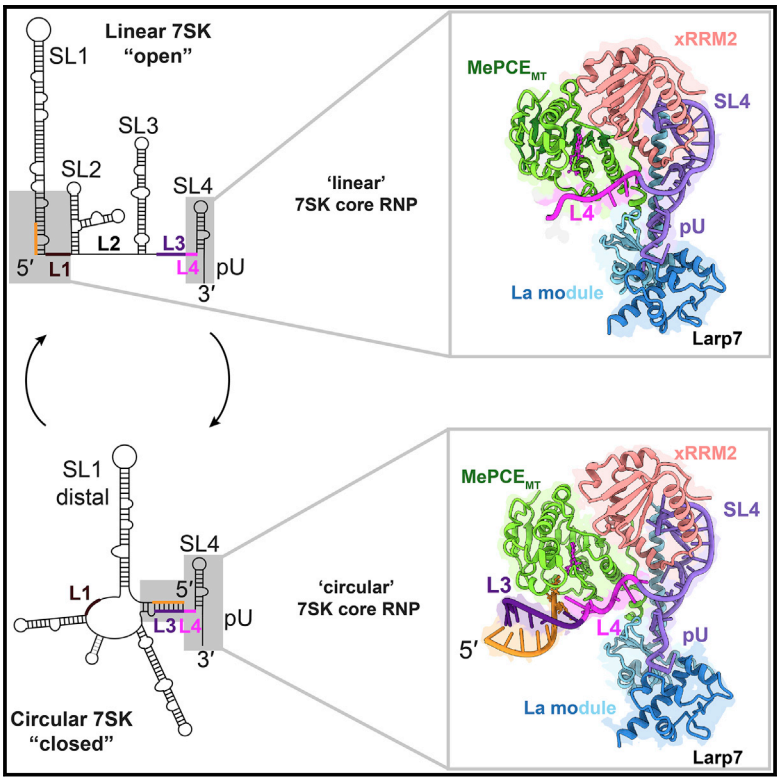
Molecular Cell

Article

Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP

Graphical abstract



Highlights

- Cryo-EM structures of MePCE-7SK-Larp7 core RNP complex for linear and circular 7SK
- MePCE is in a methyltransferase-inhibited conformation
- MePCE interaction with Larp7 is RNA dependent and reveals unexpected motifs
- 7SK RNA transitions from nascent linear to circular conformation upon core assembly

Authors

Yuan Yang, Shiheng Liu, Sylvain Egloff, ..., Tamás Kiss, Z. Hong Zhou, Juli Feigon

Correspondence

hong.zhou@ucla.edu (Z.H.Z.), feigon@mbi.ucla.edu (J.F.)

In brief

Yang et al. report cryo-EM structures of MePCE-7SK-Larp7 core RNP complex formed by linear and circular 7SK RNA, revealing full-length Larp7 structure and mechanism of MePCE methyltransferase inactivation. A common MePCE-7SK-Larp7 platform allows large-scale RNA conformational change to take place during RNP biogenesis and P-TEFb regulation.

Yang et al., 2022, Molecular Cell *82*, 1724–1736
May 5, 2022
https://doi.org/10.1016/j.molcel.2022.03.001

CellPress

CellPress

Molecular Cell

Article

Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP

Yuan Yang,[1,7] Shiheng Liu,[2,3,7] Sylvain Egloff,[4] Catherine D. Eichhorn,[1,6] Tanya Hadjian,[1] James Zhen,[1,3] Tamás Kiss,[4,5] Z. Hong Zhou,[2,3,*] and Juli Feigon[1,8,*]

[1]Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA 90095, 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
[4]Molecular, Cellular, and Developmental Biology Department (MCD), Centre de Biologie Intégrative (CBI), University of Toulouse, CNRS, UPS, 31062 Toulouse, France
[5]Biological Research Centre, Szeged, Temesvári krt. 62, 6726, Hungary
[6]Present address: Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588, USA
[7]These authors contributed equally
[8]Lead contact
*Correspondence: hong.zhou@ucla.edu (Z.H.Z.), feigon@mbi.ucla.edu (J.F.)
https://doi.org/10.1016/j.molcel.2022.03.001

SUMMARY

7SK non-coding RNA (7SK) negatively regulates RNA polymerase II (RNA Pol II) elongation by inhibiting positive transcription elongation factor b (P-TEFb), and its ribonucleoprotein complex (RNP) is hijacked by HIV-1 for viral transcription and replication. Methylphosphate capping enzyme (MePCE) and La-related protein 7 (Larp7) constitutively associate with 7SK to form a core RNP, while P-TEFb and other proteins dynamically assemble to form different complexes. Here, we present the cryo-EM structures of 7SK core RNP formed with two 7SK conformations, circular and linear, and uncover a common RNA-dependent MePCE-Larp7 complex. Together with NMR, biochemical, and cellular data, these structures reveal the mechanism of MePCE catalytic inactivation in the core RNP, unexpected interactions between Larp7 and RNA that facilitate a role as an RNP chaperone, and that MePCE-7SK-Larp7 core RNP serves as a scaffold for switching between different 7SK conformations essential for RNP assembly and regulation of P-TEFb sequestration and release.

INTRODUCTION

For most eukaryotic gene expression events, RNA polymerase II (RNA Pol II) pauses at 25–50 nt downstream of the transcription start site, creating a rate-limiting step after transcription is initiated upon RNA Pol II binding to the promoter (Core and Adelman, 2019). The release of promoter-proximal paused RNA Pol II into productive elongation is triggered by the kinase activity of P-TEFb, a heterodimer of Cyclin-dependent kinase 9 (Cdk9) and Cyclin T (Price, 2000). P-TEFb phosphorylates pausing regulators DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF), as well as the C-terminal domain (CTD) of the RNA Pol II Rpb1 subunit. This leads to the rearrangement of the RNA Pol II complex into a productive elongation complex (Vos et al., 2018). The kinase activities of P-TEFb were initially discovered in studies of the human immunodeficiency virus 1 (HIV-1) (Zhu et al., 1997). HIV-1 transcriptional trans-activator (Tat) binds the trans-activation response (TAR) element during early transcription of the viral genome and recruits P-TEFb to release the HIV-1 promoter-proximal paused RNA Pol II. Cells contain two main pools of P-TEFb: active P-TEFb found at gene promoters and inhibited P-TEFb sequestered in a 7SK ribonucleoprotein (RNP) complex comprising MePCE, Larp7, Hexim, P-TEFb, and the abundant ~332 nt nuclear non-coding (nc) RNA 7SK. Various protein factors have been associated with the release of active P-TEFb from the 7SK RNP in response to cell signaling and stress, but the molecular mechanism of P-TEFb release remains elusive (Egloff et al., 2018). Interestingly, for a number of genes, including the HIV-1 viral genome, P-TEFb-bound 7SK RNP is directly recruited to the promoter region for subsequent release of the paused RNA Pol II (D'Orso and Frankel, 2010; Ji et al., 2013). 7SK was first identified over 40 years ago (Gurney and Eliceiri, 1980; Zieve and Penman, 1976), but its roles in the regulation of eukaryotic transcription (Nguyen et al., 2001; Yang et al., 2001), cellular stress response (Nguyen et al., 2001; Yang et al., 2001; Yik et al., 2003), and HIV replication (Sedore et al., 2007) were more recently recognized and remain only partially characterized (Peterlin et al., 2012).

7SK is monomethylated at the 5′-γ-phosphate (mpppG) by MePCE (Jeronimo et al., 2007), whose only other substrate in humans is U6 snRNA. This unique cap is likely formed co-transcriptionally, as MePCE is enriched at the 7SK and U6 promoter

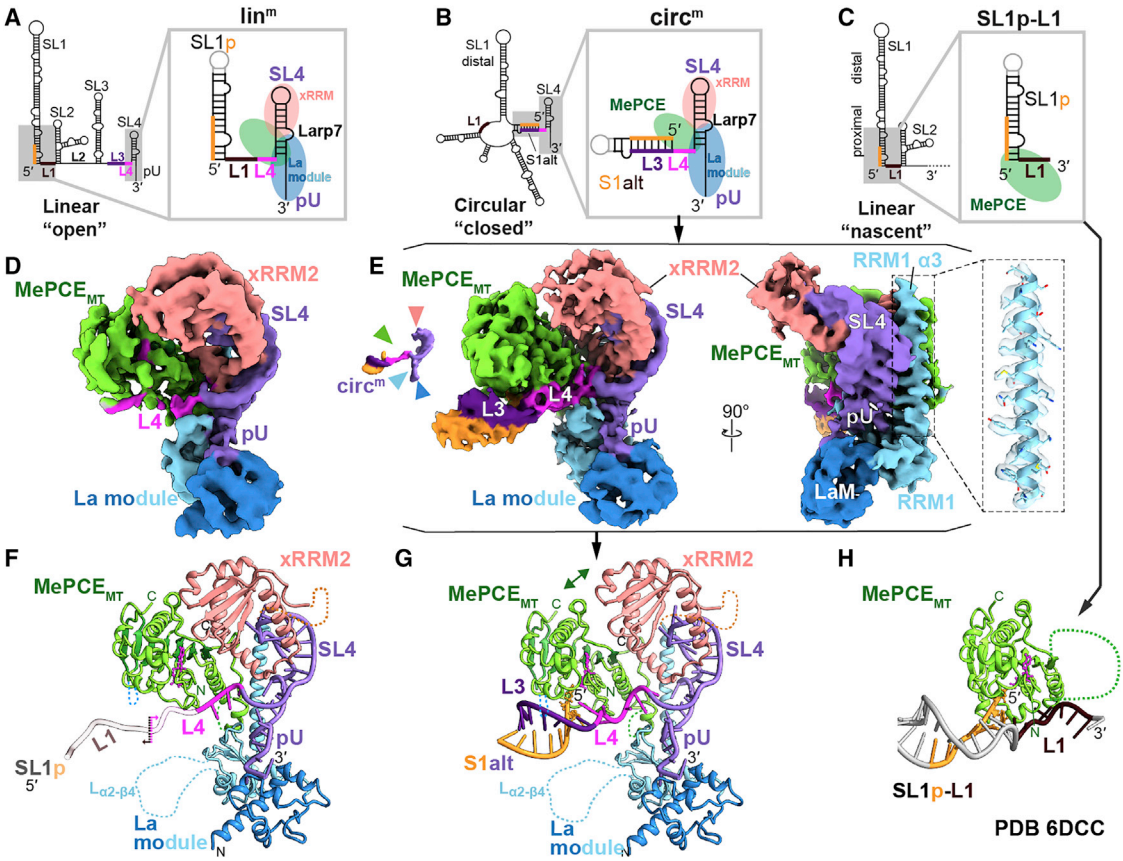


Figure 1. 7SK core RNP cryo-EM structures

(A–C) Secondary structure schematics of 7SK RNA linear (A), circular (B), and nascent linear (C). Minimal RNA constructs (lin^m, circ^m, and SL1p-L1) are highlighted by gray shadings and shown in insets with bound proteins.

(D) Cryo-EM density map of lin^m RNP.

(E) Front and side views of cryo-EM density map of circ^m RNP. SL4 and pU RNA is purple, 5′ nt 1–7 is gold, L1 is brown, L3 is dark purple, and L4 is magenta. MePCE$_{MT}$ is green, and Larp7 La motif, RRM1, and xRRM2 are blue, sky blue, and coral, respectively. Inset shows RRM1 α3 model overlayed on transparent cryo-EM density to illustrate the quality of the data. Additional overlays are in Figure S2.

(F and G) Cryo-EM structures of lin^m RNP (F) and circ^m RNP (G).

(H) Crystal structure of MePCE$_{MT}$ in complex with SL1p-L1 (PDB: 6DCC; Yang et al., 2019).

See also Figures S1–S3.

regions (Xue et al., 2010). MePCE remains bound to 7SK after capping its 5′ end, and together with the Larp7 that protects the 3′ end, it forms a stress-resistant and constitutively assembled 7SK core RNP (He et al., 2008; Krueger et al., 2008; Muniz et al., 2013; Xue et al., 2010). The LARP7 family proteins are closely related to genuine La protein, an RNA chaperone and general protector of nascent RNA Pol III transcripts. The La and LARP7 proteins bind the RNA 3′ terminal polyuridine stretch (UUU-3′-OH) using a specialized RNA-binding unit called the La module, which consists of a La motif (LaM) and an RNA recognition motif (RRM1) (Maraia et al., 2017). Little is understood about the chaperone function of La for diverse RNAs and LARP7 proteins for their specific substrates, e.g., human Larp7 for 7SK (Krueger et al., 2008) and ciliate Tetrahymena p65 and fission yeast Pof8 for telomerase RNA (Aigner et al., 2000; Collopy et al., 2018; Mennie et al., 2018; Páez-Moscoso et al., 2018; Witkin and Collins, 2004). Biochemical and structural studies suggest that LARP7 RRM2, an xRRM, largely determines target RNA specificity (Eichhorn et al., 2016, 2018; Singh et al., 2012). An RNA-dependent protein-protein interaction between Larp7 and MePCE has been proposed as a key factor in the assembly of 7SK core RNP, but the molecular details of this potential interaction are unknown (Brogie and Price, 2017; Xue et al., 2010).

A main challenge in understanding 7SK RNP structure and regulation is the existence of at least two functional 7SK conformations. A linear conformation of 7SK was initially proposed based on chemical and enzymatic probing experiments (Wassarman and Steitz, 1991); however, phylogenetic analysis suggested an alternative RNA "circular" conformation (Marz et al., 2009) (Figures 1A and 1B), which is partially supported by recent SHAPE data (Bubenik et al., 2020). An internally truncated 7SK, which can only adopt the linear conformation, is sufficient for Hexim and P-TEFb association (Egloff et al., 2006), whereas the circular conformation is proposed to correspond to the

Table 1. Cryo-EM data collection, refinement, and validation statistics

	circ^{m} RNP (EMDB-25198) (PDB: 7SLQ)	lin^{m} RNP (EMDB-25197) (PDB: 7SLP)
Data collection and processing		
Magnification	130,000	130,000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	50	50
Defocus range (μm)	−1.5 to −2.5	−1.5 to −3.0
Pixel size (Å)	1.07	1.07
Symmetry imposed	C1	C1
Initial particle images (no.)	5,598,553	11,324,367
Final particle images (no.)	294,504	159,349
Map resolution (Å) FSC threshold	3.7 0.143	4.1 0.143
Refinement		
Initial model used (PDB code)	*ab initio*	*ab initio*
Model resolution (Å) FSC threshold	4.7 0.5	7.1 0.5
Model resolution range (Å)	4.7	7.1
Map sharpening *B* factor (Å^2)	−207.1	−255.7
Model composition		
Non-hydrogen atoms	5,448	5,176
Protein residues	608	597
Ligands	2	1
B factors (Å^2)		
Protein	51.4	86.0
Ligand	68.7	57.6
RMS deviations		
Bond angles (°)	0.006	0.008
Bond lengths (Å)	1.120	1.180
Validation		
MolProbity score	1.70	1.90
Clashscore	6.44	6.98
Poor rotamers (%)	1.35	1.54
Ramachandran plot		
Favored (%)	96.34	95.26
Allowed (%)	3.66	4.19
Disallowed (%)	0.00	0.55

Hexim and P-TEFb released 7SK RNP (Brogie and Price, 2017). Consistent with this, chemical probing provides evidence for large-scale conformational changes in 7SK when bound to different proteins (Flynn et al., 2016; Krueger et al., 2010). Interestingly, MePCE can cap both linear 7SK at the proximal stem-loop 1 (SL1p) and circular 7SK at an alternative stem 1 (S1alt) (Yang et al., 2019). It remains an open question whether one or both conformations can assemble functional 7SK RNPs *in vivo*.

Here, we present cryo-EM structures of 7SK core RNP in two conformations, which provide the first view of full-length Larp7 (or any La or La-related protein), its RNA recognition, and assembly with MePCE. Unexpectedly, MePCE transitions from catalytically productive binding at the nascent linear 5′ hairpin that forms during transcription to a methyltransferase inhibited MePCE–7SK–Larp7 complex at the 3′ hairpin, which is present in both linear and circular 7SK conformations. *In vitro* methyltransferase assays show that MePCE preferably recognizes and caps the linear conformation of 7SK RNA, but Larp7 binding promotes the transition of 7SK RNA into the circular conformation, thus indicating that a conformational switch occurs during 7SK core RNP biogenesis. In cells, a mutated 7SK trapped in the linear conformation is no longer able to release P-TEFb in response to transcriptional stress, revealing the circular 7SK as a critical conformation responsible for P-TEFb release. Our study highlights 7SK core RNP as a platform that allows multiple linear-circular RNA conformational switches to control RNP assembly during biogenesis and P-TEFb binding and release during RNA Pol II regulation.

RESULTS

Architecture of the 7SK core RNP

To assemble the 7SK core RNP for structural studies, we considered the biochemically mapped MePCE and Larp7 binding sites determined in the context of the linear 7SK, which are the 5′ 3-bp of SL1p followed by linker 1 (L1) and SL4 plus the polyU tail U_{325}UUCUUUU-3′-OH (pU), respectively (Krueger et al., 2008; Muniz et al., 2013) (Figures 1A and 1C). Since 7SK with internal deletion of SL2-SL3 (lin^{14}; Figure S1D) pulls down all components of the transcription inhibitory 7SK RNP, i.e., MePCE, Larp7, Hexim1, and P-TEFb (Egloff et al., 2006; Muniz et al., 2013), we designed a minimal linear 7SK (lin^{m}; Figure 1A) by shortening SL1 to SL1p (Figure S1A), which includes the binding site for MePCE but not Hexim. Based on the phylogenetically proposed circular conformation (Marz et al., 2009), we also designed a minimal circular 7SK comprising the conserved base pairs in S1alt, L4, SL4, and pU (circ^{m}; Figure 1B), which includes the alternative MePCE S1alt-L4 binding site (Figure S1A). S1alt-L4 has an 8-fold weaker binding affinity for MePCE$_{MT}$ (MT: methyltransferase domain) than SL1p-L1 but can still be capped (Yang et al., 2019). MePCE$_{MT}$ also caps lin^{m} and circ^{m}, with similar methyltransfer behavior to SL1p-L1 and S1alt-L4 constructs, respectively (Figures S1A–S1C). Core RNPs were assembled with MePCE$_{MT}$, Larp7, and either lin^{m} or circ^{m} RNA in the presence of methyltransferase substrate S-adenosylmethionine (see STAR Methods). We obtained cryo-EM structures at 4.1 and 3.7 Å resolution for the RNPs with lin^{m} or circ^{m}, respectively. In the cryo-EM map of circ^{m} RNP, we could unambiguously trace the polypeptide chains and native RNA bases, except for the distal bulge and 3 bp in S1alt, and build an atomic model *de novo* (Figures 1E, 1G, and S2; Table 1). This also facilitated atomic modeling of lin^{m} RNP (Figures 1D, 1F, and S3; Table 1).

In circ^{m} RNP, the 5′ S1alt-L4 and 3′ SL4-pU are positioned perpendicular to each other, forming a T-shape scaffold (Figure 1E). SL4 and pU form the top arm of the T, with Larp7

xRRM2 and La module coating distal SL4 and proximal SL4-pU, respectively. $MePCE_{MT}$ binds along L4 that forms the spine of the T and the 5′ mpppG in S1alt and also with Larp7. Surprisingly, $MePCE_{MT}$ in lin^{m} RNP is also bound at L4 (Figures 1A and 1D) rather than at the higher affinity site SL1p-L1 observed in the "nascent" linear $MePCE_{MT}$–SL1p-L1 crystal structure (Yang et al., 2019) (Figures 1C and 1H). Consequently, lin^{m} and $circ^{m}$ RNPs form highly similar $MePCE_{MT}$–L4-SL4-pU–Larp7 conformations. However, while S1alt in $circ^{m}$ RNP was clearly visible, in lin^{m} RNP, there was no visible density for SL1p, and the 5′ half of L4 and L1 can only be traced with low-threshold density, suggesting flexibility (Figure S3I) and indicating that the 5′ mpppG of lin^{m} RNA is distant from $MePCE_{MT}$. $MePCE_{MT}$ and Larp7 xRRM2 are also ~1.7 Å closer together in lin^{m} RNP compared to $circ^{m}$ RNP, although they still do not directly interact (Figures 1F and 1G). To verify that these minimal RNAs are biologically relevant, we also assembled core RNPs with full-length SL1-containing lin^{14ex} or full-length 7SK (wild type, WT) (Figure S1D). Although these samples dissociate upon cryo-EM sample preparation, negative stain EM class averages showed the same $MePCE_{MT}$-7SK-Larp7 interaction seen in lin^{m} and $circ^{m}$ RNPs (Figures S1D–S1F). The long SL1 in lin^{14ex} can be observed in 2D classes (Figure S1F), but appears highly flexible relative to the RNP core, consistent with the lack of SL1p density in lin^{m} RNP. These results show that in the presence of Larp7, $MePCE_{MT}$ preferentially binds at L4, even when (for lin^{m}) mpppG is not at the binding site and despite the higher affinity of $MePCE_{MT}$ for SL1p-L1 in the absence of Larp7. The structural basis for this is described below.

Larp7 structure and assembly with 7SK

The complete structure of Larp7 bound to its target RNA is resolved in our cryo-EM maps, whereas previous structures of genuine La and Larps were limited to individual domains. Since lin^{m} and $circ^{m}$ RNPs have slightly different overall structures but share a common $MePCE_{MT}$–L4-SL4-pU–Larp7 core, for the high-resolution features discussed below, we use the model of $circ^{m}$ RNP. Larp7 is considerably more structured and the Larp7-RNA interface is significantly more extensive in 7SK core RNPs than expected, based on previous structural data for isolated domains (Eichhorn et al., 2018; Uchikawa et al., 2015). Between the La module and xRRM2 bound at pU and loop 4, respectively, are two previously unobserved structured regions at the C termini of RRM1 and xRRM2 that span most of SL4 (Figure 2A). Based on the crystal structure of a La module-pU complex, Larp7 RRM1 was reported to lack the characteristic β4 strand (Uchikawa et al., 2015) (Figure S4A). Here, we find that RRM1 in fact has a conventional 4-stranded β sheet but with an exceptionally long α2-β4 loop (189 residues), for which there is no density in the cryo-EM maps, followed by a long α3 that extends along one side of SL4 (Figures 1E, 2A, and S2I).

Surprisingly, the La module binds all 8 pU nucleotides, rather than just the expected three terminal U_{330}–U_{332} that form a "V-shape" sandwiched between LaM and RRM1 (Uchikawa et al., 2015), thus providing the first complete polyU-tail-La module interface (Figures 2A and 2C). The 5 additional pU nucleotides form an S-turn that interacts with RRM1, but not LaM, on the β3 (RNP-1) and β2-β3 loop (Figures 2A–2D). Although the RNP-1 and RNP-2 regions are canonical RNA-binding motifs in RRM folds, La module crystal structures for both Larp7 and genuine La protein with pU showed interactions with non-canonical RNA-binding surfaces only (Kotik-Kogan et al., 2008; Teplova et al., 2006; Uchikawa et al., 2015). Here, in the presence of the native RNA and protein, we find that RNP-1 of Larp7 RRM1 participates in binding the 5 additional pU nucleotides. Interestingly, in genuine La protein, the canonical RRM sequences RNP-1 and RNP-2 (on β1) (Maris et al., 2005) have been proposed to be important for its tRNA chaperone activity (Bayfield and Maraia, 2009), which suggests that La protein could bind RNA in a similar way as observed here for Larp7. Following the β sheet, RRM1 α3 extends out perpendicular to β4, fixed at its N-terminal end by interactions with RNP-1, RNP-2, and $U_{325}U_{326}$, and at its center with U_{319} and the $C_{320}U_{321}$ bulge (Figures 2E and 2F). The latter interaction explains why the $C_{320}U_{321}$ bulge, which does not contribute to xRRM2 binding to SL4 (Eichhorn et al., 2016), is highly conserved among vertebrates (Egloff et al., 2006).

For xRRM2, the unique RNP-3 sequence and α3 interact extensively with the apical loop of SL4, including base-specific interactions with G312, U313, and G314, as observed in the xRRM2-SL4 crystal structure (Eichhorn et al., 2018) (Figure S4B). However, the xRRM2 is followed by an ordered α3-α4 loop ($L_{\alpha3\text{-}\alpha4}$) and α4, defined here as the Larp7 CTD, that interact with and widen the major groove of SL4 near the apical loop (Figures 2A, 2G, and S4B). This positions α4 to interact with the C-terminal half of RRM1 α3 (Figures 2A and 2G). Together, Larp7 RRM1 α3 and CTD span the length of stem 4, linking these two domains and pre-ordering their interface with MePCE.

Larp7 domain extensions are pre-formed

To help elucidate the role of the newly identified Larp7 structured regions (RRM1 β4 and α3, and CTD) in core RNP assembly, we investigated by nuclear magnetic resonance (NMR) spectroscopy whether the entire RRM1 is pre-formed in solution or whether β4 and α3, which are connected to the rest of RRM1 by a 189-residue unstructured loop, only fold in the presence of 7SK. First, we labeled all 5 Larp7 Trp residues, which span its three domains, with 5-fluorotryptophan (5F-Trp) (Figures 3A and 3B) and obtained ^{19}F 1D spectra (Figure 3C). All 5 ^{19}F resonances were observed and assigned by comparing ^{19}F spectra of various protein constructs as described in STAR Methods. The 5F-Trp resonances display different linewidths for LaM, RRM1, and xRRM2, suggesting that these three domains tumble independently in free Larp7 (Figure 3C; Table S1). We compared ^{19}F NMR spectra of full-length Larp7 with La module constructs truncated before RRM1 β4 ($LaMod_{196}$ and $LaMod_{208}$) or after the first third of α3 ($LaMod_{390}$ and $LaMod_{390\Delta L}$, where ΔL is the α2-β4 loop deletion) (Figure 3A). 1D ^{19}F spectra of $LaMod_{390}$ and $LaMod_{390\Delta L}$ overlaid well with the three corresponding 5F-Trp peaks in the spectrum of full-length Larp7, indicating that these La module constructs have the same fold as in full-length Larp7 (Figure 3C). In contrast, for $LaMod_{208}$, Trp140 exhibits two low-intensity peaks riding on a very broad peak, indicating that at least two sub-conformations are present, one of which may correspond to the crystal structure (Uchikawa et al., 2015). $LaMod_{196}$ has only one very broad peak for Trp140, suggesting that its (partial) RRM1 is even less stable than $LaMod_{208}$.

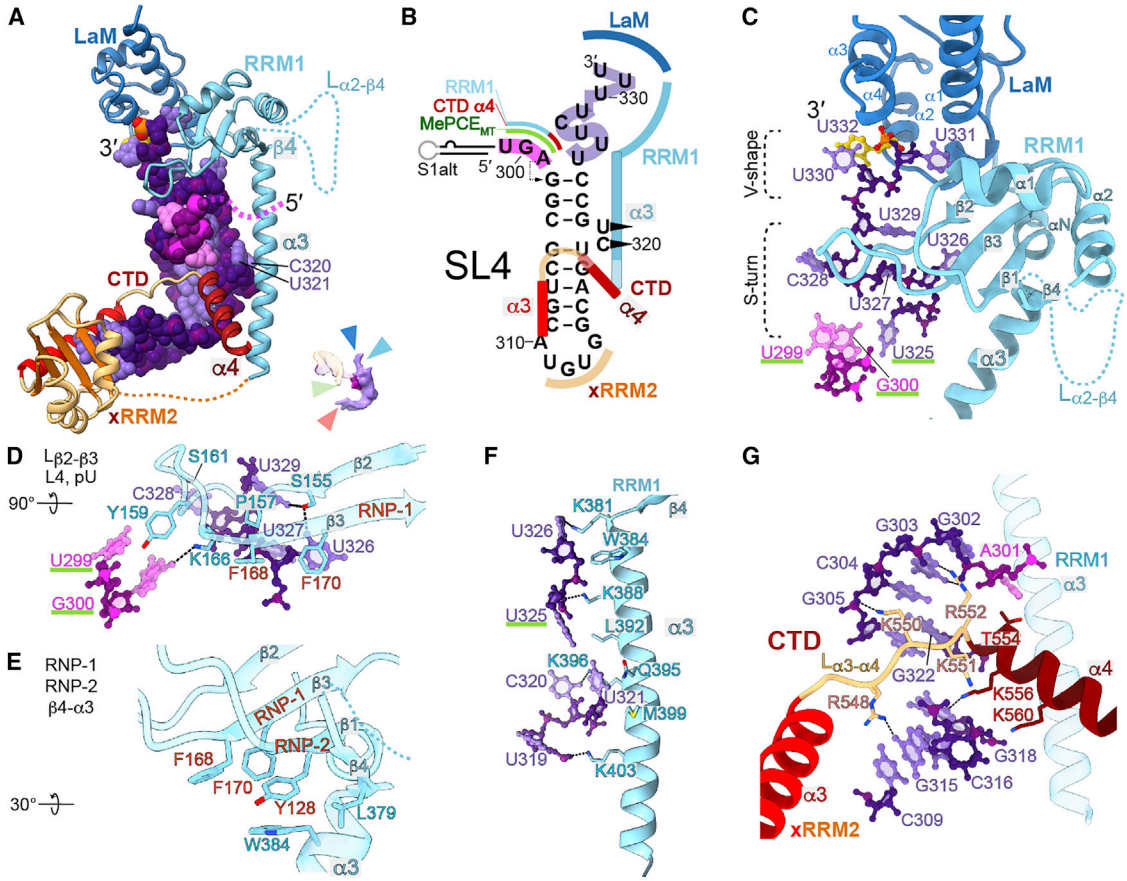


Figure 2. Larp7 coordinates SL4-pU recognition with extended motifs from C termini of both RRMs

(A) Larp7-RNA interactions in circ^{m} RNP, with two newly uncovered Larp7-RNA interfaces indicated by squares and shown in detail in (C)–(G). The circ^{m} RNA is colored purple (SL4 and pU) and magenta (L4) and is shown as sphere representation.

(B) Schematic of L4-SL4-pU with Larp7-RNA interactions.

(C) pU U_{330}–U_{332} clamped into a V-shape by LaM-RRM1 and U_{325}–U_{329} forms an S-turn on the RRM1 RNP-1, RNP-2, and β2-β3 loop.

(D) Detailed interactions between RNP-1, β2-β3 loop, and RNA.

(E) Hydrophobic network involving RRM1 RNP-1, RNP-2, β4, and α3.

(F) Interaction between the Larp7 RRM1 helix α3 and the SL4 $C_{320}U_{321}$ bulge as well as 3′ pU.

(G) Interaction of CTD α3-α4 loop and α4 with the major groove of SL4 that positions α4 toward RRM1 α3.

See also Figure S4.

Consistent with this, RRM1—but not LaM—resonances in the ^{15}N-HSQC spectrum of $LaMod_{196}$ are line-broadened (Figure 3D). Additionally, we purified $LaMod_{22–390\Delta L}$ and β4-$xRRM2_{370–582}$ (from RRM1 β4 to the C terminus) constructs and obtained TALOS-N secondary structure scores (Shen and Bax, 2013), which predict α helices and β strands based on chemical shifts (Figures 3E and S5A), and ^{1}H-^{15}N heteronuclear nuclear overhauser effect (NOE) analysis, which provides information on the local flexibility of the protein backbone (Figures 3F and S5B). These data indicate that β4 and α3 are already formed in the absence of RNA. We conclude that the RRM1 β4 and α3 regions are indispensable for the stable RRM1 fold and are integral parts of RRM1 in the absence of RNA.

Although RRM1 α3 and CTD α4 are present in free Larp7, TALOS-N scores show that they are 2.5 and 1.5 turns shorter, respectively, than in the cryo-EM core RNP structures (Figures 3E, 3F, S5A, and S5B). In the presence of RNA, heteronuclear NOE values show small but significant increases for the spectrally resolved residues of CTD $L_{\alpha3–\alpha4}$ and α4 N terminus (Figure S5B). For both RRM1 α3 and CTD residues, RNA-binding results in the broadening of most resonances, and the remaining visible peaks have above-average chemical shift perturbation (CSP) values (Figure S5C), indicating that these regions participate in RNA-binding and transition from partially disordered to ordered, as seen in the cryo-EM structures. Isothermal calorimetry (ITC) experiments reported here (Figure S5D) and previously (Eichhorn et al., 2016) show that the CTD $L_{\alpha3–\alpha4}$ and α4 regions contribute moderately to RNA-binding. In summary, Larp7 RRM1 and xRRM2-CTD have preformed C-terminal α helices and tumble independently in the absence of RNA. When bound to RNA, RRM1 α3 and CTD α4 extend further to bring them in contact with each other

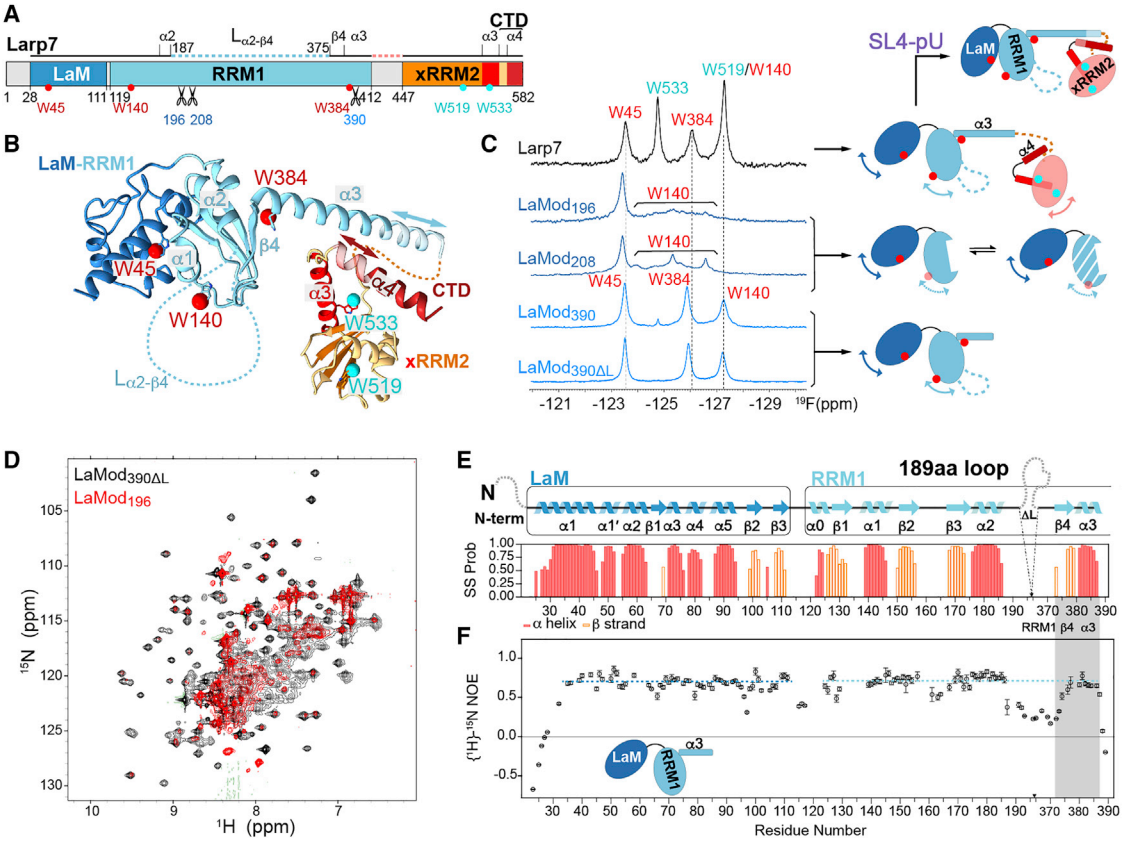


Figure 3. Extended Larp7 motifs are pre-formed

(A) Domain diagram of Larp7 with locations of native Trp and truncation constructs indicated by scissors.

(B) Larp7 structure in circ^{m} RNP, with native Trp highlighted by spheres (red in La module and cyan in xRRM2). Transparent ribbons and arrows indicate helical regions in RRM1 α3 and CTD α4 that extend upon binding RNA.

(C) ^{19}F-1D spectra of Larp7 (black, top) and La module constructs, LaMod$_{196}$, LaMod$_{208}$, LaMod$_{390}$, and LaMod$_{390\Delta L}$, with cartoons illustrating domain movements to the right.

(D) ^{15}N-HSQC spectral overlay of LaMod$_{390\Delta L}$ (black) and LaMod$_{196}$ (red).

(E) TALOS-N secondary structure scores of LaMod$_{22–390\Delta L}$ based on chemical shift assignment values.

(F) ^{1}H–^{15}N heteronuclear NOE values for LaMod$_{22–390\Delta L}$.

See also Figure S5 and Table S1.

and, together with the LaM, rigidify the entire protein on the RNA. These interactions provide the structural platform for docking with MePCE, as detailed below.

Structural basis of MePCE inactivation

Comparison of circ^{m} RNP structure with the MePCE$_{MT}$-SL1p-L1 crystal structure (Yang et al., 2019), which represents the nascent linear RNP, reveals major differences in MePCE-RNA interactions and the MePCE active site. In the crystal structure, MePCE$_{MT}$ is bound to the two 5′ base-paired Gs and six L1 nts (A_{109}–C_{114}) (Figures 4A and 4B). Five of six L1 nucleotides are stacked continuously, with Lys420 intercalated between U_{110} and G_{111} (Yang et al., 2019) (Figure 4A). In contrast, in circ^{m} RNP, MePCE$_{MT}$ is bound to the 5′ base-paired G in S1alt and six L4 nucleotides A_{296} to A_{301}, where L4 adopts an extended non-helical conformation with five non-A-form sugar puckers, and most nucleotides are not stacked on each other (Figures 4C and 4D). In lin^{m} RNP, the visible linker L4 nucleotides U_{299}–A_{301} adopt similar conformations and MePCE$_{MT}$ interactions as those in circ^{m} RNP (Figures 4E and 4F).

Even more striking differences are seen within the MePCE$_{MT}$ active site. Compared with RNA-free MePCE$_{MT}$, upon binding SL1p-L1, two regions of MePCE$_{MT}$ become ordered and form short helices (α0 and α7), and together with the pre-existing helices α5′ and α6′ form the 5′-γ-phosphate binding tunnel (Yang et al., 2019) (Figures 4A and 4B). These disordered-to-ordered transitions represent the conformational switch of MePCE$_{MT}$ from RNA-free to nascent linear RNA-bound state (Yang et al., 2019). Then, the interaction of MePCE with Larp7 in the core RNPs induces a second conformational change in MePCE$_{MT}$. The MePCE$_{MT}$ N-terminal region adopts an extended conformation sandwiched in between L4 and MePCE$_{MT}$ α3, and the β6-β7 loop becomes disordered instead of forming the α7 present in the nascent linear MePCE-7SK RNP (Figures 4C and 4D). This disrupts half of the γ-phosphate binding tunnel (Figure 4D), where mutations in α0 and α7 residues are detrimental to

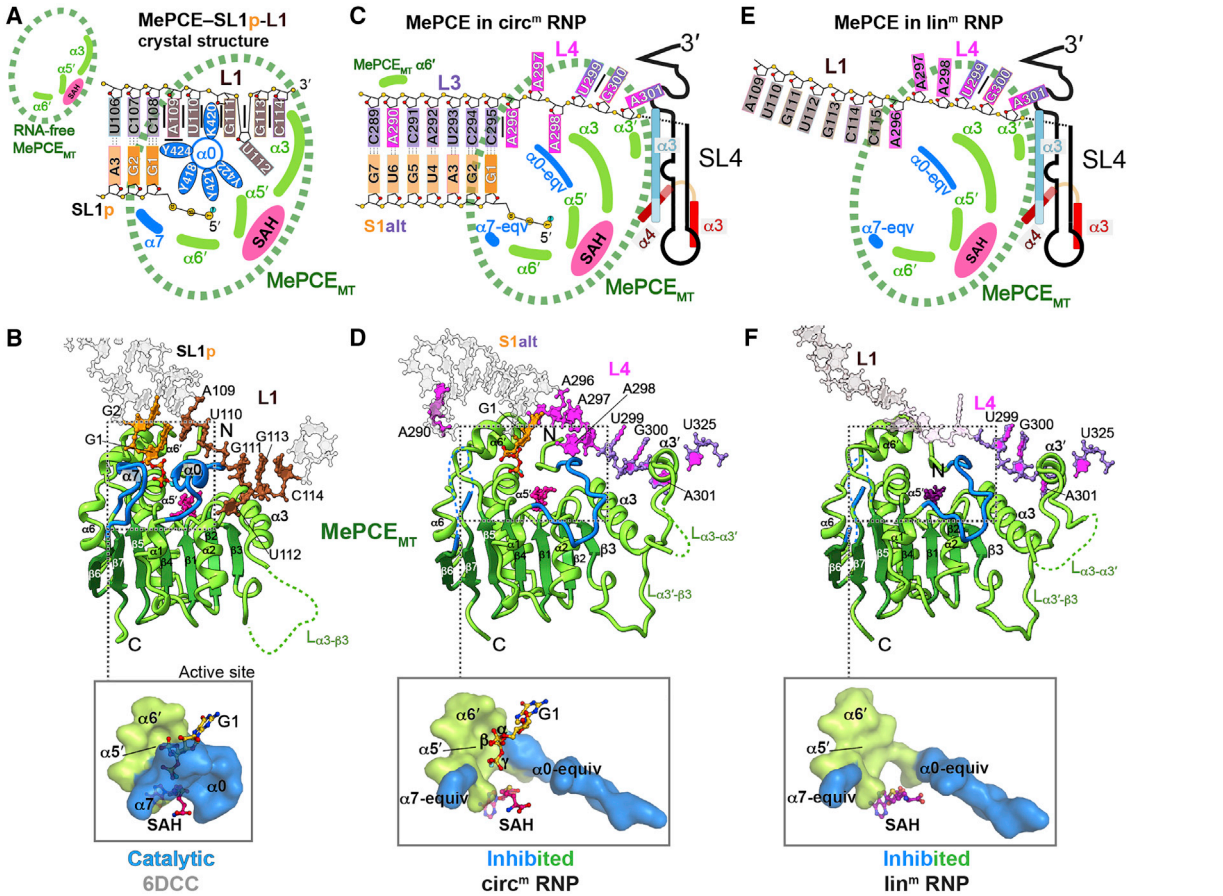


Figure 4. MePCE is inactivated in core RNP

(A and B) Cartoon representation and MePCE-SL1p-L1 interactions in the nascent linear complex crystal structure (PDB: 6DCC Yang et al., 2019). The inset small cartoon represents the MePCE conformation in the absence of RNA substrate (PDB: 5UNA).

(C and D) Cartoon representation and MePCE-circ^{m} interactions in circ^{m} RNP.

(E and F) Cartoon representation and MePCE-lin^{m} interactions in lin^{m} RNP. Insets in (B), (D), and (F) show surface representations of the triphosphate binding tunnel in each structure.

methyltransferase activity (Yang et al., 2019). These structural differences explain why MePCE$_{MT}$ is catalytically incompetent in 7SK core RNP (Brogie and Price, 2017; Xue et al., 2010): since (1) the methyltransferase active site and the single-stranded recognition sites of MePCE$_{MT}$ remain inaccessible as they are occupied by the methylated circ^{m} RNA and (2) the γ-phosphate binding tunnel is partially disrupted by the conformational changes of the catalytically important α0 and α7 residues in both circ^{m} and lin^{m} RNPs.

RNA-dependent MePCE-Larp7 interaction

In the structures of MePCE$_{MT}$ free and bound to SL1p-L1, the α3-β3 loop is disordered (Yang et al., 2019) (Figure 4B). In contrast, for circ^{m} RNP, cryo-EM density is observed for the majority of α3-β3 loop, and it contains an additional short α3′ followed by an extended α3′-β3 loop (Figures 4D and S2H). These two regions become ordered due to interactions with Larp7 RRM1 α3 and CTD α4 (Figures 5A and 5D). Outside the MePCE$_{MT}$ RNA-binding pocket described above, there are two additional nucleotides that contact the MePCE: S1alt A$_{290}$ that contacts α6′ and pU U$_{325}$ that stacks between α3′ Arg531 (α3′) and SL4 first bp (G$_{302}$•C$_{324}$), thereby bridging α3′ and Larp7 RRM1 (Figure 5B). The last L4 nucleotide, A$_{301}$, which is anchored by interactions with MePCE$_{MT}$ α3 and Larp7 CTD residues, forms a base triple with G$_{302}$•C$_{324}$ (Figure 5B, inset).

The extensive MePCE-Larp7 interface consists primarily of interactions between MePCE$_{MT}$ α3′ and α3′-β3 loop, Larp7 RRM1 α3, and Larp7 CTD α4 (Figures 5A and 5D). The Larp7 residues involved are far more extensive than the previously identified MePCE interacting domain (MID) residues (Brogie and Price, 2017), which corresponds to the CTD. MePCE$_{MT}$ α3′ and the N-terminal half of Larp7 RRM1 α3 interact through a cluster of hydrophobic interactions, further assisted by interactions between MePCE$_{MT}$ α3′ and RRM1 β sheet, including RNP-2 and RNP-1 residues (Figure 5E), while the MePCE$_{MT}$ extended α3′-β3 loop, Larp7 CTD α4, and RRM1 α3 C-terminal half all interact with each other through a network of hydrophobic interactions (Figure 5F). The opposite face of Larp7 CTD α4 also interacts with the side of MePCE$_{MT}$ at α3, β3, and β2 (Figure S5F). Consistent with the observed CTD interactions, the deletion of the last

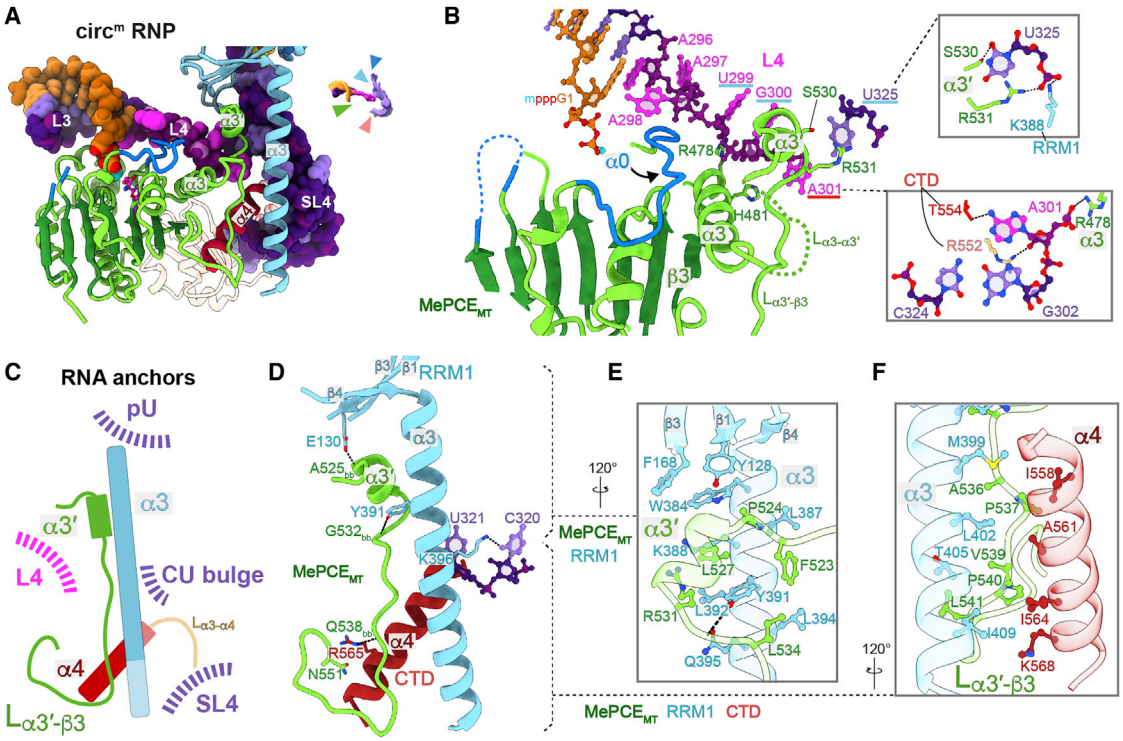


Figure 5. MePCE-Larp7 interaction is RNA dependent

(A) circ^m RNP cryo-EM structure, with regions of assembly interfaces indicated by gray boxes and shown in detail in (B) and (D).

(B) Active site conformation of MePCE$_{MT}$ in circ^m RNP (protein is green and blue). Insets show U_{325} (top) and A_{301} (bottom) RNA-protein interactions.

(C) Cartoon of the RNA anchors at the MePCE-Larp7 interface.

(D) MePCE-Larp7 interface with interdomain hydrogen bonds shown. The $C_{320}U_{321}$ bulge interaction that anchors RRM1 α3 is also shown.

(E) Hydrophobic interactions between MePCE$_{MT}$ α3′ and Larp7 RRM1.

(F) Hydrophobic interactions between the extended MePCE α3′-β3 loop, RRM1 α3, and CTD α4.

See also Figures S5 and S6.

21 amino acids of Larp7 impairs 7SK and P-TEFb/Hexim1 association (He et al., 2008), and the deletion of the last 33 amino acids of Larp7 abolishes RNA-dependent MePCE inhibition (Brogie and Price, 2017). Most of the MePCE-Larp7 interface residues have consensus identities in vertebrates, and most other residues are moderately conserved but maintain identity of long hydrophobic sidechains (Figure S6).

An important feature of the MePCE-Larp7 interface is that the relative positions of all three interacting motifs are anchored by binding to RNA: MePCE$_{MT}$ α3′ interacts with U_{325}; Larp7 RRM1 interacts with L4 nucleotides $U_{299}G_{300}$ and pU that position RRM1 α3 along one side of SL4, which is further anchored by interaction with the $C_{320}U_{321}$ bulge; and Larp7 CTD $L_{α3–α4}$ interacts along the SL4 major groove to position α4 in the tripartite MePCE-Larp7 interface (Figure 5C). The $C_{320}U_{321}$ bulge interactions explain previous biochemical data that showed that the inhibition of MePCE activity by Larp7 is CU bulge dependent (Brogie and Price, 2017). In summary, MePCE transitions from catalytically functional to structurally essential in complexes with Larp7 and 7SK.

A 7SK conformational switch during core RNP assembly

To determine which conformation(s) the full-length 7SK adopts during MePCE capping and core RNP assembly, we used *in vitro* methyltransferase assays to compare capping activity on WT 7SK with various RNA constructs that are locked into either linear or circular conformations (Figure S7D). As previously established, MePCE$_{MT}$ exhibits two distinct kinetic steps for its cognate substrates: a fast single-turnover that completes in around 2 min followed by a slow product release step (Yang et al., 2019) (Figures S1B and S1C). Based on this, we chose two time points, 2 and 25 min, to represent the single-turnover and multiple-turnover efficiency. We first interrogated the kinetic behavior of lin^m and circ^m RNA in the absence of Larp7, using SL1p-L1 as a reference. Under single-turnover conditions, circ^m has 22% lower methylation than lin^m (Figure S7A). In contrast, for multiple-turnovers circ^m has ~2-fold higher methylation than lin^m (Figure 6A). These results indicate that lin^m is both a slightly better enzymatic substrate for MePCE (single-turnover results) and a much better retained product upon methylation (multiple-turnover results) than circ^m.

We then asked what effect the assembly of core RNP has on these two kinetic steps. Surprisingly, incubating MePCE with Larp7 prior to adding RNA substrates has a very limited effect on the single-turnover methyltransfer step (Figure S7A), in contradiction to a previous proposal that Larp7 induces complete MePCE inactivation (Brogie and Price, 2017). Instead,

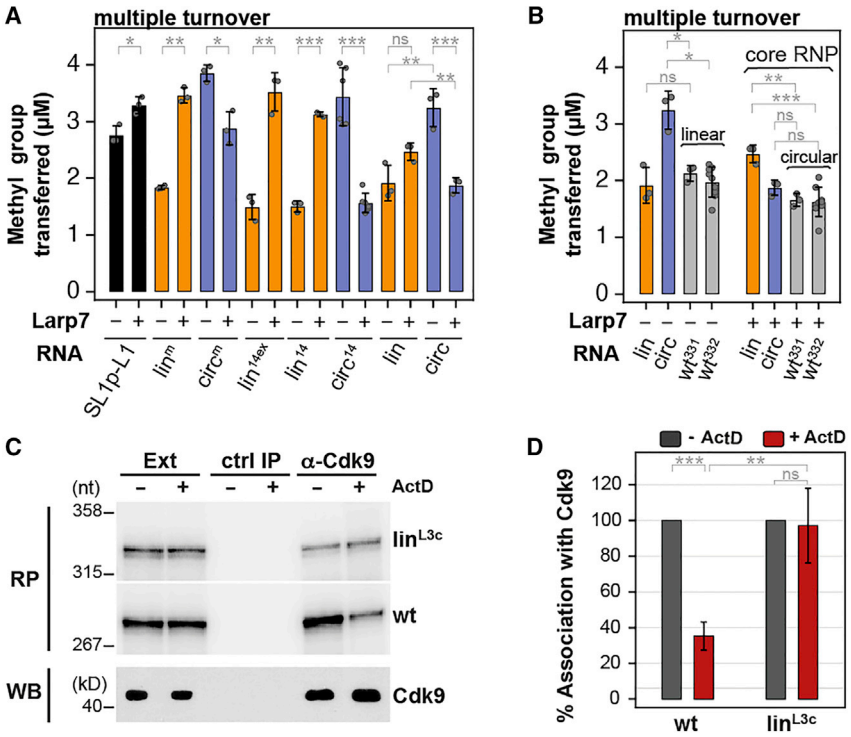


Figure 6. Circular conformation is formed by Larp7 RNP chaperone activity during biogenesis and is required for P-TEFb release

(A) Methyltransferase assays showing that preincubation with Larp7 increases multiple-turnover toward linear 7SK constructs (lin^{m}, $lin^{14–ms2}$, lin^{14}, and lin) (orange) but decreases toward circular 7SK constructs ($circ^{m}$, $circ^{14}$, and circ) (purple).

(B) Methyltransferase assays showing that wild-type (WT^{331} and WT^{332}) 7SK behave like linear 7SK (lin) in the absence of Larp7 but like circular 7SK (circ) in the presence of Larp7 (core RNP). Data are represented as mean ± STD.

(C) Association of lin^{L3c} with Cdk9 tested by co-immunoprecipitation in HeLa cells with or without ActD treatment. Endogenous 7SK (WT) and lin^{L3c} were detected by RNase protection assay. The amount of Cdk9 was analyzed by western blot (WB).

(D) Quantification of Cdk9-associated RNA reveals that lin^{L3c} does not release P-TEFb in response to ActD treatment like the endogenous 7SK. Error bars represent standard deviation from four independent experiments. Statistical p values from t test between two data sets are denoted by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and ns, no significant.

See also Figure S7.

MePCE is only inhibited once stoichiometric methylation (single-turnover) is complete and the core RNP is fully assembled. Given these results, for multiple-turnover conditions, a higher methylation efficiency would indicate an increased product release rather than increased single-turnover enzymatic activity. The addition of Larp7 only has limited effects on multiple-turnover for SL1p-L1 (Figure 6A). However, preincubation with Larp7 results in a significant increase in multiple-turnover for lin^{m} but a decrease with $circ^{m}$ (Figure 6A). This result suggests that for lin^{m}, which is the optimally retained product for $MePCE_{MT}$ in the absence of Larp7, the formation of the core RNP results in a more rapidly vacated methyltransferase active site than with $MePCE_{MT}$ alone, consistent with the lack of stable MePCE-mpppG-L4 interaction in the lin^{m} RNP structure. In contrast, for $circ^{m}$, when Larp7 is present, $MePCE_{MT}$ is retained longer on the product. Similar results were observed if Larp7 is added during the methyltransferase reactions instead of prior incubation, suggesting the effect is RNA dependent (Figure S7B). This pattern of activities is fully recapitulated for five longer versions of locked linear and circular 7SK constructs (Figures 6A and S7). For full-length 7SKs (nt 1–331 and 1–332, two lengths reported in the literature; Gürsoy et al., 2000; Wassarman and Steitz, 1991), the multiple-turnover values in the absence of Larp7 are the same as that for lin, but in the presence of Larp7 are close to circ (Figure 6B). Taken together, these methyltransferase studies, although indirect, support a model for 7SK RNP biogenesis where in the absence of Larp7, the linear 7SK conformation is the preferred MePCE methyltransferase substrate and retains MePCE better than the circular conformation, but assembly of $MePCE_{MT}$-7SK-Larp7 favors a conformational change in 7SK to the circular conformation, resulting in a more stable core RNP in terms of RNA degradation since both ends of the RNA are protected (Figure 1).

The circular conformation is required for P-TEFb release

The above results suggest that a nascent linear-to-circular conformational switch takes place during biogenesis, resulting in 7SK-dependent MePCE-Larp7 interaction. As proposed, the circular conformation would not be expected to stably bind Hexim/P-TEFb since it contains only one of the two binding sites for the homodimeric Hexim (Brogie and Price, 2017; Egloff et al., 2006; Marz et al., 2009). To test for and gain insight into the *in vivo* function of the circular conformation, we investigated the dynamic interaction of P-TEFb with a 7SK construct locked into a linear conformation. Specifically, we substituted the linker 3 with its complementary sequence (lin^{L3c}), which prevents the circular conformation due to the disruption of S1alt. (We note that substitutions in the 5′ end sequence would disrupt initial MePCE recognition during 7SK biogenesis.) Hela cells transiently expressing 7SK lin^{L3c} were treated with ActD, and P-TEFb was immunoprecipitated with an anti-Cdk9 antibody from the cellular extracts. It has been established that the inhibition of global cellular RNA synthesis with Actinomycin D (ActD) treatment increases the nuclear level of active P-TEFb through inducing Hexim/P-TEFb release from the 7SK RNP (Nguyen et al., 2001; Yang et al., 2001). Co-purified 7SK amounts were quantified by RNase A/T1 protection analysis for both the transiently expressed lin^{L3c} and the endogenous WT 7SK. Consistent with previous studies, ActD treatment reduced the amount of

endogenous 7SK-associated with Cdk9 by 60% (Figures 6C and 6D). In contrast, the amount of lin^{L3c}-associated with Cdk9 did not change upon ActD treatment. This indicates that 7SK core/Hexim/P-TEFb RNP assembled with lin^{L3c} is resistant to ActD-induced disassembly. The same results were obtained when comparing transiently expressed WT with lin^{L3c} in 7SK knockout (KO) cells (Figure S7). We conclude that the formation of the long-range S1alt in the circular 7SK conformation is essential for efficient stress-induced release of active P-TEFb from the 7SK core/Hexim/P-TEFb transcriptional regulatory RNP.

DISCUSSION

The structures of 7SK core RNPs with linear and circular RNA conformations reported here, together with biochemical data, reveal a stepwise mechanism for the early RNP biogenesis events, where MePCE preferably caps and retains the nascent linear 7SK during RNA Pol III transcription, but the binding of Larp7 triggers the formation of circular 7SK with MePCE bound to Larp7 at the 3′-hairpin. Significantly, this MePCE-7SK-Larp7 core RNP structure is maintained with full-length 7SK, which adopts a circular conformation in the core RNP, and with a locked linear 7SK conformation. Thus, MePCE and Larp7 can remain constitutively associated with 7SK during and after conformational switching between linear and circular RNA. In contrast, the reversibly assembled components (Hexim, P-TEFb, heterogeneous nuclear RNP or hnRNPs, etc.) of 7SK RNPs may require specific binding sites only present in the linear or circular conformation. Specifically, Hexim1 has been shown to interact with two sites in the linear SL1 (Muniz et al., 2010), but only one of the two sites is present in the SL1-distal region of the circular conformation (Figure 7A). It has also been well established by chemical probing (Brogie and Price, 2017; Wassarman and Steitz, 1991) and cellular functional assays (Egloff et al., 2006) that the Hexim/P-TEFb-bound 7SK is in the linear conformation. *In vitro* binding assays also indicated that Hexim binds only weakly to the circular alternative SL1, which has only one GAUC motif (Czudnochowski et al., 2010; Muniz et al., 2010). This is consistent with the higher affinity and specificity of homodimeric Hexim for two binding sites, as commonly seen for dimeric DNA-binding proteins. Thus, we propose that another switch (shown here as circular-to-linear) must occur that exposes the complete Hexim1 homodimer binding sites for the association of Hexim/P-TEFb with the core RNP (Figure 7A). Consistent with this, 7SK linkers L3 and L4 are more chemically reactive when Hexim1 is bound (Flynn et al., 2016). Our in-cell experiments support the model that the reverse switching must also occur, since the circular conformation is required for P-TEFb release in response to transcriptional stress. Since around half of cellular P-TEFb is bound to 7SK (Nguyen et al., 2001; Yang et al., 2001), but a significant portion of 7SK remains P-TEFb-free (Haaland et al., 2003), hnRNPs and other transcription factors may stabilize the P-TEFb-free 7SK through selective binding to the circular conformation, thus establishing a regulatory mechanism using 7SK conformational remodeling. While this manuscript was under review, Olson et al. (2021) reported in BioRxiv a new method, DANCE-MaP, to deconvolute the ensemble SHAPE (Merino et al., 2005) data for the underlying 7SK RNA conformations. Their data support an equilibrium between two predominant 7SK conformations, "A" and "B," for both free 7SK *in vitro* and 7SK RNPs in cell, which they propose correspond to P-TEFb-bound (A state; Figure 7A, linear dashed line) and P-TEFb-free (B state; Figure S7E), consistent with our results. We note that while both their secondary structure models include a short helix between the 5′ end and L3, there is no probing data for the 5′ end, and L3 is only partially protected. Importantly, our results do not preclude such alternative conformations as long as the "linear" conformation contains an SL1 that includes the homodimeric Hexim-binding sites (Brogie and Price, 2017; Egloff et al., 2006; Wassarman and Steitz, 1991) (Figure 7A). Our results establish that a transition from nascent linear to a circular conformation is required for formation of the core RNP and show how MePCE and Larp7 interact and can remain constitutively assembled during switching between P-TEFb-free and P-TEFb-bound RNPs, even during extensive RNA conformational remodeling processes.

Interestingly, MePCE undergoes two distinct conformational changes in its RNA-binding site during 7SK RNP assembly: first, the binding of the nascent linear 7SK leads to the disorder-to-order transition of two helices to form the active site (Yang et al., 2019); second, RNA-dependent MePCE interaction with Larp7 at the 3′-hairpin disrupts the methyltransferase active site, leading to MePCE inhibition in the core RNP for both circular and linear 7SK (Figure 4). The structure of full-length Larp7 revealed that nearly the entire protein is structured when bound to RNA. In contrast to a previous report (Uchikawa et al., 2015), RRM1 adopts a conventional fold followed by a C-terminal α3. RRM1 α3 and a CTD α4, which are partially disordered in the absence of RNA, participate in both RNA recognition and MePCE-Larp7 interaction, revealing that Larp7 functions as both an RNA and RNP chaperone. Most of the conserved hydrophobic Larp7 residues involved in MePCE binding are not present in human genuine La protein or *Tetrahymena* p65, consistent with Larp7 replacing La protein to form the 7SK core RNP and Larp7 specificity for 7SK RNP (Muniz et al., 2013) (Figure 7B).

In conclusion, the 7SK core RNP structures presented here provide a new paradigm for understanding biogenesis and dynamic regulation processes of 7SK RNPs as well as non-coding RNPs in general. In other RNPs, LARP7s might also have dual roles, both as RNA chaperones and in RNP assembly. The structures reveal that the constitutively assembled 7SK core RNP provides a flexible scaffold for protein assembly and disassembly during nucleoplasmic P-TEFb regulation. The flexible scaffold could also help the direct recruitment of 7SK RNPs to certain cellular gene promoters and the HIV viral genome prior to the stimulation of transcription elongation (D'Orso and Frankel, 2010; Ji et al., 2013).

Limitations of the study

This study provides evidence that circular 7SK conformation is required for P-TEFb release during cellular stress response; our data do not directly address the conformation of SL1-SL3 outside of the MePCE and Larp7 binding sites. Although our

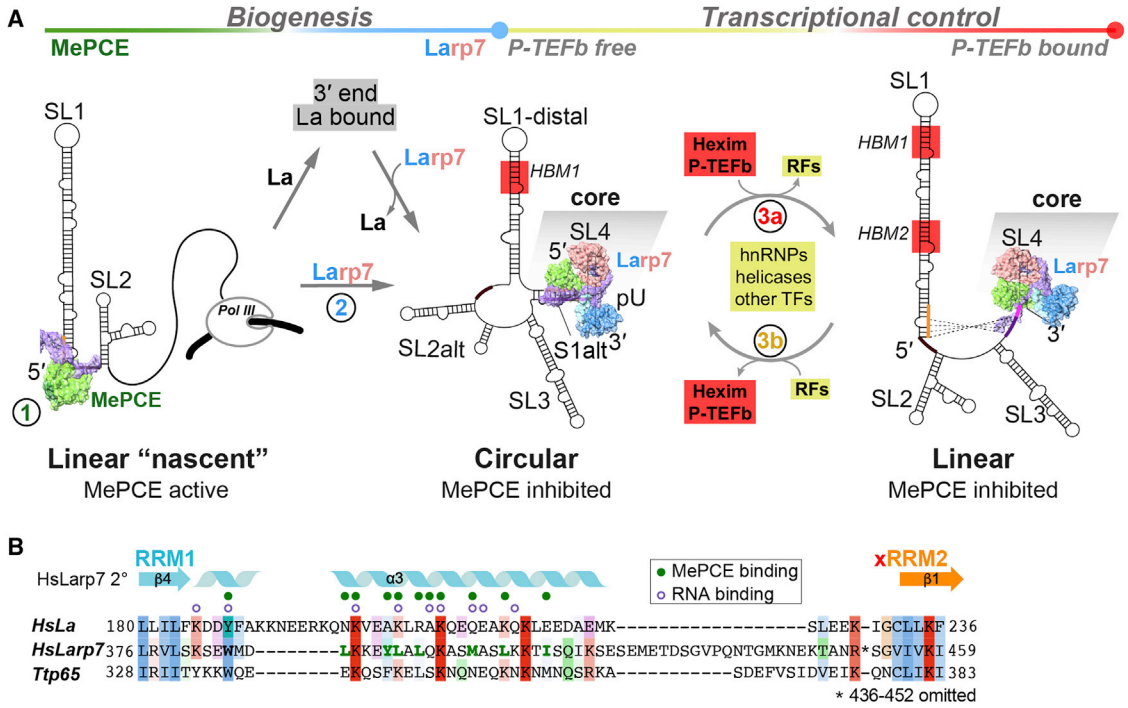


Figure 7. A model for 7SK conformational switches in RNP biogenesis and transcriptional regulation

(A) Proposed 7SK RNP biogenesis assembly pathway and transcription control conformations. The two core RNP structures from this study are shown and highlighted by gray shading. 1. During biogenesis, MePCE binds and methylates the nascent linear 7SK during transcription by RNA Pol III, and genuine La is thought to initially protect the 3′ end. Larp7 subsequently replaces La protein. 2. Interaction of Larp7 with MePCE triggers circularization of 7SK RNA. 3a. For RNA Pol II pause release transcriptional control, Hexim1 binding prefers a linear conformation of 7SK (P-TEFb bound), where two high-affinity binding sites exist in SL1 for each of the monomers in the Hexim homodimer. HBM: Hexim-binding motif. 3b. Switching back to a circular conformation allows for P-TEFb release, which is triggered and/or stabilized by release factors, including hnRNPs, helicases, and other transcription factors. The dashed lines in the P-TEFb-bound linear conformation denotes an alternative conformation that is still compatible with Hexim1 and P-TEFb binding. Full-length secondary structures of circular and linear 7SK RNA are from previous studies (Bubenik et al., 2020; Marz et al., 2009; Wassarman and Steitz, 1991).

(B) Structure-based sequence alignment of human La protein, human Larp7, and *Tetrahymena thermophila* p65. Secondary structural elements of human Larp7 are labeled at the top. Residues interacting with SL4 RNA are marked by purple open circles and with $MePCE_{MT}$ are marked with green closed circles.

See also Figure S7.

extensive structural and biochemical analyses support the cryo-EM structures presented in this study, we cannot rule out the possibility that other forms of the core complex exist in the discarded particles due to the dynamic nature of 7SK RNA. Future structural studies of 7SK core RNP with Hexim/P-TEFb bound are required to resolve this question for the P-TEFb-bound 7SK RNP. The potential roles of the mutually exclusive P-TEFb release factors on conformational switching are also not addressed here.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2022.03.001.

ACKNOWLEDGMENTS

This work was supported by NIH R35GM131901 and NIH R01AI155170 grants to J.F., R01GM071940 grant to Z.H.Z., INSERM Plan Cancer and ANRS grants to T.K., American Heart Association Postdoctoral Fellowship 20POST35210850 to Y.Y., and American Cancer Society Postdoctoral Fellowship 126777-PF-14-179-01-DMC to C.D.E. We acknowledge the use of instruments at the Electron Imaging Center for Nanomachines, supported by UCLA and by instrumentation grants from NIH (1S10RR23057, 1S10OD018111, and U24GM116792), NSF (DBI-1338135 and DMR-1548924), and the use of NMR equipment, supported by grants from NIH (S10OD016336 and S10OD025073) and from DOE (DE-FC02-02ER63421).

AUTHOR CONTRIBUTIONS

J.F. and Z.H.Z. supervised the project. Y.Y. and C.D.E. prepared the RNA and EM samples. Y.Y. performed methyltransferase assays. S.L. collected and processed the EM data and built the atomic models. Y.Y., C.D.E., T.H., and J.Z. performed and analyzed NMR experiments. S.E. performed cellular stress response experiments. J.Z. helped with ITC experiments and negative stain EM data collection. Y.Y., S.L., J.F., and C.D.E. analyzed the models. Y.Y., S.L., and S.E. prepared the illustrations. Y.Y. and J.F. wrote the manuscript. All authors contributed to the editing of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 26, 2021
Revised: December 27, 2021
Accepted: February 28, 2022
Published: March 22, 2022

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. *74*, 531–544.

Aigner, S., Lingner, J., Goodrich, K.J., Grosshans, C.A., Shevchenko, A., Mann, M., and Cech, T.R. (2000). Euplotes telomerase contains an La motif protein produced by apparent translational frameshifting. EMBO J. *19*, 6230–6239.

Bayfield, M.A., and Maraia, R.J. (2009). Precursor-product discrimination by La protein during tRNA metabolism. Nat. Struct. Mol. Biol. *16*, 430–437.

Brogie, J.E., and Price, D.H. (2017). Reconstitution of a functional 7SK snRNP. Nucleic Acids Res. *45*, 6864–6880.

Bubenik, J.L., Hale, M., McConnell, O., Wang, E.T., Swanson, M.S., Spitale, R.C., and Berglund, J.A. (2020). RNA structure probing to characterize RNA–protein interactions on a low abundance pre-mRNA in living cells. RNA *27*, 343–358.

Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H., and Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy *135*, 24–35.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Collopy, L.C., Ware, T.L., Goncalves, T., Í Kongsstovu, S., Yang, Q., Amelina, H., Pinder, C., Alenazi, A., Moiseeva, V., Pearson, S.R., et al. (2018). LARP7 family proteins have conserved function in telomerase assembly. Nat. Commun. *9*, 557.

Core, L., and Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. Genes Dev. *33*, 960–982.

Crowley, P.B., Kyne, C., and Monteith, W.B. (2012). Simple and inexpensive incorporation of ^{19}F-tryptophan for protein NMR spectroscopy. Chem. Commun. (Camb) *48*, 10681–10683.

Czudnochowski, N., Vollmuth, F., Baumann, S., Vogel-Bachmayr, K., and Geyer, M. (2010). Specificity of Hexim1 and Hexim2 complex formation with cyclin T1/T2, Importin α and 7SK snRNA. J. Mol. Biol. *395*, 28–41.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on Unix pipes. J. Biomol. NMR *6*, 277–293.

D'Orso, I., and Frankel, A.D. (2010). RNA-mediated displacement of an inhibitory snRNP complex activates transcription elongation. Nat. Struct. Mol. Biol. *17*, 815–821.

Egloff, S., Studniarek, C., and Kiss, T. (2018). 7SK small nuclear RNA, a multifunctional transcriptional regulatory RNA with gene-specific features. Transcription *9*, 95–101.

Egloff, S., Van Herreweghe, E., and 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.

Eichhorn, C.D., Chug, R., and Feigon, J. (2016). hLARP7 C-terminal domain contains an xRRM that binds the 3′ hairpin of 7SK RNA. Nucleic Acids Res. *44*, 9977–9989.

Eichhorn, C.D., Yang, Y., Repeta, L., and Feigon, J. (2018). Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7. Proc. Natl. Acad. Sci. USA *115*, E6457–E6466.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Flynn, R.A., Do, B.T., Rubin, A.J., Calo, E., Lee, B., Kuchelmeister, H., Rale, M., Chu, C., Kool, E.T., Wysocka, J., et al. (2016). 7SK-BAF axis controls pervasive transcription at enhancers. Nat. Struct. Mol. Biol. *23*, 231–238.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25.

Goodall, G.J., and Filipowicz, W. (1990). The minimum functional length of pre-mRNA introns in monocots and dicots. Plant Mol Biol *14*, 727–733. https://doi.org/10.1007/BF00016505.

Guillerez, J., Lopez, P.J., Proux, F., Launay, H., and Dreyfus, M. (2005). A mutation in T7 RNA polymerase that facilitates promoter clearance. Proc. Natl. Acad. Sci. USA *102*, 5958–5963.

Gurney, T., Jr., and Eliceiri, G.L. (1980). Intracellular distribution of low molecular weight RNA species in HeLa cells. J. Cell Biol. *87*, 398–403.

Gürsoy, H.C., Koper, D., and Benecke, B.J. (2000). The vertebrate 7S K RNA separates hagfish (*Myxine glutinosa*) and lamprey (*Lampetra fluviatilis*). J. Mol. Evol. *50*, 456–464.

Haaland, R.E., Herrmann, C.H., and Rice, A.P. (2003). Increased association of 7SK snRNA with Tat cofactor P-TEFb following activation of peripheral blood lymphocytes. AIDS *17*, 2429–2436.

He, N., Jahchan, N.S., Hong, E., Li, Q., Bayfield, M.A., Maraia, R.J., Luo, K., and Zhou, Q. (2008). A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. Mol. Cell *29*, 588–599.

Jády, B.E., and Kiss, T. (2000). Characterisation of the U83 and U84 small nucleolar RNAs: two novel 2′-O-ribose methylation guide RNAs that lack complementarities to ribosomal RNAs. Nucleic Acids Res. *28*, 1348–1354.

Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Thérien, C., Bergeron, D., Bourassa, S., Greenblatt, J., 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.

Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C.Y., Xiao, R., Burge, C.B., and Fu, X.D. (2013). SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell *153*, 855–868.

Kotik-Kogan, O., Valentine, E.R., Sanfelice, D., Conte, M.R., and Curry, S. (2008). Structural analysis reveals conformational plasticity in the recognition of RNA 3′ ends by the human La protein. Structure *16*, 852–862.

Krueger, B.J., Jeronimo, C., Roy, B.B., Bouchard, A., Barrandon, C., Byers, S.A., Searcey, C.E., Cooper, J.J., Bensaude, O., Cohen, E.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.

Krueger, B.J., Varzavand, K., Cooper, J.J., and Price, D.H. (2010). The mechanism of release of P-TEFb and HEXIM1 from the 7SK snRNP by viral and cellular activators includes a conformational change in 7SK. PLoS One *5*, e12335.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nat. Methods *11*, 63–65.

Lee, W., Tonelli, M., and Markley, J.L. (2015). NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics *31*, 1325–1327.

Maraia, R.J., Mattijssen, S., Cruz-Gallardo, I., and Conte, M.R. (2017). The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives. Wiley Interdiscip. Rev. RNA *8*, e1430.

Maris, C., Dominguez, C., and Allain, F.H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS Journal *272*, 2118–2131.

Marz, M., Donath, A., Verstraete, N., Nguyen, V.T., Stadler, P.F., and Bensaude, O. (2009). Evolution of 7SK RNA and its protein partners in metazoa. Mol. Biol. Evol. *26*, 2821–2830.

Mastronarde, D.N. (2003). SerialEM: a program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. Microsc. Microanal. *9*, 1182–1183.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

Mennie, A.K., Moser, B.A., and Nakamura, T.M. (2018). LARP7-like protein Pof8 regulates telomerase assembly and poly(A)+TERRA expression in fission yeast. Nat. Commun. *9*, 586.

Merino, E.J., Wilkinson, K.A., Coughlan, J.L., and Weeks, K.M. (2005). RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). J. Am. Chem. Soc. *127*, 4223–4231.

Muniz, L., Egloff, S., and Kiss, T. (2013). RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. Nucleic Acids Res. *41*, 4686–4698.

Muniz, L., Egloff, S., Ughy, B., Jády, B.E., and Kiss, T. (2010). Controlling cellular P-TEFb activity by the HIV-1 transcriptional transactivator Tat. PLoS Pathog. *6*, e1001152.

Nguyen, V.T., Kiss, T., Michels, A.A., and Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. Nature *414*, 322–325.

Olson, S.W., Turner, A.-M.W., Arney, J.W., Saleem, I., Weidmann, C.A., Margolis, D.M., Weeks, K.M., and Mustoe, A.M. (2021). Large-scale allosteric switch in the 7SK RNA regulates transcription in response to growth and stress. bioRxiv, 2021.2009.2016.460563.

Páez-Moscoso, D.J., Pan, L., Sigauke, R.F., Schroeder, M.R., Tang, W., and Baumann, P. (2018). Pof8 is a La-related protein and a constitutive component of telomerase in fission yeast. Nat. Commun. *9*, 587.

Peterlin, B.M., Brogie, J.E., and Price, D.H. (2012). 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. Wiley Interdiscip. Rev. RNA *3*, 92–103.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera – a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Price, D.H. (2000). P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. Mol. Cell. Biol. *20*, 2629–2634.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. *333*, 721–745.

Sedore, S.C., Byers, S.A., Biglione, S., Price, J.P., Maury, W.J., and Price, D.H. (2007). Manipulation of P-TEFb control machinery by HIV: recruitment of P-TEFb from the large form by Tat and binding of HEXIM1 to TAR. Nucleic Acids Res. *35*, 4347–4358.

Shen, Y., and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J. Biomol. NMR *56*, 227–241.

Singh, M., Wang, Z., Koo, B.K., Patel, A., Cascio, D., Collins, K., and Feigon, J. (2012). Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme la family protein p65. Mol. Cell *47*, 16–26.

Studniarek, C., Tellier, M., Martin, P.G.P., Murphy, S., Kiss, T., and Egloff, S. (2021). The 7SK/P-TEFb snRNP controls ultraviolet radiation-induced transcriptional reprogramming. Cell Rep. *35*, 108965.

Teplova, M., Yuan, Y.R., Phan, A.T., Malinina, L., Ilin, S., Teplov, A., and Patel, D.J. (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.

Uchikawa, E., Natchiar, K.S., Han, X., Proux, F., Roblin, P., Zhang, E., Durand, A., Klaholz, B.P., and Dock-Bregeon, A.C. (2015). Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. Nucleic Acids Res. *43*, 3373–3388.

Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., and Cramer, P. (2018). Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature *560*, 607–612.

Wassarman, D.A., and Steitz, J.A. (1991). Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. Mol. Cell. Biol. *11*, 3432–3445.

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

Xue, Y., Yang, Z., Chen, R., and 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.

Yang, Y., Eichhorn, C.D., Wang, Y., Cascio, D., and Feigon, J. (2019). Structural basis of 7SK RNA 5′-gamma-phosphate methylation and retention by MePCE. Nat. Chem. Biol. *15*, 132–140.

Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. Nature *414*, 317–322.

Yik, J.H., Chen, R., Nishimura, R., Jennings, J.L., Link, A.J., and Zhou, Q. (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.

Zhang, K. (2016). Gctf: real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M.B., and Price, D.H. (1997). Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. Genes Develop. *11*, 2622–2632.

Zieve, G., and Penman, S. (1976). Small RNA species of the HeLa cell: metabolism and subcellular localization. Cell *8*, 19–31.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*, e42166.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-Cdk9	Abcam	Cat#ab6544; RRID: AB_305557
Rabbit monoclonal anti-Cdk9	Cell Signaling Technology	Cat#2316; RRID: AB_2291505
Anti-rabbit IgG HRP-linked antibody	Promega	Cat#W401B; RRID: AB_430833
Bacterial and virus strains		
E. coli BL21-Gold(DE3) competent cells	Agilent Technologies	Cat #230132
E. coli DH5α competent cells	ThermoFisher	Cat #18265017
Chemicals, peptides, and recombinant proteins		
Q5 mutagenesis kit	NEB	Cat # E0554S
$^{15}NH_4Cl$	Cambridge Isotope Laboratories	Cat # NLM-467-1
[^{13}C-6]-D-glucose	Cambridge Isotope Laboratories	Cat # CLM-1396-1
S-(5′-adenosyl)-l-methionine p-toluenesulfonate salt (SAM)	Sigma	Cat # A2408
S-(5′-adenosyl)-l-[methyl-^{3}H]methionine ([me^{3}H]SAM)	PerkinElmer Life Sciences	Cat # NET155V001MC
Protease Inhibitor Cocktail	ThermoFisher	Cat # A32965
TEV protease	In house	N/A
HiTrap Q HP, 5 mL	GE Healthcare	Cat #17115401
Superdex 75 HiLoad 26/600 column	GE Healthcare	Cat #28989334
T7 RNA polymerase P266L mutant	Guillerez et al. (2005)	N/A
Spermidine	Sigma-Aldrich	Cat. # S2626
BamHI-HF restriction enzyme	NEB	Cat # R3136M
DraI restriction enzyme	NEB	Cat # R0129S
T4 Polynucleotide Kinase	NEB	Cat # M0201L
Dulbecco's Modified Eagle Medium (DMEM)	GIBCO, Invitrogen	Cat # 61965-026
Iscove's Modified Dulbecco Medium (IMDM)	GIBCO, Invitrogen	Cat # 12440-053
Fetal Bovine Serum (FBS)	PAN-Biotech	Cat # P30-3306
Penicillin and Streptomycin	GIBCO, Invitrogen	Cat # 15140-122
Actinomycin D	Sigma	Cat # A9415
[α-^{32}P]CTP	Hartmann Analytic	Cat # SRP-809
[γ-^{32}P]ATP	Hartmann Analytic	Cat # SRP-501
Deposited data		
Cryo-EM structure of 7SK lin^m RNP complex	This paper	EMDB: EMD- 25197, PDB: 7SLP
Cryo-EM structure of 7SK circ^m RNP complex	This paper	EMDB: EMD-25198, PDB: 7SLQ
NMR chemical shift assignments of Larp7 $LaMod_{22\text{-}390\Delta L}$	This paper	BMRB: 51253
NMR chemical shift assignments of Larp7 β4-$xRRM2_{370\text{-}582}$	This paper	BMRB: 51254
Experimental models: Cell lines		
Hela cells	Jády and Kiss (2000)	N/A
HAP1 7SK KO cells	Studniarek et al. (2021)	N/A
Oligonucleotides		
Primers for cloning and point mutation, see Table S2	This paper	N/A
DNA templates for in vitro RNA transcription, see Table S3	This paper	N/A
Recombinant DNA		
pET28a-His_6-$MePCE_{MT}$	Addgene	plasmid #25304
pETDuet-1-His_6-MBP-Larp7	Eichhorn et al. (2018)	N/A
pET30-His_6-Larp7(1-196)	Eichhorn et al. (2018)	N/A

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
pET30-His$_6$-Larp7(445-561)	Eichhorn et al. (2018)	N/A
pET30-His$_6$-Larp7(445-582)	Eichhorn et al. (2018)	N/A
pETDuet-1-His$_6$-MBP-Larp7(1-208)	This paper	N/A
pETDuet-1-His$_6$-MBP-Larp7(1-390)	This paper	N/A
pETDuet-1-His$_6$-MBP-Larp7(1-390) ΔL	This paper	N/A
pETDuet-1-His$_6$-MBP-Larp7(22-390) ΔL	This paper	N/A
pETDuet-1-His$_6$-MBP-Larp7(370-582)	This paper	N/A
pUC57-7SK-wt(1-331)	This paper	N/A
pUC57-7SK-wt(1-332)	This paper	N/A
pUC57-7SK-lin^{14}	This paper	N/A
pUC57-7SK-lin^{14ex}	This paper	N/A
pUC57-7SK-circ^{14}	This paper	N/A
pUC57-7SK-lin	This paper	N/A
pUC57-7SK-circ	This paper	N/A
p7SK-lin^{L3c}	This paper	N/A
Software and algorithms		
RELION 3.0	Zivanov et al. (2018)	https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page
cryoSPARC v2	Punjani et al. (2017)	https://cryosparc.com/
CTFFIND4	Rohou and Grigorieff (2015)	https://grigoriefflab.umassmed.edu/ctffind4
MotionCor2	Zheng et al. (2017)	https://emcore.ucsf.edu/ucsf-software
Phenix	Afonine et al. (2018)	https://phenix-online.org/
Coot	Emsley et al. (2010)	https://www2.mrc-lmb.cam.ac.uk/
MolProbity	Chen et al. (2010)	http://molprobity.biochem.duke.edu
SerialEM	Mastronarde (2003)	https://bio3d.colorado.edu/SerialEM/
UCSF Chimera	Pettersen et al. (2004)	https://www.cgl.ucsf.edu/chimera/
UCSF ChimeraX	Goddard et al. (2018)	https://www.cgl.ucsf.edu/chimerax/
PyMol	Schrödinger, LLC	https://pymol.org/2/
TopSpin	Bruker	
nmrPipe	Delaglio et al. (1995)	https://www.ibbr.umd.edu/nmrpipe/
NMRFAM-Sparky	Lee et al. (2015)	https://nmrfam.wisc.edu/nmrfam-sparky-distribution/
IgorPro	WaveMetrics	https://www.wavemetrics.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Juli Feigon (feigon@mbi.ucla.edu).

Materials availability

We will distribute the plasmids freely to the scientific community upon request.

Data and code availability

- Cryo-EM maps for 7SK lin^m and circ^m RNPs have been deposited in the Electron Microscopy Data Bank (EMDB) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Coordinates for models of 7SK lin^m and circ^m RNPs have been deposited I the Protein Data Bank (PDB) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. NMR chemical shift assignments of Larp7 LaMod$_{22\text{-}390\Delta L}$ and Larp7 β4-xRRM2$_{370\text{-}582}$ have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.

- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

E. coli DH5α strain was incubated at 37°C for 0.5 day during cloning experiments. E. coli BL21-Gold(DE3) strain was incubated at 37°C for 0.5 day during transformation or initial growth, and incubated at 18°C for 18-22hrs during protein expression. Hela cells and HAP1 7SK KO cells were grown in Dulbecco's Modified Eagle's medium (DMEM) and (IMDM), respectively, supplemented with 10% (v/v) fetal bovine serum and 100 ug/ml of penicillin and streptomycin (GIBCO, Invitrogen). The cells were routinely tested for the presence of mycoplasma using MycoAlert (Lonza).

METHOD DETAILS

Plasmids and mutagenesis

The pET28a-LIC recombinant plasmid with gene encoding the hexahistidine-tagged (His_6-tag) human MePCE (FLJ20257) methyltransferase domain (residues 400–689) was a gift from C. Arrowsmith (Addgene plasmid 25304). Full-length human Larp7 (aa 1-582) was cloned into a pETDuet-1 vector with N-terminal His_6-tagged maltose binding protein (MBP) fusion followed by a Tobacco Etch Virus (TEV) protease cleavage site between MBP and Larp7. Minimum La module construct $LaMod_{196}$ (aa 1-196) and minimum xRRM2-CTD constructs (aa 445-561 and 445-582) was cloned into a pET30 Xa/LIC vector with an N-terminal His_6-tag and a TEV protease cleavage site. Additional La module and xRRM2 constructs of Larp7, including $LaMod_{208}$ (aa 1-208), $LaMod_{390}$ (aa 1-390), β4−xRRM2 (aa 370-582), were cloned by site-directed mutagenesis using the pETDuet-MBP-Larp7 as template. The α2−β4 loop deletion (aa 197-364) La module constructs, $LaMod_{390\Delta L}$ and $LaMod_{22\text{-}390\Delta L}$, were cloned by site-directed mutagenesis using the pETDuet-MBP-$LaMod_{390}$ as template.

Full-length wild-type 7SK gene and locked linear 7SK (lin^{14}) gene were cloned into pUC19 vector with a Hammerhead ribozyme gene at the 3′ end or into pUC57 vector with HDV ribozyme at the 3′ end. The pUC57-HDV vector was a gift from Samuel Butcher. DraI restriction site was inserted between 7SK and HDV ribozyme gene for full-length 7SK in vitro transcription. All site-directed mutagenesis was performed using Q5 kit (New England Biolabs).

Protein expression and purification

Expression and purification of MePCE methyltransferase domain ($MePCE_{MT}$), Larp7 xRRM2 constructs and full-length Larp7 were performed as described previously (Eichhorn et al., 2018; Yang et al., 2019). Larp7 La module constructs in either pET30 Xa/LIC or pETDuet-MBP vectors were expressed using BL21-GOLD(DE3) competent cells. The cells were grown at 37°C in M9 minimal media to OD_{600} of 0.6-0.8, induced with 0.5 mM IPTG and expressed for 18-22 hrs at 18°C. For uniformly [^{15}N, ^{13}C] or [^{15}N] enriched proteins, M9 minimal media containing 1 g/L of $^{15}NH_4Cl$ and 3 g/L of [^{13}C-6]-D-glucose (or natural abundance glucose) (Cambridge Isotope Laboratories) were used. Pelleted cells were resuspended in lysis buffer (20 mM HEPES, pH 8, 1 M NaCl, 15 mM Imidazole, 5% w/v glycerol, 1 mM TCEP and 1 mM PMSF) supplemented with lysozyme and lysed by sonication on ice. Cell lysate was clarified by centrifugation and filtration, and was loaded on to a 5-mL Ni Sepharose affinity column (HisTrap-HP; GE Healthcare). The His_6-tagged proteins were eluted with a linear Imidazole gradient to 500 mM, pooled and dialyzed (buffer R, 20 mM HEPES, pH 7.5, 100 mM NaCl, 5 mM βME) with 1 mg of TEV protease to cleave off the His_6-tag or the His_6-MBP-tag. The dialyzed protein mixture was run on a 5-mL HiTrap Heparin HP column (GE) to separate the fusion tags and TEV protease from the target proteins. La module proteins were eluted with linear KCl gradient to 1 M and further purified with size-exclusion chromatography (SEC; HiLoad 26/600 Superdex 75; GE Healthcare) in 7SK binding buffer (buffer B, 20 mM HEPES, pH 7.5, 250 mM KCl, and 1 mM TCEP) or ITC buffer (buffer C, 20 mM HEPES, pH 7.5, 150 mM KCl, and 1 mM TCEP) at 4°C. Proteins were concentrated using Amicon devices (Millipore) and concentration was measured by absorbance at 280 nm.

In vitro transcription of RNA

RNA samples of SL4-ex, SL1p-L1, $circ^m$ and lin^m were prepared by in vitro transcription with T7 RNAP P266L mutant (Guillerez et al., 2005) and chemically synthesized DNA oligonucleotides (Integrated DNA Technologies) as previously described (Eichhorn et al., 2016). Transcription mixtures (40 mM Tris, pH 8, 1 mM spermidine, 0.01% Triton-X100, 2.5 mM DTT, 25 mM $MgCl_2$, 4 mM each rATP, rUTP, rGTP and rCTP, 0.5 μM DNA template, T7 RNAP P266L mutant, for lin^m 15% v/v DMSO was added in the transcription mixture) were incubated at 37°C for 4-6 hrs. Transcribed RNAs were purified on 10-20% denaturing polyacrylamide gel electrophoresis (Urea-PAGE) and electroeluted from the gel pieces using an Elutrap device (GE Whatman) in 1X TBE buffer (90 mM Tris-borate,2 mM EDTA, pH 8.3). Collected RNA eluents were concentrated and buffer exchanged into sterilized nanopure water using Amicon devices (Millipore), supplied with counterion in high salt buffer (10 mM sodium phosphate, pH 7.6, 1 mM EDTA, 1.5 M KCl) and buffer exchanged back into sterilized nanopure water. Diluted RNA (<100 μM) solutions were heated at 95°C for 5 min and snap

cool on ice for 1 hr. RNAs were then concentrated and stock solutions were stored in -20°C freezer. For ITC experiments, RNAs were buffer exchanged extensively into buffer C using Amicon devices.

RNA samples of wild-type 7SK, lin^{14}, lin^{14ex}, circ^{14}, lin and circ were prepared by the same in vitro transcription procedure as above, but with plasmid DNA template and modified condition. The pUC19-HH, pUC57-HDV and pUC57-DraI plasmids containing the respective DNA templates were purified by QIAGEN Plasmid Maxi Kit or Invitrogen PureLink HiPure Maxiprep kit, washed extensively with 70% EtOH and dissolved to 900-2000 ng/μL DNA concentration. The purified plasmids were digested with BamHI-HF or DraI restriction enzyme (New England Biolabs) at 37°C overnight and verified digestion on a 0.8% Agarose gel. 200 μL of digested DNA templates were added to 10 mL of transcription mixture. 15-20 mM $MgCl_2$ were used instead of 25 mM. T4 PNK (New England Biolabs) was added to the transcription mixtures in order to turn the 2′,3′-cyclic phosphate resulted from Hammerhead or HDV ribozyme self-cleavage into a native 3′-OH for the transcribed RNA. Steps for purification follow those described above.

Assembly of 7SK core RNP complex samples

To assemble 7SK core RNP complex samples for EM studies, MePCE$_{MT}$ and full-length Larp7 proteins are purified freshly, where fivefold molar excess of S-(5′-adenosyl)-l-methionine *p*-toluenesulfonate salt (SAM) (Sigma A2408; ≥ 80% purity) was added to stabilize the purified MePCE$_{MT}$. RNA solution was mixed with 1.2 or 1.5 molar equivalent of MePCE$_{MT}$ and incubated at 37°C for 5 min. After cooling on ice, 1.2 molar equivalent of Larp7 was added to the RNA/MePCE$_{MT}$ mixture to form the core RNP complex. Electrophoresis mobility shift assays (EMSA) were used to identify optimal protein-to-RNA ratios for homogenous and stoichiometric assembly. For cryo-EM samples, the assembled RNP complex was buffer exchanged into low salt buffer (20mM HEPES, pH 7.5, 1 mM TCEP and 55 mM KCl) using Amicon devices.

Methyltransferase assays

In vitro RNA methyltransferase assays were described previously (Yang et al., 2019) with modifications. Briefly, all experiments were performed in final reaction buffer of 50 mM HEPES, pH 7.5, 150 mM KCl and 1 mM TCEP and contained 1 μM MePCE$_{MT}$, 10 μM RNA substrates and 20 μM SAM mixture. SAM mixture is a 4.6:1 molar ratio mixture of SAM *p*-toluenesulfonate salt (Sigma) to S-(5′-adenosyl)-l-[*methyl*-^{3}H]methionine ([me^{3}H]SAM) (stock solution of 55.9 μM (17.9 Ci/mmol) in 10 mM H_2SO_4/EtOH (9:1, vol/vol); PerkinElmer Life Sciences). Thin-walled PCR tubes holding reaction mixtures of 25 μL with or without 1 μM Larp7 in the absence of RNA substrates were incubated at 37°C for 5 min prior to reaction initiation by addition of 5 μL of RNA 60 μM stock solutions (prewarmed to 37°C). For reactions where Larp7 was added in later time points, the reaction mixtures of 20 μL without Larp7 or RNA were incubated and initiated by addition of 5 μL of RNA stock solutions, and 5 μL Larp7 stock solutions (6-30 μM) or 1X reaction buffer were added to target and control experiments, respectively, 9 min after reaction initiation. Reactions were quenched by prompt heating to 98°C for 2 min. 25 μL of the 30 μL reaction mixtures were spotted on Amersham Hybond-N+ membrane (GE Healthcare), washed and counted as described previously (Yang et al., 2019). All individual data points from independent reactions are shown on the presented data plots, with bars representing average values and whiskers representing standard deviations.

Isothermal titration calorimetry

The dissociation constants (K_d) for binding between Larp7 constructs and RNA constructs were measured using a MicroCal iTC200 instrument (GE Health Sciences). Protein constructs were purified in buffer C for the final SEC step, concentrated and buffer exchange again in buffer C using Amicon devices. ITC experiments were performed at 25°C, with 9-12 μM RNA in the cell and 100-130 μM protein in the syringe.

Calorimetric data were fit using ORIGIN7 (MicroCal) with association constant K, stoichiometry N, entropy change (ΔS) and enthalpy change (ΔH) all kept as floating variables. Experiments were performed in five independent repeats, data fitted individually and binding parameters averaged from each fit with standard deviations shown.

Nuclear magnetic resonance analysis of Larp7 proteins

For NMR assignments of Larp7 β4−xRRM2$_{370-582}$ construct, 0.16 mM sample uniformly enriched with ^{15}N and ^{13}C buffer exchanged in NMR buffer (20 mM sodium phosphate, pH 6.05, 50 mM KCl, 1 mM TCEP, 5% v/v D_2O) was used. Data acquisition was performed at 298.15 K on Avance Neo 800MHz spectrometer equipped with TCI HCN cryoprobe. Backbone assignments were carried out using the following NMR experiments: HNCACB, CBCA(CO)NH, HNCA and C(CO)NH. For NMR assignments of Larp7 LaMod$_{22-390\Delta L}$ construct, 0.8 mM sample uniformly enriched with ^{15}N and ^{13}C buffer exchanged in the above NMR buffer was used. Data acquisition was performed at 298.15 K on Avance Neo 800MHz spectrometer equipped with TCI H&F cryoprobe and Avance III HD 600MHz spectrometer equipped with QCI HCNP cryoprobe. Backbone assignments were carried out using the following NMR experiments: HNCACB, CBCA(CO)NH, HNCA, HN(CO)CA, HNCO, HN(CA)CO and C(CO)NH. NMR data was collected using TopSpin (Bruker), processed with NMRPipe (Delaglio et al., 1995) and analyzed using NMRFAM-Sparky (Lee et al., 2015). Assignments of LaMod$_{22-390\Delta L}$ were further verified using individual LaM (aa 1-122) and RRM1 (aa 116-390ΔL) constructs. Heteronuclear ^{1}H-^{15}N Nuclear Overhauser Effect (NOE) enhancement factors were measured for spectrally resolved backbone resonances at 298.15 K on Avance Neo 800MHz spectrometer equipped with TCI H&F cryoprobe for 0.20 mM [U-^{15}N] β4−xRRM2$_{370-582}$ alone or bound to SL4-ex RNA at 1:1.2 protein:RNA molar ratio, and Avance III HD 600MHz spectrometer equipped with QCI HCNP cryoprobe for 0.29 mM [U-^{15}N] LaMod$_{22-390\Delta L}$. SL4-ex bound β4−xRRM2$_{370-582}$ is assigned by comparison to RNA-bound xRRM2$_{445-556}$

(Eichhorn et al., 2016) and tracing peak trajectories of RNA titration. The NOE data were acquired using the Bruker standard sensitivity enhanced HSQC experiment with saturation during the 5 s recycle delay in an interleaved manner to record the peak intensities for the NOE saturated vs. unsaturated signal. Various Larp7 La module constructs were screened with ^{1}H-^{15}N-HSQC to arrive at the optimal construct for backbone assignment.

For ^{19}F-1D NMR experiments of various Larp7 constructs, 0.03-0.24 mM protein samples uniformly enriched with ^{15}N and 5-fluorotryptophan or 5-fluorotryptophan alone buffer exchanged in NMR buffer were used. The expressions of the [^{15}N, 5F-Trp]-labeled or 5F-Trp-labeled protein were performed in M9 minimal media following established protocol (Crowley et al., 2012) and purification follows the same procedure described above. Data acquisition was performed at 298.15 K on Avance Neo 800MHz spectrometer equipped with TCI H&F cryoprobe with the ^{1}H channel tuned to ^{19}F. The assignment of the 5F-Trp signals was achieved through comparison of ^{19}F-1D spectra between various domain constructs and confirmed by constructs bearing Trp-to-Phe mutagenesis.

Electron microscopy of both stained and frozen-hydrated samples

For negative stain electron microscopy (EM), samples were stained with uranyl acetate and air-dried. A total of 300 and 239 micrographs were recorded at a nominal magnification of 70,000× in an FEI Tecnai TF20 microscope for WT and lin^{14ex} 7SK core RNP complex samples, respectively. The defocus value of each micrograph was determined by CTFFIND4 (Rohou and Grigorieff, 2015). Particles were automatically picked without reference using Gautomatch (Zhang, 2016). The particles were boxed out in dimensions of 64 × 64 square pixels (pixel size at the sample level: 4.29 Å), and then subjected to reference-free 2D classification using RELION3.0 (Zivanov et al., 2018).

For cryo-EM sample optimization, an aliquot of 2.5 μL of sample at a concentration of 1.8 μM was applied onto a glow-discharged holey carbon-coated copper grid (300 mesh, QUANTIFOIL® R 2/1). The grid was blotted with Grade 595 filter paper (Ted Pella) and flash-frozen in liquid ethane with an FEI Mark IV Vitrobot. An FEI TF20 cryo-EM instrument was used to screen grids. Cryo-EM grids with optimal particle distribution and ice thickness were obtained by varying the gas source (air using PELCO easiGlowTM, target vacuum of 0.37 mbar, target current of 15 mA; or H_2/O_2 using Gatan Model 950 advanced plasma system, target vacuum of 70 mTorr, target power of 50 W) and time for glow discharge, the volume of applied samples, chamber temperature and humidity, blotting time and force, as well as drain time after blotting. Our best grids for circ^{m} 7SK core RNP complex were obtained with 20 s glow discharge using H_2/O_2 and with the Vitrobot sample chamber set at 10 °C temperature, 100% humidity, 8 s blotting time, 2 blotting force, and 0 s drain time. The best grids for lin^{m} 7SK core RNP complex were obtained with 20 s glow discharge using H_2/O_2 and with the Vitrobot sample chamber set at 10 °C temperature, 100% humidity, 9 s blotting time, 4 blotting force, and 0 s drain time.

Optimized cryo-EM grids were loaded into an FEI Titan Krios electron microscope with a Gatan Imaging Filter (GIF) Quantum LS device and a post-GIF K2 Summit direct electron detector. The microscope was operated at 300 kV with the GIF energy-filtering slit width set at 20 eV. Movies were acquired with SerialEM (Mastronarde, 2005) by electron counting in super-resolution mode at a pixel size of 0.535 Å/pixel. 40 frames were acquired in 8 s for each movie, giving a total dose of ~50 e^{-}/Å^2/movie. Frames in each movie were aligned for drift correction with the GPU-accelerated program MotionCor2 (Zheng et al., 2017). The first frame was discarded during drift correction due to concern of more severe drift/charging of this frame. Two averaged micrographs, one with dose weighting and the other without, were generated for each movie after drift correction. The averaged micrographs have a calibrated pixel size of 1.07 Å at the specimen scale. The averaged micrographs without dose weighting were used only for defocus determination and the averaged micrographs with dose weighting were used for all other steps of image processing.

Structure determination for circ^{m} RNP complex

For circ^{m} RNP complex, the defocus values of the averaged micrographs were determined by CTFFIND4 to be ranging from −1.5 to −2.5 μm. Initially, a total of 5,598,553 particles were automatically picked from 8260 averaged micrographs without reference using Gautomatch. The particles were boxed out in dimensions of 200 × 200 square pixels and binned to 100 × 100 square pixels (pixel size of 2.14 Å) before further processing by RELION3.0. Several iterations of reference-free 2D classification were subsequently performed to remove "bad" particles (*i.e.*, classes with fuzzy or un-interpretable features, including junk, dissociated particles and contaminations), yielding 1,582,722 good particles. The resulting particles were subjected to ab initio reconstruction with three classes in cryoSPARC v2 (Punjani et al., 2017). One class exhibiting good model features (intact features as shown in representative 2D classes plus visible secondary structural elements like α helices) was kept, which was then used as initial model for 3D classification with five classes in RELION3.0. The 3D class (520,122 particles) with the best model features was kept and treated with another round of 3D classification. 3 of the 5 3D classes showed good model features, thus the particles of the 3 classes (348,171 particles, 20.0% of the particles from 2D classification) were combined together for a 3D auto-refinement in RELION3.0. The refinement map has an average resolution of 4.28 Å, which hits the nyquist resolution of the current dataset with 2.14 Å pixel size. For final clearance, duplicated particles were removed, and the resulting particles were un-binned to 200 × 200 square pixels (pixel size of 1.07 Å) for one round of 2D classification (Figure S2A).

The 294,504 un-binned, unique particles (18.6% of the particles from 2D classification) resulting from multiple classifications were subjected to a final step of 3D auto-refinement in RELION3.0 (Figure S2A). The two half maps from this auto-refinement step were subjected to RELION's standard post-processing procedure. The final map of circ^{m} RNP complex has an average resolution of 3.7 Å based on RELION's gold-standard Fourier shell correlation (FSC).

Structure determination for lin^m RNP complex

For lin^m RNP complex, the defocus values of micrographs were determined by CTFFIND4 to be ranging from −1.5 to −3 μm. Initially, a total of 11,324,367 particles were automatically picked from 16,932 micrographs with Gautomatch (Zhang, 2016) without using references. The particles were boxed out in dimensions of 200 × 200 square pixels (pixel size of 1.07 Å) for and binned to 100 × 100 square pixels (pixel size of 2.14 Å) before further processing. Several iterations of reference-free 2D classification were subsequently performed to remove "bad" particles (*i.e.*, particles in 2D classes with fuzzy or un-interpretable features, including junk, dissociated particles and contaminations), yielding 2,031,928 good particles. The final 3D map of circ^m RNP complex was binned to a pixel size of 2.14 Å and served as the initial model for 3D classification of lin^m RNP complex. Two (854,476 particles, 42.1% of the particles from 2D classification) of the five 3D classes exhibiting good model features (intact features as shown in representative 2D classes plus visible secondary structural elements like α helices) were kept (Figure S3A). The resulting particles were un-binned to 200 × 200 square pixels (pixel size of 1.07 Å) before further processing. The 854,476 un-binned particles underwent iterations of 2D and 3D classifications, followed by duplication removing, yielding 159,349 good particles (Figure S3A).

The 159,349 un-binned, unique particles (7.8% of the particles from 2D classification) resulting from the multiple classification were subjected to a final step of 3D auto-refinement in RELION3.0 (Figure S3A). The two half maps from this auto-refinement step were subjected to RELION's standard post-processing procedure. The final map of lin^m RNP complex has an average resolution of 4.1 Å based on RELION's gold-standard FSC. We noted that final map has slight orientation bias, nonetheless, interpretability of the final map is not impacted due to the inclusion of all possible views.

Resolution assessment

All resolutions reported above are based on the "gold-standard" FSC 0.143 criterion (Chen et al., 2013). FSC curves were calculated using soft spherical masks and high-resolution noise substitution was used to correct for convolution effects of the masks on the FSC curves (Chen et al., 2013). Prior to visualization, all maps were sharpened by applying a negative B factor which was estimated using automated procedures (Rosenthal and Henderson, 2003). Local resolution was estimated using ResMap (Kucukelbir et al., 2014). The overall quality of the maps for circ^m and lin^m RNP complexes is presented in Figures S2D–S2G and S3D–S3G, respectively. Data collection and reconstruction statistics are summarized in Table 1.

Model building, refinement, and validation

To aid subunit assignment and model building, we took advantage of the reported MePCE and Larp7 structures (PDB 6DCC for MePCE; PDBs 4WKR and 6D12 for Larp7), which were fitted into the circ^m RNP complex density map by UCSF Chimera (Pettersen et al., 2004). We manually adjusted its side chain conformation and, when necessary, moved the main chains to match the density map using COOT (Emsley et al., 2010). This allowed us to identify extra densities for aa. 376-413 & 546-571 of Larp7, aa. 522-551 of MePCE and identify the entire circ^m RNA with the exception of SL4 upper stem and apical loop. Next, we built the atomic model for these regions *de novo*. Protein sequence assignment was mainly guided by visible densities of amino acid residues with bulky side chains, such as Trp, Tyr, Phe, and Arg. Other residues including Gly and Pro also helped the assignment process. Unique patterns of sequence segments containing such residues were utilized for validation of residue assignment. Well-defined nucleotide densities, along with the base pairs, facilitated the RNA model building process. For the upper stem and apical loop of SL4, the previous RNA model from Larp7 structure (PDB 6D12) was adjusted for their base conformation and, when necessary, for their main chains to match the density map.

For the lin^m RNP complex, the circ^m RNP complex model was rigidly fitted into the density map of lin^m RNP complex and manually adjusted using COOT. This enabled us to identify extra densities for nucleotides close to the SL1p-L1 with a different conformation from S1alt in circ^m RNA. As the map resolution in this region is insufficient for *de novo* atomic modeling, we traced the main chain using COOT for nt 109–115 (L1) and 296–298 (L4) of lin^m RNA.

The atomic models were refined using PHENIX in real space (Adams et al., 2010) with secondary structure and geometry restraints. Refinement statistics of circ^m and lin^m RNP complexes were summarized in Table 1. These two models were also evaluated based on Molprobity scores (Chen et al., 2010) and Ramachandran plots (Table 1). Model/map FSC validation was shown in Figures 3G and S2G. Representative densities for the proteins and RNA are shown in Figures 3H and S2H. Structure-related images in this paper were generated using PyMol v2.3.5, UCSF Chimera and ChimeraX (Goddard et al., 2018).

Cellular transcription stress response

The p7SK lin^{L3c} construct encoded an altered version of 7SK snRNA in which the 289-CACAUCC-295 WT linker 3 sequences were substituted for the GUGUAGG complementary sequences preventing helix formation with the 5′-terminal 1-GGAUGUG-7 nucleotides of 7SK. To generate p7SK lin^{L3c}, the 7SK gene was PCR-amplified by using the (TGGAGACTGCAGTATTTAGC) and the (ATAGA ATTCGGGTCAAAAGAAAGGCAGACTGCCACATGCAGCGCCTCATTTCCTACACTCTGGAGTCTTGGAAGCTTG) oligonucleotides as 5′ and mutagenic 3′ primers and the p7SK plasmid carrying the WT 7SK gene as a template (Egloff et al., 2006). The amplified DNA fragment was digested with PstI and EcoRI and inserted into the same sites of pBluescript II KS(+) (Stratagene).

For cell culture and transfection, HeLa cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum and 100 μg/ml of penicillin and streptomycin (GIBCO, Invitrogen). For transfection, the p7SK lin^{L3c} expression

plasmid was transfected into HeLa cells using LipoFectamine 2000 (Invitrogen) according to the manufacturer's instructions. 48h after transfection, cells were treated or non-treated with Actinomycin D (1μg/mL, Sigma) for 1 hour.

For immunoprecipitation, about 8 × 10^6 cells were washed in ice-cold PBS, resuspended in 500 μL of cold NET-2 buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.05% NP-40) and sonicated for 5 × 15 s with a Branson sonifier at setting 4. Cell debris were removed by centrifugation at 16,000 x g for 10 min, and the supernatant was incubated with 20 μL of protein A-agarose beads (Sigma) swollen in NET-2 buffer and saturated with anti-CDK9 antibody (ab6544; Abcam). After immunoprecipitation, the beads were washed six times with 1 ml of NET-2 buffer, RNAs were extracted with Trizol reagent and proteins were eluted with 1x SDS-PAGE loading buffer.

For western blot analysis, rabbit anti-CDK9 antibody (#2316, Cell Signaling Technology) was diluted to 1:1,000, and horseradish peroxidase-conjugated anti-rabbit IgGs (W401B, Promega) was used as secondary antibody.

For RNAse protection assays, the RNase A/T1 mapping procedure was used as described earlier (Goodall and Filipowicz, 1990). To synthesize the sequence-specific antisense RNA probe, the 7SK lin^{L3c} expression construct was linearized by PstI and used as template for transcription by T7 RNA polymerase in the presence of [α-^{32}P]CTP (3000 Ci/mmol).

For Northern blot analysis, RNAs were fractionated on 6% sequencing gel, transferred onto Hybond-N nylon membrane and probed with the 7SK (GTGTCTGGAGTCTTGGAAGC) and 7SL (GTCTCGCTATGTTGCCCAGGCT) radiolabeled antisense oligonucleotides.

QUANTIFICATION AND STATISTICAL ANALYSIS

Refer to Method details, Figures S2 and S3 for cryo-EM processing, and Table 1 for model building and refinement. Data for quantification analyses are presented as mean ± standard deviation (SD). The statistical analyses were performed by Student's t test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; ****$p < 0.0001$; ns, no significant, $p \geq 0.05$) using the IgorPro 8 software. The RNA levels in the RNAse protection assays were quantified by Multi Gauge software. The number of replicates is shown as individual date points on the figure and indicated in the figure legends.