



## OPEN ACCESS

## EDITED BY

Loïc Salmon,  
Centre de Résonance Magnétique  
Nucleaire A Tres Hauts Champs (CRMN),  
France

## REVIEWED BY

Aditi Borkar,  
University of Nottingham,  
United Kingdom  
Isabelle Lebars,  
UPR9002 Architecture et Réactivité de l'  
arN, France

## \*CORRESPONDENCE

Catherine D. Eichhorn,  
✉ ceichhor@unl.edu

## SPECIALTY SECTION

This article was submitted to Genome  
Organization and Dynamics,  
a section of the journal  
Frontiers in Molecular Biosciences

RECEIVED 30 January 2023

ACCEPTED 16 March 2023

PUBLISHED 27 March 2023

## CITATION

Camara MB, Sobeh AM and Eichhorn CD  
(2023), Progress in 7SK ribonucleoprotein  
structural biology.  
*Front. Mol. Biosci.* 10:1154622.  
doi: 10.3389/fmolb.2023.1154622

## COPYRIGHT

© 2023 Camara, Sobeh and Eichhorn.  
This is an open-access article distributed  
under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#).  
The use, distribution or reproduction in  
other forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Progress in 7SK ribonucleoprotein structural biology

Momodou B. Camara<sup>1</sup>, Amr M. Sobeh<sup>1</sup> and  
Catherine D. Eichhorn<sup>1,2\*</sup>

<sup>1</sup>Department of Chemistry, University of Nebraska, Lincoln, NE, United States, <sup>2</sup>Nebraska Center for Integrated Biomolecular Communication, Lincoln, NE, United States

The 7SK ribonucleoprotein (RNP) is a dynamic and multifunctional regulator of RNA Polymerase II (RNAPII) transcription in metazoa. Comprised of the non-coding 7SK RNA, core proteins, and numerous accessory proteins, the most well-known 7SK RNP function is the sequestration and inactivation of the positive transcription elongation factor b (P-TEFb). More recently, 7SK RNP has been shown to regulate RNAPII transcription through P-TEFb-independent pathways. Due to its fundamental role in cellular function, dysregulation has been linked with human diseases including cancers, heart disease, developmental disorders, and viral infection. Significant advances in 7SK RNP structural biology have improved our understanding of 7SK RNP assembly and function. Here, we review progress in understanding the structural basis of 7SK RNA folding, biogenesis, and RNP assembly.

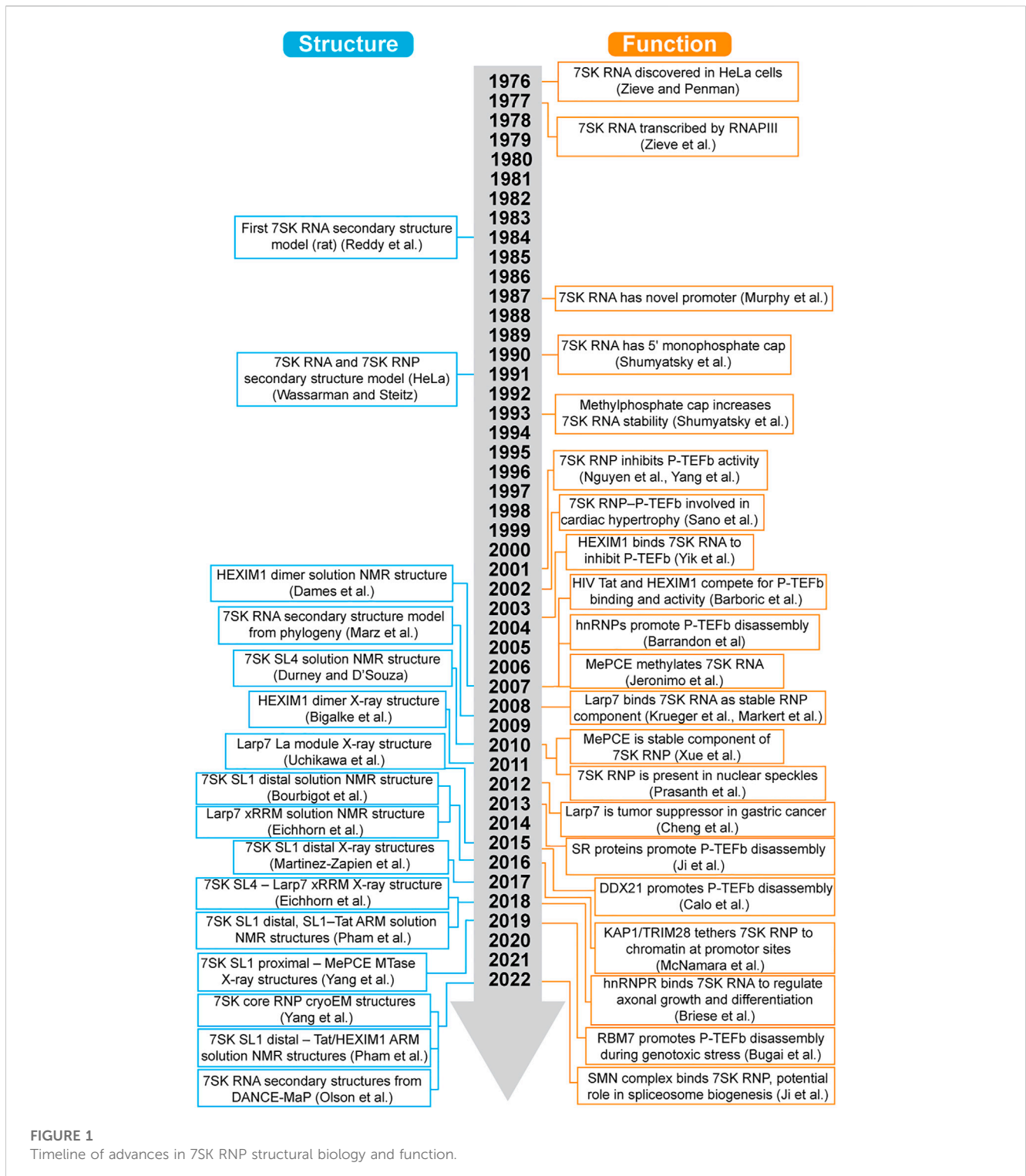
## KEYWORDS

cryoEM, solution state NMR, X-ray crystallography, RNA structural dynamics, chemical probing, RNA-protein interactions, gene regulation

## Introduction

7SK RNA is a highly abundant non-coding RNA with an estimated  $2 \times 10^5$  copies per cell (Gurney and Eliceiri, 1980), first identified in 1976 when it sedimented in the 7S fraction (Zieve and Penman, 1976). However, a specific function was not identified until 2001, when two research groups independently reported that 7SK RNA interacts with the positive elongation factor b (P-TEFb) to regulate RNA polymerase II (RNAPII) transcription elongation (Nguyen et al., 2001; Yang et al., 2001) (Figure 1). Shortly after, hexamethylene bis-acetamide (HMBA) induced in human vascular smooth muscle cells 1 (HEXIM1, also known as MAQ1 and CLP1) was shown to be required to recruit P-TEFb to the 7SK ribonucleoprotein (RNP) (Michels et al., 2003; Yik et al., 2003) (Figure 1). Core proteins methylphosphate capping enzyme (MePCE, also known as BCDIN3) and La-related protein 7 (Larp7, also known as PIP7S and HDCMA18P) were not identified until 2007 (Jeronimo et al., 2007) and 2008 (Krueger et al., 2008; Markert et al., 2008), respectively (Table 1). The first high-resolution structure of a domain component of 7SK RNA was reported in 2010 (Durney and D'Souza, 2010), two decades after establishing the secondary structure (Reddy et al., 1984; Wassarman and Steitz, 1991).

The past decade has seen a dramatic increase in high-resolution structures determined of domains of 7SK RNA, individual core and accessory proteins, and RNA-protein complexes (Table 2). Likewise, rapid developments in chemical mapping techniques and ensemble modeling approaches have provided a wealth of data on 7SK RNA secondary structure. These advances in 7SK RNP structural biology have provided a growing understanding of 7SK RNA structure and protein assembly to form a functional RNP. However, in the cellular environment a heterogeneous mixture of pools of 7SK RNPs are present, comprised of different sets of



proteins each of which has a unique regulatory function (Table 1). This complexity challenges a detailed understanding of the structural basis for transcriptional regulation. Several excellent reviews discuss 7SK RNP functions and mechanisms in gene expression regulation (Peterlin et al., 2012; McNamara et al., 2016a; Briese and Sendtner, 2021; Studniarek et al., 2021; Fujinaga et al., 2022). Here, we describe progress in 7SK RNP structural biology, discuss outstanding challenges, and present perspectives for the future.

## 7SK RNP: A master regulator of eukaryotic transcription

7SK RNA is transcribed by RNAPIII and is primarily retained in the nucleus (Zieve et al., 1977; Murphy et al., 1987). Phylogenetic studies have identified 7SK RNA in metazoa with high sequence conservation, particularly among vertebrates (Gursoy et al., 2000; Gruber et al., 2008; Marz et al., 2009). In humans, 7SK RNA is a

TABLE 1 7SK RNP interacting proteins and associated functions.

| Function   | Protein/complex             | Name   | Association with 7SK RNP   | References   |
|--|-----------------------------|--|----------------------------|--|
| <b>Core proteins</b>   |                             |  |                            |  |
| 7SK RNA capping  | MePCE                       | Methyl phosphate capping enzyme                            | SL1, Larp7                 | <a href="#">Jeronimo et al. (2007)</a>   |
| 7SK RNP stabilization  | Larp7                       | La-related protein 7                                       | SL4, 3' end, MePCE, P-TEFb | <a href="#">He et al. (2008)</a> , <a href="#">Krueger et al. (2008)</a> , <a href="#">Uchikawa et al. (2015)</a>  |
| <b>Accessory proteins</b>  |                             |  |                            |  |
| P-TEFb kinase activity inhibition  | HEXIM1/2                    | Hexamethylene bisacetamide-induced protein 1 or 2          | SL1, P-TEFb                | <a href="#">Yik et al. (2003)</a> , <a href="#">Michels et al. (2004)</a>  |
| RNAPII regulation  | P-TEFb                      | Positive transcription elongation factor b                 | HEXIM1, Larp7              | <a href="#">Nguyen et al. (2001)</a> , <a href="#">Yang et al. (2001)</a>  |
| P-TEFb release   | hnRNP A1, A2, B1, R, K, Q   | Heterogeneous nuclear ribonucleoprotein A1, A2/B1, R, K, Q | SL3                        | <a href="#">Barrandon et al. (2007)</a> , <a href="#">Van Herreweghe et al. (2007)</a> , <a href="#">Briese et al. (2018)</a>                                  |
| P-TEFb release <i>via</i> CDK9 T186 dephosphorylation                      | PPM1 $\alpha$ , PP2B, PPM1G | Protein phosphatase 1 $\alpha$ , $\gamma$ , 2B             | 7SK RNA, P-TEFb            | <a href="#">Chen et al. (2008)</a> , <a href="#">Gudipaty et al. (2015)</a>  |
| P-TEFb release   | SRSF2                       | Serine/arginine-rich splicing factor 2                     | SL3                        | <a href="#">Ji et al. (2013)</a>   |
| P-TEFb release through helicase activity                                   | DDX5, 6, 9, 21              | DEAD box protein 5, 6, 9, 21                               | unknown                    | <a href="#">Van Herreweghe et al. (2007)</a> , <a href="#">Calo et al. (2015)</a> , <a href="#">Muck et al. (2016)</a> , <a href="#">Sithole et al. (2020)</a> |
| P-TEFb release during cell stress  | RBM7                        | RNA-binding motif 7 protein                                | SL3                        | <a href="#">Bugai et al. (2019)</a>  |
| P-TEFb dissociation from 7SK RNP and/or HEXIM1, kinase activation          | Brd4                        | Bromodomain-containing protein 4                           | unknown                    | <a href="#">Jang et al. (2005)</a> , <a href="#">Schroder et al. (2012)</a>  |
| P-TEFb release by MePCE cleavage, 7SK RNA uncapping                        | JMJD6                       | Jumonji C-domain-containing protein 6                      | MePCE, P-TEFb, 5' end      | <a href="#">Liu et al. (2013)</a> , <a href="#">Lee et al. (2020)</a>  |
| Chromatin remodeling, enhancer RNA transcription                           | BAF                         | BRG1/BRM-associated factor                                 | SL3                        | <a href="#">Flynn et al. (2016)</a>  |
| Transcription regulation of snRNA and snoRNA genes                         | LEC                         | Little elongation complex                                  | unknown                    | <a href="#">Egloff et al. (2017)</a>   |
| P-TEFb binding enhancement to Tat  | AFF1                        | AF4/FMR2 protein 1   | Cyclin T1, Larp7, HEXIM1   | <a href="#">Lu et al. (2014)</a>   |
| 7SK RNP tethering to chromatin   | KAP1                        | Kruppel-associated box (KRAM)-interacting protein 1        | Larp7                      | <a href="#">McNamara et al. (2016b)</a> , <a href="#">McNamara et al. (2016c)</a>  |
| snRNP biogenesis regulation <i>via</i> association with 7SK-hnRNPs complex | SMN                         | Survival motor neuron                                      | Larp7, MePCE               | <a href="#">Ji et al. (2021)</a>   |
| 7SK RNA pseudouridylation  | DKC1                        | Dyskerin   | SL3 (U250)                 | <a href="#">Zhao et al. (2016)</a>   |
| 7SK RNA m6A modification   | METTL3, METTL16             | Methyltransferase-like protein 3, 16                       | 7SK RNA, Larp7, MePCE      | <a href="#">Warda et al. (2017)</a> , <a href="#">Covelo-Molares et al. (2021)</a> , <a href="#">Leger et al. (2021)</a>                                       |

331–334 nucleotide (nt) transcript expressed by a single gene on chromosome 6 ([Driscoll et al., 1994](#)). 7SK RNA has several posttranscriptional modifications including pseudouridylation ([Zhao et al., 2016](#)) and m6A ([Warda et al., 2017](#); [Leger et al., 2021](#)) (Table 1). The nascent 7SK RNA is capped by MePCE with a monomethyl group at the 5'  $\gamma$ -phosphate ([Gupta et al., 1990](#); [Shumyatsky et al., 1990](#)). MePCE associates at promoter sites of 7SK and U6 (its other primary substrate) genes, suggesting MePCE is proximally located for co-transcriptional capping ([Jeronimo et al., 2007](#); [Xue et al., 2010](#)). Characteristic of RNAPIII transcripts, 7SK RNA has a UUU-OH 3' terminus, and

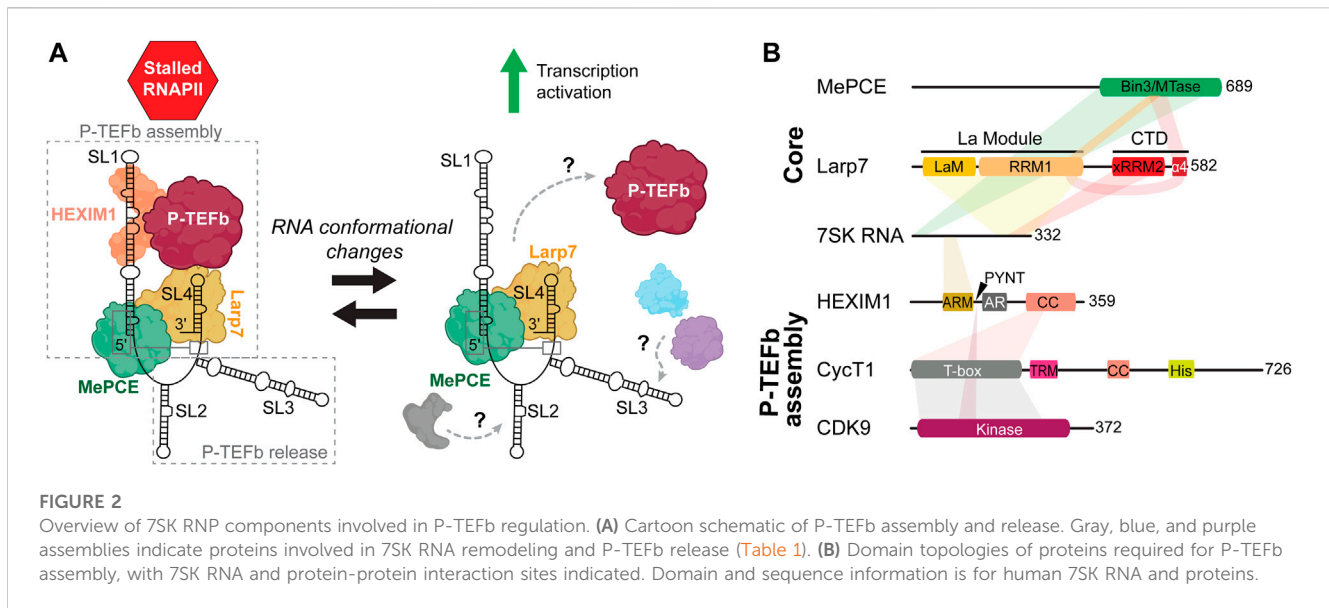
a small fraction of 7SK RNA associates with genuine La protein ([Stefano, 1984](#); [Van Herreweghe et al., 2007](#)). A converging model of 7SK RNA biogenesis is that MePCE first binds and caps the nascent 7SK RNA; next, La binds to the 3' end after transcription termination; subsequently, Larp7 displaces La ([Bayfield et al., 2010](#); [Hasler et al., 2021](#)) (Figure 2). Both MePCE and Larp7 are required to protect 7SK RNA from degradation ([He et al., 2008](#); [Krueger et al., 2008](#); [Xue et al., 2010](#)). Together 7SK RNA, MePCE, and Larp7 comprise the constitutively assembled ternary complex, called the 7SK core RNP (Figure 2), that is resistant to exonucleases, proteases, and high salt concentrations ([Xue et al., 2010](#)).

TABLE 2 Experimentally determined structures of 7SK RNP domains.

| PDB ID          | RNA domain(s)       | Protein domain(s)  | Year released | Method | Resolution | Comments                              | References                    |
|-----------------|---------------------|--------------------|---------------|--------|------------|---------------------------------------|-------------------------------|
| RNA             |                     |                    |               |        |            |                                       |                               |
| 2KX8            | SL4                 | n/a                | 2010          | NMR    | n/a        | Bound to arginine                     | Durney and D'Souza (2010)     |
| 5IEM            | SL1 distal          | n/a                | 2016          | NMR    | n/a        |                                       | Bourbigot et al. (2016)       |
| 5LYS            | SL1 distal          | n/a                | 2017          | X-ray  | 2.32 Å     | Gold derivative                       | Martinez-Zapien et al. (2017) |
| 5LYU            | SL1 distal          | n/a                | 2017          | X-ray  | 2.20 Å     |                                       | Martinez-Zapien et al. (2017) |
| 5LYV            | SL1 distal          | n/a                | 2017          | X-ray  | 2.35 Å     | Osmium derivative                     | Martinez-Zapien et al. (2017) |
| 6MCI            | SL1 distal          | n/a                | 2018          | NMR    | n/a        |                                       | Pham et al. (2018)            |
| 6MCF            | SL1 distal          | Tat ARM            | 2018          | NMR    | n/a        | Peptide sequence: GISYGRKKRRQRRRAHQ   | Pham et al. (2018)            |
| 7T1N            | SL1 distal          | HEXIM1 ARM         | 2022          | NMR    | n/a        | Peptide sequence: GISYGRQLGKKKHRRRAHQ | Pham et al. (2022)            |
| 7T1O            | SL1 distal          | Tat subtype G ARM  | 2022          | NMR    | n/a        | Peptide sequence: GISYGRKKRRHRRRAHQ   | Pham et al. (2022)            |
| 7T1P            | SL1 distal          | Tat Finland ARM    | 2022          | NMR    | n/a        | Peptide sequence: GISYGRKKRKHRRRAHQ   | Pham et al. (2022)            |
| Core RNP        |                     |                    |               |        |            |                                       |                               |
| 5UNA            | n/a                 | MePCE MTase        | 2017          | X-ray  | 2.55 Å     | Initial deposition 3G07 in 2009       | n/a                           |
| 4WKR            | 3' UUU-OH           | Larp7 La module    | 2015          | X-ray  | 3.20 Å     | construct residues 1-208              | Uchikawa et al. (2015)        |
| 5KNW            | n/a                 | Larp7 xRRM2        | 2016          | NMR    | n/a        |                                       | Eichhorn et al. (2016)        |
| 6D12            | SL4 upper stem      | Larp7 xRRM2        | 2018          | X-ray  | 2.20 Å     |                                       | Eichhorn et al. (2018)        |
| 6DCB            | SL1 proximal        | MePCE MTase        | 2019          | X-ray  | 2.00 Å     | Substrate RNA                         | Yang et al. (2019)            |
| 6DCC            | SL1 proximal        | MePCE MTase        | 2019          | X-ray  | 2.10 Å     | Capped product RNA                    | Yang et al. (2019)            |
| 7SLP            | Minimal 7SK SL1+SL4 | Larp7, MePCE MTase | 2022          | CryoEM | 4.10 Å     | "Linear" model                        | Yang et al. (2022)            |
| 7SLQ            | Minimal 7SK SL1+SL4 | Larp7, MePCE MTase | 2022          | CryoEM | 3.70 Å     | "Circular" model                      | Yang et al. (2022)            |
| P-TEFb assembly |                     |                    |               |        |            |                                       |                               |
| 2GD7            | n/a                 | HEXIM1             | 2007          | NMR    | n/a        |                                       | Dames et al. (2007)           |
| 3S9G            | n/a                 | HEXIM1 coiled-coil | 2011          | X-ray  | 2.10 Å     | Δstammer                              | Bigalke et al. (2011)         |

To this 7SK core RNP, assembly of numerous accessory proteins guides discrete mechanisms of transcription regulation (Quaresma et al., 2016; Briese and Sendtner, 2021) (Table 1). The most well-established 7SK RNP function is its role in P-TEFb regulation. The pausing of RNAPII ~20–60 nt downstream of the transcription start site is a major transcriptional checkpoint in metazoa, resulting in a stalled, transcriptionally inactive, RNAPII complex (Dollinger and Gilmour, 2021). P-TEFb, a heterodimer comprised of cyclin-dependent kinase 9 (CDK9) and cyclin T1 (CycT1), phosphorylates serine 2 in the C-terminal domain of RNAPII in addition to negative elongation factors to promote the transition of RNAPII from a paused to a productive elongating state (Bowman

and Kelly, 2014; Jonkers and Lis, 2015). To perform this task, P-TEFb associates with proteins including bromodomain-containing protein 4 (Brd4) and is part of the super elongation complex (SEC) (Bres et al., 2008; Bacon and D'Orso, 2019). 7SK RNP acts as a negative regulator against RNAPII transcription elongation by sequestering P-TEFb in a catalytically inactive pool. To recruit P-TEFb to the 7SK RNP, HEXIM1 (or, less commonly, its paralog HEXIM2) is minimally required to first assemble onto 7SK core RNP (Yik et al., 2003; Yik et al., 2005). Given the important role in transcription regulation, dysregulation of 7SK RNP and/or P-TEFb has been linked to myriad human diseases including cancers (Cheng et al., 2012; Thorne et al., 2022),



developmental disorders (Alazami et al., 2012; Ling and Sorrentino, 2015), heart disease (Sano et al., 2002; Krystof et al., 2010; Catalucci and Condorelli, 2013), and neurodegenerative disease (Santoro et al., 2016; Schneeberger et al., 2019).

50%–90% of P-TEFb is associated with 7SK RNP, depending on cell type (Peterlin and Price, 2006). The 7SK RNA 5' and 3' ends are required for P-TEFb assembly (Egloff et al., 2006), whereas central residues appear to be required for P-TEFb release (Van Herreweghe et al., 2007; Briese et al., 2018) (Figure 2A). HEXIM1 binds directly to the distal region of the 5' SL1. On binding 7SK RNA, HEXIM1 is able to stably recruit P-TEFb and inactivate CDK9 kinase activity. P-TEFb release is a highly complex process with dozens of identified proteins, many of which are involved in RNA processing (Table 1) (Barrandon et al., 2007; Flynn et al., 2016; Egloff et al., 2017). P-TEFb release may require dismantling of the 7SK core RNP: in one study, MePCE was shown to be cleaved by JMJD6 as part of the P-TEFb release mechanism (Lee et al., 2020). Several of these proteins associate with SL3 (Ji et al., 2013; Bugai et al., 2019; Luo et al., 2021), and a 7SK RNA construct lacking SL3 was able to bind but not release P-TEFb (Van Herreweghe et al., 2007; Briese et al., 2018), suggesting that SL3 may act as a hub to recruit proteins for P-TEFb release (Figure 2A). Rather than act as a static scaffold, 7SK RNA undergoes secondary structural remodeling as part of the mechanism of P-TEFb release (Van Herreweghe et al., 2007; Krueger et al., 2010; Calo et al., 2015).

Poised for action, 7SK–P-TEFb complexes are tethered to chromatin at promoter and enhancer sites (Ji et al., 2013; Wu et al., 2013; McNamara et al., 2016a; Flynn et al., 2016). There is growing consensus that rather than act as a global transcription regulator, P-TEFb-stimulated gene activation is localized and coordinated by cellular cues such as proliferative status and/or environmental cues such as DNA damage, stress, or small molecules (e.g., flavopiridol) (McNamara et al., 2016a; Bugai et al., 2019; Studniarek et al., 2021). In addition, infectious agents leverage 7SK RNP to reprogram host cell transcription to evade detection and promote survival. The most well-established example

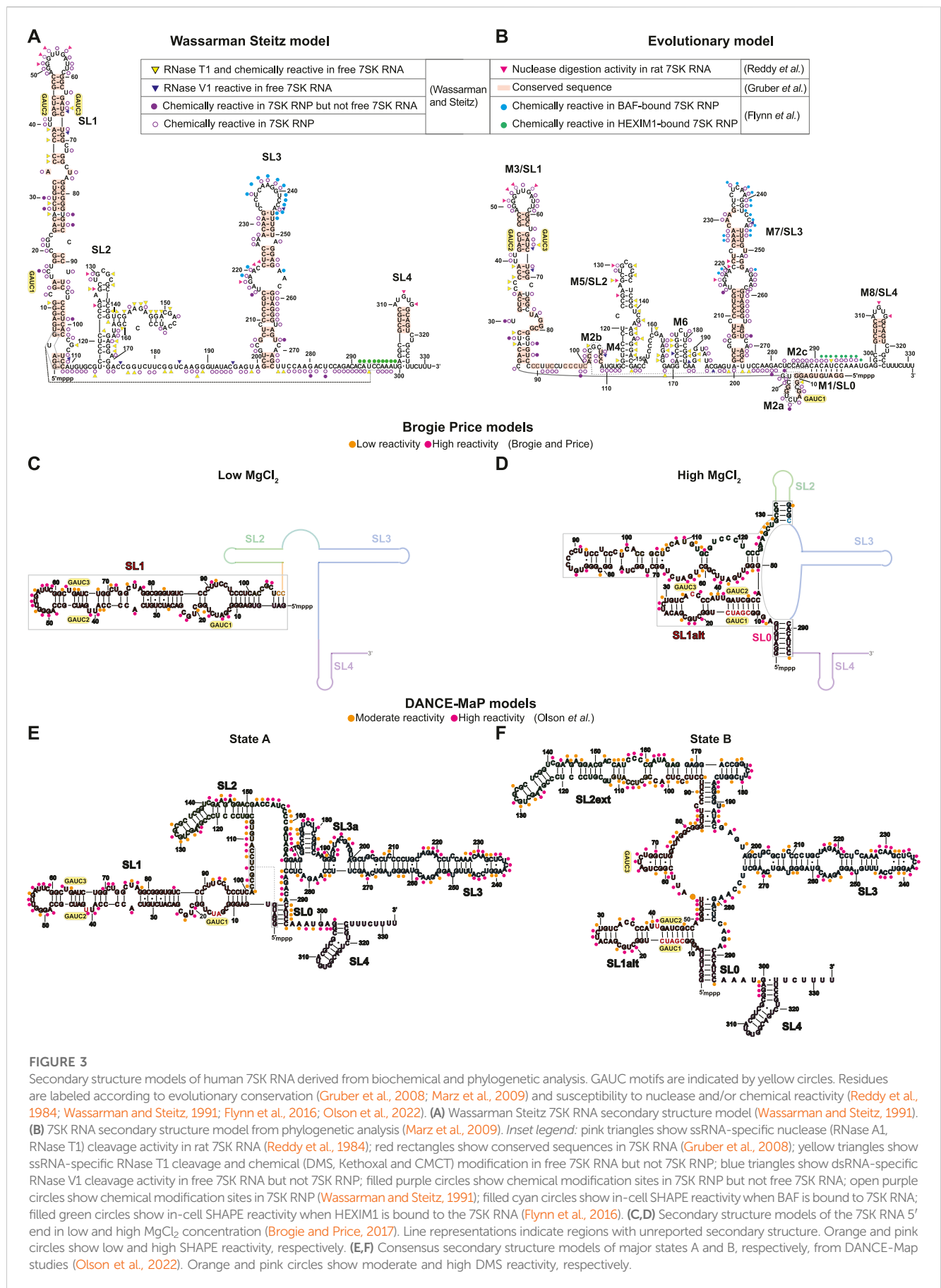
is HIV-1, where the HIV transcriptional transactivator (Tat) protein competes with HEXIM1 to bind 7SK RNA and recruit the 7SK–P-TEFb complex to the nascently transcribing HIV RNA genome (He and Zhou, 2011). Other infectious agents including SARS-CoV-2 (Gordon et al., 2020) and bacteria such as *Legionella* (Von Dwingelo et al., 2019; Cordsmeier et al., 2022) have also been shown to maneuver 7SK RNP to reprogram the host transcriptome.

Beyond P-TEFb regulation, 7SK RNP has been shown to regulate RNAPII function in pathways independent of P-TEFb. For example, 7SK RNP has been identified to regulate transcription initiation of enhancer RNAs through association with the chromatin-remodeling BRG1/BRM-associated factor (BAF) complex (Flynn et al., 2016); transcription elongation of small nuclear and nucleolar RNAs, localized in Cajal bodies (Egloff et al., 2017); and axonal growth and differentiation in motoneurons, localized in the cytosol (Briese et al., 2018; Briese and Sendtner, 2021) (Table 1). However, the limited biochemical data and lack of high-resolution structures of 7SK RNP accessory proteins in complex with 7SK RNA or core proteins preclude a detailed understanding of 7SK RNP-mediated transcription regulation.

## 7SK RNA secondary and tertiary structure

### 7SK RNA secondary structure

Several secondary structure models of 7SK RNA have been proposed based on experimental data from chemical probing techniques as well as evolutionary studies. Although differences are observed among studies, the 7SK RNA secondary structure generally consists of four stem-loop domains (SL1, SL2, SL3, SL4) with intervening single-stranded RNA (ssRNA) linker sequences. The first secondary structure model was reported in 1984, where 7SK RNA was extracted from rat cells and subjected to nuclease digestion (Reddy et al., 1984). Four regions of 7SK RNA



were highly susceptible to nuclease activity: SL1 apical loop, SL2, SL3 internal loop, and SL4 apical loop, which were used in combination with maximal base-pairing to construct a secondary structure model (Figure 3A) (Reddy et al., 1984). A landmark study in 1991 (Wassarman and Steitz, 1991) used a combinatorial chemical and enzymatic probing approach with 7SK RNA samples isolated from HeLa cells in the presence and absence of at-the-time unknown protein cofactors. Overall, a “linear” secondary structure was identified consisting of four stem-loops (Figure 3A) that differed from the previous secondary structure model (Reddy et al., 1984) in SL1 and SL2 regions. Significant differences in reactivity were observed when comparing the 7SK RNA and 7SK RNP (Figure 3A). For example, SL1 distal, SL2, and SL2-SL3 linker regions were particularly reactive in free 7SK RNA but unreactive in 7SK RNP (Figure 3A). In contrast, SL1 proximal end and SL3 were reactive in 7SK RNP but unreactive in free 7SK RNA (Figure 3A). Here, differences in reactivity between 7SK RNA and 7SK RNP may be attributed to differences in secondary structure or protection from nuclease activity when bound to protein.

Sequence homology studies showed that SL1 and SL4 are highly conserved across both vertebrate and invertebrate 7SK RNA sequences (Figures 3A, B) (Gruber et al., 2008). The central stem-loops, particularly SL3, are conserved in vertebrates but less conserved in invertebrate species (Gruber et al., 2008; Yazbeck et al., 2018). A hallmark feature of 7SK RNA is the presence of 5' GAUC sequences in SL1 distal (Egloff et al., 2006; Marz et al., 2009). This sequence was later found to be the binding site for HEXIM1/2 and HIV-1 Tat proteins (Lebars et al., 2010). Phylogenetic analysis of 7SK RNA sequences led to the proposal of a distinct secondary structure model, named ‘circular’ for a long-range base-pairing interaction between 5' and 3' residues to form SL0 (alternately called M1) (Figure 3B) (Marz et al., 2009). The “circular” model contains more stem-loops than the “linear” model, with multiple potential base-pairing arrangements at the proximal end of SL1 (Figure 3B). With the exception of the SL1 proximal end, this “circular” model shares similarities with the Wassarman Steitz “linear” model at SL1 distal end, SL3, and SL4 (Figures 3A, B), suggesting these regions are somewhat structurally conserved.

Recently, DMS and/or SHAPE chemical probing approaches have been used to investigate 7SK RNA secondary structure *in vitro* and *in cellulo* (Brogie and Price, 2017; Luo et al., 2021; Olson et al., 2022; Van Damme et al., 2022). Price and coworkers found that the 7SK RNA secondary structure, particularly SL1, was sensitive to MgCl<sub>2</sub> concentration (Brogie and Price, 2017). In addition to the two GAUC motifs previously identified in SL1 distal end, a third GAUC motif near the 5' end was identified (GAUC1, GAUC2, and GAUC3) allowing an alternate pairing arrangement (SL1alt) leading to the formation of a “circular” conformation at high MgCl<sub>2</sub> concentrations (Figures 3C, D). The authors propose that 7SK RNA adopts a “linear conformation” in the presence of HEXIM1-P-TEFb and rearranges to the “circular” conformation on P-TEFb release (Brogie and Price, 2017). In a separate study, in-cell SHAPE experiments showed different chemical reactivities in 7SK RNA SL3 and SL3-SL4 linker regions when bound to HEXIM1 (P-TEFb dependent) or BAF (P-TEFb independent), suggesting different RNA secondary structures when bound to different protein cofactors (Figure 3A) (Flynn et al., 2016). Consistent with the model proposed by Price and coworkers, the chemical reactivities indicate that, in the HEXIM1-bound 7SK RNP, 7SK RNA adopts a “linear” conformation (Figures 3A,B). Additional secondary

structure models in support of “linear” (Luo et al., 2021) or “circular” (Van Damme et al., 2022) conformations have been proposed using chemical mapping approaches.

Recent advances in chemical probing combined with ensemble detection methods have helped to untangle 7SK RNA secondary structure heterogeneity. Olson et al. (2022) used deconvolution and annotation of ribonucleic conformational ensembles (DANCE-Map), a technique that relies on DMS modification and a maximum likelihood deconvolution strategy to generate secondary structure ensembles. These experiments revealed a structural ensemble of two major conformational states, where state A resembles the “linear” model and state B resembles the “circular” model (Figures 3E, F). Both A and B states contain the SL0 stem observed in the “circular” model (Figures 3E, F), although the SL0 stem in state A showed an increased reactivity relative to state B suggesting a “linear-like” conformation. The most significant difference between the two major states is found in SL1 and SL2. In state B, this region undergoes rearrangement resulting in the formation of an alternate SL1 (SL1alt) and an extended SL2 (SL2ext), which has not been previously observed in literature (Figure 3F). Interestingly, the populations of states A and B differed in different cell types and upon treatment of P-TEFb inhibitors, suggesting that 7SK RNA secondary structure is sensitive to cellular and environmental conditions. Together, the past three decades of biochemical and bioinformatics studies on 7SK RNA secondary structure have advanced understanding of 7SK RNA folding and begun to resolve ensembles of secondary structures, leading to models of secondary structure switching coupled to functional states.

## High-resolution structures of SL1

The highly conserved SL1 distal end (nts 27–84) has a near-identical secondary structure for experimentally and evolutionarily determined secondary structures (Figure 3). Several solution NMR and X-ray crystallographic structures have been determined for the SL1 distal region with a modified apical loop (Table 2; Figure 4). The SL1 distal secondary structure features several short stems separated by pyrimidine-rich asymmetric internal loops capable of forming base triples with nearby base-pairs (Figures 4A, B). D'Souza and coworkers named this motif the “arginine sandwich motif” (ASM) (Figures 4A, 8B) (Pham et al., 2018). This motif has been previously observed in other biologically important RNAs, in particular the HIV-1 transactivation response (TAR) element (Tao et al., 1997; Brodsky et al., 1998; Huthoff et al., 2004).

Dock-Bregeon and coworkers determined one solution NMR structural ensemble and three X-ray crystal structures of a SL1 distal construct (Table 2) (Bourbigot et al., 2016; Martinez-Zapien et al., 2017). In the absence of divalent cations, the RNA adopts an extended conformation where base triples do not form (Figure 4C). In contrast, X-ray crystal structures determined in the presence of divalent cations revealed four coaxially stacked subdomains and one base triple at each loop (Figure 4C). Differences in base triple pairing arrangements were observed among structures obtained for native, osmium-, and gold-soaked crystals, indicating conformational variability at these sites. Independently, D'Souza and coworkers determined a solution NMR structure of a SL1 distal construct and observed formation of all ASM motifs (Figure 4C; Table 2) (Pham et al., 2018).

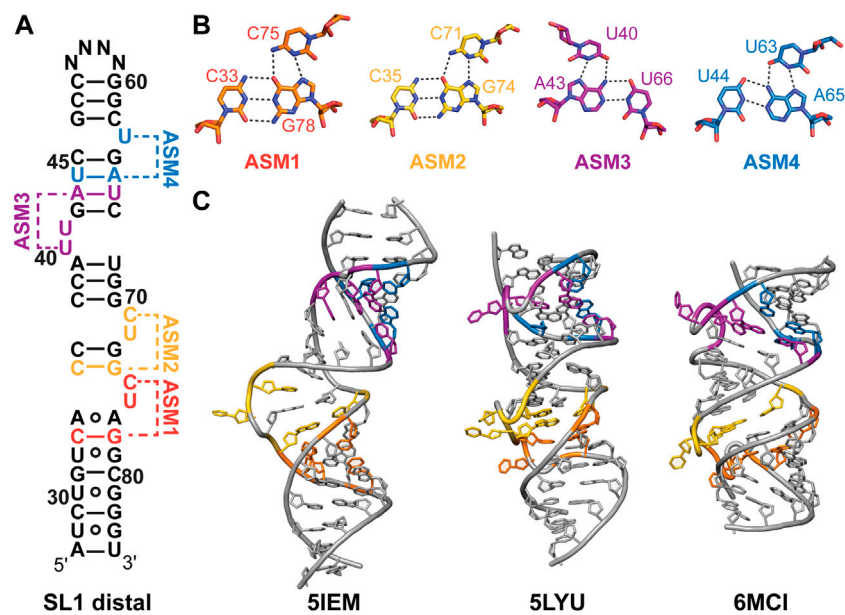


FIGURE 4

Structures of 7SK RNA SL1 distal region. (A) Secondary structure model of the SL1 distal constructs used to determine solution NMR or X-ray crystal structures. N indicates variable apical loop sequence. ASMs are shown as defined by D'Souza and coworkers (Pham et al., 2018). (B) Base triples of arginine sandwich motifs (ASM) for solution NMR structure (PDB ID 7T1N). (C) Structures of SL1 distal constructs from solution NMR (PDB IDs 5IEM, 6MCI) or X-ray crystallography (PDB ID 5LYU).

Comparison of this solution NMR structure (PDB ID 6MCI) to the X-ray structure of native crystals (PDB ID 5LYU) show similar base triple formation for ASM1, ASM2, and ASM3. For ASM4, U<sub>63</sub> forms a base triple to either G<sub>42</sub>-C<sub>67</sub> (native crystal, PDB ID 5LYU) or U<sub>44</sub>-A<sub>65</sub> (solution NMR, PDB ID 6MCI). A detailed comparison of X-ray and solution NMR structural differences is discussed elsewhere (Brillet et al., 2020).

All SL1 distal structures have globally similar conformations, although differences are observed in helical compaction, loop backbone conformation, and interhelical bending (Figure 4C). Molecular dynamics simulations indicate that ASM base triples are weak and transient (Roder et al., 2020). Sample conditions differed for all determined structures in particular pH, temperature, and presence of divalent cations. To reconcile these differences, Dock-Bregeon and coworkers examined the effect of pH and MgCl<sub>2</sub> on SL1 distal conformation (Brillet et al., 2020). Small-angle X-ray scattering (SAXS) and NMR studies showed that addition of MgCl<sub>2</sub> resulted in a more compact molecule, consistent with base triple formation. A metal binding site was identified in ASM3, which resulted in metal-induced bending about ASM3 on addition of MgCl<sub>2</sub>. Compaction was shown to be driven by addition of divalent cations or protonation of C<sub>71</sub> (ASM2), C<sub>75</sub> (ASM1), and A<sub>77</sub> (ASM1) which may explain the observed structural differences.

## High-resolution structures of SL4

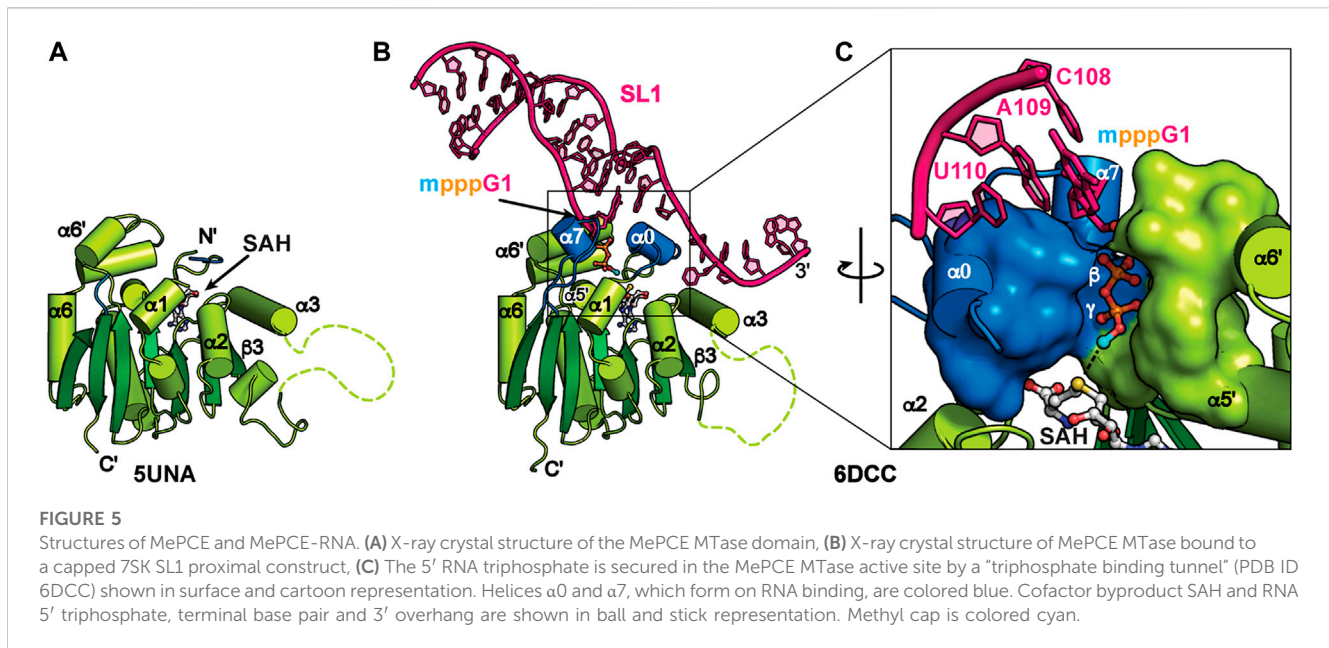
In addition to SL1, SL4 is highly conserved among metazoa and consists of a short stem-loop with a two-nt asymmetric internal loop (Figure 6A). A solution NMR structure was determined for a

SL4 construct in complex with arginine (Figure 6B; Table 2) (Durney and D'Souza, 2010). In this arginine-bound structure at low pH, C<sub>320</sub> in the C<sub>320</sub>U<sub>321</sub> loop was found to form a base triple with the G<sub>303</sub>-C<sub>323</sub> base-pair (Durney and D'Souza, 2010), similar to ASM motifs observed in SL1 described above. In a separate study, solution NMR relaxation experiments of SL4 showed that in the absence of amino acid or protein substrates the C<sub>320</sub>U<sub>321</sub> loop is dynamic, suggesting that the base triple does not form (Eichhorn et al., 2018). An X-ray crystal structure of the SL4 upper stem (nts 305–319) bound to the Larp7 extended RNA recognition motif 2 (xRRM2) was determined (Eichhorn et al., 2018) (Table 2). When bound to Larp7, the SL4 upper stem was observed to have widening of the major groove, attributed to insertion of the xRRM2 helix α3 (Figure 6C). In addition, significant differences in the apical loop were observed compared to the previously determined NMR structure (Eichhorn et al., 2018) (Figures 6B, C). The apical loop backbone makes an S-turn about U<sub>313</sub> and G<sub>314</sub>, stabilized by protein-RNA interactions. These structural features explain the high specificity of 7SK RNA recognition provided by the Larp7 xRRM2.

## 7SK core RNP

Prior to structural studies, several lines of biochemical evidence suggested that 7SK RNA, MePCE, and Larp7 engage in a complex interplay of interactions in the 7SK core RNP. The MePCE-Larp7 interaction is RNA-dependent: MePCE can be assembled on a Larp7-7SK RNA complex, even in the absence of a MePCE binding site, indicating that Larp7 is scaffolded onto 7SK





RNA to provide a MePCE interaction surface (Muniz et al., 2013). Larp7 assembly results in loss of MePCE methyltransferase activity (Xue et al., 2010; Brogie and Price, 2017) through interactions with the C-terminal helix  $\alpha 4$ , also called the MePCE interacting domain (MID) (Figure 2B) (Brogie and Price, 2017). High-resolution structures of individual domains of core proteins in isolation or bound to 7SK RNA domains have provided critical insights into RNA recognition. Further, cryoEM structures of the 7SK core RNP ternary complex have aided understanding of 7SK RNA biogenesis and assembly into a stably associated core RNP (Yang et al., 2022).

## MePCE and MePCE-RNA

MePCE is a mammalian ortholog of the *Drosophila* bicoid-interacting protein 3 (Bin3) (Jeronimo et al., 2007). In addition to 7SK RNA, MePCE methylates a subset of RNAPIII small nuclear RNAs including U6, B2, and U3 (Gupta et al., 1990; Shumyatsky et al., 1990). MePCE contains a variable N-terminal domain with predicted low structural complexity and a C-terminal methyltransferase (MTase) domain (Figure 2B) (Cosgrove et al., 2012). The MTase domain uses cofactor S-adenosyl methionine (SAM) as the methyl donor to transfer the methyl group to the RNA 5'  $\gamma$ -phosphate. For efficient binding and catalysis, a terminal G-C base-pair and 3' ssRNA overhang are required (Singh et al., 1990; Yang et al., 2019). An X-ray crystal structure of the human MTase domain in complex with S-adenosyl homocysteine (SAH) byproduct, determined by the Structural Genomics Consortium in Toronto (Table 2), identified a classical Rossman fold from residues 431–685 with a disordered  $\alpha 3$ - $\beta 3$  loop (residues 492–538) (Figure 5A).

X-ray crystal structures of the human MePCE MTase domain in complex with RNA constructs of the 7SK RNA 5' proximal hairpin revealed the structural basis for capping and retention after catalysis (Figure 5B) (Yang et al., 2019). Structures were determined of

MTase-SAH bound to either uncapped substrate RNA or capped product RNA (Table 2). When bound to RNA, the MTase has substantial conformational changes surrounding the active site including ordering of N-terminal residues 411–430 to form new helix  $\alpha 0$  (Figure 5). Helix  $\alpha 0$  is enriched in tyrosine residues, which lie underneath the terminal base-pair in the RNA helix and interact with the RNA helix-overhang interface (Figures 5B, C). Residues in the  $\beta 6$ - $\beta 7$  loop become ordered, forming new helix  $\alpha 7$ , and interact with both RNA and N-terminal residues (Figures 5B, C). The RNA 5' triphosphate is encased within the active site via a "triphosphate binding tunnel" comprised of residues in  $\alpha 0$ ,  $\alpha 7$ ,  $\alpha 5'$ , and  $\alpha 6'$  (Figure 5C). An extensive network of hydrogen bonds fixes the 5' triphosphate in place and positions the terminal oxygen on the 5'  $\gamma$ -phosphate in line for methyl transfer (Figure 5C). Alanine substitution of residues in the triphosphate binding tunnel significantly impaired catalysis (Yang et al., 2019).

Kinetics experiments showed poor enzymatic turnover, which was explained by ITC binding experiments showing that the methylated RNA product had 3-fold improved binding affinity compared to substrate RNA. Catalytic turnover was found to be inhibited both by addition of SAH and methylated RNA (Yang et al., 2019). An RNA construct of the "linear" SL1 proximal hairpin and 3' overhang sequence had higher affinity binding compared to an RNA construct of the "circular" SL0 hairpin and 3' overhang sequence (Yang et al., 2019). Together, these structural and biochemical data explain how MePCE is retained on 7SK RNA after catalysis as part of the core 7SK RNP.

## Larp7 and Larp7-RNA

Larp7 is a member of the LARP7 class of the La and La-related protein superfamily (Bayfield et al., 2010; Maraia et al., 2017). In ciliates and fission yeast, which lack 7SK RNA, LARP7 proteins bind to telomerase RNA to function in telomerase biogenesis and activity

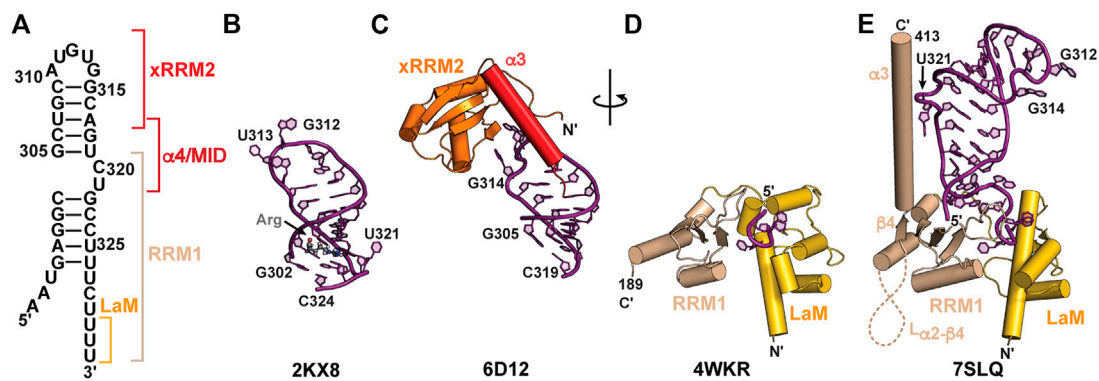


FIGURE 6

Structures of 7SK RNA 3' end and Larp7 recognition of 7SK RNA. (A) Secondary structure model of the 7SK RNA 3' end (nts 297–332), with Larp7 interaction sites indicated in brackets. (B) Solution NMR structure of 7SK SL4 bound to arginine (Arg). Arginine ligand is shown in stick representation. (C) X-ray crystal structure of Larp7 xRRM2 bound to SL4 upper stem. (D) X-ray crystal structure of Larp7 La module bound to UUU RNA sequence. (E) 7SK core RNP cryoEM structure (PDB ID 7SLQ) reveals cryptic  $\beta 4$  and  $\alpha 3$  in Larp7 RRM1 and extensive La module-RNA binding surface.

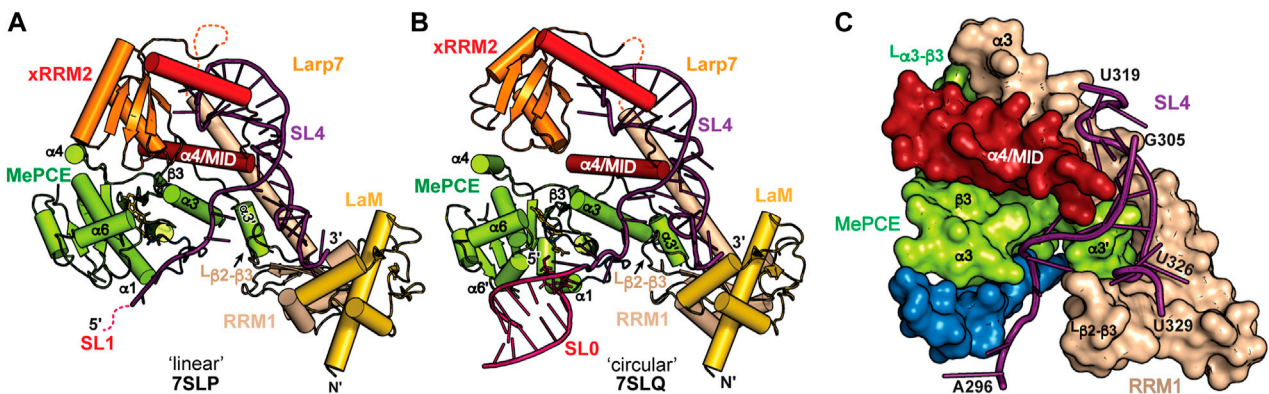


FIGURE 7

CryoEM structures of 7SK core RNP ternary complex. RNA constructs model the (A) "linear" or (B) "circular" conformations. (C) SL4-Larp7-MePCE interface features extensive interactions (PDB ID 7SLQ). MePCE MTase is colored green and residues corresponding to helix  $\alpha 0$ , which becomes unstructured in the ternary complex, are colored blue. Larp7 RRM1 is colored apricot and helix  $\alpha 4$  is colored maroon.

(Aigner and Cech, 2004; Singh et al., 2012; Collopy et al., 2018; Mennie et al., 2018; Paez-Moscoso et al., 2018). In metazoa, Larp7 primarily binds to 7SK RNA (He et al., 2008; Krueger et al., 2008; Markert et al., 2008) although more recently Larp7 has been shown to bind Lin28 mRNA (Dai et al., 2014) as well as U6 snRNA and U6-specific snoRNAs (Wang et al., 2020; Hasler et al., 2021). Larp7 has two RNA-binding domains: an N-terminal La module comprised of a La motif (LaM) followed by an RRM1 and C-terminal atypical xRRM2, with a variable-length intervening linker sequence (Figure 2B). The La module binds to the UUU-OH 3' terminus whereas the xRRM2 binds to the SL4 apical loop and upper stem (Figure 6).

An X-ray crystal structure of the human La module (construct residues 1–208) in complex with an RNA UUU trimer showed a similar RNA recognition mode to genuine La (Uchikawa et al., 2015) (Figure 6D; Table 2). RNA substrate is sandwiched between LaM and RRM1 domains, which are arranged in a "V" orientation, where

the LaM consisted of a classical winged helix-turn-helix and the RRM1 contained three  $\beta$ -strands rather than the four-five typically observed in RRMs (Figure 6D). CryoEM structures of an *in vitro* reconstituted 7SK core RNP, which included recombinantly expressed full-length human Larp7, revealed new insights into Larp7 structure and recognition of 7SK RNA (Yang et al., 2022) (Table 2). The La module has a larger binding surface than previously observed for any La or LARP, in which the La module binds to all eight ssRNA nts. Aromatic residues in  $\beta 2$  and  $\beta 3$  strands on the RRM1  $\beta$ -sheet surface interact with ssRNA nts. The RRM1  $\beta 2$ - $\beta 3$  loop lies at the base of SL4 and contacts both the 5' AAAU and 3' ssRNA overhangs (Figures 6E, 7C). Surprisingly, the Larp7 RRM1 was found to contain a cryptic  $\beta 4$  strand and helix  $\alpha 3$ , with a 189-residue  $\alpha 2$ - $\beta 4$  loop (Yang et al., 2022). Elegant NMR experiments demonstrated that the RRM1  $\beta 4$  and  $\alpha 3$  are stable components of the La module, even in the absence of RNA (Yang et al., 2022). Helix  $\alpha 3$  is positioned perpendicular to the  $\beta$ -sheet and

traverses up SL4 (Figure 6E). In contrast to the solution NMR structure of SL4 bound to arginine, the C<sub>320</sub>•G<sub>303</sub>-C<sub>323</sub> base triple is not formed; rather, C<sub>320</sub>U<sub>321</sub> loop residues are extruded to interact with RRM1 helix α3 (Yang et al., 2022), explaining biochemical assays demonstrating the importance of C<sub>320</sub>U<sub>321</sub> loop residues for Larp7 recognition of 7SK RNA (Muniz et al., 2013; Uchikawa et al., 2015; Brogie and Price, 2017).

The C-terminal atypical xRRM2 is required for specific recognition of 7SK RNA and binds to the SL4 apical loop (Uchikawa et al., 2015; Eichhorn et al., 2018). The xRRM2 is a signature of LARP7 proteins, initially identified in the *Tetrahymena thermophila* telomerase protein p65 (Singh et al., 2012). A solution NMR structure of the human Larp7 xRRM2 showed a similar global structure to the *Tetrahymena* p65 xRRM2 (Eichhorn et al., 2016). Like canonical RRM2s, the xRRM2 contains a βαββαβ topology with α-helices packing underneath an antiparallel β-sheet surface, which in canonical RRM2s serves as the RNA binding surface. However, unlike canonical RRM2s, the xRRM2 contains an additional helix α3 that is stably fixed atop the β-sheet, occluding this surface from RNA binding. In addition, the RRM-conserved RNP1 and RNP2 sequences on β3 and β1, respectively, are absent, and the xRRM2 contains a conserved RNP3 sequence on β2 (Singh et al., 2012; Basu et al., 2021). An X-ray crystal structure of the human Larp7 C-terminal domain showed that, consistent with the mode of xRRM2–RNA recognition, RNA bound to the side of the β-sheet rather than the surface (Figure 6C). Helix α3 has extensive contacts to RNA and lies along the major groove to direct the C-terminal residues down the RNA helical axis. Helix α3 residue W<sub>533</sub> formed a stacking interaction with N-terminal residue F<sub>451</sub> to position the N-terminus alongside helix α3 and down the SL4 helical axis (Figure 6C). This interaction may help to position the RRM1 helix α3 in close proximity to xRRM2. CryoEM structures of the 7SK core RNP showed that the xRRM2 interacts with the SL4 apical loop in the same manner as observed in the previously determined X-ray structure. The crystal construct lacked α4/MID residues, and in the cryoEM structure helix α4 crosses SL4 and engages in hydrophobic interactions with the Larp7 RRM1 helix α3 and MePCE MTase (Figure 7). Together, Larp7 interacts with the entire length of SL4, predominantly in the major groove and loops, explaining how Larp7 recognition of 7SK RNA protects the 3' end from exonucleolytic cleavage and promotes Larp7–MePCE binding.

## MePCE–7SK RNA–Larp7 ternary complex

7SK RNA secondary structure heterogeneity has been a major obstacle to obtaining high-resolution structural information of 7SK RNP macromolecular complexes. One solution to this problem was presented by Feigon and coworkers (Yang et al., 2022). Minimal 7SK RNA constructs were rationally designed that included 5' and 3' regions, required to bind MePCE and Larp7, and lacked the central RNA stem-loops. A similar SL1–SL4 construct was previously shown to be competent for P-TEFb assembly (Van Herreweghe et al., 2007; Briese et al., 2018), indicating this construct is biologically relevant. To address the question of “linear” and “circular” 7SK RNA secondary structure models, minimal 7SK RNA constructs representing either “linear” or “circular” models were *in vitro* reconstituted with recombinantly expressed human MePCE

MTase and full-length human Larp7 to form minimal 7SK core RNP ternary complexes. CryoEM structures were determined for both complexes and found to be overall similar (Figures 7A, B), suggesting the core RNP can accommodate different 7SK RNA conformations (Yang et al., 2022). The most substantial difference is observed at the 5' hairpin: in the “linear” 7SK core RNP SL1 showed no visible density and could not be modeled in the structure (Figure 7A) whereas in the “circular” 7SK core RNP SL0 had clear density and could be modeled (Figure 7B).

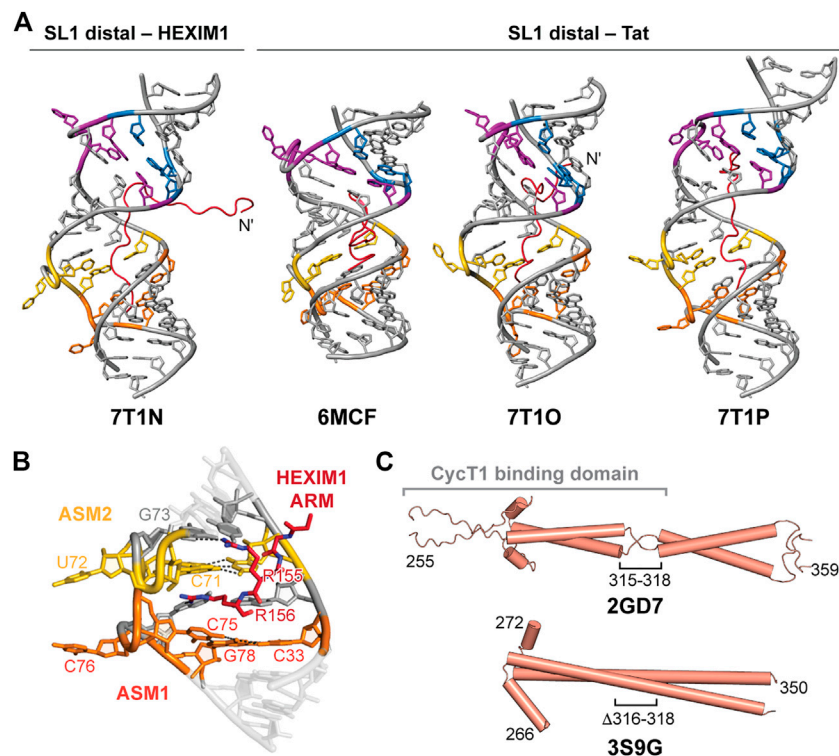
The ternary structures show distinct conformational differences when compared to individual domains, suggesting that 7SK core RNP assembly induces structural changes to both RNA and protein. RNA binding to Larp7 fixes the orientations of the three Larp7 domains (La module, xRRM2, α4/MID), providing a binding surface for MePCE–Larp7 interaction. In addition to protein–RNA interactions, there is an extensive MePCE–Larp7 interaction interface at MePCE α3–β3 loop, Larp7 RRM1 α3, and α4/MID (Figure 7C). For X-ray crystal structures of MePCE in the absence and presence of RNA, the 61 aa α3–β3 loop is disordered (Figures 5A, B). However, in the ternary complex an additional helix α3' and ordered loop form to interact with Larp7 RRM1 α3 and α4/MID (Figure 7C). The previously determined X-ray crystal structure of the MePCE–SL1 complex (Yang et al., 2019) showed specific interactions with the terminal base-pair and 3' ssRNA overhang (Figure 5C). In the representative “circular” 7SK RNP structure, MePCE also binds to the SL0 terminal base-pair and 3' ssRNA overhang (Figures 7B, C). However, the RNA binding pocket is remodeled such that the triphosphate binding tunnel is disorganized, particularly helices α0 and α7, thereby disrupting active site geometry. These structures explain enzymatic inactivation of MePCE induced by Larp7 and provide the first structural insights into 7SK RNP macromolecular assembly.

## 7SK RNP in P-TEFb sequestration and inactivation

The prominence of P-TEFb in human health and disease has made it a promising therapeutic target for wide-ranging diseases such as cancer (Anshabo et al., 2021), cardiac hypertrophy (Krystof et al., 2010), and HIV-1 (Massari et al., 2013). Over two dozen X-ray crystal structures have been determined of P-TEFb, the majority of which are bound to small molecule inhibitors, HIV-1 Tat, or SEC domain components. More recently, cryoEM structures of the RNAPII–SEC complex have provided key insights into delivery of P-TEFb to RNAPII and mechanisms of transcription elongation (Schier and Taatjes, 2020; Chen et al., 2021). In sharp contrast, there are at present no structures of P-TEFb bound to any component of 7SK RNP. Here, we summarize current understanding of HEXIM1 structure, HEXIM1–7SK RNA recognition, and biochemical evidence for HEXIM1–P-TEFb recognition.

## HEXIM1 and HEXIM1–7SK RNA

HEXIM1 is a promiscuous double-stranded RNA (dsRNA) binding protein (Li et al., 2007; Michels and Bensaude, 2018)



**FIGURE 8**

HEXIM1 and ARM recognition of 7SK SL1 RNA. **(A)** Solution NMR structures of SL1 distal bound to Tat or the first segment of the HEXIM1 ARM peptides. **(B)** ASM1-2 interacts with arginine residues 155–156 in HEXIM1 ARM (PDB ID 7T1N). **(C)** Solution NMR (PDB ID 2GD7) and X-ray crystal (PDB ID 3S9G) structures of HEXIM1 coiled-coil domain, with the putative CycT1 binding domain indicated in brackets.

that regulates differentiation, inflammation, development, and heart physiology (Michels et al., 2004). HEXIM1/2 contains an unstructured N-terminus and coiled-coil domain at the C-terminus that mediates dimerization (Dames et al., 2007; Schonichen et al., 2010). HEXIM1 contains several functional domains: a variable N-terminal region (residues 1–120); basic region (BR), alternately called an arginine rich motif (ARM, residues 150–165), and entwined bipartite nuclear localization signal (NLS, residues 150–177) (Yik et al., 2004); conserved PYNT motif (residues 202–205); acidic region (AR, residues 210–250); and C-terminal coiled-coil domain (Figure 2B). The N-terminal region appears to have an autoinhibitory effect to prevent HEXIM1–P-TEFb association in the absence of 7SK RNA (Michels et al., 2004; Li et al., 2005). In HEXIM1 proteins the ARM contains two basic segments, KKKHRRR and KKKRHWK, linked by a conserved PS sequence (Yik et al., 2004). The ARM binds to the SL1 distal region of 7SK RNA (Lebars et al., 2010). Phosphorylation of HEXIM1 ARM residue S<sub>158</sub> by PKC abolishes interactions with 7SK RNA, likely through electrostatic repulsion (Fujinaga et al., 2012). As a dimer, HEXIM1 has up to two RNA binding sites. One well-established binding site is the GAUC2-GAUC3 motif on the distal 5' SL1 (Figures 3, 4, 8) (Lebars et al., 2010). A second site has been identified at the proximal end of SL1 for the “linear” 7SK RNA model (Muniz et al., 2010). Although SL1 is remodeled in the “circular” conformation, SL1alt contains a single putative

HEXIM1 binding site (GAUC1-GAUC2) (Figures 3D, F), and an RNA construct of SL1alt was able to bind HEXIM1 (Li et al., 2007).

D'Souza and coworkers determined four solution NMR structures of the 7SK RNA SL1 distal construct, used previously for structure determination, in complex with the first segment of the HEXIM1 ARM RNA-binding motif (residues 145–156) or Tat ARM variants (residues 44–60) (Figure 8A; Table 2). Tat and HEXIM1 ARM peptides bound in the major groove between ASM1 and ASM4, but did not bind to ASM motif base triples. Rather, ARM peptide residues were primarily found to interact with base-pairs adjacent to base triples, acting as a “sandwich” around the base triple, as well as to the phosphodiester backbone (Figures 8A, B). Peptide-bound structures showed wide variance in binding modes, potentially due to degeneracy in available hydrogen bonding sites within the major groove and backbone (Figure 8A). Tat ARM, which is enriched in arginine residues compared to HEXIM1 (Table 2), appears to have increased RNA interactions compared to HEXIM1 ARM particularly at the N-terminus (Figure 8A). Four arginine residues in Tat peptides (R<sub>52</sub>, R<sub>53</sub>, R<sub>56</sub>, R<sub>57</sub>) intercalate into the major groove and interact with ASM1, ASM2, ASM3, and ASM4; In contrast, only two arginine residues in HEXIM1 peptide (R<sub>155</sub> and R<sub>156</sub>) interacted with ASM1 and ASM2 (Figure 8B). In binding assays, Tat peptide preferentially bound to SL1 distal compared to HEXIM1 peptide (Pham et al., 2018), consistent with previous studies showing that Tat could displace HEXIM1 from 7SK RNA (Muniz et al., 2010).

In addition to serving as a dimer interface, the C-terminal coiled-coil domain contains the CycT1 binding domain (TBD) (Dames et al., 2007). Structures have been determined of the HEXIM1 TBD domain by solution NMR and X-ray crystallography (Table 2). The solution NMR structure of the HEXIM1 TBD (residues 255–339) (Table 2) showed that the coiled-coil motif forms a dimer *via* leucine zipper (Figure 8C). In addition, a “stammer” motif (residues 315–318) was identified, which is a short unstructured linker that disrupts the formation of a continuous  $\alpha$ -helix (Figure 8C). An X-ray crystal structure of the HEXIM1 TBD lacking the stammer motif (TBD- $\Delta$ stammer,  $\Delta$ 316–318) showed that the coiled-coil motif forms a continuous  $\alpha$ -helix and repositions the second coiled-coil segment (Figure 8C).

## 7SK RNP assembly with P-TEFb

While domain structures of HEXIM1 ARM or coiled-coil motifs have been determined, structural information on HEXIM1 binding to P-TEFb and recruitment to 7SK RNP are lacking. HEXIM1 domains are involved in binding both CycT1 and CDK9. The stammer motif in the C-terminal TBD was shown to be important for CycT1 binding, and deletion resulted in a 5-fold reduction in HEXIM1–CycT1 association compared to wild-type (Schonichen et al., 2010). A model of HEXIM1-mediated P-TEFb inhibition has been proposed in which on binding to 7SK RNA, the HEXIM1 PYNT site is exposed to enable the assembly of P-TEFb (Michels et al., 2004). HEXIM1 directly interacts with the T-loop in the CDK9 active site through the central PYNT motif, masking the substrate-binding site to inhibit kinase activity (Michels et al., 2004; Kobbi et al., 2016; Michels and Bensaude, 2018). Assembly of P-TEFb with 7SK RNP requires that CDK9 T-loop residue T<sub>186</sub> is phosphorylated (Chen et al., 2004). Several phosphatases have been identified that dephosphorylate CDK9 T<sub>186</sub> (Chen et al., 2008; Gudipaty et al., 2015) as part of the P-TEFb release mechanism Table 1. In order to have P-TEFb catalytic activity restored, CDK9 T<sub>186</sub> must be re-phosphorylated upon P-TEFb release.

In addition to HEXIM1 requirement for P-TEFb assembly, 7SK SL4 and Larp7 are also required for stable association of P-TEFb onto 7SK RNP. In the HIV-1 TAR element, CycT1 and Tat cooperate to bind to the TAR apical loop (Schulze-Gahmen and Hurley, 2018). A similar interaction has been proposed for 7SK RNP, where CycT1–Tat interacts with SL4 (Durney and D’Souza, 2010). However, experimental evidence to support this model is lacking. Given the stable association of SL4 with Larp7 and MePCE, it is more likely that P-TEFb interacts with Larp7 rather than SL4. A C-terminal Larp7 construct (residues 375–589) was able to co-immunoprecipitate CDK9, supporting a Larp7–P-TEFb interaction (Markert et al., 2008). Although a direct interaction between CycT1 and 7SK RNA has not yet been demonstrated, it is possible that CycT1 may bind 7SK RNA in addition to HEXIM1 and Larp7.

## Conclusion and perspectives

Advances in 7SK RNP structural biology over the past decade have significantly improved our understanding of 7SK RNP assembly and maturation. A consensus model has started to develop that describes

early 7SK RNA biogenesis. Likewise, a nucleotide-resolution model is beginning to emerge that describes 7SK RNA secondary structure switching in response to different functional states. High-resolution structures of individual domains, RNA-protein complexes, and ternary complexes have increased understanding of HEXIM1 recognition of 7SK RNA, as well as assembly into a stable 7SK core RNP. Rich biochemical data have provided insights into P-TEFb sequestration and kinase inactivation and have revealed alternate pathways of 7SK RNP function in RNAPII transcription regulation. However, despite these advances, several outstanding questions remain unresolved, particularly regarding the structural basis for P-TEFb assembly, kinase inactivation, and release.

What is the 7SK RNA secondary structural landscape, and how is 7SK RNA secondary structure remodeling achieved? Several secondary structure models have been proposed using biochemical or bioinformatics approaches. While some 7SK RNA regions have common secondary structures across models, residues spanning SL1–SL2 vary substantially. Several helicases have been loosely identified to bind to the central region of 7SK RNA (Table 1). It has been proposed that helicase recognition leads to secondary structure switching, enabling the assembly or release of downstream accessory proteins such as P-TEFb. However, at present there is little mechanistic or structural understanding of how these proteins alter 7SK RNA secondary structure. Recent chemical probing-based approaches have shown that 7SK RNA secondary structure differs based on protein composition (Flynn et al., 2016; Brogie and Price, 2017), and that 7SK RNA exists as an ensemble of multiple secondary structures (Luo et al., 2021; Olson et al., 2022). However, direct evidence connecting a specific secondary structure with a specific protein-bound state remains tenuous. More studies combining proteomics with chemical probing experiments are needed to validate the proposed secondary structural models and switching mechanism. It is plausible that 7SK RNA forms alternate secondary structures that are not yet accounted for, and it will be interesting to see how future studies continue to shape our understanding of the complexity of 7SK RNA structure. Beyond protein-mediated RNA structural dynamics, there is a growing acknowledgement of the impact of posttranscriptional modifications on RNA structure and function (Lewis et al., 2017). How these modifications (Table 1) impact 7SK RNA structure and protein recognition is unexplored territory. Finally, although progress has been made toward modeling 7SK RNA secondary structure, it remains unclear how RNA domains are organized in 3D space to enable protein recognition.

How is P-TEFb assembled onto 7SK RNP, and what is the mechanism of kinase inactivation and re-activation? A picture is starting to emerge regarding HEXIM1 recognition of 7SK RNA, and HEXIM1 recognition of P-TEFb. However, a lack of structural information has impeded detailed understanding of 7SK RNP sequestration and inactivation of P-TEFb. In addition to HEXIM1, evidence suggests that P-TEFb interacts with Larp7 (Markert et al., 2008). Moreover, 7SK RNA may potentially interact with P-TEFb to further promote stable association. Beyond P-TEFb, other components of the SEC may also associate with 7SK RNP (Table 1). When bound to 7SK RNP, P-TEFb is sequestered in a catalytically inactive state. HEXIM1 has been proposed to act as a steric block to prevent CDK9-substrate interactions. Several studies have shown that CDK9 phosphorylation status, particularly of active site residue T<sub>186</sub>,

changes as part of the release mechanism from 7SK RNP. The basis for P-TEFb kinase inactivation, and re-activation upon 7SK RNP release, remains unclear. Additional studies are needed to gain an atomistic understanding of P-TEFb assembly with 7SK RNP, P-TEFb inactivation when bound to 7SK RNP, and re-activation upon release from 7SK RNP.

How do accessory proteins bind 7SK RNP to direct P-TEFb release and alternate functional pathways? Numerous proteins have been identified to be involved in P-TEFb release from 7SK RNP, including RNA processing factors, transcription factors, and helicases (Table 1). Rather than act as a universal transcriptional activator, it appears that specific cellular signals promote P-TEFb release and re-activation to stimulate expression of specific genes. 7SK RNP function appears to be particularly critical for regulating cell proliferation (Zhou and Yik, 2006; Musavi et al., 2019) and DNA damage response (Bugai et al., 2019; Lans et al., 2019; Studniarek et al., 2021). In addition, 7SK RNP appears to regulate RNAPII in multiple pathways. Intriguingly, 7SK RNP is found in different subcellular compartments in each of these pathways including the nucleolus (He et al., 2006), Cajal bodies (Egloff et al., 2017), nuclear speckles (Barboric et al., 2005; Prasanth et al., 2010; Ji et al., 2013), and cytoplasm (Briese et al., 2018; Ji et al., 2021). It is tempting to speculate that assembly of specific accessory proteins directs the 7SK RNP to different intracellular locations to carry out diverse functions.

Although substantial progress has been made in 7SK RNP structural biology, the sparsity of determined structures has remained an obstacle toward gaining a detailed understanding of how 7SK RNP assembles and functions in transcription regulation. The inherent structural plasticity of 7SK RNA and heterogeneity of 7SK RNP particles has posed a significant challenge toward structural elucidation. Overcoming these challenges will be critical to advance understanding of 7SK RNP structure, particularly for macromolecular complexes. Ultimately, structural biology coupled with biochemical, -omics, and cellular approaches will enable a comprehensive understanding of 7SK RNP cellular function and role in human health and disease.

## References

- Aigner, S., and Cech, T. R. (2004). The Euplotes telomerase subunit p43 stimulates enzymatic activity and processivity *in vitro*. *RNA* 10 (7), 1108–1118. doi:10.1261/rna.7400704
- Alazami, A. M., Al-Owain, M., Alzahrani, F., Shuaib, T., Al-Shamrani, H., Al-Falki, Y. H., et al. (2012). Loss of function mutation in LARP7, chaperone of 7SK ncRNA, causes a syndrome of facial dysmorphism, intellectual disability, and primordial dwarfism. *Hum. Mutat.* 33 (10), 1429–1434. doi:10.1002/humu.22175
- Anshabo, A. T., Milne, R., Wang, S., and Albrecht, H. (2021). CDK9: A comprehensive review of its biology, and its role as a potential target for anti-cancer agents. *Front. Oncol.* 11, 678559. doi:10.3389/fonc.2021.678559
- Bacon, C. W., and D'Orso, I. (2019). CDK9: A signaling hub for transcriptional control. *Transcription* 10 (2), 57–75. doi:10.1080/21541264.2018.1523668
- Barboric, M., Kohoutek, J., Price, J. P., Blazek, D., Price, D. H., and Peterlin, B. M. (2005). Interplay between 7SK snRNA and oppositely charged regions in HEXIM1 direct the inhibition of P-TEFb. *EMBO J.* 24 (24), 4291–4303. doi:10.1038/sj.emboj.7600883
- Barrandon, C., Bonnet, F., Nguyen, V. T., Labas, V., and Bensaude, O. (2007). The transcription-dependent dissociation of P-TEFb-HEXIM1-7SK RNA relies upon formation of hnRNP-7SK RNA complexes. *Mol. Cell Biol.* 27 (20), 6996–7006. doi:10.1128/MCB.00975-07
- Basu, R., Eichhorn, C. D., Cheng, R., Peterson, R. D., and Feigon, J. (2021). Structure of *S. pombe* telomerase protein Pof8 C-terminal domain is an xRRM conserved among LARP7 proteins. *RNA Biol.* 18 (8), 1181–1192. doi:10.1080/15476286.2020.1836891
- Bayfield, M. A., Yang, R., and Maraia, R. J. (2010). Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochimica biophysica acta* 1799 (5-6), 365–378. doi:10.1016/j.bbagr.2010.01.011
- Bigalke, J. M., Dames, S. A., Blankenfeldt, W., Grzesiek, S., and Geyer, M. (2011). Structure and dynamics of a stabilized coiled-coil domain in the P-TEFb regulator Hexim1. *J. Mol. Biol.* 414 (5), 639–653. doi:10.1016/j.jmb.2011.10.022
- Bourbigot, S., Dock-Bregeon, A. C., Eberling, P., Coutant, J., Kieffer, B., and Lebars, I. (2016). Solution structure of the 5'-terminal hairpin of the 7SK small nuclear RNA. *RNA* 22 (12), 1844–1858. doi:10.1261/rna.056523.116
- Bowman, E. A., and Kelly, W. G. (2014). RNA polymerase II transcription elongation and pol II ctd Ser2 phosphorylation: A tail of two kinases. *Nucleus* 5 (3), 224–236. doi:10.4161/nucl.29347
- Bres, V., Yoh, S. M., and Jones, K. A. (2008). The multi-tasking P-TEFb complex. *Curr. Opin. Cell Biol.* 20 (3), 334–340. doi:10.1016/j.cob.2008.04.008
- Briese, M., Saal-Bauernschubert, L., Ji, C., Moradi, M., Ghanawi, H., Uhl, M., et al. (2018). hnRNP R and its main interactor, the noncoding RNA 7SK, coregulate the axonal transcriptome of motoneurons. *Proc. Natl. Acad. Sci. U. S. A.* 115 (12), E2859–E2868. doi:10.1073/pnas.1721670115
- Briese, M., and Sendtner, M. (2021). Keeping the balance: The noncoding RNA 7SK as a master regulator for neuron development and function. *Bioessays* 43 (8), e2100092. doi:10.1002/bies.202100092

## Author contributions

CDE conceived the article and CDE and MBC wrote the manuscript. CDE, MBC, and AMS prepared the figures and edited the manuscript. All authors have approved the work for publication.

## Funding

We acknowledge funding support from the National Institutes of Health (1R35GM143030), the National Science Foundation (2047328), the Nebraska Center for Integrated Biomolecular Communication (P20 GM113126), and UNL startup funds to CDE.

## Acknowledgments

We thank members of the Eichhorn lab for insightful discussions and feedback.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Brillet, K., Martinez-Zapien, D., Bec, G., Ennifar, E., Dock-Bregeon, A. C., and Lebars, I. (2020). Different views of the dynamic landscape covered by the 5'-hairpin of the 7SK small nuclear RNA. *RNA* 26 (9), 1184–1197. doi:10.1261/rna.074955.120
- Brodsky, A. S., Erlacher, H. A., and Williamson, J. R. (1998). NMR evidence for a base triple in the HIV-2 TAR C-G-C+ mutant-argininamide complex. *Nucleic Acids Res.* 26 (8), 1991–1995. doi:10.1093/nar/26.8.1991
- Brogie, J. E., and Price, D. H. (2017). Reconstitution of a functional 7SK snRNP. *Nucleic Acids Res.* 45 (11), 6864–6880. doi:10.1093/nar/gkx262
- Bugai, A., Quaresma, A. J. C., Friedel, C. C., Lenasi, T., Duster, R., Sibley, C. R., et al. (2019). P-TEFb activation by RBM7 shapes a pro-survival transcriptional response to genotoxic stress. *Mol. Cell* 74 (2), 254–267.e10. doi:10.1016/j.molcel.2019.01.033
- Calo, E., Flynn, R. A., Martin, L., Spitale, R. C., Chang, H. Y., and Wysocka, J. (2015). RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* 518 (7538), 249–253. doi:10.1038/nature13923
- Catalucci, D., and Condorelli, G. (2013). HEXIM1: A new player in myocardial hypertrophy? *Cardiovasc Res.* 99 (1), 1–3. doi:10.1093/cvr/cvt134
- Chen, R., Liu, M., Li, H., Xue, Y., Ramey, W. N., He, N., et al. (2008). PP2B and PP1alpha cooperatively disrupt 7SK snRNP to release P-TEFb for transcription in response to Ca<sup>2+</sup> signaling. *Genes & Dev.* 22 (10), 1356–1368. doi:10.1101/gad.1636008
- Chen, R., Yang, Z., and Zhou, Q. (2004). Phosphorylated positive transcription elongation factor b (P-TEFb) is tagged for inhibition through association with 7SK snRNA. *J. Biol. Chem.* 279 (6), 4153–4160. doi:10.1074/jbc.M310044200
- Chen, Y., Vos, S. M., Dienemann, C., Ninov, M., Urlaub, H., and Cramer, P. (2021). Allosteric transcription stimulation by RNA polymerase II super elongation complex. *Mol. Cell* 81 (16), 3386–3399.e10. doi:10.1016/j.molcel.2021.06.019
- Cheng, Y., Jin, Z., Agarwal, R., Ma, K., Yang, J., Ibrahim, S., et al. (2012). LARP7 is a potential tumor suppressor gene in gastric cancer. *Laboratory investigation; a J. Tech. methods pathology* 92, 1013–1019. doi:10.1038/labinvest.2012.59
- Collopy, L. C., Ware, T. L., Goncalves, T., S. I. K., Yang, Q., Amelina, H., et al. (2018). LARP7 family proteins have conserved function in telomerase assembly. *Nat. Commun.* 9 (1), 557. doi:10.1038/s41467-017-02296-4
- Cordsmeier, A., Rinkel, S., Jenning, M., Schulze-Luehrmann, J., Olke, M., Schmid, B., et al. (2022). The *Coxiella burnetii* T4SS effector protein AnkG hijacks the 7SK small nuclear ribonucleoprotein complex for reprogramming host cell transcription. *PLoS Pathog.* 18 (2), e1010266. doi:10.1371/journal.ppat.1010266
- Cosgrove, M. S., Ding, Y., Rennie, W. A., Lane, M. J., and Hanes, S. D. (2012). The Bin3 RNA methyltransferase targets 7SK RNA to control transcription and translation. *Wiley Interdiscip. Rev. RNA* 3 (5), 633–647. doi:10.1002/wrna.1123
- Covelo-Molares, H., Obrdlík, A., Postulkova, I., Dohnalkova, M., Gregorova, P., Ganji, R., et al. (2021). The comprehensive interactomes of human adenosine RNA methyltransferases and demethylases reveal distinct functional and regulatory features. *Nucleic Acids Res.* 49 (19), 10895–10910. doi:10.1093/nar/gkab900
- Dai, Q., Luan, G., Deng, L., Lei, T., Kang, H., Song, X., et al. (2014). Primordial dwarfism gene maintains *Lin28* expression to safeguard embryonic stem cells from premature differentiation. *Cell Rep.* 7 (3), 735–746. doi:10.1016/j.celrep.2014.03.053
- Dames, S. A., Schonichen, A., Schulte, A., Barboric, M., Peterlin, B. M., Grzesiek, S., et al. (2007). Structure of the Cyclin T binding domain of Hexim1 and molecular basis for its recognition of P-TEFb. *Proc. Natl. Acad. Sci. U. S. A.* 104 (36), 14312–14317. doi:10.1073/pnas.0701848104
- Dollinger, R., and Gilmour, D. S. (2021). Regulation of promoter proximal pausing of RNA polymerase II in metazoans. *J. Mol. Biol.* 433 (14), 166897. doi:10.1016/j.jmb.2021.166897
- Driscoll, C. T., Darlington, G. J., and Maraia, R. J. (1994). The conserved 7SK snRNA gene localizes to human chromosome 6 by homolog exclusion probing of somatic cell hybrid RNA. *Nucleic Acids Res.* 22 (5), 722–725. doi:10.1093/nar/22.5.722
- Durney, M. A., and D'Souza, V. M. (2010). Preformed protein-binding motifs in 7SK snRNA: Structural and thermodynamic comparisons with retroviral TAR. *J. Mol. Biol.* 404 (4), 555–567. doi:10.1016/j.jmb.2010.08.042
- 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 (2), 630–642. doi:10.1128/MCB.26.2.630-642.2006
- Egloff, S., Vitali, P., Tellier, M., Raffel, R., Murphy, S., and Kiss, T. (2017). The 7SK snRNP associates with the little elongation complex to promote snRNA gene expression. *EMBO J.* 36 (7), 934–948. doi:10.15252/embj.201695740
- 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 (20), 9977–9989. doi:10.1093/nar/gkw833
- 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. U. S. A.* 115 (28), E6457–E6466. doi:10.1073/pnas.1806276115
- Flynn, R. A., Do, B. T., Rubin, A. J., Calo, E., Lee, B., Kuchelmeister, H., et al. (2016). 7SK-BAF axis controls pervasive transcription at enhancers. *Nat. Struct. Mol. Biol.* 23 (3), 231–238. doi:10.1038/nsmb.3176
- Fujinaga, K., Barboric, M., Li, Q., Luo, Z., Price, D. H., and Peterlin, B. M. (2012). PKC phosphorylates HEXIM1 and regulates P-TEFb activity. *Nucleic Acids Res.* 40 (18), 9160–9170. doi:10.1093/nar/gks682
- Fujinaga, K., Huang, F., and Peterlin, B. M. (2022). P-TEFb: The master regulator of transcription elongation. *Mol. Cell* 83, 393–403. doi:10.1016/j.molcel.2022.12.006
- Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 583 (7816), 459–468. doi:10.1038/s41586-020-2286-9
- Gruber, A. R., Koper-Emde, D., Marz, M., Tafer, H., Bernhart, S., Obernosterer, G., et al. (2008). Invertebrate 7SK snRNAs. *J. Mol. Evol.* 66 (2), 107–115. doi:10.1007/s00239-007-9052-6
- Gudipaty, S. A., McNamara, R. P., Morton, E. L., and D'Orso, I. (2015). PPM1G binds 7SK RNA and Hexim1 to block P-TEFb assembly into the 7SK snRNP and sustain transcription elongation. *Mol. Cell Biol.* 35 (22), 3810–3828. doi:10.1128/MCB.00226-15
- Gupta, S., Busch, R. K., Singh, R., and Reddy, R. (1990). Characterization of U6 small nuclear RNA cap-specific antibodies. Identification of gamma-monomethyl-GTP cap structure in 7SK and several other human small RNAs. *J. Biol. Chem.* 265 (31), 19137–19142. doi:10.1016/s0021-9258(17)30635-x
- Gurney, T., Jr., and Eliceiri, G. L. (1980). Intracellular distribution of low molecular weight RNA species in HeLa cells. *J. Cell Biol.* 87 (2), 398–403. doi:10.1083/jcb.87.2.398
- Gursoy, 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 (5), 456–464. doi:10.1007/s002390010048
- Hasler, D., Meister, G., and Fischer, U. (2021). Stabilize and connect: The role of LARP7 in nuclear non-coding RNA metabolism. *RNA Biol.* 18 (2), 290–303. doi:10.1080/15476286.2020.1767952
- He, N., Jahchan, N. S., Hong, E., Li, Q., Bayfield, M. A., Maraia, R. J., et al. (2008). A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol. Cell* 29 (5), 588–599. doi:10.1016/j.molcel.2008.01.003
- He, N., and Zhou, Q. (2011). New insights into the control of HIV-1 transcription: When Tat meets the 7SK snRNP and super elongation complex (SEC). *J. Neuroimmune Pharmacol.* 6 (2), 260–268. doi:10.1007/s11481-011-9267-6
- He, W. J., Chen, R., Yang, Z., and Zhou, Q. (2006). Regulation of two key nuclear enzymatic activities by the 7SK small nuclear RNA. *Cold Spring Harb. Symp. Quant. Biol.* 71, 301–311. doi:10.1101/sqb.2006.71.019
- Huthoff, H., Girard, F., Wijmenga, S. S., and Berkhout, B. (2004). Evidence for a base triple in the free HIV-1 TAR RNA. *RNA* 10 (3), 412–423. doi:10.1261/rna.5161304
- Jang, M. K., Mochizuki, K., Zhou, M., Jeong, H. S., Brady, J. N., and Ozato, K. (2005). The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell* 19 (4), 523–534. doi:10.1016/j.molcel.2005.06.027
- Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., et al. (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol. Cell* 27 (2), 262–274. doi:10.1016/j.molcel.2007.06.027
- Ji, C., Bader, J., Ramanathan, P., Hennlein, L., Meissner, F., Jablonka, S., et al. (2021). Interaction of 7SK with the Snn complex modulates snRNP production. *Nat. Commun.* 12 (1), 1278. doi:10.1038/s41467-021-21529-1
- Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C. Y., et al. (2013). SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* 153 (4), 855–868. doi:10.1016/j.cell.2013.04.028
- Jonkers, I., and Lis, J. T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 16 (3), 167–177. doi:10.1038/nrm3953
- Kobbi, L., Demey-Thomas, E., Braye, F., Proux, F., Kolesnikova, O., Vinh, J., et al. (2016). An evolutionary conserved Hexim1 peptide binds to the Cdk9 catalytic site to inhibit P-TEFb. *Proc. Natl. Acad. Sci. U. S. A.* 113 (45), 12721–12726. doi:10.1073/pnas.1612331113
- Krueger, B. J., Jeronimo, C., Roy, B. B., Bouchard, A., Barrandon, C., Byers, S. A., et al. (2008). LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res.* 36 (7), 2219–2229. doi:10.1093/nar/gkn061
- 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 (8), e12335. doi:10.1371/journal.pone.0012335
- Krystof, V., Chamrad, I., Jorda, R., and Kohoutek, J. (2010). Pharmacological targeting of CDK9 in cardiac hypertrophy. *Med. Res. Rev.* 30 (4), 646–666. doi:10.1002/med.20172
- Lans, H., Hoeijmakers, J. H. J., Vermeulen, W., and Marteijn, J. A. (2019). The DNA damage response to transcription stress. *Nat. Rev. Mol. Cell Biol.* 20 (12), 766–784. doi:10.1038/s41580-019-0169-4

- Lebars, I., Martinez-Zapien, D., Durand, A., Coutant, J., Kieffer, B., and Dock-Bregeon, A. C. (2010). HEXIM1 targets a repeated GAUC motif in the riboregulator of transcription 7SK and promotes base pair rearrangements. *Nucleic Acids Res.* 38 (21), 7749–7763. doi:10.1093/nar/gkq660
- Lee, S., Liu, H., Hill, R., Chen, C., Hong, X., Crawford, F., et al. (2020). JMJD6 cleaves MePCE to release positive transcription elongation factor b (P-TEFb) in higher eukaryotes. *Elife* 9, e53930. doi:10.7554/eLife.53930
- Leger, A., Amaral, P. P., Pandolfini, L., Capitanchik, C., Capraro, F., Miano, V., et al. (2021). RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat. Commun.* 12 (1), 7198. doi:10.1038/s41467-021-27393-3
- Lewis, C. J., Pan, T., and Kalsotra, A. (2017). RNA modifications and structures cooperate to guide RNA-protein interactions. *Nat. Rev. Mol. Cell Biol.* 18 (3), 202–210. doi:10.1038/nrm.2016.163
- Li, Q., Cooper, J. J., Altwerger, G. H., Feldkamp, M. D., Shea, M. A., and Price, D. H. (2007). HEXIM1 is a promiscuous double-stranded RNA-binding protein and interacts with RNAs in addition to 7SK in cultured cells. *Nucleic Acids Res.* 35 (8), 2503–2512. doi:10.1093/nar/gkm150
- Li, Q., Price, J. P., Byers, S. A., Cheng, D., Peng, J., and Price, D. H. (2005). Analysis of the large inactive P-TEFb complex indicates that it contains one 7SK molecule, a dimer of HEXIM1 or HEXIM2, and two P-TEFb molecules containing Cdk9 phosphorylated at threonine 186. *J. Biol. Chem.* 280 (31), 28819–28826. doi:10.1074/jbc.M502712200
- Ling, T. T., and Sorrentino, S. (2015). Compound heterozygous variants in the LARP7 gene as a cause of Alazami syndrome in a Caucasian female with significant failure to thrive, short stature, and developmental disability. *Am. J. Med. Genet. A* 170A, 217–219. doi:10.1002/ajmg.a.37396
- Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., et al. (2013). Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* 155 (7), 1581–1595. doi:10.1016/j.cell.2013.10.056
- Lu, H., Li, Z., Xue, Y., Schulze-Gahmen, U., Johnson, J. R., Krogan, N. J., et al. (2014). AFF1 is a ubiquitous P-TEFb partner to enable Tat extraction of P-TEFb from 7SK snRNP and formation of SECs for HIV transactivation. *Proc. Natl. Acad. Sci. U. S. A.* 111 (1), E15–E24. doi:10.1073/pnas.1318503111
- Luo, L., Chiu, L. Y., Sugarman, A., Gupta, P., Rouskin, S., and Tolbert, B. S. (2021). HnRNP A1/A2 proteins assemble onto 7SK snRNA via context dependent interactions. *J. Mol. Biol.* 433 (9), 166885. doi:10.1016/j.jmb.2021.166885
- 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 (6). doi:10.1002/wrna.1430
- Markert, A., Grimm, M., Martinez, J., Wiesner, J., Meyerhans, A., Meyuhos, O., et al. (2008). The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep.* 9 (6), 569–575. doi:10.1038/embor.2008.72
- Martinez-Zapien, D., Legrand, P., McEwen, A. G., Proux, F., Cragnolini, T., Pasquali, S., et al. (2017). The crystal structure of the 5 functional domain of the transcription riboregulator 7SK. *Nucleic Acids Res.* 45 (6), 3568–3579. doi:10.1093/nar/gkw1351
- 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 (12), 2821–2830. doi:10.1093/molbev/msp198
- Massari, S., Sabatini, S., and Tabarrini, O. (2013). Blocking HIV-1 replication by targeting the Tat-hijacked transcriptional machinery. *Curr. Pharm. Des.* 19 (10), 1860–1879. doi:10.2174/1381612811319100010
- McNamara, R. P., Bacon, C. W., and D'Orso, I. (2016a). Transcription elongation control by the 7SK snRNP complex: Releasing the pause. *Cell Cycle* 15 (16), 2115–2123. doi:10.1080/15384101.2016.1181241
- McNamara, R. P., Guzman, C., Reeder, J. E., and D'Orso, I. (2016b). Genome-wide analysis of KAP1, the 7SK snRNP complex, and RNA polymerase II. *Genom Data* 7, 250–255. doi:10.1016/j.gdata.2016.01.019
- McNamara, R. P., Reeder, J. E., McMillan, E. A., Bacon, C. W., McCann, J. L., and D'Orso, I. (2016c). KAP1 recruitment of the 7SK snRNP complex to promoters enables transcription elongation by RNA polymerase II. *Mol. Cell* 61 (1), 39–53. doi:10.1016/j.molcel.2015.11.004
- 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 (1), 586. doi:10.1038/s41467-018-02874-0
- Michels, A. A., and Bensaude, O. (2018). Hexim1, an RNA-controlled protein hub. *Transcription* 9 (4), 262–271. doi:10.1080/21541264.2018.1429836
- Michels, A. A., Fraldi, A., Li, Q., Adamson, T. E., Bonnet, F., Nguyen, V. T., et al. (2004). Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.* 23 (13), 2608–2619. doi:10.1038/sj.emboj.7600275
- Michels, A. A., Nguyen, V. T., Fraldi, A., Labas, V., Edwards, M., Bonnet, F., et al. (2003). MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. *Mol. Cell Biol.* 23 (14), 4859–4869. doi:10.1128/MCB.23.14.4859-4869.2003
- Muck, F., Bracharz, S., and Marschalek, R. (2016). DDX6 transfers P-TEFb kinase to the AF4/AF4N (AFF1) super elongation complex. *Am. J. Blood Res.* 6 (3), 28–45.
- 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 (8), 4686–4698. doi:10.1093/nar/gkt159
- 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 (10), e1001152. doi:10.1371/journal.ppat.1001152
- Murphy, S., Di Liegro, C., and Melli, M. (1987). The *in vitro* transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. *Cell* 51 (1), 81–87. doi:10.1016/0092-8674(87)90012-2
- Musavi, M., Kohram, F., Abasi, M., Bolandi, Z., Ajoudanian, M., Mohammadi-Yeganeh, S., et al. (2019). Rn7SK small nuclear RNA is involved in cellular senescence. *J. Cell Physiol.* 234 (8), 14234–14245. doi:10.1002/jcp.28119
- 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 (6861), 322–325. doi:10.1038/35104581
- Olson, S. W., Turner, A. W., Arney, J. W., Saleem, I., Weidmann, C. A., Margolis, D. M., et al. (2022). Discovery of a large-scale, cell-state-responsive allosteric switch in the 7SK RNA using DANCE-MaP. *Mol. Cell* 82 (9), 1708–1723.e10. doi:10.1016/j.molcel.2022.02.009
- Paez-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 (1), 587. doi:10.1038/s41467-017-02284-8
- 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 (1), 92–103. doi:10.1002/wrna.106
- Peterlin, B. M., and Price, D. H. (2006). Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell* 23 (3), 297–305. doi:10.1016/j.molcel.2006.06.014
- Pham, V. V., Gao, M., Meagher, J. L., Smith, J. L., and D'Souza, V. M. (2022). A structure-based mechanism for displacement of the HEXIM adapter from 7SK small nuclear RNA. *Commun. Biol.* 5 (1), 819. doi:10.1038/s42003-022-03734-w
- Pham, V. V., Salguero, C., Khan, S. N., Meagher, J. L., Brown, W. C., Humbert, N., et al. (2018). HIV-1 Tat interactions with cellular 7SK and viral TAR RNAs identifies dual structural mimicry. *Nat. Commun.* 9 (1), 4266. doi:10.1038/s41467-018-06591-6
- Prasanth, K. V., Camiola, M., Chan, G., Tripathi, V., Denis, L., Nakamura, T., et al. (2010). Nuclear organization and dynamics of 7SK RNA in regulating gene expression. *Mol. Biol. Cell* 21 (23), 4184–4196. doi:10.1091/mbc.E10-02-0105
- Quaresma, A. J. C., Bugai, A., and Barboric, M. (2016). Cracking the control of RNA polymerase II elongation by 7SK snRNP and P-TEFb. *Nucleic Acids Res.* 44 (16), 7527–7539. doi:10.1093/nar/gkw585
- Reddy, R., Henning, D., Subrahmanyam, C. S., and Busch, H. (1984). Primary and secondary structure of 7-3 (K) RNA of Novikoff hepatoma. *J. Biol. Chem.* 259 (19), 12265–12270. doi:10.1016/s0021-9258(20)71349-9
- Roder, K., Stirnemann, G., Dock-Bregeon, A. C., Wales, D. J., and Pasquali, S. (2020). Structural transitions in the RNA 7SK 5' hairpin and their effect on HEXIM binding. *Nucleic Acids Res.* 48 (1), 373–389. doi:10.1093/nar/gkz1071
- Sano, M., Abdellatif, M., Oh, H., Xie, M., Bagella, L., Giordano, A., et al. (2002). Activation and function of cyclin T-Cdk9 (positive transcription elongation factor-b) in cardiac muscle-cell hypertrophy. *Nat. Med.* 8 (11), 1310–1317. doi:10.1038/nm778
- Santoro, M., Nociti, V., Lucchini, M., De Fino, C., Losavio, F. A., and Mirabella, M. (2016). Expression profile of long non-coding RNAs in serum of patients with multiple sclerosis. *J. Mol. Neurosci.* 59 (1), 18–23. doi:10.1007/s12031-016-0741-8
- Schier, A. C., and Taatjes, D. J. (2020). Structure and mechanism of the RNA polymerase II transcription machinery. *Genes Dev.* 34 (7-8), 465–488. doi:10.1101/gad.335679.119
- Schneeberger, P. E., Bierhals, T., Neu, A., Hempel, M., and Kutsche, K. (2019). De novo MEPCE nonsense variant associated with a neurodevelopmental disorder causes disintegration of 7SK snRNP and enhanced RNA polymerase II activation. *Sci. Rep.* 9 (1), 12516. doi:10.1038/s41598-019-49032-0
- Schonichen, A., Bigalke, J. M., Urbanke, C., Grzesiek, S., Dames, S. A., and Geyer, M. (2010). A flexible bipartite coiled coil structure is required for the interaction of Hexim1 with the P-TEFb subunit cyclin T1. *Biochemistry* 49 (14), 3083–3091. doi:10.1021/bi902072f
- Schroder, S., Cho, S., Zeng, L., Zhang, Q., Kaehlecke, K., Mak, L., et al. (2012). Two-pronged binding with bromodomain-containing protein 4 liberates positive transcription elongation factor b from inactive ribonucleoprotein complexes. *J. Biol. Chem.* 287 (2), 1090–1099. doi:10.1074/jbc.M111.282855
- Schulze-Gahmen, U., and Hurley, J. H. (2018). Structural mechanism for HIV-1 TAR loop recognition by Tat and the super elongation complex. *Proc. Natl. Acad. Sci. U. S. A.* 115 (51), 12973–12978. doi:10.1073/pnas.1806438115



- Shumyatsky, G. P., Tillib, S. V., and Kramerov, D. A. (1990). B2 RNA and 7SK RNA, RNA polymerase III transcripts, have a cap-like structure at their 5' end. *Nucleic Acids Res.* 18 (21), 6347–6351. doi:10.1093/nar/18.21.6347
- Singh, M., Wang, Z., Koo, B. K., Patel, A., Cascio, D., Collins, K., et al. (2012). Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol. Cell* 47 (1), 16–26. doi:10.1016/j.molcel.2012.05.018
- Singh, R., Gupta, S., and Reddy, R. (1990). Capping of mammalian U6 small nuclear RNA *in vitro* is directed by a conserved stem-loop and AUAUAC sequence: Conversion of a noncapped RNA into a capped RNA. *Mol. Cell Biol.* 10 (3), 939–946. doi:10.1128/mcb.10.3.939-946.1990
- Sithole, N., Williams, C. A., Abbink, T. E. M., and Lever, A. M. L. (2020). DDX5 potentiates HIV-1 transcription as a co-factor of Tat. *Retrovirology* 17 (1), 6. doi:10.1186/s12977-020-00514-4
- Stefano, J. E. (1984). Purified lupus antigen La recognizes an oligouridylylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* 36 (1), 145–154. doi:10.1016/0092-8674(84)90083-7
- 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 (2), 108965. doi:10.1016/j.celrep.2021.108965
- Tao, J., Chen, L., and Frankel, A. D. (1997). Dissection of the proposed base triple in human immunodeficiency virus TAR RNA indicates the importance of the Hoogsteen interaction. *Biochemistry* 36 (12), 3491–3495. doi:10.1021/bi962259t
- Thorne, L. G., Bouhaddou, M., Reuschl, A. K., Zuliani-Alvarez, L., Polacco, B., Pelin, A., et al. (2022). Evolution of enhanced innate immune evasion by SARS-CoV-2. *Nature* 602 (7897), 487–495. doi:10.1038/s41586-021-04352-y
- Uchikawa, E., Natchiar, K. S., Han, X., Proux, F., Roblin, P., Zhang, E., et al. (2015). Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. *Nucleic Acids Res.* 43 (6), 3373–3388. doi:10.1093/nar/gkv173
- Van Damme, R., Li, K., Zhang, M., Bai, J., Lee, W. H., Yesselman, J. D., et al. (2022). Chemical reversible crosslinking enables measurement of RNA 3D distances and alternative conformations in cells. *Nat. Commun.* 13 (1), 911. doi:10.1038/s41467-022-28602-3
- Van Herreweghe, E., Egloff, S., Goiffon, I., Jady, B. E., Froment, C., Monsarrat, B., et al. (2007). Dynamic remodelling of human 7SK snRNP controls the nuclear level of active P-TEFb. *EMBO J.* 26 (15), 3570–3580. doi:10.1038/sj.emboj.7601783
- Von Dwingelo, J., Chung, I. Y. W., Price, C. T., Li, L., Jones, S., Cygler, M., et al. (2019). Interaction of the ankyrin H core effector of Legionella with the host LARP7 component of the 7SK snRNP complex. *mBio* 10 (4), e01942–19. doi:10.1128/mBio.01942-19
- Wang, X., Li, Z. T., Yan, Y., Lin, P., Tang, W., Hasler, D., et al. (2020). LARP7-Mediated U6 snRNA modification ensures splicing fidelity and spermatogenesis in mice. *Mol. Cell* 77 (5), 999–1013.e6. doi:10.1016/j.molcel.2020.01.002
- Warda, A. S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Hobartner, C., et al. (2017). Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* 18 (11), 2004–2014. doi:10.15252/embr.201744940
- 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 (7), 3432–3445. doi:10.1128/mcb.11.7.3432-3445.1991
- Wu, P., Zuo, X., Deng, H., Liu, X., Liu, L., and Ji, A. (2013). Roles of long noncoding RNAs in brain development, functional diversification and neurodegenerative diseases. *Brain Res. Bull.* 97, 69–80. doi:10.1016/j.brainresbull.2013.06.001
- 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 (2), 360–369. doi:10.1093/nar/gkp977
- 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 (2), 132–140. doi:10.1038/s41589-018-0188-z
- Yang, Y., Liu, S., Egloff, S., Eichhorn, C. D., Hadian, T., Zhen, J., et al. (2022). Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP. *Mol. Cell* 82 (9), 1724–1736.e7. doi:10.1016/j.molcel.2022.03.001
- 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 (6861), 317–322. doi:10.1038/35104575
- Yazbeck, A. M., Tout, K. R., and Stadler, P. F. (2018). Detailed secondary structure models of invertebrate 7SK RNAs. *RNA Biol.* 15 (2), 158–164. doi:10.1080/15476286.2017.1412913
- 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 (4), 971–982. doi:10.1016/s1097-2765(03)00388-5
- Yik, J. H., Chen, R., Pezda, A. C., Samford, C. S., and Zhou, Q. (2004). A human immunodeficiency virus type 1 Tat-like arginine-rich RNA-binding domain is essential for HEXIM1 to inhibit RNA polymerase II transcription through 7SK snRNA-mediated inactivation of P-TEFb. *Mol. Cell Biol.* 24 (12), 5094–5105. doi:10.1128/MCB.24.12.5094-5105.2004
- Yik, J. H., Chen, R., Pezda, A. C., and Zhou, Q. (2005). Compensatory contributions of HEXIM1 and HEXIM2 in maintaining the balance of active and inactive positive transcription elongation factor b complexes for control of transcription. *J. Biol. Chem.* 280 (16), 16368–16376. doi:10.1074/jbc.M500912200
- Zhao, Y., Karijovich, J., Glaunsinger, B., and Zhou, Q. (2016). Pseudouridylation of 7SK snRNA promotes 7SK snRNP formation to suppress HIV-1 transcription and escape from latency. *EMBO Rep.* 17 (10), 1441–1451. doi:10.15252/embr.201642682
- Zhou, Q., and Yik, J. H. (2006). The Yin and Yang of P-TEFb regulation: Implications for human immunodeficiency virus gene expression and global control of cell growth and differentiation. *Microbiol. Mol. Biol. Rev. MMBR* 70 (3), 646–659. doi:10.1128/MMBR.00011-06
- Zieve, G., Benecke, B. J., and Penman, S. (1977). Synthesis of two classes of small RNA species *in vivo* and *in vitro*. *Biochemistry* 16 (20), 4520–4525. doi:10.1021/bi00639a029
- Zieve, G., and Penman, S. (1976). Small RNA species of the HeLa cell: Metabolism and subcellular localization. *Cell* 8 (1), 19–31. doi:10.1016/0092-8674(76)90181-1