activity.

The p65 RRM1 helix α3 appears to be common to genuine La and LARP7 RRM1s, as indicated by the sequence alignment (Figure S3). A helix α3 is present in the crystal structure of human genuine La protein in complex with UUU[4] and in the cryo-EM structure of human Larp7 in complex with its 7SK RNA substrate. In contrast, helix αN may be unique to *Tetrahymena* p65, where it helps determine protein specificity for telomerase RNA. Helix αN wedges between the LaM and RRM1, while in genuine La protein and human Larp7 these two domains directly interact via hydrogen bond interactions when bound to polyU. Intriguingly, in yeast telomerase, Pof8, a homologue of *Tetrahymena* p65, has also recently been shown to interact with both PK and UUU-3′-OH, but the latter interaction is via Lsm2-8.[20]

To date, there are only two structures of a full-length La related protein bound to RNA, p65 reported here and hLarp7 that is a core component of 7SK RNP.[35] Both proteins utilize linker helices and extensions to augment specific interactions from their globular LaM, RRM1, and RRM2 domains, but in significantly different ways. In p65, NTD helix αN wedges the two domains of La module apart and helps to stabilize the pseudoknot of TER and linker helix αL stabilizes RRM2 α3x residues to assist in S4 interactions. In hLarp7, a C-terminal helix α4 following xRRM2 interacts with RRM1 helix α3, via 7SK RNA SL4 anchoring for both helices. These two helices form the protein–protein interface with MePCE, the other 7SK core RNP protein component. In addition, Larp7 RRM1 interacts with six out of eight polyU tail (UUU-CUUUU) nucleotides vs one in p65. The two structures illustrate a general concept that protein and RNA regions that are flexible linkers in the free subunits often become ordered in the protein–RNA complex, highlighting their importance for highly specialized complex formation and RNA specificity.

The p65–TER interactions provide new insights into assembly of TER with TERT (Figure 7). In

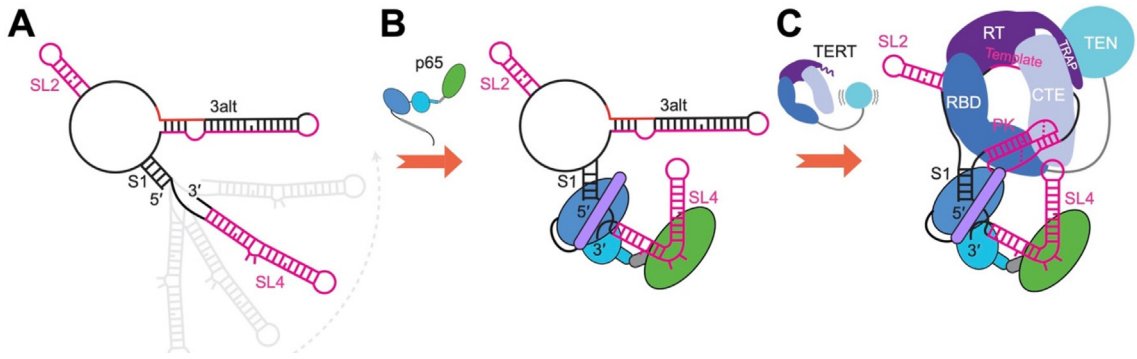

Figure 7. Steps in p65–TER–TERT assembly. (**A**) Secondary structure of free TER based on NMR data (54). The t/PK of free TER forms stem 3alt, sequestering the residues that form the PK. STE SL4 is connected stem 1, which closes the t/PK circle, through a short single-stranded linker (CAA). Gray shadows indicate that its position is not fixed relative to the t/PK. (**B**) p65 αN-La module binds S1 and UUUU-3′-OH while αL-xRRM2 binds and bends SL4. (**C**) The assembled p65–TER–TERT complex. p65 orients SL4 to insert into the RBD-CTE interface to lock the TERT ring, and αN-La module interacts with stem 1, UUUU-3′OH, and PK. αN-La module may facilitate formation of the PK and/or stabilization of its interaction with TERT.

Tetrahymena, telomerase RNP assembly begins with binding of chaperone protein p65 to TER,[22] where it may displace *Tetrahymena* genuine La protein. Previous studies have suggested that the free TER favors a lower energy Stem 3 alternative conformation (S3alt) over the PK, generating a small circle linked to SL4 through a short single-stranded linker (Figure 7(A)).[41] In the first step, the p65 La module interacts with both ends of TER to fix the S1-S4 angle and protect the TER ends from nucleolytic degradation, and p65 xRRM2 together with αL specifically recognizes the S4 GA bulge and bends SL4 (Figure 7(B)). In one possible scenario for the second step, the t/PK circle is opened through the unfolding of S3alt by TERT binding and/or an endogenous helicase, so that the TERT ring can enter the t/PK circle prior to pseudoknot formation. Nucleotides adjacent to S2 are anchored on TERT TRBD and the template interacts with the RT domain, and the TERT ring closure is stabilized by p65-oriented SL4 inserting between CTE and TRBD.[31–33] Finally, p65 interactions stabilize the pseudoknot fold and help to lock it onto the TERT ring (Figure 7(C)). In summary, using an integrative structural biology approach we solved the structure of full-length p65 in complex with TER in telomerase, which provides insights into its roles as an RNA chaperone and integral component of telomerase RNP and sheds light on similar functions for other La and LARP family proteins.

Materials and methods

Telomerase sample preparation

Tetrahymena telomerase was purified following the previously described protocol,[33] with minor modifications. Sixteen liters of *Tetrahymena* TERT-FZZ cells were grown in PPYS media. The cells were harvested by centrifugation, washed with 20 mM HEPES·NaOH pH 8.0, and then lysed in lysis buffer (20 mM HEPES·NaOH pH 8.0, 50 mM NaCl, 1 mM Ethylenediaminetetraacetic acid (EDTA), 1 mM tris(2-carboxyethyl)phosphine (TCEP), 10% glycerol, 0.2% IGEPAL CA-630, 0.1% Triton X-100) at 4 °C for 30 min. After ultracentrifugation at 230,000 × g for 1 hour, the supernatant of the cell lysate was incubated with Rabbit-IgG agarose slurry (Sigma) overnight at 4 °C. The resin was washed with wash buffer (20 mM HEPES·NaOH pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$, 1 mM TCEP, 10% glycerol, 0.1% IGEPAL CA-630). TEV protease was added into the solution to elute telomerase from IgG resin. During this step, 5 μM single-stranded telomeric DNA (sstDNA) was added to saturate telomerase. The elution fraction was incubated with anti-Flag M2 affinity gel (Sigma) for 1 hour at 4 °C. After that, the anti-Flag resin was washed extensively with wash buffer and eluted using glycerol-free buffer (20 mM HEPES·NaOH pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$, 1 mM TCEP, 0.1% IGEPAL CA-630) supplemented with 1 mg/mL 3 × FLAG peptide.

Cryo-EM specimen preparation and data collection

Lacey carbon grids with a supporting ultrathin carbon film (Ted Pella) were used for cryo-EM sample preparation. 3 μL of sample was applied to glow-discharged grid, then blotted with filter paper and flash-frozen in liquid ethane using an FEI Vitrobot Mark IV. Vitrified cryo-EM grids were loaded into an FEI Titan Krios electron microscope at 300 kV for automated image acquisition with SerialEM.[48] Movies of dose-fractionated frames were acquired with a Gatan K2 Summit direct electron detector operated in super-resolution mode, yielding a pixel size of 0.68 Å on the sample level. A Gatan Imaging Filter

(GIF) was installed between the electron microscope and the K2 camera with the slit width setting to 20 eV. The dose rate on the detector was set to ~6 electrons/Å^2/s and the total exposure time of each movie was 8 s, which fractionated into 40 frames of images with 0.2 s exposure time for each frame. In total, 12,922 movies for telomerase bound with $(GTTGGG)_2GTTGGG^LG^LT^LT$ sstDNA (dataset1, T^L/G^L represents locked nucleic acid oligonucleotide), and 13,097 movies for telomerase bound with $(GTTGGG)_5$ sstDNA (dataset2) were collected in separate imaging sessions.

Cryo-EM data processing

Dose-fractionated frames except for the first of each movie were 2 × binned (pixel size of 1.36 Å) and aligned for correction of beam-induced drift using MotionCor2.[49] The contrast transfer function (CTF) parameters were determined by CTFFIND4.[50] All micrographs after motion correction and their power spectra were visually inspected, and the micrographs with too much thick carbon area, ice contamination or defocus value outside the range from −0.8 to −4.0 μm were discarded. Gautomatch was used for particle picking with template projections obtained from a previous reconstruction (EMD-7821).[31]

The cryo-EM data processing procedure is outlined in Figure S1(A). Two datasets were initially processed separately in two batches using RELION 3.0.[51] For dataset1, 3,951,762 particles were initially picked and screened as previously described.[34] The resulted 924,041 particles were refined with a soft mask (mask1) to obtain orientation parameters of each particle, which were used as inputs for the following 3D classification with local angular search (RELION options: --sigma_ang 12). 487,859 particles in two good classes were combined and subjected to another round of 3D classification (RELION options: --skip_align --tau2_fudge 10) with a mask (mask2) encompassing p65 La module. 193,669 particles with strong density of p65 La module were eventually selected from dataset1 for the subsequent processing. Dataset2 was processed separately in a similar way, and 106,811 particles were selected.

Since sstDNA and p65 bind to separate regions of TER without any crosstalk, selected particles from two datasets were combined together to improve the overall resolution of p65. Refinement of the combined 300,480 particles generated an overall 3.5 Å resolution reconstruction. Notably, this resolution was obtained mainly based on high-resolution features of TERT, and the local resolution of p65 was only around 4–5 Å. We have also tried signal subtraction and focused refinement of p65 region (p65 together with TER PK, S1 and SL4) alone, but the result was worse than 3D refinement of the whole complex with mask1, probably because the p65 region is too small (stable volume ~70 kDa) to generate enough signal for 3D alignment. Alternatively, we extracted signals of p65 LaM from each particle with mask3 and performed focused 3D classification without alignment (RELION options: --skip_align --tau2_fudge 10). 162,358 particles were selected, followed by CTF refinement and Bayesian polishing. The resulting "shiny" particles were refined with mask1, resulting in a final 3.8 Å resolution reconstruction with improved densities of p65 LaM.

Resolutions of the cryo-EM maps were estimated on the basis of the "gold-standard" Fourier Shell Correlation (FSC) = 0.143 criterion.[52] The cryo-EM maps were corrected for the modulation transfer function (MTF) of the detector, sharpened with a negative B-factor and low-pass filtered to the stated resolution using the *relion_postprocess* program in RELION.[51] Local resolution evaluations were determined by ResMap[53] with two independently refined half-maps. Data collection and processing statistics are given in Table S1.

Model building and refinement

The atomic coordinates of p65 and TER were built and refined using Coot[54] and UCSF Chimera[55]. For the xRRM2 domain (residues 377–417, 453–542), modeling was initiated by fitting of the crystal structure of p65 xRRM2-SL4 (PDB: 4ERD) into the cryo-EM map, and further manually adjusted in Coot. A homology model of La module was initially generated by the SWISS-MODEL server using the crystal structure of human La protein (PDB: 2VOO) as template. The generated LaM model (residues 112–231) shows there is flexible p65-specific loop between α4 and α5 (residues 174–197), which was deleted during the subsequent modeling. LaM alone was unambiguously fitted into the cryo-EM density in Chimera. The orientation and location of RRM1 model (residues 240–331) was manually fitted into the density using the conserved interfaces among RRM1, LaM, and UUU-3′-OH of human La protein. For NTD, the secondary structure information was obtained from NMR chemical shift, the C-terminus (residues 87–104) was modeled as an ideal helix. These generated individual models were served as starting coordinates for rebuilding the p65 (residues 83–542) in RosettaCM[56]. 3000 models were calculated, and the top scoring hits were manually inspected in Coot.

The initial TER model was adopted from previous model (PDB: 6D6V). The residues in PK and SL4 parts with high resolution cryo-EM features were manually adjusted for their base conformation against the density map. The secondary structure of S1 was obtained from NMR data, and then was manually fitted into the density map. The UUU157-159 were initially adopted from the crystal structure of human La protein with UUU-3′-OH (PDB: 2VOO), and then were manually adjusted against the density. The remaining linker and single-stranded regions connecting the

secondary structure elements discussed above were modeled into the density map using Coot.

The entire p65–TER complex was refined in Phenix[57] using "phenix.real_space_refine" with secondary structure, Ramachandran, and rotamer restraints, and was further refined by ISOLDE in ChimeraX[58]. Structural models were validated using Molprobity.[59]

p65 expression and purification

The DNA encoding NTD, LaM, RRM1, La module, αN-La module and αL-xRRM2 constructs were cloned into pETDuet vector, which was further modified to include a hexahistidine (His_6) tag followed by a tobacco etch virus (TEV) protease cleavage site at the N terminus before the protein sequence. For αL-xRRM2, a β2-β3 loop truncation was included as described previously.[25] Protein point substitutions/deletions were carried out with a Q5 site-directed mutagenesis kit (NEB) and verified by DNA sequencing. All the constructs were transformed into *Escherichia coli* BL21-Gold (DE3) for protein expression. Bacterial cultures were grown in M9 minimal medium at 37 °C to an OD_{600} of 0.6–0.8, then transferred to 18 °C for 1 h before induction with 0.5 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) for 18 h. Cells were harvested by centrifugation and the pellets were resuspended with resuspension buffer (20 mM HEPES, pH 8, 1 M NaCl, 1 mM TCEP, 15 mM imidazole, 5% glycerol, and 1 mM PMSF) supplemented with lysozyme, then sonicated on ice. Cell lysate was clarified by centrifugation and filtration, and the His_6-tagged proteins were purified with a nickel affinity column and further purified on an S75 column attached to an AKTA FPLC (GE Healthcare) using NMR buffer.

In vitro transcription and purification of RNA

All TER RNA constructs were prepared by *in vitro* transcription using T7 RNA polymerase (P266L mutant)[60] with synthetic DNA templates (Integrated DNA Technologies) as described previously.[61] After *in vitro* transcription, the RNA samples were ethanol precipitated, purified using 15%-20% denaturing polyacrylamide gels, electroeluted with Elutrap system (Whatman), and further purified by anion exchange with a 5-mL Hi-Trap Q column (GE Healthcare). All purified RNA was desalted and exchanged into nanopure water using an Amicon filtration system with 3,000 molecular weight cutoff membranes (Millipore). RNA samples were diluted to concentrations of ~10 μM in the desired buffer containing 10 mM potassium phosphate, heated at 95 °C for 3 min, snap-cooled on ice for 30 min, and then concentrated for NMR and gel shift experiments.

NMR sample preparation and data collection

For the assignment of NTD, 0.3 mM sample uniformly enriched with ^{15}N and ^{13}C in 20 mM sodium phosphate, pH 7.0, 50 mM NaCl, and 2 mM DTT, 0.02% NaN_3, 5% v/v D_2O) was used. NMR spectra were recorded on Bruker DRX 600 MHz and Avance 800 MHz spectrometers equipped with HCN cryogenic probes at 298 K. The backbone assignment (83% completion) was achieved from analysis of HNCACB, CBCA(CO)NH, HNCA and C(CO)NH spectra. The NMR data of ^{13}C, ^{15}N-labelled αL-xRRM2 with helix αL was collected on Bruker DRX 600 MHz spectrometers in NMR buffer (20 mM sodium phosphate, pH 6.1, 50 mM KCl, 1 mM TCEP, 5% v/v D_2O) at 298 K. The backbone assignment of ^{13}C, ^{15}N-labelled αL-xRRM2 (69% completion) was achieved from analysis of HNCACB, CBCA(CO)NH, HNCA and CC(CO)NH, aided by previous NMR assignments of xRRM2. Spectra of all RNA samples (0.5 mM to 1 mM in 10 mM sodium phosphate, pH 6.4, 50 mM KCl, 10% v/v D_2O) were collected on Avance 800-MHz spectrometer at 283 K. To assign the 1D imino proton spectra, exchangeable-proton NOESY spectra were recorded. All other p65 domain samples were in 20 mM sodium phosphate, pH 7.0, 50 mM KCl, 1 mM TCEP, 10% v/v D_2O and were collected at 298 K on Avance 800-MHz spectrometer. The NTD and LaMod(108–355) binding experiments were conducted by adding LaMod(108–355) to NTD by varying the [NTD]:[LaMod(108–355)] concentration ratio from 1:0 to 1:2.0. The protein domains were mixed under diluted conditions and then concentrated using Amicon filtration system (Millipore, 3 K MWCO). The final protein concentration of NTD in the NMR samples was 50 μM. NMR spectra were processed and analyzed with Topspin (Bruker), NMRPipe, and NMRFAM-Sparky.

Electrophoretic mobility shift assay (EMSA)

TER (S1hp-S4hp, SL4hp, and SL4hpΔU117) and p65 (αN-LaMod(67–355), LaMod(108–355), αL-xRRM2 and xRRM2) samples were prepared separately in binding buffer (pH 7.0, 50 mM KCl, and 1 mM TCEP). RNA (100 nM for S1hp-S4hp and 150 nM for SL4hp and SL4hpΔU117, respectively) and protein were mixed at various ratios in a total volume of 10 μl. The complexes were incubated on ice for 1 h prior to gel electrophoresis on 5.7% polyacrylamide gel (37.5:1 crosslinking ratio) with 1X TBE buffer at 4 °C. Gels were subsequently stained with SYBR Gold (Invitrogen) and imaged by Pharos FX Plus scanner (Bio-Rad). The signal intensity of the free DNA was quantified with ImageJ software. The binding affinity (*Kd*) was further determined by fitting the fraction bound values to the following equation

$$\theta = \frac{([Pt] - \frac{[Rt]+[Pt]+K_{Dapp}-\sqrt{([Rt]+[Pt]+K_{Dapp})^2-4[Rt][Pt]}}{2})^n}{K_{Dapp}^n + ([Pt] - \frac{[Rt]+[Pt]+K_{Dapp}-\sqrt{([Rt]+[Pt]+K_{Dapp})^2-4[Rt][Pt]}}{2})^n}$$

where *[Rt], [Pt], K_{Dapp}, n* and θ are the total RNA concentration, protein concentration, global dissociation constant, the Hill coefficient, and fraction bound, respectively.

Accession Numbers

Cryo-EM density maps and atomic coordinates for p65-focused refinement of *Tetrahymena* telomerase have been deposited in the Electron Microscopy Data Bank and wwPDB, respectively, under accession codes EMD-29903 and PDB 8GAP. Backbone chemical shifts assignments for NTD (free and SDS-bound) and free αL-xRRM2 and imino assignments for S1hp and S1hp-ACAA have been deposited in the Biological Magnetic Resonance Data Bank, under accession IDs 51798, 51797, 51808, 51809, respectively.

Funding

This work was supported by the National Institutes of Health (NIH) R35GM131901 and R01AI155170, and the National Science Foundation (NSF) MCB2016540 to J.F, and NIH grant R01GM071940 to Z.H.Z. The NMR facility was supported in part by NIH instrumentation grants S10OD016336 and S10OD025073 and Department of Energy DE-FC03-02ER63421. The Electron Imaging Center for Nanomachines was supported by UCLA and by instrumentation grants from NIH [S10RR23057, S10OD018111 and U24GM116792], NSF [DBI-1338135 and DMR-1548924].

CRediT authorship contribution statement

Yaqiang Wang: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Yao He:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Yanjiao Wang:** Validation, Formal analysis, Investigation, Data curation, Writing – review & editing, Visualization. **Yuan Yang:** Validation, Formal analysis, Investigation, Data curation, Writing – review & editing, Visualization. **Mahavir Singh:** Investigation. **Catherine D. Eichhorn:** Data curation, Investigation, Writing – review & editing. **Xinyi Cheng:** Investigation. **Yi Xiao Jiang:** Investigation. **Z. Hong Zhou:** Supervision, Funding acquisition. **Juli Feigon:** Conceptualization, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

DATA AVAILABILITY

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2023.168044.

Received 25 January 2023;
Accepted 3 March 2023;

Keywords:
La protein;
La module;
RRM;
telomerase;
pseudoknot

† Equal contribution.
‡ Present Address: Molecular Biophysics Unit, Indian Institute of Science, Bengaluru, 560012, India (Mahavir Singh); Department of Chemistry, University of Nebraska, Lincoln, NE 68508, USA (Catherine D. Eichhorn).

References

1. Maraia, R.J., Mattijssen, S., Cruz-Gallardo, I., Conte, M.R., (2017). The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives. *Wiley Interdiscip Rev RNA* **8**
2. Wolin, S.L., Cedervall, T., (2002). The La protein. *Annu. Rev. Biochem* **71**, 375–403.
3. Bayfield, M.A., Yang, R., Maraia, R.J., (2010). Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *BBA* **1799**, 365–378.
4. Kotik-Kogan, O., Valentine, E.R., Sanfelice, D., Conte, M. R., Curry, S., (2008). Structural analysis reveals conformational plasticity in the recognition of RNA 3' ends by the human La protein. *Structure* **16**, 852–862.
5. Naeeni, A.R., Conte, M.R., Bayfield, M.A., (2012). RNA chaperone activity of human La protein is mediated by variant RNA recognition motif. *J. Biol. Chem.* **287**, 5472–5482.
6. Jacks, A., Babon, J., Kelly, G., Manolaridis, I., Cary, P.D., Curry, S., et al., (2003). Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure* **11**, 833–843.

7. Hasler, D., Meister, G., Fischer, U., (2021). Stabilize and connect: the role of LARP7 in nuclear non-coding RNA metabolism. *RNA Biol.* **18**, 290–303.
8. Markert, A., Grimm, M., Martinez, J., Wiesner, J., Meyerhans, A., Meyuhas, O., et al., (2008). The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep.* **9**, 569–575.
9. Krueger, B.J., Jeronimo, C., Roy, B.B., Bouchard, A., Barrandon, C., Byers, S.A., et al., (2008). LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res.* **36**, 2219–2229.
10. Muniz, L., Egloff, S., Kiss, T., (2013). RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. *Nucleic Acids Res.* **41**, 4686–4698.
11. Wang, Y., Feigon, J., (2017). Structural biology of telomerase and its interaction at telomeres. *Curr. Opin. Struct. Biol.* **47**, 77–87.
12. Chan, H., Wang, Y., Feigon, J., (2017). Progress in Human and Tetrahymena Telomerase Structure Determination. *Annu. Rev. Biophys.* **46**, 199–225.
13. Wang, Y., Susac, L., Feigon, J., (2019). Structural Biology of Telomerase. *In: Cold Spring Harbor perspectives in biology*, p. 11.
14. Aigner, S., Postberg, J., Lipps, H.J., Cech, T.R., (2003). The Euplotes La motif protein p43 has properties of a telomerase-specific subunit. *Biochemistry* **42**, 5736–5747.
15. Aigner, S., Lingner, J., Goodrich, K.J., Grosshans, C.A., Shevchenko, A., Mann, M., et al., (2000). Euplotes telomerase contains an La motif protein produced by apparent translational frameshifting. *EMBO J.* **19**, 6230–6239.
16. Paez-Moscoso, D.J., Pan, L., Sigauke, R.F., Schroeder, M. R., Tang, W., Baumann, P., (2018). Pof8 is a La-related protein and a constitutive component of telomerase in fission yeast. *Nature Commun.* **9**, 587.
17. Mennie, A.K., Moser, B.A., Nakamura, T.M., (2018). LARP7-like protein Pof8 regulates telomerase assembly and poly(A)+TERRA expression in fission yeast. *Nature Commun.* **9**, 586.
18. Collopy, L.C., Ware, T.L., Goncalves, T., S, I.K., Yang, Q., Amelina, H., et al., (2018). LARP7 family proteins have conserved function in telomerase assembly. *Nature Commun.* **9**, 557.
19. Basu, R., Eichhorn, C.D., Cheng, R., Peterson, R.D., Feigon, J., (2021). Structure of *S. pombe* telomerase protein Pof8 C-terminal domain is an xRRM conserved among LARP7 proteins. *RNA Biol.* **18**, 1181–1192.
20. Hu, X., Kim, J.K., Yu, C., Jun, H.I., Liu, J., Sankaran, B., et al., (2020). Quality-Control Mechanism for Telomerase RNA Folding in the Cell. *Cell Rep.* **33**, 108568
21. Witkin, K.L., Collins, K., (2004). Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes Dev.* **18**, 1107–1118.
22. Stone, M.D., Mihalusova, M., O'Connor, C.M., Prathapam, R., Collins, K., Zhuang, X., (2007). Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458–461.
23. Prathapam, R., Witkin, K.L., O'Connor, C.M., Collins, K., (2005). A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nature Struct. Mol. Biol.* **12**, 252–257.
24. O'Connor, C.M., Lai, C.K., Collins, K., (2005). Two purified domains of telomerase reverse transcriptase reconstitute sequence-specific interactions with RNA. *J. Biol. Chem.* **280**, 17533–17539.
25. Singh, M., Wang, Z., Koo, B.K., Patel, A., Cascio, D., Collins, K., et al., (2012). Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol. Cell* **47**, 16–26.
26. O'Connor, C.M., Collins, K., (2006). A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. *Mol. Cell Biol.* **26**, 2029–2036.
27. Uchikawa, E., Natchiar, K.S., Han, X., Proux, F., Roblin, P., Zhang, E., et al., (2015). Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. *Nucleic Acids Res.* **43**, 3373–3388.
28. Eichhorn, C.D., Yang, Y., Repeta, L., Feigon, J., (2018). Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7. *PNAS* **115**, E6457–E6466.
29. Eichhorn, C.D., Chug, R., Feigon, J., (2016). hLARP7 C-terminal domain contains an xRRM that binds the 3' hairpin of 7SK RNA. *Nucleic Acids Res.* **44**, 9977–9989.
30. Akiyama, B.M., Loper, J., Najarro, K., Stone, M.D., (2012). The C-terminal domain of Tetrahymena thermophila telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA. *RNA* **18**, 653–660.
31. Jiang, J., Wang, Y., Susac, L., Chan, H., Basu, R., Zhou, Z. H., et al., (2018). Structure of Telomerase with Telomeric DNA. *Cell* **173** (1179–1190), e1113.
32. Jiang, J., Chan, H., Cash, D.D., Miracco, E.J., Ogorzalek Loo, R.R., Upton, H.E., et al., (2015). Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science* **350**, aab4070.
33. Jiang, J., Miracco, E.J., Hong, K., Eckert, B., Chan, H., Cash, D.D., et al., (2013). The architecture of Tetrahymena telomerase holoenzyme. *Nature* **496**, 187–192.
34. He, Y., Wang, Y., Liu, B., Helmling, C., Susac, L., Cheng, R., et al., (2021). Structures of telomerase at several steps of telomere repeat synthesis. *Nature* **593**, 454–459.
35. Yang, Y., Liu, S., Egloff, S., Eichhorn, C.D., Hadjian, T., Zhen, J., et al., (2022). Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP. *Mol. Cell* **82** (1724–1736), e1727.
36. Dong, G., Chakshusmathi, G., Wolin, S.L., Reinisch, K.M., (2004). Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *EMBO J.* **23**, 1000–1007.
37. Drozdetskiy, A., Cole, C., Procter, J., Barton, G.J., (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* **43**, W389–W394.
38. Marsh, J.A., Singh, V.K., Jia, Z., Forman-Kay, J.D., (2006). Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci.* **15**, 2795–2804.
39. Sanfelice, D., Kelly, G., Curry, S., Conte, M.R., (2008). NMR assignment of the N-terminal region of human La free and in complex with RNA. *Biomol. NMR Assign.* **2**, 107–109.
40. Cruz-Gallardo, I., Martino, L., Kelly, G., Atkinson, R.A., Trotta, R., De Tito, S., et al., (2019). LARP4A recognizes

polyA RNA via a novel binding mechanism mediated by disordered regions and involving the PAM2w motif, revealing interplay between PABP, LARP4A and mRNA. *Nucleic Acids Res.* **47**, 4272–4291.
41. Cash, D.D., Feigon, J., (2017). Structure and folding of the Tetrahymena telomerase RNA pseudoknot. *Nucleic Acids Res.* **45**, 482–495.
42. Berman, A.J., Gooding, A.R., Cech, T.R., (2010). 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.
43. Podlevsky, J.D., Bley, C.J., Omana, R.V., Qi, X., Chen, J. J., (2008). The telomerase database. *Nucleic Acids Res.* **36**, D339–D343.
44. Singh, M., Choi, C.P., Feigon, J., (2013). xRRM: a new class of RRM found in the telomerase La family protein p65. *RNA Biol.* **10**, 353–359.
45. Richards, R.J., Wu, H., Trantirek, L., O'Connor, C.M., Collins, K., Feigon, J., (2006). Structural study of elements of Tetrahymena telomerase RNA stem-loop IV domain important for function. *RNA* **12**, 1475–1485.
46. Stefano, J.E., (1984). Purified lupus antigen La recognizes an oligouridylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* **36**, 145–154.
47. Dock-Bregeon, A.C., Lewis, K.A., Conte, M.R., (2021). The La-related proteins: structures and interactions of a versatile superfamily of RNA-binding proteins. *RNA Biol.* **18**, 178–193.
48. Mastronarde, D.N., (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51.
49. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., Agard, D.A., (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **14**, 331–332.
50. Rohou, A., Grigorieff, N., (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221.
51. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., et al., (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**
52. Rosenthal, P.B., Henderson, R., (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745.
53. Kucukelbir, A., Sigworth, F.J., Tagare, H.D., (2014). Quantifying the local resolution of cryo-EM density maps. *Nature Methods* **11**, 63–65.
54. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501.
55. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., et al., (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
56. Song, Y., DiMaio, F., Wang, R.Y., Kim, D., Miles, C., Brunette, T., et al., (2013). High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735–1742.
57. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., et al., (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221.
58. Croll, T.I., (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519–530.
59. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., et al., (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315.
60. Guillerez, J., Lopez, P.J., Proux, F., Launay, H., Dreyfus, M., (2005). A mutation in T7 RNA polymerase that facilitates promoter clearance. *PNAS* **102**, 5958–5963.
61. Wang, Y., Yesselman, J.D., Zhang, Q., Kang, M., Feigon, J., (2016). Structural conservation in the template/pseudoknot domain of vertebrate telomerase RNA from teleost fish to human. *PNAS* **113**, E5125–E5134.