Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7

Catherine D. Eichhorn[a], Yuan Yang[a], Lucas Repeta[a], and Juli Feigon[a,1]

[a]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095-1569

Edited by Michael F. Summers, Howard Hughes Medical Institute, University of Maryland, Baltimore County, Baltimore, MD, and approved May 30, 2018 (received for review April 12, 2018)

The La and the La-related protein (LARP) superfamily is a diverse class of RNA binding proteins involved in RNA processing, folding, and function. Larp7 binds to the abundant long noncoding 7SK RNA and is required for 7SK ribonucleoprotein (RNP) assembly and function. The 7SK RNP sequesters a pool of the positive transcription elongation factor b (P-TEFb) in an inactive state; on release, P-TEFb phosphorylates RNA Polymerase II to stimulate transcription elongation. Despite its essential role in transcription, limited structural information is available for the 7SK RNP, particularly for protein–RNA interactions. Larp7 contains an N-terminal La module that binds UUU-3′OH and a C-terminal atypical RNA recognition motif (xRRM) required for specific binding to 7SK and P-TEFb assembly. Deletion of the xRRM is linked to gastric cancer in humans. We report the 2.2-Å X-ray crystal structure of the human La-related protein group 7 (hLarp7) xRRM bound to the 7SK stem-loop 4, revealing a unique binding interface. Contributions of observed interactions to binding affinity were investigated by mutagenesis and isothermal titration calorimetry. NMR ^{13}C spin relaxation data and comparison of free xRRM, RNA, and xRRM–RNA structures show that the xRRM is preordered to bind a flexible loop 4. Combining structures of the hLarp7 La module and the xRRM–7SK complex presented here, we propose a structural model for Larp7 binding to the 7SK 3′ end and mechanism for 7SK RNP assembly. This work provides insight into how this domain contributes to 7SK recognition and assembly of the core 7SK RNP.

RNP structure | RNP assembly | RNA recognition motif | X-ray crystallography | NMR

The La and the La-related protein (LARP) superfamily is a diverse class of RNA binding proteins that are involved in RNA processing, chaperone activity, and function (1, 2). Genuine La proteins contain an N-terminal La module composed of an La motif and RNA recognition motif (RRM) (3, 4) that act in concert to broadly bind to RNA Polymerase III (RNAPIII) transcripts at the single-stranded UUU-3′OH (5–7). In eukaryotes, except in yeasts, genuine La has a second C-terminal atypical RRM with features distinct from the canonical RRM fold (8) that has RNA folding and chaperone activity (9, 10). Although all LARPs contain an N-terminal La module, the remaining domain structure diverges significantly among the different LARP families depending on their substrate RNA and function (1, 2). Members of the LARP7 family are the closest relatives to genuine La and like La, contain an N-terminal La module and C-terminal atypical RRM (xRRM), but unlike genuine La, they bind a single cognate RNA. In ciliates *Tetrahymena thermophila* and *Euplotes aediculatus* and fission yeast, the LARP7 family proteins p65, p43, and Pof8, respectively, bind telomerase RNA (TER) and promote hierarchal assembly of the telomerase ribonucleoprotein (RNP) particle (11–15). In metazoa, the LARP7 family protein Larp7 binds the 7SK long noncoding RNA (lncRNA) and is required to protect 7SK from degradation and for hierarchal assembly of the 7SK RNP (16, 17).

7SK is a highly abundant nuclear lncRNA that regulates RNAPII transcription, primarily by sequestering and inactivating the positive transcription elongation factor b (P-TEFb) (18, 19). P-TEFb, a heterodimer composed of cyclin-dependent kinase 9 and cyclin T1 (20), is an integral component of the super elongation complex that phosphorylates negative transcription elongation factors and the C-terminal domain (CTD) of RNAPII to clear promoter-proximal paused RNAPII complexes and stimulate productive elongation of mRNA transcripts (21). P-TEFb assembles onto the 7SK RNP in an inactive state in a dynamic and reversible equilibrium (18, 19). Although Larp7 and another protein cofactor, HEXIM1/2 (22, 23), are required to assemble P-TEFb onto the 7SK RNP (17, 24, 25), multiple protein cofactors are required to disassemble P-TEFb from the 7SK RNP, including CTIP2 (26); hnRNPs A, Q, and R (27); HMGA1 (28–32); KAP1 (33); PPM1G (34); and DDX21 helicase (35). Control of P-TEFb activity is essential for cell maintenance, and several diseases are linked to improper P-TEFb regulation, including cardiac hypertrophy (36), leukemia (37–39), and lymphoma (40). Furthermore, P-TEFb is an essential host cofactor for HIV replication (41, 42), where it is hijacked from the 7SK RNP by the transactivator of transcription protein and recruited to stalled HIV genomic transcripts (43–47). Beyond regulation of P-TEFb, 7SK also regulates the RNAPII transcription of small nuclear RNAs (48) as well as enhancer RNAs (49) and interacts with hnRNP R to regulate its function in axon maintenance (50).

7SK folds into four stem loops (SLs) as identified by footprinting and chemical mapping experiments (51, 52) (Fig. 1*A*), although only the 5′ SL1 and 3′ SL4 are required for P-TEFb assembly (24). The methylphosphate capping enzyme (MePCE)

Significance

The 7SK ribonucleoprotein (RNP) complex regulates the activity of the kinase positive transcription elongation factor b (P-TEFb), an essential activator of RNA Polymerase II transcription. The human La-related protein group 7 (hLarp7) protein is an essential and constitutively assembled component of the 7SK RNP and is required for 7SK RNA stability and P-TEFb recruitment. We report the structure of the hLarp7 C-terminal RNA recognition motif bound to the 7SK stem-loop 4, revealing a unique binding interface. From this and other available structures, we generate a structural model of hLarp7 bound to the 7SK 3′ end. This work provides seminal insights into the unique recognition of 7SK RNA by hLarp7 and a working model for how hLarp7 assembles with 7SK to form the 7SK RNP.

Author contributions: C.D.E., Y.Y., and J.F. designed research; C.D.E., Y.Y., L.R., and J.F. performed research; C.D.E., Y.Y., and J.F. analyzed data; and C.D.E., Y.Y., and J.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The crystallography, atomic coordinates, and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID code 6D12).

[1]To whom correspondence should be addressed. Email: feigon@mbi.ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1806276115/-/DCSupplemental.

Published online June 26, 2018.

and Larp7 are constitutively assembled with 7SK and bind the 5′ and 3′ ends, respectively, to form a stable core 7SK RNP (16, 53). In humans, MePCE associates with 7SK and U6 RNA transcripts in vivo (54) and methylates the 5′ γ-phosphate of 7SK (52, 53). Larp7 interacts with MePCE in a 7SK-dependent manner and inactivates its capping activity by an unknown mechanism (52, 53). Larp7 requires both SL4 and the UUU-3′ OH for binding in vivo (Fig. 1*A*) (55). Larp7 is essential for 7SK stability in vivo and the stable association of P-TEFb to the 7SK RNP (16, 17). Down-regulation of Larp7 results in high levels of transcriptional activity (17, 56). Consistent with these observations, mutations resulting in loss of Larp7 expression in humans are correlated with primordial dwarfism (57–60), cervical cancer (61), and breast cancer metastasis (56).

Human La-related protein group 7 (hLarp7) comprises an N-terminal La module, a 256-aa linker, C-terminal xRRM, and a 31-aa CTD (Fig. 1*B*). The La module, structurally similar to the La module of genuine La, binds 7SK at the UUU-3′OH (55, 62). The linker between the La module and C-terminal RRM is predicted to have low structural complexity (62). The C-terminal 21 amino acids are required for assembly of the 7SK RNP in vivo (25) but do not bind 7SK (63); rather, these residues seem to be required for MePCE capping inactivation (52).

The hLarp7 C-terminal xRRM is required for specific recognition of 7SK (25, 55), and mutations resulting in deletion of this domain are linked to gastric cancer (64, 65). In vivo mutagenesis (55), chemical footprinting (62), and NMR chemical shift mapping (63) experiments showed that this RRM binds to 7SK SL4. The hLarp7 xRRM structure in the absence of RNA is globally similar to the xRRMs in genuine La and p65 (8, 63, 66). All three contain an additional helix α3 that is across the surface of the canonical β1-α1-β2-β3-α2-β4 β-sheet and lack the RNA binding RNP1 and RNP2 sequences normally located on β3 and β1, respectively. The structure of p65 C-terminal RRM bound to TER led to its identification as a new class of RRM named xRRM (66, 67). Rather than bind across the β-sheet surface like canonical RRMs (68, 69), the xRRM recognizes a combination of single-stranded and base-paired nucleotides with an RNP3 surface on β2, conserved Arg on β3, and helix α3 (63, 66, 67). Sequence and structural homology suggested that other LARP7 family proteins contained xRRMs, although this could not be confirmed without structural information of their RNA–protein complexes.

Despite the central role of 7SK in transcription regulation and its discovery over 30 y ago (70), relatively little is known at a structural or mechanistic level about how proteins assemble and disassemble with 7SK to form a functional 7SK RNP. Structures of individual 7SK RNA or protein components are sparse (62, 63, 71–74), and the structure of the hLarp7 La module bound to 5′-UUU-3′ is the only example of a 7SK RNP protein domain in complex with RNA (62). However, this structure does not reveal the orientation of the La module on 7SK or explain how hLarp7 specifically recognizes 7SK over other RNAs with UUU-3′OH.

Here, we report the 2.2-Å X-ray crystal structure of the hLarp7 C-terminal xRRM bound to 7SK SL4. The structure reveals that the xRRM binds to the SL4 loop and upper stem in a unique manner, providing insights into the mode of interaction between these two domains. Mutagenesis of both protein and RNA residues confirms the importance of the observed interactions for high-affinity binding. We compare the hLarp7 and p65 xRRMs in complex with their cognate RNAs and propose how other LARP7 family proteins bind to their cognate RNA. Combining the structures of the hLarp7 La module (62), 7SK SL4 RNA (71), and the xRRM–SL4 structure reported here, we generate a model for the hLarp7–7SK complex using HADDOCK and propose a mechanism for the association of hLarp7 with MePCE to assemble the core 7SK RNP.

Results and Discussion

Structure of hLarp7 xRRM Bound to 7SK SL4 RNA. The hLarp7 xRRM and 7SK SL4 domain boundaries were determined using a combination of EMSAs and isothermal titration calorimetry (ITC), where the upper stem and loop 4 of SL4 were determined to be necessary and sufficient for high-affinity binding to the xRRM (63). Using this information, we designed numerous protein and RNA constructs for crystallization of a complex of the hLarp7 xRRM with 7SK SL4. Crystals that diffracted to 2.2 Å were ultimately obtained using a protein with two point substitutions (E501L and Q504L) in helix α2 to reduce surface entropy and promote crystal contacts (75) (*SI Appendix*, Fig. S1). These mutations did not impair binding to 7SK SL4 (*SI Appendix*, Fig. S1). The RNA construct contains two hairpins A and

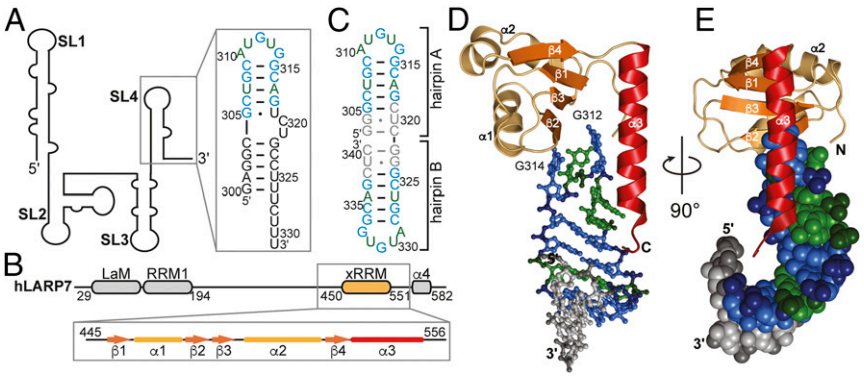

Fig. 1. Structure of 7SK RNA and hLarp7 xRRM complex. (*A*) Cartoon schematic of the 7SK RNA with the sequence of SL4 and UUU-3′OH end shown. (*Inset*) A and U residues are colored green, G and C residues are colored blue, and residues that are not part of the crystal construct are colored black. (*B*) Domain structure of hLarp7 with the secondary structure topology of the C-terminal xRRM. (*Inset*) Residue numbers reflect the boundary domains for the xRRM and crystal construct. (*C*) RNA crystal construct has two identical hairpins. Residues in each hairpin are colored as shown in *A*, and the designed crystal construct residues are colored gray. Each hairpin contains the 7SK SL4 5-bp upper stem and loop 4 (residues 306–319 in hairpin A, residues 325–338 in hairpin B), with U319 substituted with a C to form a G305-C319 base pair. Two additional base pairs (G303-C321 and G304-U320 in hairpin A, G322-C340 and G323-U339 in hairpin B) were added to stabilize the hairpin. (*D*) Side view of the xRRM–SL4 complex. RNA is shown in ball and stick view, with protein shown as cartoon. Helices α1 and α2 and loops are shown in light orange, β-strands are shown in orange, helix α3 is shown in red, and residues that become stabilized on binding SL4 are shown in maroon. RNA residues are colored as shown in *B*, with phosphates colored in darker shades of blue (G, C) or green (A, U). (*E*) Top view of the xRRM–SL4 complex with SL4 shown as spheres. Helix α3 inserts along the RNA major groove.

B, identical in sequence, that were designed to form a coaxially stacked "dumbbell" with two independent protein binding sites (Fig. 1*C* and *SI Appendix*, Fig. S1). In the crystal structure, the RNA has the expected global fold with one xRRM bound to each RNA hairpin at loop 4, resulting in symmetric binding sites (*SI Appendix*, Fig. S1). Electron density was visible at contour levels up to 2σ for RNA residues that interact with the xRRM (nucleotides 307–317 and 326–336) and for xRRM residues 446–545 (*SI Appendix*, Fig. S1). Electron density was visible at contour levels up to 1σ for RNA residues at the stem–stem junction, which do not bind to protein, and for xRRM C-terminal residues 546–548 (*SI Appendix*, Fig. S1). No electron density was observed for N-terminal residues 444–445 and C-terminal residues 549–556.

The hLarp7 xRRM has the expected βαββαβα topology, consistent with the structure of the free hLarp7 xRRM determined by solution NMR (Fig. 1 *B*, *D*, and *E*) (63). The front face of the xRRM features a four-strand antiparallel β-sheet composed of β4 (residues 519–523), β1 (residues 456–461), β3 (residues 492–497), and β2 (residues 482–486) with helix α3 (residues 526–546) lying on the surface of β-sheet perpendicular to the β-strand axis (Fig. 1 *D* and *E*). The back of the xRRM consists of helices α1 (residues 468–478) and α2 (residues 500–517), which are under β2 and β4, respectively. There are five structured loops [β1-α1 (residues 462–467), α1-β2 (residues 479–481), β2-β3 (residues 487–491), β3-α2 (residues 498–499), α2-β4 (residue 518), and β4-α3 (residues 524–525)] and a structured N terminus (residues 446–455) and C terminus (residues 547–548).

Each hairpin of the SL4 RNA crystal construct comprises a 7-bp A-form helix (nucleotides 303–309 and 315–321 in hairpin A, nucleotides 322–328 and 334–340 in hairpin B) and a 5-nt loop (residues 310–314 in hairpin A, residues 329–333 in hairpin B) (Fig. 1 *C*–*E*). The hLarp7 xRRM binds to loop 4 of 7SK SL4, in agreement with NMR chemical shift mapping (63), in vivo pull-down (55), and small angle X-ray scattering (62) experiments. Loop residues are sandwiched between the β-sheet surface at β2 and helix α3 on the major groove side and the β-sheet surface and helix α1 on the minor groove side. Helix α3 contacts loop residues and extends down into the major groove at the top of the stem (Fig. 1*E*).

Interlacing xRRM–RNA Contacts Define the Binding Interface. In loop 4, A310 and U311 stack on the 5′ side of the RNA helix. The backbone makes an S turn about G312, U313, and G314, and these residues are flipped out of loop 4 to interact with the protein. A key feature of the xRRM–RNA binding interface is a five-residue stacking interaction composed of alternating protein and RNA residues on the 3′ side of the loop (Fig. 2 *A* and *B*). Loop 4 residues G312 and G314 are both oriented toward β2; G312 inserts between helix α3 and β2, and G314 is in the same plane as the xRRM β-sheet (Fig. 2*A*). U313 is oriented perpendicular to G312 and G314. To determine the contribution of protein–RNA interactions to formation of a stable xRRM–SL4 complex, we performed point substitutions for both protein and RNA residues and determined the equilibrium dissociation constant (K_d) using ITC (Table 1).

β2, containing the RNP3 motif, has extensive hydrogen bonding (H bonding) and stacking interactions with loop 4 residues G312, U313, and G314 (Fig. 2 *A*–*D*). RNP3 residue Y483 stacks between G312 and G314, and its hydroxyl group H bonds to the U313 phosphate OP1 (Fig. 2 *A* and *D*). A Y483A substitution resulted in a fourfold reduction in binding (K_d = 498 nM) (Table 1 and *SI Appendix*, Fig. S3). In contrast, the more conservative Y483F substitution did not reduce binding compared with the wild-type xRRM (K_d = 150 nM) (Table 1 and *SI Appendix*, Fig. S3). Sequence alignment of Larp7 proteins shows that Y483 is 95% conserved, and other residues at this position are usually His and Phe (*SI Appendix*, Fig. S4), suggesting that residues that retain π–π stacking between G312 and G314 are important for RNA recognition. Other RNP3 residues V484 and D485 H bond to the Watson–Crick (WC) faces of G314 (Fig. 2 *A* and *E*) and G312 (Fig. 2 *A* and *C*), respectively. On β3, R496 has two H bonds to the Hoogsteen (HG) face of G312 (Fig. 2*C*) and one H bond to the U313 base O4 (Fig. 2*D*). R496A substitution resulted in a sevenfold reduction in binding to SL4 (K_d = 920 nM) (Table 1 and *SI Appendix*, Fig. S3). Consistent with its contribution to high-affinity binding to RNA, R496 is 100% conserved among Larp7 proteins (*SI Appendix*, Fig. S4).

Of helix α3 residues that are across the surface of the β-sheet (residues 526–537), Y532 and I536 contact loop 4 residue G312. Y532 has one H bond to the imino of G312 (Fig. 2*C*), and I536 stacks on G312 (Fig. 2*A*). Helix α3 has extensive interactions with loop 4 at the C-terminal end (residues 538–546), where helix α3 is over G312, U311, and A310 and inserts in the major groove of the RNA helix (Figs. 1*E* and 2*A*). The two amino acids following helix α3 are in the RNA major groove, and

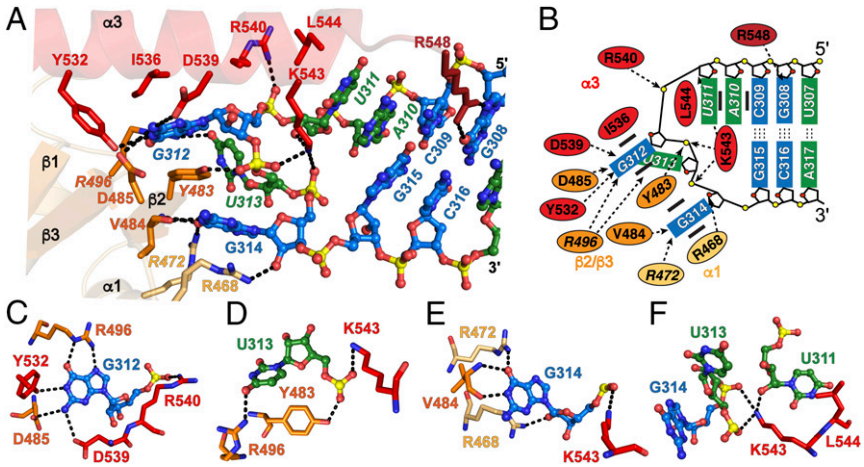

Fig. 2. hLarp7 xRRM recognition of 7SK loop 4. (*A*) Overview of the protein–RNA binding interface shows that the mode of interaction features interlacing protein and RNA residues. Protein is shown in cartoon view, with residues that interact with RNA shown as sticks. RNA is shown in ball and stick view and colored by element (nitrogen is blue, oxygen is red, phosphorus is yellow, carbon is blue for G and C or green for A and U). Italicized residues indicate point substitutions. (*B*) Cartoon schematic of RNA–protein interactions. H bonds are shown as dashes, and stacking interactions are shown as solid bars. (*C*–*F*) Protein recognition of individual loop 4 residues shows an extensive interaction network for residues (*C*) G312, (*D*) U313, and (*E*) G314 by the xRRM. (*F*) K543 inserts in loop 4 and has H bonds to U311, U313, and G314.

Table 1. Thermodynamic parameters for binding of hLarp7 to 7SK SL4 by ITC

Construct	K_d (nM)	*N*
xRRM		
WT*	129 ± 29	1.0 ± 0.01
F451A	120 ± 70	0.7 ± 0.03
R472A	91 ± 0.6	1.1 ± 0.08
R496A	920 ± 200	1.0 ± 0.1
Y483A	498 ± 95	1.0 ± 0.02
Y483F	150 ± 43	1.1 ± 0.07
7SK SL4		
A310C	322 ± 69	0.9 ± 0.2
U311A	51 ± 33	1.0 ± 0.1
U311G	121 ± 69	1.1 ± 0.3
G312C*	607 ± 130	1.5 ± 0.3
G312U	113 ± 33	1.0 ± 0.01
U313A	74 ± 26	0.9 ± 0.1
U313C	125 ± 53	1.0 ± 0.07
hLarp7		
1–582	6.6 ± 4	1.1 ± 0.1
1–190	230 ± 83	1.6 ± 0.08
445–582*	97 ± 13	0.8 ± 0.02

All protein constructs contained residues 445–561. SL4 RNA contained residues 300–325, with a U326C substitution and additional G-C base pair added (sequence and secondary structure are shown in Fig. 5). hLarp7 domain constructs are full-length 1–582, La module (1–190), and xRRM-CTD (445–582). The isolated xRRM domain (445–561) is labeled WT.
*Values have been previously reported in ref. 63.

R548 has two H bonds to the HG face of G308 (Fig. 2*A*). D539 and R540 splay over G312 and H bond to its WC face and phosphate OP1, respectively (Fig. 2*C*). The stacked A310 and U311 on the 5′ side of loop 4 create a pocket on the 3′ side of loop 4, where K543 inserts and H bonds to the U311 ribose 2′ OH, U313 phosphate OP2, and G314 OP2 (Fig. 2 *A*, *D*, and *F*). Although A310 does not have any direct interactions with the xRRM, an A310C substitution reduced the binding affinity threefold (K_d = 322 nM) (Table 1 and *SI Appendix*, Fig. S2). This substitution sequesters G314 in a base pair (*SI Appendix*, Fig. S5), turning the pentaloop into a triloop and requiring this base pair to be broken for G314 to flip out and interact with the xRRM and K543 to insert in the loop. In addition to the H bond between the U311 ribose 2′OH and K543, U311 stacks between A310 and L544 (Fig. 2 *A* and *F*). U311 is the least conserved residue in loop 4 but is highly conserved to be a pyrimidine (76–78). However, a U311G substitution did not disrupt binding to the xRRM (K_d = 121 nM), and surprisingly, a U311A substitution improved the binding affinity 2.5-fold (K_d = 51 nM) (Table 1 and *SI Appendix*, Fig. S2). Substitution from a pyrimidine to a purine may increase stacking between A310 and L544 and would not disrupt the H bond between the ribose 2′OH and K543.

On the 3′ side of loop 4, G312 forms extensive interactions with xRRM residues in β2, β3, and helix α3 (Fig. 2 *B* and *C*). G312 stacks between I536 (α3) and Y483 (β2) (Fig. 2*A*). There are three H bonds on its WC face to D485 (β2) and to Y532 and D539 (α3), two H bonds on the HG face to R496 (β3), and one H bond between the phosphate OP1 and R540 (α3) (Fig. 2*C*). Nearly all H-bond donor/acceptor sites are occupied, explaining why a G312C substitution reduces the binding affinity to the xRRM fivefold in vitro (Table 1) and abolishes binding to hLARP7 in vivo (55, 62, 63). In contrast to the G312C substitution, no change in K_d was observed with a G312U substitution (K_d = 113 nM) (Table 1 and *SI Appendix*, Fig. S2). A G312C substitution would eliminate every H bond, except the single H bond between its phosphate OP1 and R540, and the substitution of the O6 carbonyl to an NH_2 would result in an electrostatic repulsion with R496 (β3). However, a G312U substitution would be able to maintain H bonding to R496 (β3), Y532 (α3), and R540 (α3), although H bonding to D485 (β2) and D539 (α3) would be abolished.

There is a kink in the phosphate backbone at U313, which is stabilized by H bonds to Y483 (β2) and K543 (α3) (Fig. 2 *D* and *F*). The U313 base is oriented perpendicular relative to G312, Y483, and G312, and the U313 O4 has an H bond to R496 (β3) (Fig. 2*D*). Neither U313A nor U313C substitutions had a deleterious effect on binding to xRRM, and surprisingly, U313A had an approximately twofold higher binding affinity compared with the wild type (K_d = 74 nM) (Table 1 and *SI Appendix*, Fig. S2). The adenine base could potentially have two H bonds to the R496 sidechain compared with the single H bond to uridine. However, in vivo, a U313A substitution resulted in an ∼50% loss of binding to hLarp7 (55), indicating that the in vivo reduction in 7SK–hLARP7 association is due to factors other than binding affinity. The G314 base stacks between Y483 (β2) and R468 (α1) and has two H bonds between its WC face and the V484 backbone, one H bond between the base O6 and R472 (α1), and an H bond between the ribose 2′OH and R468 (α1) (Fig. 2*E*). An R472A substitution, which would disrupt the single H bond to the G314 O6, did not affect binding (K_d = 91 nM) (Table 1 and *SI Appendix*, Fig. S3).

In summary, we find that the highly conserved xRRM β-sheet residues Y483 (RNP3) and R496 (β3) as well as the 7SK SL4 residue G312 make up the core of the xRRM–RNA binding interface and contribute the most significantly to RNA–protein complex formation. Helix α3 is across loop 4, and K543 inserts into loop 4 to stabilize the S turn in the RNA backbone.

The xRRM–RNA Binding Pocket Is Preorganized to Bind Loop 4. Comparison of the free (63) and bound structures of the hLarp7 xRRM shows that it does not undergo large conformational changes on binding 7SK SL4 (all-atom rmsd of 1.6 Å for residues 445–546) (Fig. 3 *A* and *B*). In the absence of RNA,

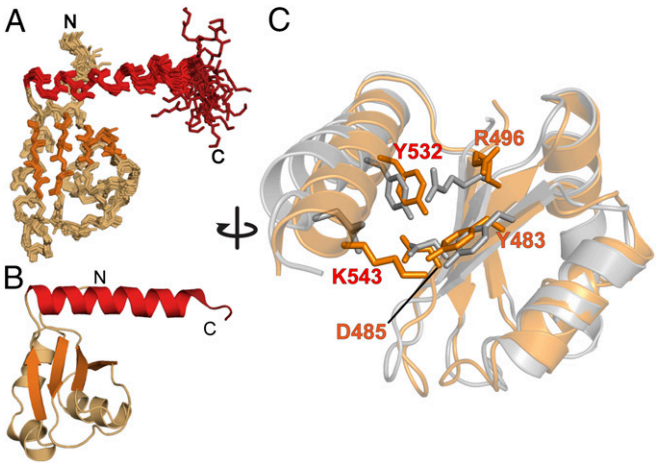

Fig. 3. hLarp7 xRRM binding pocket is preformed. (*A*) Solution NMR ensemble of the 20 lowest-energy states of the free xRRM (PDB ID code 5KNW) shows that helix α3 becomes increasingly mobile near the C terminus. Residues 545–548, shown in maroon, are unstructured. (*B*) In complex with RNA, helix α3 is on the β-sheet surface and contains an additional two residues (545–546; shown in maroon) that were disordered in the absence of RNA. (*C*) Overlay of free (gray) and bound (orange) xRRM shows similar orientations of sidechains involved in RNA recognition, particularly Y483 and D485 in the conserved RNP3 and Y532 (α3) that contact G312. Helix α3 is closer to the β-sheet in the RNA-bound xRRM structure and has minor sidechain rearrangements compared with the free xRRM.

residues 541–544 at the C-terminal end of helix α3 are increasingly disordered in the solution NMR structural ensemble, with helix α3 ending at residue 544 (Fig. 3*A*). On binding to SL4, helix α3 is stabilized and reorients ∼14°, positioning it closer to the β-sheet (Fig. 3*C*). In complex with RNA, helix α3 is extended half a turn, ending at residue 546, with additional electron density to residue 548 (Fig. 3*B*). In the RNA binding pocket, sidechains of RNP3 residues Y483 and D485 (β2) and Y532 (α3) that interact with loop 4 are in similar orientations when free or bound to RNA (Fig. 3*C*). In contrast, the sidechains of R496 (β3) and K543 (α3) move on binding SL4 (Fig. 3*C*).

Some of the largest differences between the free and bound xRRM structures occur at the N terminus. In the solution NMR structure of the free hLarp7 xRRM, N-terminal residues 446–454 are flexible (63), but in the X-ray crystal structure of the xRRM bound to SL4, these residues are structured (Fig. 1*E*). Chemical shift perturbation experiments to map the RNA–protein interface showed large chemical shift changes for N-terminal residues Q450 and F451, although they were predicted to be far from the RNA binding site, and heteronuclear NOEs showed reduced motions for the N terminus in the xRRM–RNA complex compared with free, indicating that it is stabilized on binding RNA (63). Comparison of the ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectra of the xRRM and a construct with an F451A substitution shows chemical shift differences in α3, β3, and β2, indicating that this residue may affect helix α3 positioning on the β-sheet (*SI Appendix*, Fig. S6). In the crystal structure of the RRM–RNA complex, F451 stacks with W533 (α3), which may explain why the loss of this interaction in the F451A substitution results in chemical shift changes across the β-sheet and helix α3 (*SI Appendix*, Fig. S6). Although the xRRM N terminus is flexible, F451 is positioned close to W533 in the solution NMR structural ensemble (*SI Appendix*, Fig. S6). However, the F451A substitution does not impair RNA binding (K_d = 120 nM) (Table 1 and *SI Appendix*, Fig. S3). In summary, the xRRM binding pocket is preorganized to bind 7SK SL4. On RNA recognition, helix α3 is extended and stabilized, and the N terminus becomes structured.

Loop 4 Is Flexible and Has Large Conformational Changes on Binding xRRM. Comparison of the structure of SL4 in our hLarp7 xRRM–SL4 complex with a previously reported solution NMR structure of SL4 (71) shows that there are significant differences in the conformation of loop 4 (Fig. 4). We note that the NMR structure of SL4 was solved in the presence of arginine, which binds to the C320, U321 (CU) bulge but does not contact or affect the structure of loop 4. While in both structures, loop residues A310 and U311 stack on the 7SK SL4 stem and G312, U313, and G314 are flipped out (71) (Fig. 4*A*), the backbone conformations of G312, U313, and G314 are strikingly different (Fig. 4 *A* and *B*). In the protein-free loop 4, the backbone turns about the G312 phosphate, while in the xRRM–RNA complex, the loop 4 backbone is more compact and makes an S turn (79, 80) about U313 (Fig. 4 *A*–*C*). This turn is stabilized by four H bonds: U313 phosphate OP1 and Y483 (β2), U313 phosphate OP2 and K543 (α3), G314 OP2 and K543 (α3), and U311 ribose 2′OH and K543 (α3) (Fig. 2). After the turn at U313, the backbone turns again at G314. G312 and G314 bases are flipped out in the *anti* conformation and interact extensively with the xRRM. However, in the NMR structure of the free loop 4, G312 and G314 bases are in the *syn* conformation (Fig. 4*B*), suggesting that the bases of these residues rotate about the glycosidic bond to bind xRRM. In addition to distortions in the loop 4 backbone, the upper three base pairs in the stem have a widened major groove compared with SL4 in the absence of xRRM (Fig. 4*C*), likely due to the insertion of helix α3.

To gain insights into the dynamic properties of the 7SK SL4 at pico- to nanosecond timescales, we measured longitudinal (R_1) and transverse (R_2) carbon relaxation data for nucleobase C2, C5, C6, C8, and ribose C1′ moieties using 2D ^{13}C spin relaxation R_1 and $R_1\rho$ NMR experiments (*SI Appendix*, Table S3) (81). The measured R_1 and R_2 values were used to calculate a relative order parameter (S^2_{rel}) for each residue, an internally normalized parameter that describes the relative degree of order within a molecule with values ranging from zero to one representing minimum and maximum order, respectively. Residues in the SL4 stem have the expected S^2_{rel} values of approximately one except for the 2-nt bulge in the middle of the stem, which has reduced S^2_{rel} values for the CU bulge but S^2_{rel} values greater than one for the ribose of base pairs adjacent to the bulge, indicating chemical exchange contributions to R_2 (Fig. 4*D*) (82). We note that, in the solution NMR structure of SL4 with arginine, C320 forms a base triple with the G303-C323 base pair (71). The base triple is likely stabilized by the arginine, which stacks on C320, and this would explain the chemical exchange observed here for the SL4 riboses adjacent to C320 and U321 in the absence of arginine.

In loop 4, residue A310 has a similar degree of order to paired residues (S^2_{rel} ∼ 1) in agreement with the solution NMR structure where it stacks on the stem (Fig. 4 *B* and *D*). U311 has reduced base and ribose S^2_{rel} values, indicating that it is mobile in solution. However, in the solution NMR structure, U311 stacks on A310 (Fig. 4*B*); in the absence of the xRRM, U311 may fluctuate between stacked and unstacked conformations. G312,

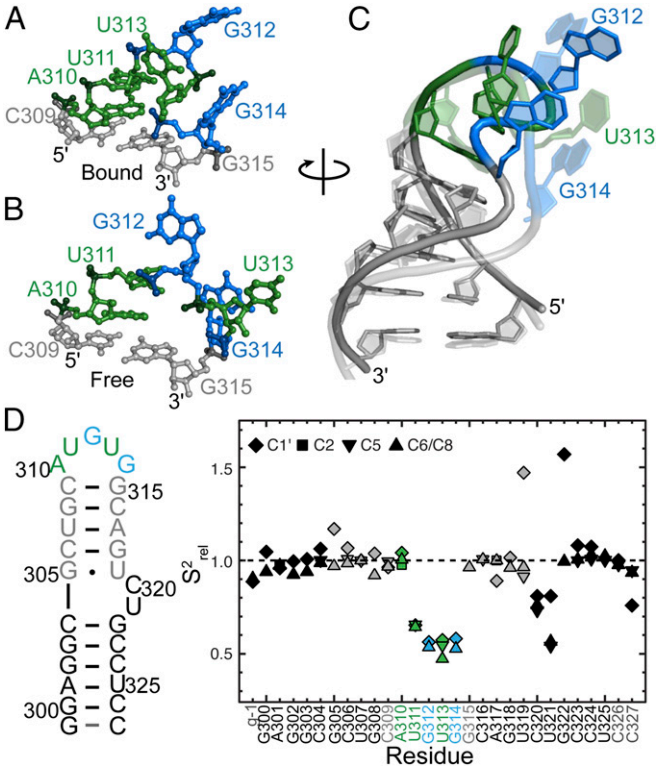

Fig. 4. 7SK loop 4 changes conformation on binding hLarp7 xRRM. Loop 4 residues of 7SK SL4 (*A*) in the crystal structure bound to hLarp7 and (*B*) in the solution NMR structure (PDB ID code 2KX8) not bound to protein shown in stick view. RNA is colored as shown in Fig. 1*D*. (*C*) Cartoon representation of free (2KX8, transparent) and bound (solid) 7SK SL4 upper stem and loop 4. (*D*) ^{13}C spin relaxation shows that loop 4 and 2-nt bulge are flexible in the absence of protein. (*Left*) Sequence of 7SK SL4 construct. Loop 4 residues are shown in color, with the upper stem shown in gray and CU bulge and lower stem residues shown in black. (*Right*) Relative order parameters (S^2_{rel}) of nucleobase (C2, C5, C6, C8) and ribose (C1′) moieties from ^{13}C spin relaxation R_1 and $R_{1\rho}$ experiments.

U313, and G314 also have reduced base and ribose S^2_{rel} values, consistent with these residues being flipped out in the solution NMR structure. The ^{13}C spin relaxation data show that loop 4 residues U311, G312, U313, and G314 are flexible in solution. These residues form the most extensive contacts with the xRRM, and their flexibility may assist in hLarp7 recognition, possibly through a conformational selection mechanism.

The xRRM Is a Conserved Motif in the LARP7 Family. The xRRM was first identified in *Tetrahymena* p65, with sequence and predicted structural homology suggesting that other members of the LARP7 family may contain an xRRM (66, 67). Compared with a canonical RRM, the xRRM has an additional helix α3 that is across the surface of the β-sheet, where RNA binds to canonical RRMs; lacks conserved RNP1 and RNP2 sequences on β3 and β1, respectively; and has conserved [(F/Y/W)–X–(D/Q/E/N)] residues (RNP3 motif) on β2 and Arg on β3. In complex with RNA, a binding pocket between β2 (including the RNP3 motif) and helix α3 interacts with unpaired nucleotides and residues at the C-terminal end of helix α3, which extend beyond the β-sheet, become ordered (helix α3x), and interact with RNA. The structure of the hLarp7 C-terminal RRM in the absence of RNA and binding determinants to 7SK SL4 RNA showed that it was an xRRM with features expected for an xRRM (63). Details of the protein–RNA interactions observed in the complex support the proposal that a C-terminal xRRM, first observed for p65, is a general feature of LARP7 family proteins. In particular, the RNP3 motif and helix α3-α3x interact with the RNA as predicted, although we note that the extension of α3 is shorter than for p65.

hLarp7 and p65 xRRMs have nearly identical binding pockets (Fig. 5 *A*–*D*). A single-stranded G residue (G312 in 7SK, G121 in TER) fits within a cleft between a conserved RNP3 Tyr on the β-sheet and Ile on helix α3 (Fig. 5 *A*–*D*). A conserved RNP3 Asp and helix α3 Tyr H bond to the WC face of this G residue, and a conserved Arg in β3 H bonds to its HG face. Both hLarp7 and p65 xRRMs interact with the single-stranded RNA residue after this conserved G (U313 in 7SK, A122 in TER), which is perpendicular to the G (Fig. 5 *A*–*D*). A residue on helix α3 (K543 in hLarp7, F521 in p65) encloses the essential G residue in the xRRM binding pocket and packs alongside the residue preceding the single-stranded G residue (U311 in 7SK, C120 in TER) to stabilize the sharp turn in the RNA backbone (Fig. 5 *A* and *B*). The hLarp7 xRRM binds an additional G residue (G314) that fits within a cleft between the β-sheet and helix α1 (Fig. 5 *A* and *C*).

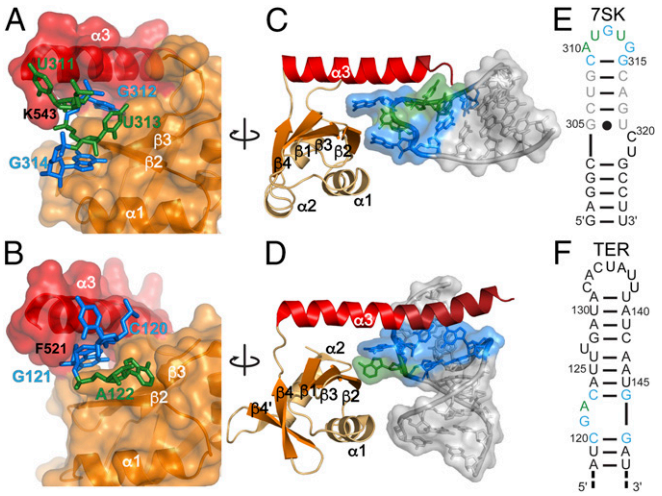

Fig. 5. LARP7 family proteins contain a conserved xRRM that binds diverse RNA secondary structure motifs. (*A*) The hLarp7 xRRM binds loop 4 UGUG nucleotides, and (*B*) p65 xRRM (PDB ID code 4ERD) binds CGA nucleotides. (*C*) The hLarp7 xRRM binds loop 4, and helix α3 inserts along the RNA major groove. (*D*) The p65 xRRM binds the 2-nt bulge, and helix α3-α3x inserts across the RNA major groove. (*E*) Human 7SK SL4 sequence, with loop 4 and the terminal base pair colored as in Fig. 1. (*F*) *Tetrahymena* TER S4 sequence, with the 2-nt bulge and adjacent base pairs colored as in Fig. 1.

Despite the similarities in recognition of single-stranded RNA nucleotides, hLarp7 and p65 xRRMs recognize highly divergent RNA secondary structure motifs (Fig. 5 *C*–*F*). The hLarp7 xRRM binds at the apical loop of an SL, while p65 binds at a 2-nt bulge (Fig. 5 *C*–*F*). For both xRRMs, helix α3 inserts in the major groove of the RNA, although the hLarp7 helix α3 is along the 5′ strand and down the helical axis, while the p65 helix α3 is across the helical axis (Fig. 5 *C* and *D*). hLarp7 and p65 have a high degree of sequence homology in the first 12 residues of helix α3, which are across the β-sheet and interact with the conserved G residue. However, the C-terminal residues of helix α3, including α3x, differ significantly in both sequence and stability. The hLarp7 helix α3 has 5.5 turns and is increasingly mobile approaching the C terminus, and it is extended by half a turn (α3x) and stabilized on binding RNA (Fig. 3 *A* and *B*) (63). In contrast, the p65 helix α3 has 4.5 turns in the absence of RNA and is extended by four turns (α3x) on binding RNA (Fig. 5*D*). In the hLarp7 xRRM, helix α3 ends at the helical turn of the 7SK SL4 and is likely to be sterically restricted from having a long α3x like p65. Helix α3-α3x in the p65 xRRM is enriched in aromatic residues that insert in between the G-C base pair on either side of the GA bulge and roll open the RNA (*SI Appendix*, Figs. S7 and S8). In contrast, helix α3-α3x in the hLarp7 xRRM is enriched in polar and basic residues and contains no aromatic residues. Differences in the sequence composition of the C-terminal end of helix α3 may confer a recognition preference for RNA secondary structure motif.

Other LARP7 proteins are predicted to contain xRRMs, although the precise mode of RNA recognition remains to be determined. In *Euplotes*, p43 binds to a two-way junction in stem 4 of TER and is proposed to have an xRRM. However, the C-terminal end of helix α3 of *Euplotes* p43 resembles the hLarp7 xRRM more than the p65 xRRM and has charged residues at the putative RNA binding interface rather than aromatic residues (*SI Appendix*, Figs. S7 and S8). While both *Tetrahymena* p65 and *Euplotes* p43 bind to stem 4 of TER, the stem 4 two-way junction differs significantly between these two ciliate species. Tetrahymena TER stem 4 has a 2-nt bulge (sequence GA), while *Euplotes* TER stem 4 has a 10-nt bulge (sequence UGAAAACCCC) and a second 9-nt bulge (83). The G and first A residue in the 10-nt bulge are protected from RNase cleavage when bound to p43 (84), in agreement with RNA residue requirement for xRRM recognition. In *Tetrahymena*, binding of the p65 xRRM to stem 4 bends the RNA to position loop 4 to contact the telomerase reverse transcriptase (TERT) protein, promoting hierarchal assembly of the telomerase RNP (66). In *Euplotes*, the significantly larger bulge likely has a greater degree of intrinsic flexibility and can sample a larger conformational space. The p43 helix α3 likely does not require aromatic residues to insert in the bulge but rather, can stabilize an existing bent conformation, positioning the TER SL4 for favorable contacts with TERT. Another LARP7 family protein, Pof8, was recently identified in fission yeast and shown to bind TER and promote RNP assembly, although its binding site has not been determined. Sequence alignment of the Pof8 CTD with other LARP7 proteins suggests that it contains an xRRM, although Pof8 has a short helix α3 compared with p65 and hLarp7 (*SI Appendix*, Figs. S7 and S8).

Despite the above differences in cognate RNA sequence and secondary structures, for both hLarp7 and p65, the xRRM is essential for RNP assembly and function. In metazoa, the xRRM in hLarp7 is required for specific binding to 7SK RNA as well as assembly with P-TEFb (25). In ciliates, the xRRM in p65 and the putative xRRM in p43 are required for assembly of TER with

TERT (12, 85, 86). In fission yeast, the putative xRRM promotes assembly of telomerase (13–15). Overall, the xRRM seems to be a required domain in the LARP7 family that promotes RNP assembly.

A Model for hLarp7 and MePCE Assembly on 7SK. hLarp7 is constitutively assembled on 7SK and is essential for 7SK stability, with a direct correlation between hLarp7 expression levels and 7SK RNA levels in vivo, and it binds 7SK as its only target (16, 17). To model hLarp7 in complex with 7SK, we docked together and refined available structures of 7SK SL4 [nucleotides 300–331; Protein Data Bank (PDB) ID code 2KX8], the hLarp7 La module bound to UUU (PDB ID code 4WKR), and hLarp7 xRRM bound to loop 4 (Fig. 6*A*) using HADDOCK (87, 88). In the 20 lowest-energy models, the xRRM stably associates with SL4, and R548 remains in contact with the SL4 stem (*SI Appendix*, Fig. S9). Steric hindrance by the RNA upper stem restricted the orientation of the hLarp7 CTD to point away from SL4 (Fig. 6*A* and *SI Appendix*, Fig. S9).

In contrast to the xRRM, the La module docked on the flexible UUU-3′OH with a wide range of conformations largely due to conformational sampling of the UUU-3′OH. The 20 lowest-energy models can be clustered into three conformations, where the orientation of the La module differs with respect to SL4 (*SI Appendix*, Fig. S9). The majority of the lowest-energy conformations, including the lowest-energy model, cluster into conformation 1, in which the RRM1 helix α1 is in the major groove of SL4 near the 2-nt bulge (*SI Appendix*, Fig. S9). The bulge forms a C320-G303-C323 base triple when bound to arginine (71), and R143 in helix α1 is positioned near C320 in the lowest-energy model. In conformations 2 and 3, the La module is positioned distal to SL4, although conformation 2 places the RRM1 β2-β3 loop near the SL4 2-nt bulge (*SI Appendix*, Fig. S9). Conformation 3 is similar to a model proposed previously based on footprinting experiments (62), in which the La motif is oriented underneath the back side of the xRRM and the RRM1 β2-β3 loop is underneath SL4 (*SI Appendix*, Fig. S9). In this model, both the RRM1 α1 and β2-β3 loop are available for interactions with other RNAs. The RRM1 β2-β3 loop has been proposed to interact with the 7SK 5′ end (62); however, MePCE binds the 5′ end and is a required component of the 7SK RNP, suggesting a different RNA binding site for the RRM1 β2-β3 loop in conformation 3.

The hLarp7 31-aa CTD, which has been proposed to bind and inactivate MePCE capping activity through its methylphosphate capping enzyme inactivating domain (MID) (52), contains an α-helix (63). The MID, SL4 2-nt bulge, and La module are all required to inactivate MePCE capping (52). Consistent with these data, in the lowest-energy structural models from HADDOCK, these three elements are all on one side of the complex, suggesting a putative MePCE binding surface that would be both RNA and hLarp7 dependent (Fig. 6*A*).

To determine the contributions of individual RNA binding domains in hLarp7 to binding 7SK, we used ITC to measure the binding of the La module, xRRM, and full-length hLarp7 to an SL4 construct including the 3′ end UUU-3′OH (named SL4u; residues 299–331) (Fig. 1*A*, Table 1, and *SI Appendix*, Fig. S10). The La module alone (residues 1–190) binds to SL4u with a K_d of 230 nM, similar to the binding affinity of the La module of human genuine La to UUU-3′OH measured by ITC (9), although higher affinities have been reported measured by EMSA using pretRNA or a 9-nt substrate (89, 90). Full-length hLarp7 bound to SL4u with a K_d of 6.6 nM: 19-fold lower than the xRRM and 35-fold lower than the La module (Table 1). Binding of hLarp7 to 7SK can be approximated as a bivalent model, where the La module and xRRM independently bind to the 7SK substrate. Assuming this model, it is apparent that there is an additive effect for the overall RNA binding affinity by linking the La module and xRRM covalently with a long linker. However, the Gibbs free energy barrier of full-length hLarp7 binding to the 7SK 3′ end vs. the individual binding events can be calculated using the equation $\Delta g = \Delta G_{FL} - (\Delta G_{La} + \Delta G_{xRRM})$, where Δg reflects the collective enthalpy (Δh) and entropy (Δs) penalties as obtained by applying the Gibbs–Helmholtz relationship and ΔG_{FL}, ΔG_{La}, and ΔG_{xRRM} are the free energies of 7SK SL4u binding hLarp7, La module, and xRRM, respectively (*SI Appendix*, Table S6) (91). In a bivalent model, an important factor that contributes to Δg is the translational and rotational entropy loss by covalently linking two binding domains (92). For hLarp7, the Δh is −7 kcal/mol [$\Delta h = \Delta H_{FL} - (\Delta H_{La} + \Delta H_{xRRM})$], indicating a favorable enthalpy for full-length hLarp7 binding to SL4u. However, the Δg is +7.4 kcal/mol, indicating a significant entropy penalty. One explanation could be that the linker may become partially structured on the N-terminal La module and C-terminal xRRM by binding to the 7SK 3′ end, although it is predicted to have low structural complexity in the free protein (62). The linker contains several basic regions and may bind RNA, consistent with the favorable enthalpy observed on binding full-length hLarp7.

In addition to binding the 7SK 3′ end, hLarp7 interacts with MePCE to assemble a stress-resistant ternary complex (52, 53, 93). hLarp7 does not bind MePCE in the absence of RNA (53, 55, 93). Taking these data and the literature into account (16, 52, 53, 55, 93, 94), we can propose a model for 7SK ternary complex formation (Fig. 6*B*). MePCE, which is concentrated at the 7SK promoter (53), binds and caps the nascent 7SK 5′ transcript. Genuine La protein is associated with a small population of 7SK

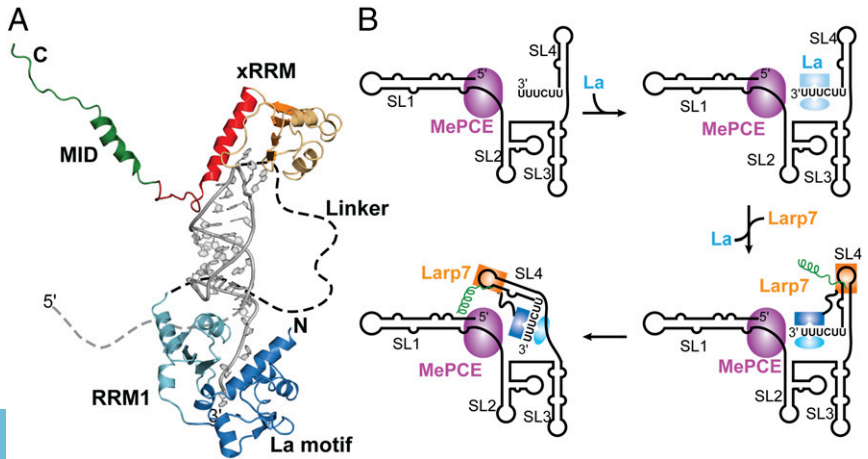

Fig. 6. Model of hLarp7 recognition of the 7SK 3′ end and mechanism of assembly of core 7SK RNP. (*A*) Cartoon representation of the lowest-energy HADDOCK model of the hLarp7 La module (PDB ID code 4WKR), xRRM, and CTD bound to the 7SK SL4u modified from the solution NMR structure of SL4 (PDB ID code 2KX8). RNA is colored gray, La motif is colored blue, RRM1 is colored cyan, xRRM is colored as shown in Fig. 1, and CTD containing the MID is colored green. The hLarp7 interdomain linker is shown as a black dashed line, and the 7SK 5′ linker is shown as a gray dashed line. (*B*) Proposed mechanism of 7SK ternary core RNP formation. MePCE (purple) binds to the nascent 7SK transcript and caps the 5′ end. Genuine La, colored blue, binds 7SK and is replaced by hLarp7 (colored the same as in *A*). On binding 7SK, hLarp7 is able to assemble with MePCE to form the core 7SK RNP.

in immunoprecipitation experiments, and it is proposed to bind nascent 7SK transcripts and subsequently, be replaced by Larp7 (16, 94). Binding of hLarp7 to 7SK may bring the MePCE binding elements in close proximity to each other, forming an interface where hLarp7 can stably bind MePCE and assemble into a stable core ternary complex. This model is consistent with the observation that hLarp7 bound to 7SK can bind to MePCE, even when 7SK lacks the MePCE binding site, but MePCE cannot recruit hLarp7 to 7SK when 7SK lacks the hLarp7 binding site (55).

In summary, the structure of the hLarp7 xRRM bound to 7SK SL4 provides seminal insights into the unique recognition of 7SK RNA by Larp7 and provides a working model for how hLarp7 interacts with MePCE to promote assembly of P-TEFb on the 7SK RNP.

Materials and Methods

Sample Preparation for Crystallography. hLarp7 xRRM crystal constructs (Gene ID 51574; residues 445–556) were prepared as described previously (63). A modified construct was used with a tobacco etch virus (tev) protease cleavage site (ENLYFQS) between the His_6 tag and xRRM start site. Two point substitutions (E501L and Q504L) were made to helix α2 to reduce surface entropy and promote crystal contacts (75). Briefly, *Escherichia coli* BL21 (DE3) cells were grown at 37 °C in M9 minimal media to an OD of 0.6–0.8, transferred to 18 °C, and grown for 18–20 h. Cells were pelleted, resuspended with buffer R (20 mM Tris, pH 8, 750 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP), 0.002% sodium azide, 5% glycerol) and lysozyme, and sonicated. Cell debris was separated from lysate by centrifugation, and protein was purified by HisTrap HP nickel affinity column (GE Healthsciences). Eluted protein was dialyzed (20 mM Tris, pH 7, 50 mM NaCl, 1 mM TCEP) with 1 mg tev protease at room temperature for 1–2 h. After tev cleavage, the protein was run again on the nickel affinity column to separate His_6-tagged from cleaved protein. Cleaved protein in the flow through was concentrated and further purified on an S75 column attached to an AKTA FPLC (GE Healthsciences). Fractions containing pure protein were pooled, buffer exchanged into xRRM–SL4 complex buffer (buffer C: 20 mM Hepes, pH 7, 50 mM KCl, 1 mM TCEP) using a 3-kDa molecular mass cutoff Amicon (EMD Millipore), and concentrated to 0.5–1 mM.

A construct with two additional mutations (L475M, L521M) was used to prepare selenomethionine-labeled samples for phase determination by multiple-wavelength anomalous dispersion (MAD). These mutations did not affect RNA binding (*SI Appendix*, Fig. S1). *E. coli* BL21 (DE3) cells were grown at 37 °C in M9 minimal media to an OD of 0.3. An amino acid mix containing 100 mg lysine, 100 mg phenylalanine, 100 mg threonine, 100 mg selenomethionine, 50 mg isoleucine, 50 mg valine, and 50 mg leucine was added to the bacterial culture, and cells were grown at 37 °C to an OD of 0.9. The amino acid mix was added again along with Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. The culture was transferred to 18 °C and grown for 18–20 h. Protein was purified as described above. The presence of selenomethionine was confirmed by comparing the masses of purified native and selenomethionine-labeled protein using MALDI-TOF mass spectrometry.

The 7SK (Gene ID 125050) SL4 RNAs used for crystallization trials were in vitro transcribed using T7 RNAP (P266L mutant) (95) and chemically synthesized DNA templates (Integrated DNA Technologies) following established protocols (63). Briefly, the transcription reaction (40 mM Tris, pH 8; 25 mM $MgCl_2$; 1 mM spermidine; 0.01% Triton X; 2.5 mM DTT; 2 mM each rATP, rCTP, rUTP, and rGTP; 0.5 μM DNA template) was incubated at 37 °C for 4–6 h. Transcribed RNA was purified by 15–20% denaturing PAGE, the band containing RNA was visualized by UV shadowing and excised from the gel, and RNA was electroeluted into 1× TBE (Tris, boric acid, EDTA) from the gel pieces using an Elutrap device (GE Waters). The RNA was buffer exchanged into ddH_2O, supplied with counterions by addition of high-salt buffer (1.5 M KCl), and briefly buffer exchanged into ddH_2O before being annealed by heating to 95 °C for 3 min followed by incubation on ice for 1 h. RNA was then buffer exchanged into buffer C using a 3-kDa molecular mass cutoff Amicon concentrator device and concentrated to 0.4–1 mM. The RNA construct that crystallized is given in Fig. 1 and includes SL4 residues 305–319.

Sample Preparation for NMR and ITC Experiments. The 7SK SL4 RNA hairpin (called SL4, residues 299–327), modified by substituting the terminal G-U base pair with a G-C base pair and adding a terminal G-C base pair (Fig. 5), was in vitro transcribed and purified as described above. For $^{13}C/^{15}N$-labeled RNA samples, uniformly labeled $^{13}C/^{15}N$ rNTPs (Cambridge Isotope Laboratories) were used in the transcription reaction. After electroelution, RNA was further purified by ion exchange chromatography using a DEAE column (GE Healthcare) and eluted into 10 mM sodium phosphate, pH 7.6, 1 mM EDTA, and 1.5 M KCl. RNA was diluted to <100 μM in ddH_2O and annealed by heating to 95 °C for 3 min followed by incubation on ice for 1 h. RNA was then buffer exchanged into appropriate buffer using a 3-kDa molecular mass cutoff Amicon and concentrated to 0.4–1 mM. For NMR studies, the RNA was exchanged into protein–RNA complex NMR buffer (Buffer N: 20 mM sodium phosphate, pH 6.05, 50 mM KCl, 1 mM TCEP). For ITC studies, the RNA was exchanged into ITC buffer (20 mM Tris, pH 7, 50 mM NaCl).

Full-length hLarp7 (1–582) was cloned into a pETDuet vector containing fusion proteins Maltose Binding Protein (MBP) at the N terminus with tev cleavage sites between MBP and hLarp7. Protein was expressed and purified as described above with the following exception: after tev cleavage and dialysis, hLarp7 was run on a 5-mL HiTrap Heparin HP column (GE) to separate MBP and GFP fusion proteins from hLarp7. As a final purification step, hLarp7 was run on an S200 column attached to an AKTA FPLC (GE Healthsciences). Fractions containing pure protein were pooled, buffer exchanged into ITC buffer, and concentrated to 100–300 μM using a 10-kDa molecular mass cutoff Amicon device. The hLarp7 La module (1–190) was cloned into a pet30 Xa/LIC vector with an His_6 tag and modified to remove the intervening sequence between the tag and La module. Protein was expressed as described for full-length hLarp7 and purified by lysis, sonication, pelleting, and nickel affinity column as described above. Protein was further purified by heparin column, run on an S75 column with fractions containing pure protein pooled, buffer exchanged into ITC buffer, and concentrated to 100 μM using a 3-kDa molecular mass cutoff Amicon device.

Crystallization. EMSA was used to determine the optimal stoichiometry of xRRM to RNA for sample preparation (*SI Appendix*, Fig. S1A). Trays that contained a 2.7:1 ratio of xRRM to SL4 crystal construct gave the best diffracting crystals. Samples were prepared by adding the xRRM and RNA under dilute conditions into buffer C and concentrating to 8 mg/mL using a 10-kDa molecular mass cutoff Amicon concentrator device. The crystallization was performed using the hanging drop vapor diffusion method by mixing xRRM–RNA complex with crystallization buffer (20 mM sodium cacodylate, 80 mM magnesium acetate, 24% PEG 4000) in a 1.3:0.7 ratio of complex to buffer. A reservoir solution of 0.7 M sodium chloride solution yielded the best diffracting crystals. Rod-shaped crystals grew to 50–100 μm in 1–2 d at ambient temperature (*SI Appendix*, Fig. S1).

Data Collection and Crystal Structure Calculation. xRRM–SL4 complex datasets (*SI Appendix*, Table S1) were collected at the Advanced Photon Source at Argonne National Laboratory on beamline 24-ID-C (NE-CAT). Datasets were collected at three different wavelengths for peak, inflection, and high remote. The crystal decayed rapidly, and only the peak and inflection data were used in molecular phasing. These data were processed with Rapid Automated Processing of Data (necat.chem.cornell.edu/). Macromolecular phasing was performed using HKL2MAP (96) using MAD. The resulting phases were improved using SHARP, and a preliminary protein model was built using autoSHARP (97). At this point, we refined the partial model against a dataset merged from two crystals, both collected at the peak wavelength. These data were indexed, integrated, and scaled using XDS (98). This dataset was anisotropic, with diffraction limits of 2.4 Å along the a* and c* directions but 2.2 Å along the b* direction. For this reason, an anisotropic scale factor was applied along with −10 $Å^2$ B-factor correction using the University of California, Los Angeles–US Department of Energy Lab Diffraction Anisotropy Server (99). RNA was manually modeled using Coot (100). There are two molecules in the asymmetric unit. Final iterative rounds of model building and refinement were performed using Coot and PHENIX with TLS refinement (101), noncrystallographic symmetry restraints, simulated annealing, and RNA H-bonding restraints. Final data collection, phasing, and refinement statistics are presented in *SI Appendix*, Tables S1 and S2.

NMR Experiments. NMR experiments were performed at 298 K on AVANCE 600- and 800-MHz Bruker spectrometers equipped with an HCN cryoprobe. NMR data were collected using Topspin (Bruker), processed using NMRPipe (102), and analyzed using NMRView or Sparky 3.110 (103). Additional NMR methods can be found in *SI Appendix*. ^{13}C spin relaxation experiments (81) were performed at 298 K at 600 MHz using uniformly $^{13}C/^{15}N$-labeled SL4 RNA. The following delays, spinlock powers, and offsets were used: C1′ {R_1 (20 ms, 600 ms × 2), $R_1\rho$ (4 ms, 60 ms, 80 ms × 2), spinlock power 2,244 Hz, offset 2,000 Hz}, C5 {R_1 (20 ms, 540 ms × 2), $R_1\rho$ (4 ms, 48 ms, 60 ms × 2), spinlock power 3,484 Hz, offset −1,750 Hz}, C2 C6 C8 {R_1 (20 ms, 540 ms × 2), $R_1\rho$ (4 ms, 48 ms, 60 ms × 2), spinlock power 3,484 Hz, offset 1,700 Hz}. R_1

and R_2 values were computed using in-house software and are reported in *SI Appendix*, Table S3. Relative order parameters were calculated by normalizing ($2R_2$–R_1) to either U307 (C5, C6, C1′) or A317 (C8, C2). Additional NMR methods can be found in *SI Appendix*.

ITC. The binding dissociation coefficient (K_d) for binding of hLarp7 xRRM constructs to 7SK SL4 constructs was determined using a MicroCal 200 ITC instrument (GE). RNA and protein were individually exchanged into ITC buffer. Protein at concentrations of 100–250 μM was titrated into 5–10 μM RNA at 295 K. Calorimetric data were fit using ORIGIN 7 (MicroCal). The binding parameters stoichiometry (N), entropy (ΔS), enthalpy (ΔH), and association constant (K) were kept as floating variables during each fit. Experiments were performed in duplicate or triplicate, with each experiment fit individually and binding parameters averaged.

HADDOCK Modeling. The model of the SL4u–La module–xRRM complex was generated using HADDOCK 2.2 (87, 88). Each calculation generated 1,000/200/200 models for the rigid body docking, semi- and fully flexible simulated annealing, and explicit solvent refinement, respectively. Input structures of the three components were protein coordinates from (*i*) the crystal structure of the La module in complex with 5′-CUUUU-3′ nucleotides (PDB ID code 4WKR; residues 29–189), (*ii*) the solution NMR structure of SL4 bound to arginine (PDB ID code 2KX8; nucleotides 300–331), and (*iii*) protein and RNA coordinates from the crystal structure of the xRRM (residues 445–548) in complex with SL4 (nucleotides 305–319) in this work. First, a model for the 7SK RNA SL4u (nucleotides 300–331) including the UUU-3′OH tail was created in PyMol by superimposing the helical stems between the RNA coordinates of crystal structure in this work and the NMR structure of SL4. After superimposition, coordinates were combined into one molecule. Second, a model of the hLarp7 C-terminal end including the xRRM and CTD (residues 445–582) was created in Coot, where residues 560–570 were modeled as an α-helix as previously determined (63). A model of this construct bound to SL4u (nucleotides 300–331) was generated in HADDOCK using unambiguous distance restraints determined from the complex crystal structure in this work and additional RNA base pairing H-bond restraints (*SI Appendix*, Table S4). 7SK nucleotides 327–331 and hLarp7 residues 545–582 were defined as fully flexible to sample spatial conformations and apply dihedral angle restraints, respectively. The 20 lowest-energy models of the SL4u–xRRM complex were used as input structures to model the ternary SL4u–xRRM–La module complex. All of the restraints used in SL4u–xRRM modeling were maintained, and additional distance restraints between La module and SL4u were determined from the crystal structure (*SI Appendix*, Table S4). Cluster analysis results of the final 200 models with rmsd cutoff of 15 Å are given in *SI Appendix*, Table S5.

Supporting Information. Supplemental information includes additional methods, X-ray crystal structure data collection and refinement statistics, electron density maps, representative ITC plots, NOESY spectra of 7SK RNA constructs, ^{13}C spin relaxation R_1 and R_2 values, hLarp7 and LARP7 family sequence alignments, HADDOCK statistics table, and HADDOCK lowest-energy structural ensemble information and can be found in *SI Appendix*.

ACKNOWLEDGMENTS. We thank M. Collazo for help with crystal optimization and Drs. D. Cascio and M. Sawaya for help with crystal optimization, data collection, and analysis. We also thank M. Capel, J. Schuermann, K. Rajashankar, N. Sukumar, F. Murphy, and I. Kourinov of Northeastern Collaborative Access Team beamline ID-24 at the Advanced Photon Source of Argonne National Laboratory, which is supported by NIH Grants P41 RR015301 and P41 GM103403. This work was supported by American Cancer Society Postdoctoral Fellowship 126777-PF-14-179-01-DMC (to C.D.E.) and NIH Grant GM107567 (to J.F.). We acknowledge NMR Equipment NIH Grant S10OD016336 and US Department of Energy (DOE) Grant DE-FC0302ER63421 for support of NMR and X-ray core facilities. The Pilatus 6M detector on 24-ID-C beamline is funded by NIH Office of Research Infrastructure Programs High-End Instrumentation Grant S10 RR029205. This research used resources of the Advanced Photon Source, DOE Office of Science User Facility operated for the DOE Office of Science by the Argonne National Laboratory under Contract DE-AC02-06CH11357.

1. Bayfield MA, Yang R, Maraia RJ (2010) Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* 1799:365–378.
2. Maraia RJ, Mattijssen S, Cruz-Gallardo I, Conte MR (2017) The La and related RNA-binding proteins (LARPs): Structures, functions, and evolving perspectives. *Wiley Interdiscip Rev RNA* 8:e1430.
3. Query CC, Bentley RC, Keene JD (1989) A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* 57: 89–101.
4. Yoo CJ, Wolin SL (1994) La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: A yeast homolog of the La autoantigen is dispensable for growth. *Mol Cell Biol* 14:5412–5424.
5. Stefano JE (1984) Purified lupus antigen La recognizes an oligouridylate stretch common to the 3′ termini of RNA polymerase III transcripts. *Cell* 36:145–154.
6. Pruijn GJ, Slobbe RL, van Venrooij WJ (1991) Analysis of protein–RNA interactions within Ro ribonucleoprotein complexes. *Nucleic Acids Res* 19:5173–5180.
7. Teplova M, et al. (2006) Structural basis for recognition and sequestration of UUU(OH) 3′ temini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* 21:75–85.
8. Jacks A, 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.
9. Martino L, et al. (2012) Analysis of the interaction with the hepatitis C virus mRNA reveals an alternative mode of RNA recognition by the human La protein. *Nucleic Acids Res* 40:1381–1394.
10. Naeeni AR, Conte MR, Bayfield MA (2012) RNA chaperone activity of human La protein is mediated by variant RNA recognition motif. *J Biol Chem* 287:5472–5482.
11. Aigner S, et al. (2000) Euplotes telomerase contains an La motif protein produced by apparent translational frameshifting. *EMBO J* 19:6230–6239.
12. Witkin KL, Collins K (2004) Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes Dev* 18:1107–1118.
13. Collopy LC, et al. (2018) LARP7 family proteins have conserved function in telomerase assembly. *Nat Commun* 9:557.
14. Mennie AK, Moser BA, Nakamura TM (2018) LARP7-like protein Pof8 regulates telomerase assembly and poly(A)+TERRA expression in fission yeast. *Nat Commun* 9: 586.
15. Páez-Moscoso DJ, et al. (2018) Pof8 is a La-related protein and a constitutive component of telomerase in fission yeast. *Nat Commun* 9:587.
16. Krueger BJ, 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.
17. Markert A, 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.
18. Nguyen VT, Kiss T, Michels AA, Bensaude O (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* 414:322–325.
19. Yang Z, Zhu Q, Luo K, Zhou Q (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* 414:317–322.
20. Kohoutek J (2009) P-TEFb- the final frontier. *Cell Div* 4:19.
21. Lenasi T, Barboric M (2010) P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. *RNA Biol* 7:145–150.
22. Michels AA, et al. (2004) Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J* 23:2608–2619.
23. Yik JH, et al. (2003) Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol Cell* 12: 971–982.
24. Egloff S, Van Herreweghe E, Kiss T (2006) Regulation of polymerase II transcription by 7SK snRNA: Two distinct RNA elements direct P-TEFb and HEXIM1 binding. *Mol Cell Biol* 26:630–642.
25. He N, et al. (2008) A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* 29:588–599.
26. Cherrier T, et al. (2013) CTIP2 is a negative regulator of P-TEFb. *Proc Natl Acad Sci USA* 110:12655–12660.
27. Hogg JR, Collins K (2007) RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. *RNA* 13:868–880.
28. Eilebrecht S, et al. (2014) HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic Acids Res* 42:4962–4971.
29. Eilebrecht S, Wilhelm E, Benecke BJ, Bell B, Benecke AG (2013) HMGA1 directly interacts with TAR to modulate basal and Tat-dependent HIV transcription. *RNA Biol* 10:436–444.
30. Eilebrecht S, Benecke BJ, Benecke A (2011) 7SK snRNA-mediated, gene-specific cooperativity of HMGA1 and P-TEFb. *RNA Biol* 8:1084–1093.
31. Eilebrecht S, Bécavin C, Léger H, Benecke BJ, Benecke A (2011) HMGA1-dependent and independent 7SK RNA gene regulatory activity. *RNA Biol* 8:143–157.
32. Eilebrecht S, et al. (2011) 7SK small nuclear RNA directly affects HMGA1 function in transcription regulation. *Nucleic Acids Res* 39:2057–2072.
33. McNamara RP, et al. (2016) KAP1 recruitment of the 7SK snRNP complex to promoters enables transcription elongation by RNA polymerase II. *Mol Cell* 61:39–53.
34. Gudipaty SA, McNamara RP, Morton EL, D'Orso I (2015) PPM1G binds 7SK RNA and Hexim1 to block P-TEFb assembly into the 7SK snRNP and sustain transcription elongation. *Mol Cell Biol* 35:3810–3828.
35. Calo E, et al. (2015) RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* 518:249–253.
36. Sano M, et al. (2002) Activation and function of cyclin T-Cdk9 (positive transcription elongation factor-b) in cardiac muscle-cell hypertrophy. *Nat Med* 8:1310–1317.
37. Shilatifard A, Lane WS, Jackson KW, Conaway RC, Conaway JW (1996) An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* 271: 1873–1876.
38. Lin C, et al. (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol Cell* 37:429–437.

39. Chen R, Keating MJ, Gandhi V, Plunkett W (2005) Transcription inhibition by flavopiridol: Mechanism of chronic lymphocytic leukemia cell death. *Blood* 106: 2513–2519.
40. Bellan C, et al. (2004) CDK9/CYCLIN T1 expression during normal lymphoid differentiation and malignant transformation. *J Pathol* 203:946–952.
41. Bieniasz PD, Grdina TA, Bogerd HP, Cullen BR (1999) Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. *Proc Natl Acad Sci USA* 96: 7791–7796.
42. Peterlin BM, Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 23:297–305.
43. Barboric M, Lenasi T (2010) Kick-sTARting HIV-1 transcription elongation by 7SK snRNP deporTATion. *Nat Struct Mol Biol* 17:928–930.
44. D'Orso I, et al. (2012) Transition step during assembly of HIV Tat:P-TEFb transcription complexes and transfer to TAR RNA. *Mol Cell Biol* 32:4780–4793.
45. He N, Zhou Q (2011) New insights into the control of HIV-1 transcription: When tat meets the 7SK snRNP and super elongation complex (SEC). *J Neuroimmune Pharmacol* 6:260–268.
46. Lu H, et al. (2014) AFF1 is a ubiquitous P-TEFb partner to enable Tat extraction of P-TEFb from 7SK snRNP and formation of SECs for HIV transactivation. *Proc Natl Acad Sci USA* 111:E15–E24.
47. Sobhian B, et al. (2010) HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol Cell* 38:439–451.
48. Egloff S, Studniarek C, Kiss T (2018) 7SK small nuclear RNA, a multifunctional transcriptional regulatory RNA with gene-specific features. *Transcription* 9:95–101.
49. Flynn RA, et al. (2016) 7SK-BAF axis controls pervasive transcription at enhancers. *Nat Struct Mol Biol* 23:231–238.
50. Briese M, et al. (2018) hnRNP R and its main interactor, the noncoding RNA 7SK, coregulate the axonal transcriptome of motoneurons. *Proc Natl Acad Sci USA* 115: E2859–E2868.
51. Wassarman DA, Steitz JA (1991) Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Mol Cell Biol* 11: 3432–3445.
52. Brogie JE, Price DH (2017) Reconstitution of a functional 7SK snRNP. *Nucleic Acids Res* 45:6864–6880.
53. Xue Y, Yang Z, Chen R, Zhou Q (2010) A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP. *Nucleic Acids Res* 38:360–369.
54. Jeronimo C, et al. (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 27:262–274.
55. 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.
56. Ji X, Lu H, Zhou Q, Luo K (2014) LARP7 suppresses P-TEFb activity to inhibit breast cancer progression and metastasis. *eLife* 3:e02907.
57. Alazami AM, et al. (2012) Loss of function mutation in LARP7, chaperone of 7SK ncRNA, causes a syndrome of facial dysmorphism, intellectual disability, and primordial dwarfism. *Hum Mutat* 33:1429–1434.
58. Ling TT, Sorrentino S (2016) Compound heterozygous variants in the LARP7 gene as a cause of Alazami syndrome in a Caucasian female with significant failure to thrive, short stature, and developmental disability. *Am J Med Genet A* 170A:217–219.
59. Hollink IH, et al. (2016) Broadening the phenotypic spectrum of pathogenic LARP7 variants: Two cases with intellectual disability, variable growth retardation and distinct facial features. *J Hum Genet* 61:229–233, and erratum (2018) 63:539.
60. Dateki S, et al. (2018) Novel compound heterozygous variants in the *LARP7* gene in a patient with Alazami syndrome. *Hum Genome Var* 5:18014.
61. Biewenga P, et al. (2008) Gene expression in early stage cervical cancer. *Gynecol Oncol* 108:520–526.
62. Uchikawa 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.
63. Eichhorn CD, 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.
64. Cheng Y, et al. (2012) LARP7 is a potential tumor suppressor gene in gastric cancer. *Lab Invest* 92:1013–1019.
65. Mori Y, et al. (2002) Instabilotyping reveals unique mutational spectra in microsatellite-unstable gastric cancers. *Cancer Res* 62:3641–3645.
66. Singh M, et al. (2012) Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell* 47:16–26.
67. Singh M, Choi CP, Feigon J (2013) xRRM: A new class of RRM found in the telomerase La family protein p65. *RNA Biol* 10:353–359.
68. Daubner GM, Cléry A, Allain FH (2013) RRM-RNA recognition: NMR or crystallography. . .and new findings. *Curr Opin Struct Biol* 23:100–108.
69. Cléry A, Blatter M, Allain FH (2008) RNA recognition motifs: Boring? Not quite. *Curr Opin Struct Biol* 18:290–298.
70. Gurney T, Jr, Eliceiri GL (1980) Intracellular distribution of low molecular weight RNA species in HeLa cells. *J Cell Biol* 87:398–403.
71. Durney MA, D'Souza VM (2010) Preformed protein-binding motifs in 7SK snRNA: Structural and thermodynamic comparisons with retroviral TAR. *J Mol Biol* 404: 555–567.
72. Bigalke JM, Dames SA, Blankenfeldt W, Grzesiek S, Geyer M (2011) Structure and dynamics of a stabilized coiled-coil domain in the P-TEFb regulator Hexim1. *J Mol Biol* 414:639–653.
73. Martinez-Zapien D, et al. (2015) Intermolecular recognition of the non-coding RNA 7SK and HEXIM protein in perspective. *Biochimie* 117:63–71.
74. Bourbigot S, et al. (2016) Solution structure of the 5′-terminal hairpin of the 7SK small nuclear RNA. *RNA* 22:1844–1858.
75. Yamada H, et al. (2007) 'Crystal lattice engineering,' an approach to engineer protein crystal contacts by creating intermolecular symmetry: Crystallization and structure determination of a mutant human RNase 1 with a hydrophobic interface of leucines. *Protein Sci* 16:1389–1397.
76. Marz M, et al. (2009) Evolution of 7SK RNA and its protein partners in metazoa. *Mol Biol Evol* 26:2821–2830.
77. Gürsoy HC, Koper D, Benecke BJ (2000) The vertebrate 7S K RNA separates hagfish (Myxine glutinosa) and lamprey (Lampetra fluviatilis). *J Mol Evol* 50:456–464.
78. Yazbeck AM, Tout KR, Stadler PF (2018) Detailed secondary structure models of invertebrate 7SK RNAs. *RNA Biol* 15:158–164.
79. Wimberly B, Varani G, Tinoco I, Jr (1993) The conformation of loop E of eukaryotic 5S ribosomal RNA. *Biochemistry* 32:1078–1087.
80. Correll CC, et al. (1998) Crystal structure of the ribosomal RNA domain essential for binding elongation factors. *Proc Natl Acad Sci USA* 95:13436–13441.
81. Hansen AL, Al-Hashimi HM (2007) Dynamics of large elongated RNA by NMR carbon relaxation. *J Am Chem Soc* 129:16072–16082.
82. Davis DG, Perlman ME, London RE (1994) Direct measurements of the dissociation-rate constant for inhibitor-enzyme complexes via the T1 rho and T2 (CPMG) methods. *J Magn Reson B* 104:266–275.
83. Lingner J, Hendrick LL, Cech TR (1994) Telomerase RNAs of different ciliates have a common secondary structure and a permuted template. *Genes Dev* 8:1984–1998.
84. Aigner S, Postberg J, Lipps HJ, Cech TR (2003) The Euplotes La motif protein p43 has properties of a telomerase-specific subunit. *Biochemistry* 42:5736–5747.
85. Akiyama BM, Loper J, Najarro K, Stone MD (2012) The C-terminal domain of Tetrahymena thermophila telomerase holoenzyme protein p65 induces multiple structural changes in telomerase RNA. *RNA* 18:653–660.
86. Aigner S, Cech TR (2004) The Euplotes telomerase subunit p43 stimulates enzymatic activity and processivity in vitro. *RNA* 10:1108–1118.
87. Dominguez C, Boelens R, Bonvin AM (2003) HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 125: 1731–1737.
88. de Vries SJ, et al. (2007) HADDOCK versus HADDOCK: New features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* 69:726–733.
89. Ohndorf UM, Steegborn C, Knijff R, Sondermann P (2001) Contributions of the individual domains in human La protein to its RNA 3′-end binding activity. *J Biol Chem* 276:27188–27196.
90. Dong G, Chakshusmathi G, Wolin SL, Reinisch KM (2004) Structure of the La motif: A winged helix domain mediates RNA binding via a conserved aromatic patch. *EMBO J* 23:1000–1007.
91. Brown A (2009) Analysis of cooperativity by isothermal titration calorimetry. *Int J Mol Sci* 10:3457–3477.
92. Jencks WP (1981) On the attribution and additivity of binding energies. *Proc Natl Acad Sci USA* 78:4046–4050.
93. Barboric M, et al. (2009) 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc Natl Acad Sci USA* 106:7798–7803.
94. Chambers JC, Kurilla MG, Keene JD (1983) Association between the 7 S RNA and the lupus La protein varies among cell types. *J Biol Chem* 258:11438–11441.
95. Guillerez J, Lopez PJ, Proux F, Launay H, Dreyfus M (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc Natl Acad Sci USA* 102: 5958–5963.
96. Pape T, Schneider TR (2004) HKL2MAP: A graphical user interface for macromolecular phasing with SHELX programs. *J Appl Crystallogr* 37:843–844.
97. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–230.
98. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
99. Strong M, et al. (2006) Toward the structural genomics of complexes: Crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proc Natl Acad Sci USA* 103:8060–8065.
100. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
101. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62: 439–450.
102. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
103. Goddard TD, Kneller DG (2004) SPARKY 3 (Univ of California, San Francisco).