
In vitro methylation of the U7 snRNP subunits Lsm11 and SmE by the PRMT5/MEP50/pICln methylosome

XIAO-CUI YANG,^{1,6} ANTHONY DESOTELL,^{2,6} MIN-HAN LIN,^{2,6} ANDREW S. PAIGE,^{2,6}
AGATA MALINOWSKA,³ YADONG SUN,^{2,7} WEI SHEN AIK,^{2,8} MICHAŁ DADLEZ,^{3,4} LIANG TONG,²
and ZBIGNIEW DOMINSKI^{1,5}

¹Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

²Department of Biological Sciences, Columbia University, New York, New York 10027, USA

³Department of Biophysics, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02-106 Warsaw, Poland

⁴Institute of Genetics and Biotechnology, Warsaw University, 02-106 Warsaw, Poland

⁵Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

ABSTRACT

U7 snRNP is a multisubunit endonuclease required for 3' end processing of metazoan replication-dependent histone pre-mRNAs. In contrast to the spliceosomal snRNPs, U7 snRNP lacks the Sm subunits D1 and D2 and instead contains two related proteins, Lsm10 and Lsm11. The remaining five subunits of the U7 heptameric Sm ring, SmE, F, G, B, and D3, are shared with the spliceosomal snRNPs. The pathway that assembles the unique ring of U7 snRNP is unknown. Here, we show that a heterodimer of Lsm10 and Lsm11 tightly interacts with the methylosome, a complex of the arginine methyltransferase PRMT5, MEP50, and pICln known to methylate arginines in the carboxy-terminal regions of the Sm proteins B, D1, and D3 during the spliceosomal Sm ring assembly. Both biochemical and cryo-EM structural studies demonstrate that the interaction is mediated by PRMT5, which binds and methylates two arginine residues in the amino-terminal region of Lsm11. Surprisingly, PRMT5 also methylates an amino-terminal arginine in SmE, a subunit that does not undergo this type of modification during the biogenesis of the spliceosomal snRNPs. An intriguing possibility is that the unique methylation pattern of Lsm11 and SmE plays a vital role in the assembly of the U7 snRNP.

Keywords: Lsm11; PRMT5 methylosome; Sm ring assembly; SmE; U7 snRNP

INTRODUCTION

Spliceosomal snRNAs generated by RNA polymerase II (U1, U2, U4, U5, U11, U12, and U4_{atac}) contain a conserved 9-nt Sm-binding site, AAUUU(U/G)UGG, that nucleates the assembly of a protein ring composed of seven Sm subunits: D1, D2, B, D3, E, F, and G (Raker et al. 1999; Will and Luhrmann 2001; Khusial et al. 2005). Each Sm protein occupies a specific position in the ring and contacts an individual nucleotide of the Sm-binding site (Kambach et al. 1999; Khusial et al. 2005; Leung et al. 2011; Li et al. 2016). The Sm proteins exist in the cytoplasm as three preformed heterooligomers, SmD1/D2, SmB/D3, and SmE/F/G (Raker et al.

1996) that are assembled around the Sm-binding site of the spliceosomal snRNAs in a multistep process mediated by the PRMT5 methylosome complex and the survival of motor neurons (SMN) complex (Battle et al. 2006a; Li et al. 2014; Gruss et al. 2017). Following the assembly of the Sm ring, spliceosomal snRNPs undergo additional maturation steps in the cytoplasm and the nucleus, including binding of specific accessory proteins, to become competent for their role in pre-mRNA splicing (Lührmann et al. 1990; Matera et al. 2007; Wilkinson et al. 2020).

The PRMT5 methylosome consists of PRMT5 (protein arginine methyltransferase 5), MEP50, and pICln (Pu et al. 1999; Friesen et al. 2001b, 2002; Meister et al. 2001; Gruss et al. 2017). Both biochemical and structural studies indicate that PRMT5 and MEP50 exist as a stable complex of four heterodimers, which recruits four pICln molecules through direct interaction with PRMT5 (Antonyamy et al. 2012; Ho et al. 2013; Timm et al. 2018; Mulvaney et al.

⁶These authors contributed equally to this work.

⁷**Present address:** School of Life Science and Technology, ShanghaiTech University, Shanghai, China

⁸**Present address:** Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China

Corresponding authors: dominski@med.unc.edu, ltong@columbia.edu

Article is online at <http://www.rnajournal.org/cgi/doi/10.1261/rna.079709.123>. Freely available online through the RNA Open Access option.

© 2023 Yang et al. This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

2021). PRMT5 (previously termed JBP1), the catalytic component of the complex, belongs to type II protein methyltransferases and generates monomethyl arginines (MMA) as intermediates and symmetric dimethyl arginines (sDMA) as final products (Bedford and Richard 2005). As most protein methyltransferases, PRMT5 has specificity for protein regions rich in glycines and arginines (GR-rich) (Najbauer et al. 1993; Stopa et al. 2015; Musiani et al. 2019). MEP50 (Meister et al. 2001; Friesen et al. 2002), also known as WD45, is a WD40-repeat protein and functions to enhance substrate specificity and catalytic activity of PRMT5 (Burgos et al. 2015). pICln is an adaptor component of the methylosome and delivers the SmD1/D2 and SmB/D3 heterodimers for PRMT5-catalyzed symmetrical arginine dimethylation of the GR-rich carboxy-terminal tails of the Sm subunits D1, B, and D3 (Chari et al. 2008; Pesiridis et al. 2009).

The seven Sm proteins, including the PRMT5-modified SmD1, SmB, and SmD3, are arranged around the spliceosomal snRNAs by the SMN complex during a multistep and a highly controlled process. One of these steps is the formation of a ring-shaped 6S intermediate consisting of the Sm proteins D1, D2, E, F, and G, and pICln. In the 6S intermediate, pICln temporarily substitutes for the missing SmB/D3 heterodimer and prevents the incorporation of illegitimate RNAs (Chari et al. 2008; Grimm et al. 2013). The SMN complex consists of the SMN protein of 30 kDa, Gemins2–8 and Unrip, with the function of only some of them being understood (Battle et al. 2006a; Gruss et al. 2017). The SMN protein is at the center of the complex and acts by directly binding PRMT5-methylated residues (Friesen et al. 2001a; Selenko et al. 2001; Cote and Richard 2005; Tripsianes et al. 2011), hence bringing multiple components of the SMN complex to the vicinity of the Sm proteins. Among these components, Gemin5 recognizes spliceosomal snRNAs as correct assembly targets by simultaneously binding to their monomethylated cap structure and Sm site (Battle et al. 2006b; Lau et al. 2009; Yong et al. 2010; Wahl and Fischer 2016). Gemin2 binds and stabilizes the SmD1/D2/E/F/G pentamer, making extensive contacts with all five subunits (Zhang et al. 2011; Grimm et al. 2013; Yi et al. 2020). This step likely occurs after the departure of pICln from the 6S intermediate and prior to the delivery of a spliceosomal snRNA by Gemin5. The assembly process is completed by the ring closure upon the recruitment of symmetrically dimethylated SmB/D3 heterodimer.

Deficiency of functional SMN protein results in spinal muscular atrophy (SMA), a genetic disorder characterized by selective degeneration of motor neurons and progressive paralysis (Lefebvre et al. 1995; Gubitz et al. 2004; Iannaccone et al. 2004; Pellizzoni 2007; Chari et al. 2009). It is believed that at the molecular level, SMA is caused by inefficient assembly of the spliceosomal snRNPs and aberrant splicing of mRNA precursors important for the devel-

opment and function of motor neurons (Winkler et al. 2005; Zhang et al. 2008; Lotti et al. 2012). Alternatively, the SMN protein, besides its universal role in snRNP biogenesis, may have a tissue-specific function in motor neurons, explaining the selective death of only this group of cells in SMA patients (Monani 2005; Burghes and Beattie 2009; Fallini et al. 2012).

In addition to the spliceosomal snRNPs, animal cells (but not plants or lower eukaryotes) contain the U7 snRNP in which Sm protein D1 and D2 are replaced by the Sm-like proteins Lsm10 and Lsm11, respectively (Schumperli and Pillai 2004; Dominski and Marzluff 2007; Gruss et al. 2017). U7 snRNP is a multisubunit RNA-guided endonuclease that functions in 3' end processing of replication-dependent histone pre-mRNAs (Sun et al. 2020; Dominski and Tong 2021), generating mature histone mRNAs that end with a conserved stem-loop structure rather than a poly(A) tail typical of the vast majority of eukaryotic mRNAs (Mandel et al. 2008; Liu and Moore 2021). While Lsm10 is relatively small, resembling in size most Sm proteins (Pillai et al. 2001), Lsm11 has an extended amino-terminal region of about 150 amino acids that is essential for the activity of U7 snRNP in 3' end processing (Pillai et al. 2003; Sun et al. 2020). This region of Lsm11 interacts with FLASH (Yang et al. 2009), facilitating the recruitment of CPSF73, the catalytic subunit of the U7 snRNP (Dominski et al. 2005; Yang et al. 2013; Sun et al. 2020). U7 snRNA, the RNA component of U7 snRNP, consists of ~60 nt and contains an unusual Sm-binding site, AAUUUGUCUAG, that differs from that in the spliceosomal snRNAs and promotes the incorporation of Lsm10 and Lsm11 instead of SmD1 and SmD2 into the U7 snRNP Sm ring (Stefanovic et al. 1995; Schumperli and Pillai 2004).

Initial studies on U7 snRNP suggested that its assembly is at least partially similar to the assembly of the spliceosomal snRNPs, involving both the PRMT5 and SMN proteins (Pillai et al. 2003; Schumperli and Pillai 2004; Azzouz et al. 2005; Li et al. 2014). More recent studies supported this conclusion by showing that pICln recruits the Lsm10/11 heterodimer to the PRMT5/MEP50 heterooctamer (Paknia et al. 2016), and that down-regulation of the SMN protein impairs the assembly of U7 snRNP, resulting in a defect in 3' end processing of histone pre-mRNAs (Tisdale et al. 2013, 2022). However, an open question was whether Lsm10 and/or Lsm11 become symmetrically dimethylated by PRMT5, or the interaction of the Lsm10/11 heterodimer with the methylosome plays other than catalytic role(s) in U7 snRNP assembly, contrasting with the assembly of the spliceosomal snRNPs (Azzouz et al. 2005). It was also unclear whether besides SMN, other components of the SMN complex have any role in the assembly of U7 snRNP. In vivo studies failed to detect the binding of Gemin5 to U7 snRNA (Battle et al. 2006b), suggesting that at least this protein is dispensable for U7 snRNP assembly, being substituted by a counterpart that recognizes the

unique Sm-binding site in U7 snRNA and promotes the assembly of the U7-specific ring containing Lsm10 and Lsm11.

Here, we show that the Lsm10/11 heterodimer tightly interacts with both endogenous and recombinant PRMT5 methylosome. Using biochemical and structural approaches, we mapped the binding site of PRMT5 to a short RG cluster within the unique amino-terminal region of Lsm11. We also show that two arginine residues in this cluster, but not in other parts of Lsm11 or Lsm10, are methylated *in vitro*. Most surprisingly, within a complex consisting of Lsm10/11 dimer, SmE/F/G heterotrimer, and pICln (which we will refer to as the U7 6S intermediate), methylation also occurs at a single arginine residue located near the amino terminus of SmE. Methylation of SmE has never been detected in the context of spliceosomal snRNPs. This residue and the following glycine are conserved in all known vertebrate SmE orthologs, suggesting that symmetric dimethylation at this site in conjunction with dimethylation of the amino-terminal Lsm11 may provide an important structural determinant that discriminates between the pathways for the assembly of the U7 snRNP and the spliceosomal snRNPs.

RESULTS

Recombinant Lsm10/11 heterodimer interacts with endogenous PRMT5 methylosome from mammalian extracts

Our recent biochemical and structural studies with recombinant U7 snRNP uncovered key interactions within this multisubunit endonuclease and demonstrated how it cleaves histone pre-mRNAs (Sun et al. 2020; Yang et al. 2020). We next focused on the biogenesis of the U7 snRNP, including the poorly understood process of assembling the U7-specific Sm ring on U7 snRNA. To facilitate this study, we generated rabbit antibodies against the baculovirus-expressed heterodimer of Lsm10-MBP fusion protein and His-tagged Lsm11. The resultant serum from immunized rabbits (α 10/11) precipitated from the cytoplasm of HeLa cells very small amounts of endogenous Lsm10 and Lsm11 that could not be visualized by silver staining (Fig. 1A, lane 1), but were readily detectable by western blotting (data not shown). Interestingly, when the same HeLa cytoplasm was supplemented with Lsm10-MBP/Lsm11 heterodimer, the α 10/11 serum in addition to the two recombinant proteins precipitated three endogenous proteins strongly stained with silver (Fig. 1A, lane 2). They were identified by mass spectrometry and western blotting as subunits of the PRMT5 methylosome: PRMT5 methyltransferase of ~70 kDa (Stopa et al. 2015), MEP50 of ~50 kDa (Friesen et al. 2002), and ~35 kDa pICln (Pu et al. 1999; Meister et al. 2001). PRMT5 and MEP50 form a heterooctamer consisting of four copies of each protein (Anto-

nysamy et al. 2012; Ho et al. 2013; Timm et al. 2018), and pICln binds stoichiometrically to PRMT5 (Friesen et al. 2001b; Pesiridis et al. 2009; Guderian et al. 2011; Mulvaney et al. 2021), explaining the abundance of the three methylosome subunits relative to the small amounts of Lsm10 and Lsm11 used in the assay.

Human Lsm11 with its 360-amino acids is the largest member of the Sm/Lsm family and contains an extended amino-terminal region of 154 residues that plays a critical role in the function of U7 snRNP in 3' end processing (Fig. 1B). Within this region, amino acids ~20–50 bind FLASH (Yang et al. 2009) to form a platform for recruiting the CPSF73 endonuclease and two other components of the cleavage and polyadenylation machinery: CPSF100 and symplekin (Yang et al. 2013; Aik et al. 2017). Between amino acids 68 and 79, Lsm11 contains a GR cluster that may serve as a primary methylation target for PRMT5. Amino acids 155–360 of Lsm11 fold into a canonical Sm domain consisting of an amino-terminal α helix (amino acids 155–165) and five β strands (Sun et al. 2020). The β 3 and β 4 strands are separated by a large loop (amino acids 207–337) that is unique to Lsm11 and contains four separate GR or RG dipeptide motifs (Fig. 1B).

To identify regions of Lsm11 that bind the PRMT5 methylosome, we expressed Lsm10 fused to MBP in complex with Lsm11 lacking the internal loop (amino acids 211–322, Lsm11 Δ L), or Lsm11 lacking both the loop and the entire amino-terminal region encompassing amino acids 1–154 (Lsm11 Δ NL) (Supplemental Fig. S1). We also bacterially expressed four amino-terminal fragments of Lsm11 fused to MBP (Supplemental Fig. S1). These fragments lack the Sm fold region of Lsm11 that interacts with Lsm10 and were expressed alone.

Deletion of the internal loop from Lsm11 had no effect on the interaction of the Lsm10/11 heterodimer with endogenous PRMT5 methylosome. Compared to the wild-type dimer containing full-length Lsm11, approximately the same amounts of PRMT5, MEP50, and pICln were coprecipitated together with Lsm10/11 Δ L dimer by α 10/11 serum (Fig. 1A, lane 3). Thus, the extensive loop despite containing several GR/RG dipeptide motifs does not bind the methylosome. Readily detectable binding of the methylosome to the Lsm10/11 Δ L dimer was also observed in a cytoplasmic extract from mouse myeloma cells (Fig. 1A, lane 4) and was not abolished by additionally deleting the entire amino-terminal region of Lsm11 (Fig. 1A, lane 5). Finally, binding of PRMT5, MEP50, and pICln to Lsm10/11 heterodimers was observed in a mouse myeloma nuclear extract (Fig. 1A, lane 7), indicating that the interaction can occur in both the cytoplasmic and nuclear fractions from different mammalian cell lines.

To further analyze the interaction between the Lsm10/11 dimer and the PRMT5 methylosome, we tested a different approach by directly collecting proteins bound to the dimer on amylose beads via the MBP tag attached to

79) that likely interacts with PRMT5 (Friesen et al. 2001b). The second site likely corresponds to the Sm fold of Lsm10 that tightly interacts with pICln, resembling the interaction of this subunit with SmD1 (Chari et al. 2008; Grimm et al. 2013; Paknia et al. 2016).

Binding of recombinant PRMT5 methylosome to Lsm10/11 heterodimer and methylation of Lsm11

To determine whether the interaction between PRMT5 methylosome and Lsm10/11 can be reconstituted from recombinant components, we co-expressed PRMT5 and MEP50 in the baculovirus system, and purified pICln as a separate protein from bacteria. The PRMT5/MEP50 complex was mixed with the Lsm10/11 heterodimer in molar ratio 1:1 either in the absence or in the presence of pICln, and the amount of the methylosome components immobilized on amylose beads via the MBP tag attached to Lsm10 was assessed by SDS-PAGE and silver staining. As seen in Figure 2A, both PRMT5 and MEP50 were readily detected and their amount increased in the presence of pICln, consistent with its role as the adaptor subunit for recruiting the Lsm10/11 substrate. We conclude that the interaction between Lsm10/11 heterodimer and the methylosome can be recapitulated using recombinant components.

Our studies with cytoplasmic extracts suggested that binding of endogenous methylosome to the Lsm10/11 heterodimer with the deletion of both the internal loop and the amino-terminal region (Lsm10/11 Δ NL), hence lacking the RG cluster between residues 58 and 79, likely depends on pICln, which is predicted to bind the Sm fold of Lsm10 (Paknia et al. 2016). With the availability of recombinant PRMT5/MEP50 heterodimer and pICln as two separate entities, we directly tested this interpretation by either adding or omitting pICln in the binding assay with the Lsm10/11 Δ NL heterodimer. Only background amounts of PRMT5 and MEP50 bound to Lsm10/11 Δ NL heterodimer when pICln was excluded and the interaction was greatly stabilized in its presence (Fig. 2B, lanes 1,2, respectively). Thus, as with endogenous methylosome, recombinant methylosome recognizes two independent

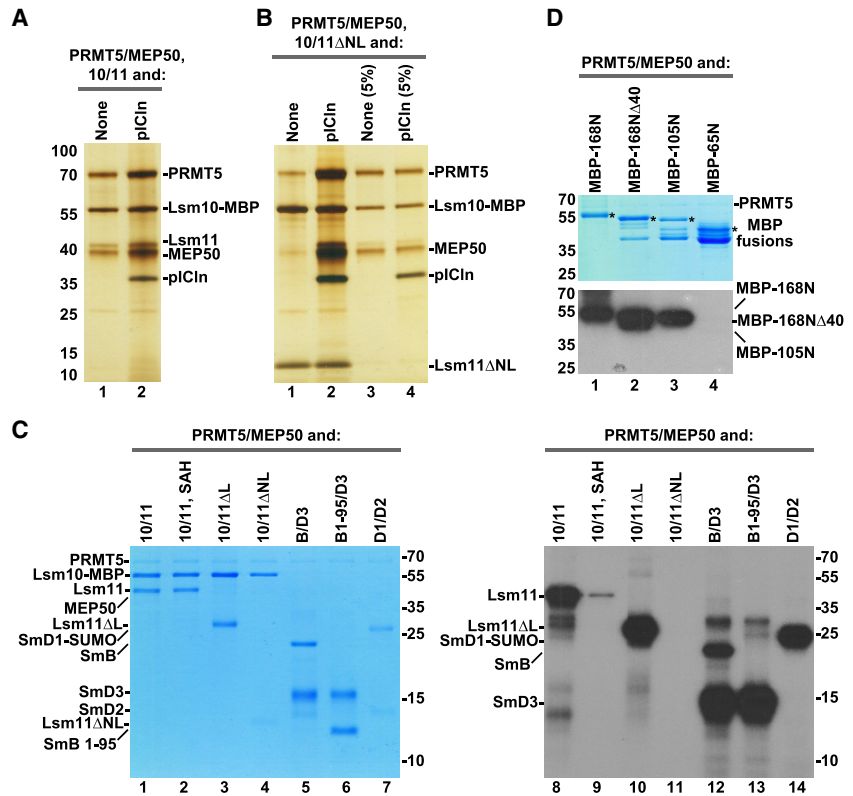


FIGURE 2. Binding and methylation activities of recombinant PRMT5 methylosome. (A) PRMT5/MEP50 complex without (lane 1) or with pICln (lane 2) bound to Lsm10/11 heterodimer was purified on amylose beads via MBP attached to Lsm10 and analyzed by silver staining. (B) Lsm10/11 Δ NL heterodimer was tested for the ability to stably bind PRMT5/MEP50 complex either in the absence (lane 1) or in the presence of pICln (lane 2). Proteins purified on amylose beads via MBP attached to Lsm10 were visualized by silver staining. Lanes 3–4 represent 5% input of proteins used in the binding assay. (C) Various Sm or Lsm heterodimers were incubated overnight in solution with recombinant complex of PRMT5 and MEP50 in the presence of 3 H SAM. Proteins were resolved by SDS-PAGE and visualized by staining with Coomassie Blue (lanes 1–7). Dried gel was used for fluorography to detect proteins labeled with radioactive methyl group (lanes 8–14). (D) Various amino-terminal fragments of Lsm11 fused at the amino terminus with MBP were tested for their ability to undergo methylation, as described in panel C. Coomassie Blue stained gel and fluorogram are shown at the top and bottom, respectively. Asterisks in the top panel indicate full-length amino-terminal Lsm11 fragments fused to MBP. Lower bands correspond to proteolytic fragments.

binding sites in the Lsm10/11 heterodimer, one likely being the RG cluster (amino acids 58–79) that interacts with PRMT5, and the other being the Sm fold of Lsm10 that interacts with pICln. These two contacts collectively contribute to the strong interaction between the methylosome and the Lsm10/11 heterodimer.

The strong association of the PRMT5 methylosome with the Lsm10/11 heterodimer and with the amino-terminal region of Lsm11 containing a GR cluster raised the possibility that either Lsm11 alone or both subunits of the heterodimer are methylated. This would be consistent with the catalytic activity of PRMT5 on the RG-rich carboxy-terminal tails of SmD1, SmB, and SmD3. To test this possibility directly, we carried out an *in vitro* methylation assay using all recombinant components and 3 H-labeled SAM as a

source of radioactively labeled methyl group. Initially, the methylation reaction was carried out only with PRMT5 and MEP50, in the absence of pICln. We tested methylation of the three variants of Lsm10/11 heterodimers, in which Lsm11 was either full length or lacked the amino-terminal region and/or the internal loop, as described above (Supplemental Fig. S1). As a control, we used spliceosomal heterodimers SmD1/D2 and SmB/D3 that are known substrates for *in vitro* symmetric arginine dimethylation by the PRMT5 methylosome (Friesen et al. 2001b, 2002; Meister et al. 2001). As an additional control, we used a variant of SmB/D3 heterodimer (SmB1-95/D3) in which the SmB subunit encompassed amino acids 1–95, lacking the entire unstructured GR-rich carboxy-terminal tail.

Following overnight incubation, the proteins used in the assay were resolved by SDS-PAGE, visualized by staining with Coomassie Blue (Fig. 2C, lanes 1–7) and their methylation status determined by fluorography (Fig. 2C, lanes 8–14). Incubation of PRMT5 and MEP50 with Lsm10/11 and Lsm10/11 Δ L heterodimers resulted in methylation of both full-length Lsm11 and Lsm11 Δ L, but not Lsm10 (Fig. 2C, lanes 8,10). As expected, radioactive labeling was almost completely blocked by excess of the competitive methylation inhibitor SAH (Fig. 2C, lane 9) (Meister et al. 2001). Thus, Lsm11, but not Lsm10, is a substrate for *in vitro* methylation by PRMT5, with methylation occurring somewhere outside the internal loop. Methylation was entirely abolished in Lsm10/11 Δ NL heterodimer (Fig. 2C, lane 11), demonstrating that the methylation site is located within the deleted amino-terminal portion of Lsm11. The spliceosomal Sm proteins, D1, B, and D3 were methylated as expected, with a deletion of the carboxy-terminal tail abolishing methylation of SmB (Fig. 2C, lanes 12–14).

To map the methylation site more precisely, we tested amino-terminal fragments of Lsm11 fused at the amino terminus to MBP (Supplemental Fig. S1). Again, the reaction was carried out in the presence of ^3H SAM and limiting amounts of recombinant PRMT5 and MEP50. Of the four amino-terminal fragments of Lsm11 fused to MBP, 168N, 168N Δ 40, and 105N, became radioactively labeled (Fig. 2D, lanes 1–3), whereas the shortest 65N fragment was not methylated (Fig. 2D, lane 4). In contrast to the three other fragments, MBP-65N does not bind the methylosome and lacks the GR cluster, strengthening the notion that the methylosome binding site and the methylation site are located between amino acids 68 and 79 of Lsm11.

Structure of the methylosome–Lsm10/11 complex

To determine how PRMT5 interacts with Lsm11, we mixed all three methylosome components with the Lsm10/11 heterodimer and purified their complex by gel filtration chromatography for cryo-EM studies, which yielded a structure of the complex at 2.86 Å resolution (Fig. 3A–D;

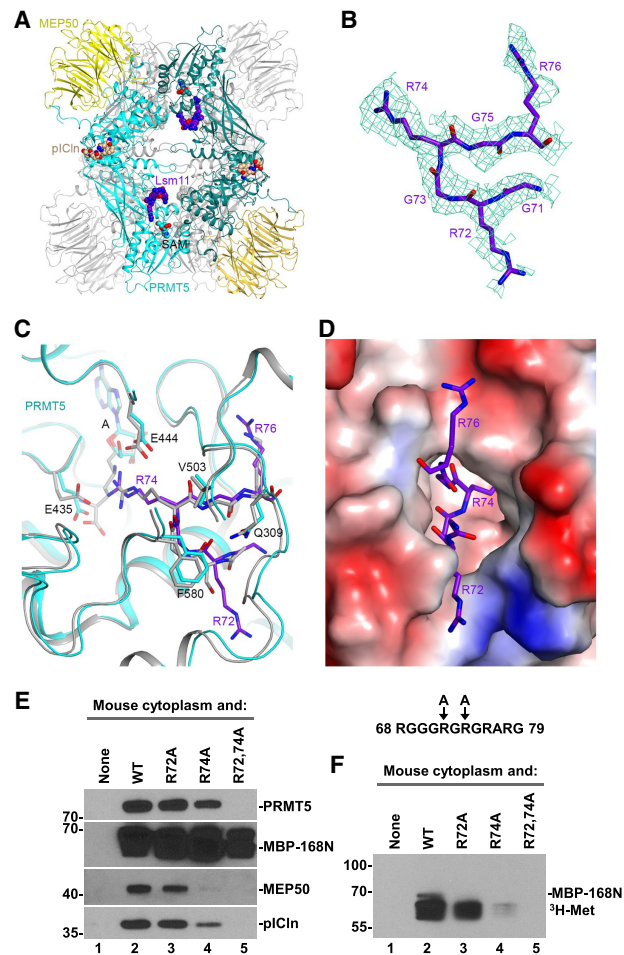


FIGURE 3. Structure of human methylosome–Lsm10/11 complex. (A) Schematic drawing of the human methylosome–Lsm10/11 complex. The molecules in the top layer are shown in color, and those in the bottom layer are in gray. Peptide segments from pICln and Lsm11, as well as the adenosine portion of the SAM cofactor, are shown in the sphere model. (B) Cryo-EM density for the Lsm11 peptide in the active site of PRMT5. The fitted atomic model is shown in the stick model. (C) Detailed interactions between the Lsm11 peptide and the active site of PRMT5. The structure of histone H4 peptide bound to the methylosome is shown in the overlay (in gray). (D) Electrostatic surface of the active site region of PRMT5. The Lsm11 peptide is shown in the stick model. The structure figures were produced with PyMOL (www.pymol.org). (E, F) WT Lsm11 encompassing the first 168 amino acids amino-terminally fused to MBP (MBP-168N) and indicated Lsm11 mutants were incubated with a mouse cytoplasmic extract. Proteins bound to Lsm11 were immobilized on amylose beads via MBP. The beads were split evenly into two halves and analyzed by western blotting (panel E) for the presence of methylosome components or resuspended in a buffer containing ^3H SAM to test methylation of Lsm11, as detected using fluorography (panel F).

Supplemental Table S1). Clear EM density was observed for PRMT5 and MEP50, and for only a few residues of the PRMT5 binding motif (PBM) in pICln (Fig. 3A; Mulvaney et al. 2021). The remaining parts of pICln, most of Lsm11 and the entire Lsm10 were disordered, providing no structural information about the interaction between the Lsm10/

Lsm11 heterodimer and pICln. PRMT5 and MEP50 form a heterooctamer of four tightly bound heterodimers (Fig. 3A), consistent with previous structural studies of the methylosome by X-ray crystallography and cryo-EM (Antonysamy et al. 2012; Ho et al. 2013; Timm et al. 2018).

We observed EM density consistent with a peptide segment in the active site of PRMT5. After examining many possibilities, residues 71-GRGRGR-76 in the amino-terminal extension of Lsm11 were chosen as the most likely interpretation of this density (Fig. 3B). The density exists in the active sites of all four PRMT5 molecules, although the quality of the density varied among them. The side chain of Arg74 is inserted deep into the active site and would be located next to the cofactor SAM (Fig. 3C,D), arguing that it is poised for methylation. Since the density for the side chain of Arg72 is relatively weak, we cannot exclude the possibility that the active site of PRMT5 accommodates the overlapping 69-GGGRGR-74 Lsm11 peptide. In this alternative model, methylation would occur at Arg72, a residue that is well conserved among vertebrate Lsm11 orthologs. Plausibly, the observed EM density is derived from a mixture of the two different peptides bound in the active site of different PRMT5 molecules.

The six-residue peptide of Lsm11 bound to the active site of PRMT5 assumes the conformation of a type II reverse turn, with the Arg residue ready for methylation located at the third position of the turn between two glycines in the primary sequence (Fig. 3C). Consistently, there is no indication of EM density for a C β atom on either side of the methylated Arg (Fig. 1B). In fact, a C β atom on the residue before the methylated Arg would clash with the side chain of Phe580 of PRMT5 (Fig. 3C). For the residue that follows the methylated Arg, a C β atom would be located ~ 3 Å from the side chains of Gln309 and Val503, and a larger side chain here would clash with PRMT5. Therefore, the current structure suggests that a GRG motif may be preferred for the methylation site, and a GRA motif could also be accommodated.

The binding mode of the Lsm11 peptide resembles that of the SGRGK histone H4 peptide (Antonysamy et al. 2012), in which the methylated arginine is also flanked by two glycines. The overall structures of the two complexes are similar, with rms distance of 0.64 Å for 609 equivalent C α atoms of PRMT5 and 0.36 Å for 295 equivalent C α atoms of MEP50. For the cofactor SAM (or SAH), EM density for only adenosine is observed in the four active sites of PRMT5, and the rest of the cofactor (the Met residue) is disordered (Fig. 3C). Possibly related to this disordering, the side chain of Glu435 does not interact with Arg74 of Lsm11 and instead assumes a different conformation that would clash with the main chain of the Met residue of SAM (Fig. 3C).

To determine which of the arginines in the RG cluster of Lsm11 is methylated, we generated three mutants of the amino-terminal Lsm11 region (amino acids 1–168) fused

amino-terminally to MBP (MBP-168N): R72A, R74A, and double mutant R72,74A. In these mutants, either one or two arginines predicted by cryo-EM to undergo methylation were replaced with alanines. Both the wild-type MBP-168N and three mutant proteins were incubated with a cytoplasmic extract from mouse myeloma cells and proteins immobilized on amylose beads via the MBP were visualized by silver staining (not shown) and western blotting (Fig. 3E). As expected, no trace of PRMT5, MEP50, and pICln was recovered from the beads in the absence of MBP-168N, and the wild-type protein readily bound all three methylosome subunits (Fig. 3E, lanes 1,2). The R72A mutant was almost as efficient in binding the methylosome, whereas the R74A mutant bound weakly, as illustrated by significantly reduced amounts of MEP50 and pICln (Fig. 3E, lanes 3,4). Importantly, the R72,74A double mutant failed to interact with the methylosome (Fig. 3E, lane 5).

We also tested the ability of each sample to methylate Lsm11. The beads were resuspended in a buffer containing ^3H SAM and incubated overnight. As expected, the results of this *in vitro* methylation reaction mirrored the results of the binding assay, with the strongest radioactive signal observed for the wild-type MBP-168N. The R72A mutant was methylated with a lower efficiency, whereas the radioactive signal for R74A and R72,74A mutants was very weak and undetectable, respectively (Fig. 3F). Based on the cryo-EM reconstruction, binding assay, and *in vitro* methylation, we conclude that arginine in position 74 of Lsm11 is the primary binding and methylation target for the methylosome. PRMT5 also binds and methylates arginine 72, but with lower efficiency than that observed for arginine 74.

Methylation of SmE

In vivo, SmD1/D2 heterodimer is believed to rapidly associate with pICln, which recruits the heterodimer to the methylosome for methylation of SmD1. The heterodimer subsequently associates with a trimer of SmE, SmF, and SmG, forming a stable 6S intermediate complex consisting of Sm proteins D1/D2/E/F/G and pICln (Chari et al. 2008; Grimm et al. 2013; Paknia et al. 2016). We tested whether methylation of Lsm11 can be affected by pICln. We also tested methylation of Lsm11 in the presence of both pICln and the SmE/F/G heterotrimer that together may form a complex equivalent to the spliceosomal-type 6S complex (Neuenkirchen et al. 2015; Paknia et al. 2016).

When the methylation reaction was carried out in the presence of either pICln, which binds Lsm10/11 (Paknia et al. 2016) or SmE/F/G heterotrimer, which based on the behavior of SmD1/D2 is unlikely to form a stable complex with a heterodimer of Lsm10/11 (Chari et al. 2008; Zhang et al. 2011; Grimm et al. 2013), methylation of Lsm11 was virtually unaffected (Fig. 4A, bottom panel,

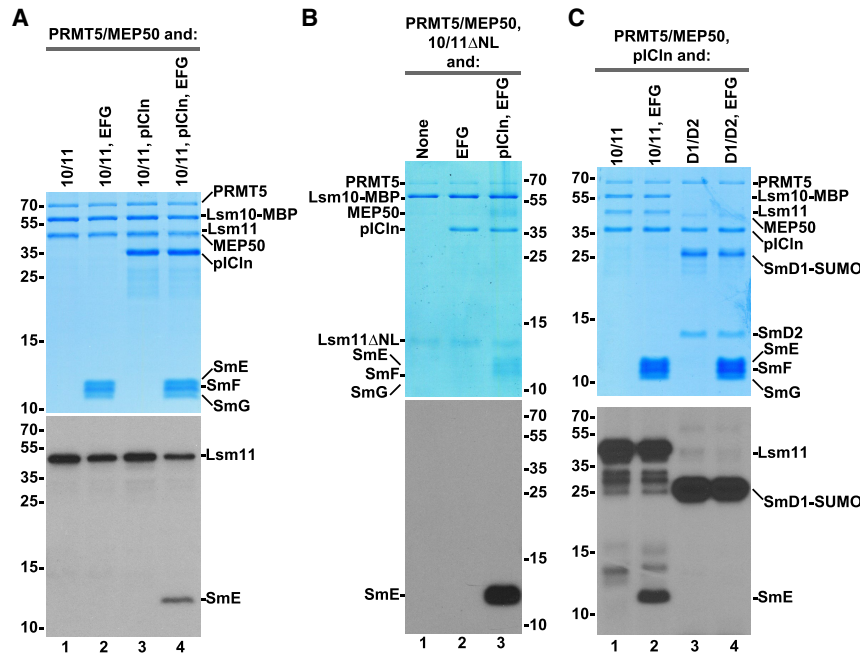


FIGURE 4. Methylation of SmE by recombinant methylosome. (A–C) Recombinant PRMT5/MEP50 complex was incubated with indicated Sm or Lsm heterodimers and ³H SAM either in the absence or presence of pICln and SmE/F/G heterotrimer, as indicated. Protein methylation was detected, as described in the legend for Figure 2. Coomassie Blue stained gels and fluorograms are shown at the top and bottom, respectively.

lanes 1–3). While the lack of any stimulatory effect by the heterotrimer was expected, the failure of pICln to increase the efficiency of Lsm11 methylation by improving the affinity of PRMT5 to Lsm10/11 was surprising. This was likely due to the high concentration of recombinant PRMT5 used in the assay. Strikingly, when both pICln and the SmE/F/G heterotrimer were included into the reaction, methylation of an additional protein migrating between 10 and 15 kDa size markers was detected (Fig. 4A, bottom panel, lane 4). Alignment of the Coomassie Blue stained gel with the fluorogram tentatively identified the methylated protein as SmE, which migrates at the top of the SmE/F/G triplet (Fig. 4A, top panel, lane 4).

The most obvious difference between the SmD1/D2 and Lsm10/11 heterodimers is the presence of the long amino-terminal region in Lsm11, which does not exist in SmD2, the spliceosomal counterpart of Lsm11. To determine whether this Lsm11 region is important for the methylation of SmE, we carried out an *in vitro* methylation assay using Lsm10/11ΔNL heterodimer, in which Lsm11 lacks both the amino-terminal extension and the internal loop within its Sm fold (Supplemental Fig. S1). As expected, due to the deletion of the RG-rich region between amino acids 68 and 79, no methylation of Lsm11 was observed (Fig. 4B, lanes 1–3). Yet, methylation of SmE in the presence of SmE/F/G and pICln was even more pronounced, being detected as a strong signal after a relatively short ex-

posure. Thus, the amino-terminal region of Lsm11 may play an inhibitory role in methylation of SmE, perhaps acting as a competitor for binding PRMT5.

We tested whether methylation of SmE can occur if the Lsm10/11 heterodimer is replaced with the spliceosomal SmD1/D2 heterodimer. Only SmD1 (fused to SUMO) was methylated, as expected, and no radioactive signal was detected for SmE or any other protein of the complex (Fig. 4C, lane 4). In a parallel reaction carried out with Lsm10/11, pICln, and SmE/F/G, SmE was readily methylated (Fig. 4C, lane 2). This result is consistent with previous studies demonstrating that only SmD1 becomes methylated by the PRMT5 methylosome in the 6S complex consisting of SmD1/D2 heterodimer, SmE/F/G heterotrimer, and pICln (Neuenkirchen et al. 2015). Thus, methylation of SmE appears to be an event specifically linked to Lsm10/11 heterodimer and may play a role in the assembly of U7 snRNP *in vivo*.

SmE is methylated by endogenous methylosome

We next tested whether methylation of SmE can be carried out by endogenous methylosome. Heterodimers of Lsm10 with either full length or ΔNL Lsm11 were incubated with a mouse cytoplasmic extract either in the absence or in the presence of the SmE/F/G heterotrimer and bound complexes collected on amylose beads via the MBP tag attached to Lsm10, as described above. Following exhaustive washes, the beads were incubated overnight in a buffer containing radioactive SAM, and the bound material was separated by SDS-PAGE, stained with Coomassie Blue (Fig. 5A, top) and subsequently analyzed by fluorography (Fig. 5A, bottom). Each recombinant Lsm10/11 heterodimer bound a number of proteins, some of which were identified as the components of endogenous methylosome (Fig. 5A, top panel, lanes 1–4, Supplemental Fig. S2), consistent with the data shown in Figure 1C. Importantly, endogenous methylosome in the presence of the SmE/F/G heterotrimer methylated SmE (Fig. 5A, bottom panel, lanes 2,4), with methylation of this subunit typically exceeding the level observed for full-length Lsm11 (Fig. 5A, bottom panel, lane 2, Supplemental Fig. S2, lane 6, and Supplemental Fig. S3, lane 3), supporting the conclusion that SmE is a legitimate PRMT5 methylation target. Lsm11ΔNL that lacks the amino-terminal RG cluster failed to undergo any detectable methylation (Fig. 5A, bottom

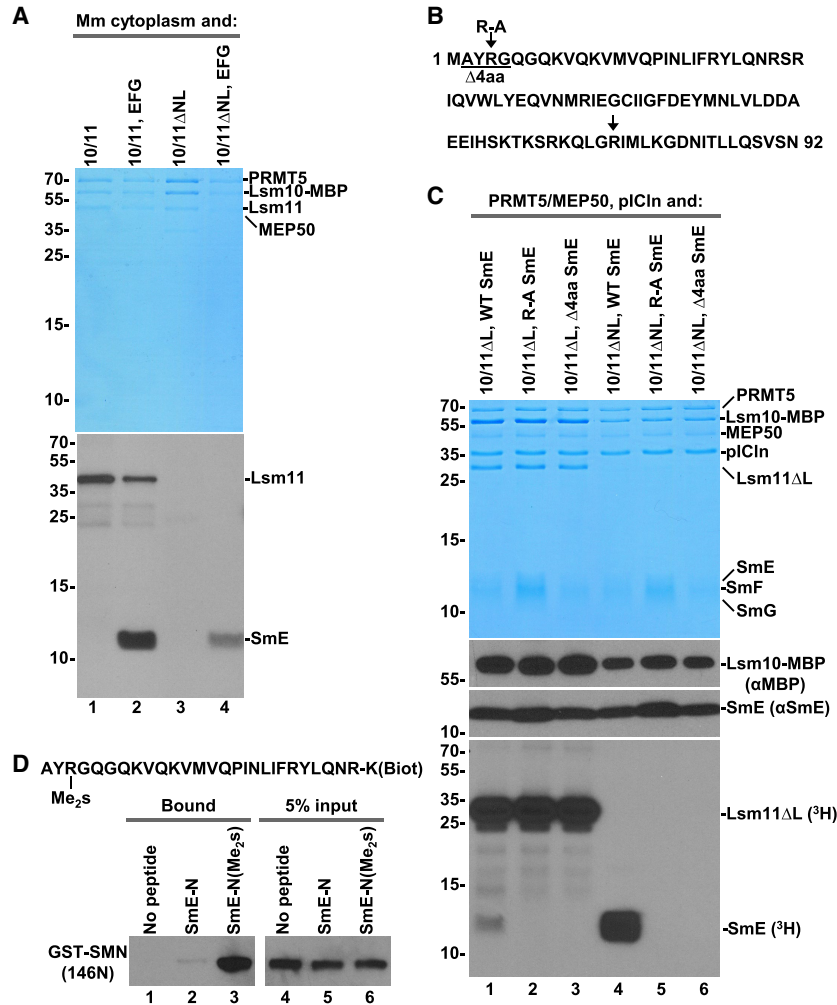


FIGURE 5. In vitro methylation of SmE by endogenous methylosome and identification of the SmE methylation site. (A) Recombinant Lsm10/11 (lanes 1,2) and Lsm10/11 Δ NL (lanes 3,4) heterodimers were incubated either in the absence or in the presence of the SmE/F/G heterotrimer with a mouse cytoplasmic extract and purified together with bound endogenous methylosome on amylose beads via MBP attached to Lsm10. The immobilized proteins were tested directly on the beads in a buffer containing 3 H SAM for the ability to methylate Lsm11 and SmE. Following overnight methylation, proteins in each sample were stained with Coomassie Blue (top panel) and their methylation status analyzed by fluorography (bottom panel). (B) Sequence of human SmE (amino acids 1–92) and mutations made near the amino terminus. The arrows indicate each of the two arginines in SmE neighboring a glycine. (C) Recombinant methylosome complex consisting of PRMT5, MEP50, and pICln was incubated with Lsm10/11 Δ NL heterodimer in the presence of 3 H SAM and SmE/F/G heterotrimer containing WT, R-A or Δ 4aa variants of SmE. Proteins used in the assay were resolved by SDS-PAGE and visualized by staining with Coomassie Blue (top panel) and their methylation status analyzed by fluorography (bottom panel). Lsm10-MBP and SmE were additionally detected by western blotting using α MBP and α SmE antibodies, respectively (middle panels). (D) Two carboxy-terminally biotinylated peptides encompassing the first 27 amino acids of SmE and either lacking or containing symmetric dimethyl group at Arg4 were incubated with a GST-tagged amino-terminal region of SMN (amino acids 1–146) containing the Tudor domain. Proteins immobilized on streptavidin beads were resolved by SDS-PAGE and probed using anti-GST antibody. Lane 1 represents the background observed in the absence of any biotinylated peptide. The input (5%) for all binding reactions is shown in lanes 4–6.

panel, lanes 3,4). Only background proteins and no components of the methylosome were purified from HeLa (Supplemental Fig. S2) and mouse (Supplemental Fig. S3)

cytoplasmic extracts in the absence of recombinant Lsm10/11 heterodimer. As expected, no protein methylation was detected in these samples.

Mutational analysis identifies the site of methylation in SmE

To prove that Lsm10/11 heterodimer promotes methylation of SmE and to identify the site of methylation, we used mutational analysis. The sequence of SmE contains two arginines directly neighboring a glycine residue: an RG dipeptide located near the amino terminus (amino acids 4–5) and a GR dipeptide (amino acids 75–76) located in the fourth β -strand of the Sm fold (Fig. 5B). The amino-terminal arginine is located within the region that is likely unstructured and this residue was considered a primary candidate for methylation in SmE. We substituted this arginine with alanine (R-A) or deleted four amino-terminal amino acids including the RG dipeptide (Δ 4aa) and tested a SmE/F/G heterotrimer containing the resultant SmE mutants for the ability to undergo methylation by recombinant PRMT5 methylosome. Importantly, both mutations tested in the context of heterodimers containing Δ L Lsm11 (Fig. 5C, bottom panel, lanes 1–3) or Δ NL Lsm11 (Fig. 5C, bottom panel, lanes 4–6) abolished methylation of SmE. The effect of these two mutations was particularly striking when tested in the presence of Lsm10/11 Δ NL, which promotes very efficient methylation of wild-type SmE (Fig. 5C, bottom panel, compare lane 4 with lanes 5,6).

Since the three subunits of the SmE/F/G often migrate as a cluster of weakly separated bands in 15% SDS/polyacrylamide gels and are difficult to identify as individual proteins, we used western blotting and anti-SmE antibody to confirm the presence of SmE in the samples that failed to show SmE methylation. We also used anti-MBP antibody to confirm the identity of the Lsm10 subunit visualized by Coomassie Blue staining (Fig. 5C, top panel). Bands corresponding to SmE and Lsm10-MBP were detected in all six

samples (Fig. 5C, middle panels). Based on these experiments, we conclude that the arginine in position four of SmE is the only residue of the SmE/F/G heterotrimer that becomes methylated by PRMT5/MEP50 in the presence of pICln and Lsm10/11 dimer.

The methylation site in SmE in both its location near the amino terminus and the presence of only one flanking glycine sharply contrasts with the methylation sites in Sm protein D1, B, and D3, which are located within carboxy-terminal GR-rich regions. Symmetrically dimethylated sites in these Sm subunits are bound by the SMN Tudor domain, facilitating the assembly of the Sm ring on spliceosomal snRNAs (Friesen et al. 2001a; Selenko et al. 2001; Cote and Richard 2005; Tripsianes et al. 2011). We tested whether the unusual methylation site in SmE is recognized by the Tudor domain of SMN. We ordered the chemical synthesis of two peptides containing biotin at the carboxyl terminus and encompassing the first 27 amino acids of SmE (Fig. 5D). Of the two peptides, one was unmodified and the other contained symmetric dimethyl group at Arg4. A binding assay was carried out in the presence of each peptide and GST-tagged SMN fragment (amino acids 1–146) containing the Tudor domain. In the absence of any peptide, no GST-SMN protein accumulated on streptavidin beads (Fig. 5D, lane 1). Importantly, the SmE-N(Me₂s) peptide, but not the unmodified peptide, readily bound the SMN fragment (Fig. 5D, lanes 2,3), indicating that symmetrically dimethylated Arg4 of SmE is recognized by the SMN Tudor domain.

Structural studies on the methylosome–U7 6S complex

Methylation of SmE by both recombinant and endogenous methylosome prompted us to assemble an entirely recombinant complex consisting of U7-specific 6S (U7 6S) intermediate and the methylosome, and to test whether it is amenable for structural studies. The U7 6S intermediate was generated by mixing Lsm10/11 heterodimer, SmE/F/G heterotrimer, and pICln, and shown to survive gel filtration chromatography (Fig. 6A), confirming that the hexameric complex is stable. The U7 6S intermediate together with recombinant PRMT5/MEP50 formed a larger complex that eluted as a single peak on gel filtration chromatography (Fig. 6A). Interestingly, gel filtration studies suggest that the methylosome heterooctamer binds only one copy of the U7 6S complex (4:1 molar ratio of PRMT5/MEP50 to U7 6S). The same ratio was observed when a large molar excess of the U7 6S complex was used, suggesting that the methylosome contains only one active binding site for the U7 6S. When applied to the cryo-EM grids, the complex became disrupted, with only the methylosome remaining stable. To stabilize the complex, we used cross-linking with 4 mM BS3 but failed to observe any density for SAM in the EM map. We repeat-

ed the cross-linking with 8 mM BS3 in the presence of the SAM analog inhibitor sinefungin (SFG). This approach allowed us to determine a structure of the methylosome at 2.69 Å resolution (Fig. 6B; Supplemental Table S1), yielding well-defined EM density for the adenosine portion of SFG. The rest of SFG had weaker density and assumed different conformations in the four PRMT5 active sites.

Weak density was observed associated with one of the PRMT5 molecules in the EM reconstruction (Fig. 6C). The density could not be conclusively interpreted as the U7 6S, although the stoichiometry was consistent with gel filtration studies. The binding of U7 6S to the methylosome is likely highly dynamic. The density binds one PRMT5 molecule and has no contact with MEP50. It is not clear based on the map why there is only one U7 6S binding site in the PRMT5 tetramer.

The structural analysis indicates that the four PRMT5 molecules in the methylosome assume two distinct conformations, with rms distance of 0.4 Å within each pair of the molecules with similar conformation. In contrast, the rms distance between two PRMT5 molecules with different conformation is 0.9 Å, and the differences are mostly located at the interface between the catalytic and TIM barrel domains of PRMT5 (Fig. 6D). This interface is also where the density for U7 6S is located, and residues 292–306 that link the two domains are ordered in the region where U7 6S contacts PRMT5 but disordered in the other PRMT5 molecule. These residues are also disordered in the structure of the methylosome–Lsm10/11 complex, although they are ordered in the complex with the histone H4 peptide (Antonyamy et al. 2012).

DISCUSSION

Spliceosomal snRNPs of the Sm-type (U1, U2, U4, U5, U11, U12, and U4_{atac}) share the same ring-shaped core domain that surrounds the Sm-binding site in their cognate snRNA component (Will and Luhrmann 2001). The ring consists of seven Sm proteins, D1, D2, D3, B, E, F, and G, which are loaded around the Sm-binding site as three preformed subcomplexes, SmD1/D2, SmB/D3, and SmE/F/G (Luhrmann et al. 1990; Khusial et al. 2005). This assembly is controlled by two large macromolecular entities: the PRMT5 methylosome consisting of three subunits (PRMT5, MEP50, pICln), and the SMN complex consisting of nine known subunits (SMN, Gemin2-8, and Unrip) (Battle et al. 2006a; Neuenkirchen et al. 2008; Li et al. 2014; Gruss et al. 2017). Initially, pICln acts outside the methylosome complex by preorganizing the Sm subunits into intermediate subcomplexes (Chari et al. 2008; Paknia et al. 2016). In the next step, pICln in conjunction with two remaining methylosome subunits mediate symmetrical dimethylation of arginines in RG-rich carboxy-terminal tails of the Sm proteins D1, B, and D3. These modifications serve to recruit the SMN protein, which together with other subunits of

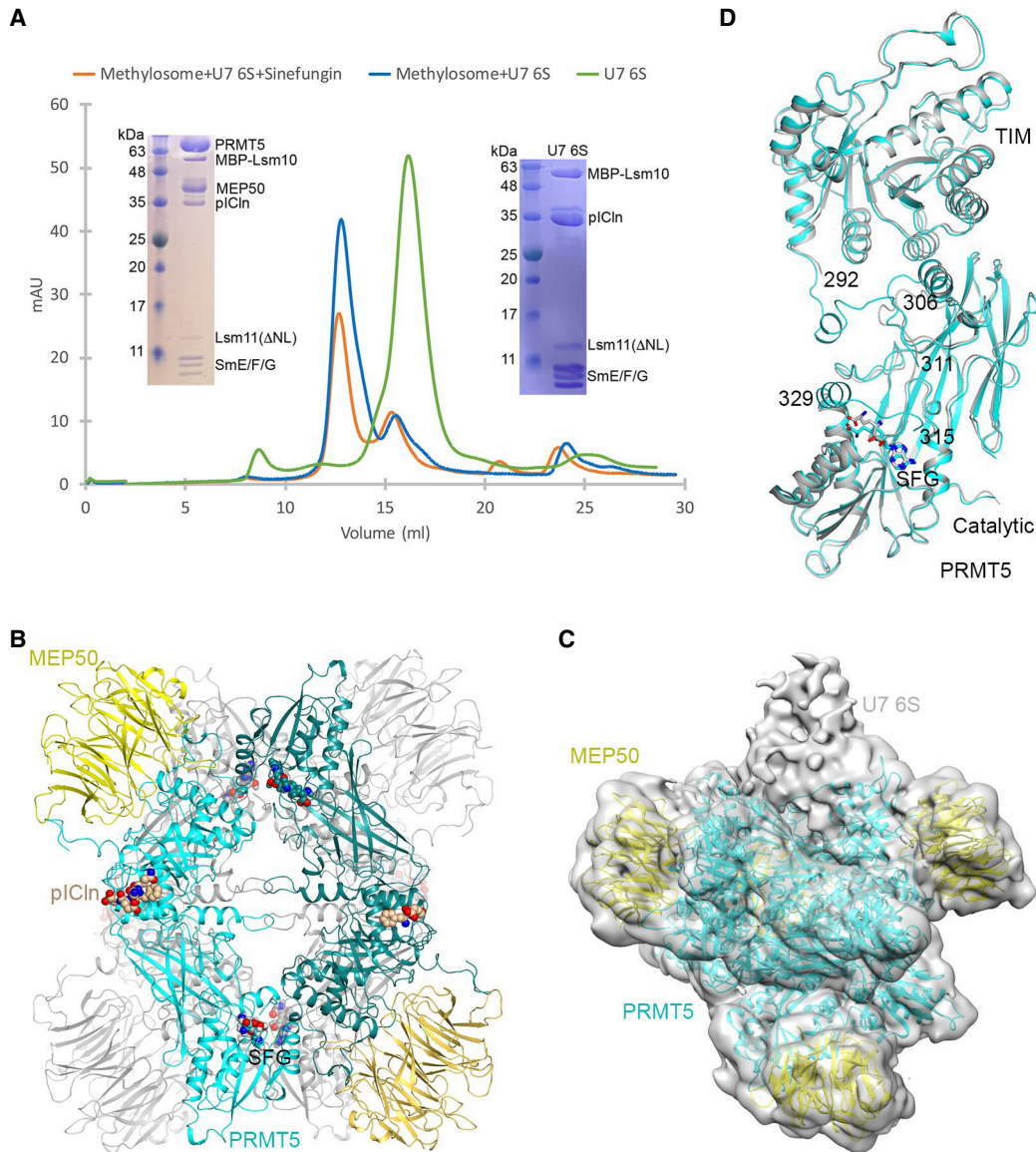


FIGURE 6. Structural studies of the human methylosome–U7 6S complex. (A) Gel filtration profiles of Lsm10/11 bound to SmE/F/G and pICln (U7 6S complex, green), methylosome bound to U7 6S (blue), and methylosome bound to U7 6S in the presence of sinefungin (SFG, orange). Inserts are SDS gels of the purified U7 6S sample (right) and methylosome–U7 6S sample (left). Lsm10 is amino-terminally fused with MBP, and the amino-terminal region and internal loop of Lsm11 are deleted (Lsm11 Δ NL). (B) Schematic drawing of the human methylosome–U7 6S complex. Only the PBM of pICln is included in the atomic model, and the rest of the U7 6S is flexible. (C) Cryo-EM density for the methylosome–U7 6S complex, low pass filtered to 7 Å resolution. The atomic model for the methylosome is shown. The extra EM density that could correspond to U7 6S is labeled. (D) Overlay of two PRMT5 molecules with different conformations in the methylosome–U7 6S complex. The PRMT5 molecule in contact with the putative U7 6S EM density is shown in cyan and the other molecule in gray. The disordered segments, 292–306 and 311–315 or 311–329, are labeled. Panel D was produced with Chimera (Goddard et al. 2007).

the SMN complex, controls subsequent steps in the assembly process, including the recognition of a spliceosomal snRNA and its orderly binding to Sm proteins.

U7 snRNP contains a unique Sm ring in which Sm proteins D1 and D2 are replaced by Lsm10 and Lsm11 (Schumperli and Pillai 2004; Sun et al. 2020). These two specific proteins and U7 snRNA are key determinants of the unusual function of U7 snRNP, which plays no recognized role in splicing and instead acts as an RNA-guided

endonuclease that cleaves replication-dependent histone pre-mRNAs during their 3' end processing (Dominski and Tong 2021). The remaining five subunits of the U7 ring are shared with the spliceosomal snRNPs. The Sm-binding site in U7 snRNA partially departs from the consensus established for the spliceosomal snRNAs, promoting the incorporation of Lsm10 and Lsm11 into the ring (Pillai et al. 2003). U7 snRNP is relatively rare in animal cells, being ~500-fold less abundant than the spliceosomal snRNPs

(Smith et al. 1991). Clearly, the assembly of U7 snRNP must be under strict cellular surveillance to prevent binding of the highly abundant SmD1/D2 heterodimer to U7 snRNA, yielding dysfunctional and potentially harmful complexes. Previous *in vitro* and *in vivo* studies established that the assembly of the U7-specific Sm ring on U7 snRNA utilizes at least some components of the PRMT5 and SMN complexes, but how exactly they work together to achieve this goal and to distinguish U7 snRNP from the spliceosomal snRNPs is unknown (Schumperli and Pillai 2004; Tisdale et al. 2013, 2022). Here, we took a closer look at the role of the PRMT5 methylosome in this process.

Interaction of the Lsm10/11 heterodimer with the PRMT5 methylosome

In our studies, we took advantage of the Lsm10/11 heterodimer that was expressed using the baculovirus system and previously used in generating catalytically active semi- and fully recombinant U7 snRNP for structural and functional studies (Bucholc et al. 2020; Sun et al. 2020; Yang et al. 2020). When incubated in cytoplasmic or nuclear extracts from mammalian cells, the heterodimer tightly interacts with an endogenous methylosome complex consisting of all three known subunits: PRMT5, MEP50, and pICln. The Lsm10/11 heterodimer also interacts with recombinant methylosome reconstituted from baculovirus-expressed PRMT5/MEP50 heterodimer and bacterially expressed pICln, facilitating structural studies by cryo-EM. In the structure, PRMT5 and MEP50 form a heterooctamer of four tightly bound heterodimers, consistent with previous structural studies of the methylosome by X-ray crystallography and cryo-EM (Antonysamy et al. 2012; Ho et al. 2013; Timm et al. 2018). In the active site of PRMT5, we identified density likely corresponding to residues 71-GRGRGR-76 of Lsm11, with the Arg74 and/or Arg72 side chain located next to the SAM methyl group donor and poised for methylation. This peptide is part of a larger RG cluster located between amino acids 68 and 79 in Lsm11, RGGGRGRGRARG, that resembles the carboxy-terminal RG clusters found in Sm proteins D1, B, and D3. Other parts of the complex were largely disordered, providing no structural information.

By using various mutants of Lsm11 in pull-down assays, we mapped the binding site for the methylosome to Arg74 (primary site) and Arg72 (secondary site), consistent with the structural studies. The methylosome strongly interacts with the Lsm10/11 heterodimer lacking the entire amino-terminal region of Lsm11, including the RGGGRGRGRARG cluster. In this case, the interaction entirely depends on the presence of pICln, suggesting the existence of a pICln-mediated contact between the methylosome and the Lsm10/11 heterodimer. Our cryo-EM structure suggests that pICln uses its PBM to stabilize the interaction between Lsm10/11 and the methylosome, consistent with a previous study

(Mulvaney et al. 2021). What remains undetectable in the structure is how pICln interacts with the Lsm10/11 heterodimer. Important clues on this interaction come from previous structural studies on the spliceosomal 6S complex consisting of pICln and the Sm proteins D1, D2, E, F, and G (Friesen et al. 2001b; Chari et al. 2008; Grimm et al. 2013; Pelz et al. 2015). In this complex, pICln uses the β 4 strand of its pleckstrin homology domain to directly contact the β 5 strand of SmD1 in an antiparallel orientation, thus mimicking the interaction mode that oligomerizes Sm proteins (Kambach et al. 1999). The most likely possibility consistent with our binding results is that the same interaction occurs in the U7 6S complex, bringing together Lsm10 (the SmD1 counterpart) and pICln (Paknia et al. 2016).

Lsm11 is methylated by PRMT5

The presence of the GR cluster of Lsm11 in the active site of PRMT5 strongly suggests that one or more residues of this protein are methylated. *In vitro* methylation assay with ^3H -labeled SAM and either endogenous or recombinant PRMT5 methylosome confirmed this prediction, demonstrating that Lsm11, but not Lsm10, is a substrate for methylation by the methylosome. Methylation of Lsm11 does not require the pICln subunit and primarily occurs at Arg74, and with lower efficiency at Arg72. Both residues are flanked by glycines, thus exist in an optimal context known to support highly efficient methylation by PRMT5 (Musiani et al. 2019). Simultaneous substitutions of Arg72 and Arg74 with alanines abolished stable interaction of the amino-terminal region of Lsm11 with the methylosome, hence preventing Lsm11 methylation. We anticipate that in the presence of Lsm10 that recruits the methylosome via pICln, the double mutant of Lsm11 may undergo methylation at other arginines within the RGGGRGRGRARG peptide and potentially elsewhere in the amino-terminal region of Lsm11 (amino acids 1–154). However, the fact that no detectable *in vitro* methylation occurs in the absence of this region (e.g., in Lsm10/11 Δ NL) suggests that other parts of Lsm11 (and Lsm10) lack suitable methylation sites.

Based on both biochemical and structural studies, we conclude that Lsm11 becomes methylated *in vitro* by PRMT5 within its amino-terminal region, contrasting with the Sm proteins D1, B, and D3, which are methylated carboxy-terminally. Our results differ from a previous report that failed to identify methylation of Lsm11, likely due to using relatively small amounts of this protein generated *in vitro* (Azzouz et al. 2005).

Methylation of SmE

Methylation of Lsm11 in a U7-specific 6S complex consisting of Lsm10, Lsm11, SmE/F/G heterotrimer, and pICln was largely unaffected, but surprisingly an additional protein

was methylated in the complex, an event not observed in the absence of pICln and/or SmE/F/G. This protein was identified as SmE, with the methylation site mapped by mutagenesis to the amino-terminal arginine that is followed by a glycine (amino acids 4–5). This RG dipeptide is located within a 15-amino acid unstructured amino-terminal extension of SmE that is highly conserved among distant vertebrates, indicative of an important biological function.

Methylation of SmE was readily catalyzed *in vitro* by both recombinant and endogenous methylosome purified from HeLa and mouse cytoplasmic extracts, making it unlikely to be an artifact of PRMT5 overexpression or misfolding. In addition, the same *in vitro* methylation reaction carried out with the spliceosome-specific 6S complex, consisting of SmD1/D2 heterodimer, SmE/F/G heterotrimer, and pICln (Grimm et al. 2013), resulted in methylation of only SmD1, consistent with previously reported data (Neuenkirchen et al. 2015). Thus, methylation of SmE specifically depends on the presence of Lsm10 and/or Lsm11 in the complex. Surprisingly, the amino-terminal region of Lsm11, the most characteristic feature distinguishing Lsm11 from all other Sm/Lsm subunits and the functional hub of the U7 snRNP, is dispensable for SmE methylation. Presumably, the U7-specific 6S intermediate containing Lsm10 and Lsm11 adopts a different conformation than the spliceosomal 6S complex with SmD1 and SmD2, exposing the amino-terminal region of SmE for the enzymatic activity of PRMT5.

To gain additional information on how PRMT5 methyltransferase accesses the amino-terminal arginine in SmE for methylation in the U7-specific 6S complex, we carried out a set of studies using cryo-EM. These studies were largely unsuccessful due to the instability and/or flexibility of the complex.

Potential roles of Lsm11 and SmE methylation in U7 snRNP biogenesis

The most important question for future studies is whether methylation of Lsm11 and SmE occurs *in vivo*. Our initial attempts with endogenous U7 snRNP partially purified from mouse and human nuclear extracts failed to detect the presence of sDMA modification on either of the two proteins by western blotting using sDMA-specific antibodies Y12 and SYM11. The same antibodies detected Sm protein B and D3 in the purified preparation of U7 snRNP (data not shown). However, compared to the multiple RG repeats in SmB and SmD3, the RG-rich track in Lsm11 is relatively short, and the methylation site in SmE contains only a single arginine, likely forming epitopes too weak to be recognized by either antibody. Due to the very limiting nature of endogenous U7 snRNP, our efforts to use mass spectrometry to determine the methylation pattern generated on Lsm11 and SmE *in vivo* were so far unsuccessful.

In conclusion, our *in vitro* results suggest that the assembly of the U7 snRNP involves specific modification of its two subunits by the PRMT5 methylosome: Lsm11 and SmE. Both subunits are methylated within short, amino-terminal motifs, contrasting with the much longer and carboxy-terminally located methylation targets in the Sm proteins D1, B, and D3. Symmetrically dimethylated arginines interact with the Tudor domain of the SMN protein (Friesen et al. 2001a; Selenko et al. 2001; Cote and Richard 2005; Tripsianes et al. 2011), facilitating the transfer of the Sm proteins to the SMN complex and efficient and faithful assembly of the Sm ring on the spliceosomal snRNAs (Gruss et al. 2017). It can be assumed that methylation of Sm proteins in the U7-specific ring serves the same purpose, with the unique pattern of methylated arginines on Lsm11 and SmE likely playing a key role in distinguishing the U7 snRNP assembly pathway from the much more general and ubiquitous assembly pathway for the spliceosomal snRNPs.

MATERIALS AND METHODS

Antibodies

Rabbit serum containing antibodies against a dimer of Lsm10-MBP and full-length Lsm11 was generated by Pacific Immunology. The following commercial antibodies were used in this study: SmB 12F5 (Sigma), SmD1 RB17341 (MyBioSource), SmD3 A303-954A (Bethyl), SmE ab229557 (Abcam), Y12 MA-1-90490 (Thermo), pICln A304-521A (Bethyl), MEP50/Mep50 A301-562A (Bethyl), PRMT5 A300-850A (Bethyl), MBP 66003-1-Ig (Proteintech), and Symmetric Dimethyl-Arginine SYM11 (Millipore).

Cell culture and preparation of cytoplasmic and nuclear extracts

HeLa and mouse myeloma cells were grown by Cell Culture Company. HeLa cells were shipped on dry ice as a frozen cell pellet and used after thawing. Myeloma cells were shipped as a concentrated cell suspension on wet ice and used immediately upon arrival. Nuclear extracts from HeLa and mouse myeloma cells were prepared as described (Dominski et al. 1995; Skrajna et al. 2018; Sun et al. 2021). The crude cytoplasmic fractions generated after collecting HeLa or mouse myeloma nuclei were used to prepare cytoplasmic extracts, as described (Mayeda and Krainer 1999), with the ultracentrifugation and dialysis steps being omitted.

Purification of pICln

Full-length mouse pICln was cloned into the pET28a (Novagen) vector with a 6xHis tag added to the amino terminus and expressed in *Escherichia coli* Rosetta strain (Novagen). The cell pellet was resuspended and lysed by sonication on ice in 100 mL of lysis buffer containing 25 mM Tris (pH 7.5), 300 mM NaCl, and 5% (v/v) glycerol. The supernatant was incubated with nickel beads

for 1 h at 4°C. The beads were washed two times with 50 bed volumes of wash buffer containing 25 mM Tris (pH 7.5), 300 mM NaCl, and 20 mM imidazole. The protein was eluted with 5 mL of elution buffer containing 25 mM Tris (pH 7.5), 300 mM NaCl, and 250 mM imidazole. Eluate was diluted 3× with 25 mM Tris (pH 7.5) and was purified by chromatography using a HiTrap MonoQ column (Cytiva) with a salt gradient. Buffer A contained 25 mM Tris (pH 7.5) and 100 mM NaCl, and Buffer B contained 25 mM Tris (pH 7.5) and 1 M NaCl. Fractions of interest were concentrated to ~1.1–1.7 mg/mL with 5% (v/v) glycerol, flash frozen using liquid nitrogen, and stored at –80°C.

Purification of the PRMT5/MEP50 complex

cDNAs encoding full-length human PRMT5 and MEP50 were cloned into pFL vector and coexpressed in insect cells. One liter of High5 cells (1.8×10^6 cells/mL) was infected with 15 mL of PRMT5/MEP50 P2 virus with the His tag on PRMT5. For purification, the cell pellet was resuspended and lysed by sonication on ice in 100 mL of lysis buffer containing 25 mM Tris (pH 7.5), 300 mM NaCl, 5% (v/v) glycerol, and one protease inhibitor cocktail tablet (Sigma). The cell lysate was centrifuged at 13,000 rpm for 45 min at 4°C. The supernatant was incubated with nickel beads for 1 h at 4°C. The beads were washed two times with 50 bed volumes of wash buffer and eluted with 5 mL of elution buffer, as described for pICln. The eluate was purified by gel filtration using a Superdex 200 Increase 10/300 GL column (Cytiva) in buffer containing 25 mM Tris (pH 7.5), 150 mM NaCl, 2 mM DTT, and 10% (v/v) glycerol. Fractions of interest were concentrated to ~1.1 mg/mL with 5% (v/v) glycerol, flash frozen using liquid nitrogen, and stored at –80°C.

Purification of the Lsm10/11 heterodimer

The expression and purification of human Lsm11 lacking residues 211–322, and full-length Lsm10, were carried out as described previously (Bucholc et al. 2020; Sun et al. 2020, 2021). Briefly, Lsm11 and Lsm10 were cloned into a pFL vector with MBP added to the amino terminus of Lsm10. One liter of High5 cells (1.8×10^6 cells/mL) was infected with 15 mL of Lsm10/11 P2 virus. The cell pellet was resuspended and lysed by sonication on ice in 100 mL buffer containing 25 mM Tris (pH 7.5), 500 mM NaCl, 5% (v/v) glycerol, and one protease inhibitor cocktail tablet (Sigma). The cell lysate was centrifuged at 13,000 rpm for 45 min at 4°C. The supernatant was incubated with nickel beads for 1 h at 4°C. The beads were washed two times with 50 bed volumes of wash buffer (25 mM Tris [pH 7.5], 500 mM NaCl, and 40 mM imidazole) and eluted with 5 mL of elution buffer containing 25 mM Tris (pH 7.5), 500 mM NaCl, 500 mM imidazole, and 5% (v/v) glycerol. The eluate was diluted 2.5× with 25 mM HEPES (pH 7.5) and 5 mM DTT. The complex was then purified by chromatography using a HiTrap Heparin column (Cytiva) with a salt gradient. Buffer A contained 20 mM Hepes (pH 7.5) and 5 mM DTT, and Buffer B contained 20 mM Hepes (pH 7.5), 5 mM DTT, and 1 M NaCl. Fractions of interest were concentrated to ~1 mg/mL with 5% (v/v) glycerol, flash frozen using liquid nitrogen, and stored at –80°C.

Purification of endogenous methylosome bound to Lsm10-MBP/Lsm11 heterodimer or Lsm11 alone

MBP-Lsm10/Lsm11 heterodimer or various amino-terminal fragments of Lsm11 fused to MBP (depending on the experiment, 25–100 pmol each) were incubated for 1 h on ice with 750 µL of cytoplasmic or nuclear extracts from mouse myeloma or HeLa cells. The samples were supplemented with 5 µL of a rabbit serum generated against MBP-Lsm10/Lsm11 heterodimer, rotated 60 min in a cold room and spun down in a microcentrifuge (10 min at 10,000g) to remove potential precipitates. The supernatants were loaded with over 30 µL of protein A plus beads (Pierce), rotated in a cold room for 75 min and gently spun to collect beads and the bound complexes. The beads were washed for 60 min with buffers matching those present in the cytoplasmic or nuclear extracts, moved to new tubes for an additional 30 min rotation, and resuspended in the SDS sample buffer. A fraction of each sample was separated in a 4%–12% SDS/polyacrylamide gel, and immunoprecipitated proteins were visualized by silver staining and identified by mass spectrometry in the Laboratory of Mass Spectrometry at the Institute of Biochemistry and Biophysics, as previously described (Yang et al. 2013; Skrajna et al. 2018). Protein identities were confirmed by western blotting using specific antibodies. In some experiments, instead of using anti-Lsm10/11 serum, proteins bound to the Lsm10/Lsm11 heterodimer were directly collected on amylose beads. In this approach, extracts mixed with MBP-Lsm10/Lsm11 heterodimer were rotated 60 min in a cold room, spun down in a microcentrifuge, as above, and rotated with 30 µL of amylose beads. The remaining steps were the same as during the immunoprecipitation protocol.

In vitro methylation by recombinant methylosome

Recombinant methylosome consisting of baculovirus-expressed PRMT5/MEP50 and bacterially expressed pICln (1 pmol each) was mixed in 20 µL of the cytoplasmic buffer with a 2- to 10-fold excess of recombinant Sm proteins, and each sample was supplemented with 2 µCi of ³H-labeled SAM (PerkinElmer) for overnight methylation at 32°C. Each sample was mixed with an equal volume of 2× SDS sample buffer, resolved in 15% or 4%–12% SDS/polyacrylamide gels. Separated proteins were visualized by Coomassie Blue staining and gel images captured. The stained gels were incubated for 30 min with Amplify solution (GE Healthcare), dried and used for 12- to 72-h fluorography at –80°C. Methylated proteins were identified by aligning fluorograms with stained gel images.

In vitro methylation by endogenous methylosome

MBP-Lsm10/Lsm11 heterodimer (100 pmol) either alone or together with 300 pmol of SmE/F/G heterotrimer was rotated for 60 min in a cold room with HeLa or mouse myeloma cytoplasmic extracts (750 µL), and spun down at 10,000g for 10 min. The heterodimer bound to endogenous methylosome was purified on amylose beads, as described above. After extensive washing, the beads were suspended in 25 µL of cytoplasmic buffer containing 2 µCi of ³H-labeled SAM (PerkinElmer) and incubated overnight at 32°C without shaking. Following incubation, the

supernatant was removed, and the beads were resuspended in 25 μ L of SDS sample buffer. Proteins collected on the beads were resolved in 15% polyacrylamide or 4%–12% SDS/polyacrylamide gels and stained with Coomassie Blue. Protein composition of bound complexes was additionally determined by silver staining and/or western blotting. Arginine methylation was detected by fluorography, as described above. In some experiments, this approach was also used to immobilize recombinant methylosome to conduct the methylation assay on amylose beads rather than in solution.

EM studies of the methylosome bound to the Lsm10/11 heterodimer

Purified PRMT5/MEP50, Lsm10/11, and pICln were mixed on ice (1:1:1.2 molar ratio) and incubated for 1 h. The resultant complex was purified by size exclusion chromatography using a Superose 6 column (Cytiva) in a buffer containing 25 mM Tris (pH 7.5), 150 mM NaCl, and 2 mM DTT. Fractions of interest were concentrated to \sim 0.5 mg/mL, flash frozen with liquid nitrogen, and stored at -80°C . Cryo-EM grids were prepared by applying 3.5 μ L of the protein sample at a concentration of 0.25 mg/mL to one side of a Quantifoil 400 mesh 1.2/1.3 gold grid with graphene oxide support film (Quantifoil). After 30 sec, the grid was blotted for 1.5 sec on the other side under 99% humidity and at 20°C using an EM GP2 Plunge Freezer (Leica) and immediately plunged into liquid ethane. A total of 4061 image stacks were collected on a Titan Krios electron microscope at the Columbia University Cryo-Electron Microscopy Center, equipped with a K3 direct electron detector (Gatan) at 300 kV with a total dose of 58.2 e^{-2} subdivided into 50 frames in 2.5 sec exposure using Leginon. The images were recorded at a nominal magnification of 105,000 \times and a calibrated pixel size of 0.83 \AA , with a defocus range from -0.8 to $-2.2 \mu\text{m}$. Image stacks were motion-corrected and dose-weighted using RELION 3.1 (Zivanov et al. 2018). The patch CTF parameters were determined with cryoSPARC (Punjani et al. 2017). First, 337,438 particles were auto-picked from 500 images and were used to generate six 3D initial models by ab initio reconstruction. The model with recognizable features by visual inspection was chosen for creating templates for template picking. A total of 1,814,708 particles were picked from 4601 micrographs by template picking. After two rounds of heterogeneous refinement, 754,047 particles were imported to RELION for CTF refinement and Bayesian polishing. The polished particles were then imported back to cryoSPARC for a homogeneous refinement, yielding a map at 2.86 \AA resolution. C1 symmetry was used throughout the reconstruction. The crystal structure of PRMT5/MEP50 bound to the PBM peptide of pICln (PDB code 6V00) was fitted as a rigid body into the EM map using Chimera (Pettersen et al. 2004). The atomic model of the Lsm11 peptide (residues 71–76) was built manually into the cryo-EM density with Coot (Emsley and Cowtan 2004). EM density was observed for the adenosine portion of the cofactor SAM, which was likely incorporated during expression, as the cofactor was not added during purification and subsequent studies. The atomic model was improved by real-space refinement with the program PHENIX (Liebschner et al. 2019). The cryo-EM information is summarized in Supplemental Table S1. The data were deposited under PDB entry code 8G1U.

Purification and EM studies of the methylosome/U7 6S complex

Human SmE/F/G heterotrimer and human MBP-Lsm10/11 Δ NL heterodimer were expressed in *E. coli* and insect cells, respectively, and purified as described earlier (Bucholc et al. 2020; Sun et al. 2020, 2021). Lsm10 carried an amino-terminal MBP and a 6xHis tag. In Δ NL Lsm11, the amino-terminal segment (residues 1–152) and the internal loop (residues 211–322) were removed. Purified SmE/F/G heterotrimer, MBP-Lsm10/ Δ NL heterodimer, and mouse pICln were mixed (molar ratio 1.2:1:1.2) in a buffer containing 20 mM Tris (pH 8.1), 100 mM NaCl, and 2 mM DTT and incubated on ice for 1 h. The assembled U7 6S complex was purified with a Superose 6 (Cytiva) gel filtration column using the same buffer. Purified PRMT5/MEP50 and U7 6S were combined (molar ratio 4:8 or 4:2) in a buffer containing 20 mM Hepes (pH 7.7), 100 mM NaCl, and 2 mM DTT and incubated on ice for 1 h. The assembled methylosome–U7 6S complex was purified with a Superose 6 (Cytiva) gel filtration column. To stabilize the methylosome–U7 6S complex, the sample was cross-linked on ice for 2 h using 4 mM BS3 allowing cryo-EM reconstruction at 3.7 \AA resolution. In some experiments, we used the SAM analog SFG at 0.5 μM and 1.5 h cross-linking with 8 mM BS3. The cryo grids were prepared following the protocols described above, and the cryo-EM data were collected using a Titan Krios microscope at the New York Structural Biology Center. The EM data processing followed the protocols described above, and the statistics are summarized in Supplemental Table S1. Weak, putative EM density for U7 6S was observed after heterogeneous refinement in cryoSPARC, while the density became much weaker after nonuniform or homogeneous refinement. Various attempts to improve the quality of the density were unsuccessful.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank J. Oledzki and A. Fabijanska (Laboratory of Mass Spectrometry, Institute of Biochemistry and Biophysics, Polish Academy of Sciences) for mass spectrometry analysis, and K. Krajewski (High-Throughput Peptide Synthesis and Array Core Facility, UNC, Chapel Hill) for peptide synthesis. We also thank L. Pellizzoni (Columbia University) for the gift of a monoclonal anti-SmB antibody. This research was funded by the National Institutes of Health (NIH) grants GM 29832 (Z.D.) and R35GM118093 (L.T.).

Received May 10, 2023; accepted July 8, 2023.

REFERENCES

- Aik WS, Lin MH, Tan D, Tripathy A, Marzluff WF, Dominski Z, Chou CY, Tong L. 2017. The N-terminal domains of FLASH and Lsm11 form a 2:1 heterotrimer for histone pre-mRNA 3'-end processing. *PLoS One* **12**: e0186034. doi:10.1371/journal.pone.0186034

- Antonyamy S, Bonday Z, Campbell RM, Doyle B, Druzina Z, Gheyi T, Han B, Jungheim LN, Qian Y, Rauch C, et al. 2012. Crystal structure of the human PRMT5:MEP50 complex. *Proc Natl Acad Sci* **109**: 17960–17965. doi:10.1073/pnas.1209814109
- Azzouz TN, Pillai RS, Dapp C, Chari A, Meister G, Kambach C, Fischer U, Schumperli D. 2005. Toward an assembly line for U7 snRNPs: interactions of U7-specific Lsm proteins with PRMT5 and SMN complexes. *J Biol Chem* **280**: 34435–34440. doi:10.1074/jbc.M505077200
- Battle DJ, Kasim M, Yong J, Lotti F, Lau CK, Mouaikel J, Zhang Z, Han K, Wan L, Dreyfuss G. 2006a. The SMN complex: an assembly machine for RNPs. *Cold Spring Harb Symp Quant Biol* **71**: 313–320. doi:10.1101/sqb.2006.71.001
- Battle DJ, Lau CK, Wan L, Deng H, Lotti F, Dreyfuss G. 2006b. The Gemin5 protein of the SMN complex identifies snRNAs. *Mol Cell* **23**: 273–279. doi:10.1016/j.molcel.2006.05.036
- Bedford MT, Richard S. 2005. Arginine methylation an emerging regulator of protein function. *Mol Cell* **18**: 263–272. doi:10.1016/j.molcel.2005.04.003
- Bucholc K, Aik WS, Yang XC, Wang K, Zhou ZH, Dadlez M, Marzluff WF, Tong L, Dominski Z. 2020. Composition and processing activity of a semi-recombinant holo U7 snRNP. *Nucleic Acids Res* **48**: 1508–1530. doi:10.1093/nar/gkz1148
- Burghes AH, Beattie CE. 2009. Spinal muscular atrophy: Why do low levels of survival motor neuron protein make motor neurons sick? *Nat Rev Neurosci* **10**: 597–609. doi:10.1038/nrn2670
- Burgos ES, Wilczek C, Onikubo T, Bonanno JB, Jansong J, Reimer U, Shechter D. 2015. Histone H2A and H4 N-terminal tails are positioned by the MEP50 WD repeat protein for efficient methylation by the PRMT5 arginine methyltransferase. *J Biol Chem* **290**: 9674–9689. doi:10.1074/jbc.M115.636894
- Chari A, Golas MM, Klingenhager M, Neuenkirchen N, Sander B, Englbrecht C, Sickmann A, Stark H, Fischer U. 2008. An assembly chaperone collaborates with the SMN complex to generate spliceosomal snRNPs. *Cell* **135**: 497–509. doi:10.1016/j.cell.2008.09.020
- Chari A, Paknia E, Fischer U. 2009. The role of RNP biogenesis in spinal muscular atrophy. *Curr Opin Cell Biol* **21**: 387–393. doi:10.1016/j.ceb.2009.02.004
- Cote J, Richard S. 2005. Tudor domains bind symmetrical dimethylated arginines. *J Biol Chem* **280**: 28476–28483. doi:10.1074/jbc.M414328200
- Dominski Z, Marzluff WF. 2007. Formation of the 3' end of histone mRNA: getting closer to the end. *Gene* **396**: 373–390. doi:10.1016/j.gene.2007.04.021
- Dominski Z, Tong L. 2021. U7 deciphered: the mechanism that forms the unusual 3' end of metazoan replication-dependent histone mRNAs. *Biochem Soc Trans* **49**: 2229–2240. doi:10.1042/BST20210323
- Dominski Z, Sumerel J, Hanson RJ, Marzluff WF. 1995. The polyribosomal protein bound to the 3' end of histone mRNA can function in histone pre-mRNA processing. *RNA* **1**: 915–923.
- Dominski Z, Yang XC, Marzluff WF. 2005. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* **123**: 37–48. doi:10.1016/j.cell.2005.08.002
- Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126–2132. doi:10.1107/S0907444904019158
- Fallini C, Bassell GJ, Rossoll W. 2012. Spinal muscular atrophy: the role of SMN in axonal mRNA regulation. *Brain Res* **1462**: 81–92. doi:10.1016/j.brainres.2012.01.044
- Friesen WJ, Massenet S, Paushkin S, Wyce A, Dreyfuss G. 2001a. SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol Cell* **7**: 1111–1117. doi:10.1016/S1097-2765(01)00244-1
- Friesen WJ, Paushkin S, Wyce A, Massenet S, Pesiridis GS, Van Duyne G, Rappsilber J, Mann M, Dreyfuss G. 2001b. The methylome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol* **21**: 8289–8300. doi:10.1128/MCB.21.24.8289-8300.2001
- Friesen WJ, Wyce A, Paushkin S, Abel L, Rappsilber J, Mann M, Dreyfuss G. 2002. A novel WD repeat protein component of the methylome binds Sm proteins. *J Biol Chem* **277**: 8243–8247. doi:10.1074/jbc.M109984200
- Goddard TD, Huang CC, Ferrin TE. 2007. Visualizing density maps with UCSF Chimera. *J Struct Biol* **157**: 281–287. doi:10.1016/j.jsb.2006.06.010
- Grimm C, Chari A, Pelz JP, Kuper J, Kisker C, Diederichs K, Stark H, Schindelin H, Fischer U. 2013. Structural basis of assembly chaperone-mediated snRNP formation. *Mol Cell* **49**: 692–703. doi:10.1016/j.molcel.2012.12.009
- Gruess OJ, Meduri R, Schilling M, Fischer U. 2017. UsnRNP biogenesis: mechanisms and regulation. *Chromosoma* **126**: 577–593. doi:10.1007/s00412-017-0637-6
- Gubitz AK, Feng W, Dreyfuss G. 2004. The SMN complex. *Exp Cell Res* **296**: 51–56. doi:10.1016/j.yexcr.2004.03.022
- Guderian G, Peter C, Wiesner J, Sickmann A, Schulze-Osthoff K, Fischer U, Grimm M. 2011. RioK1, a new interactor of protein arginine methyltransferase 5 (PRMT5), competes with pICln for binding and modulates PRMT5 complex composition and substrate specificity. *J Biol Chem* **286**: 1976–1986. doi:10.1074/jbc.M110.148486
- Ho MC, Wilczek C, Bonanno JB, Xing L, Seznec J, Matsui T, Carter LG, Onikubo T, Kumar PR, Chan MK, et al. 2013. Structure of the arginine methyltransferase PRMT5-MEP50 reveals a mechanism for substrate specificity. *PLoS One* **8**: e57008. doi:10.1371/journal.pone.0057008
- Iannaccone ST, Smith SA, Simard LR. 2004. Spinal muscular atrophy. *Curr Neurol Neurosci Rep* **4**: 74–80. doi:10.1007/s11910-004-0016-6
- Kambach C, Walke S, Young R, Avis JM, De la Fortelle E, Raker VA, Lührmann R, Li J, Nagai K. 1999. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**: 375–387. doi:10.1016/S0092-8674(00)80550-4
- Khusial P, Plaag R, Zieve GW. 2005. LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem Sci* **30**: 522–528. doi:10.1016/j.tibs.2005.07.006
- Lau CK, Bachorik JL, Dreyfuss G. 2009. Gemin5-snRNA interaction reveals an RNA binding function for WD repeat domains. *Nat Struct Mol Biol* **16**: 486–491. doi:10.1038/nsmb.1584
- Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, et al. 1995. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**: 155–165. doi:10.1016/0092-8674(95)90460-3
- Leung AK, Nagai K, Li J. 2011. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* **473**: 536–539. doi:10.1038/nature09956
- Li DK, Tisdale S, Lotti F, Pellizzoni L. 2014. SMN control of RNP assembly: from post-transcriptional gene regulation to motor neuron disease. *Semin Cell Dev Biol* **32**: 22–29. doi:10.1016/j.semcdb.2014.04.026
- Li J, Leung AK, Kondo Y, Oubridge C, Nagai K. 2016. Re-refinement of the spliceosomal U4 snRNP core-domain structure. *Acta Crystallogr D Struct Biol* **72**: 131–146. doi:10.1107/S2059798315022111
- Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, et al. 2019.

- Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**: 861–877. doi:10.1107/S2059798319011471
- Liu H, Moore CL. 2021. On the cutting edge: regulation and therapeutic potential of the mRNA 3' end nuclease. *Trends Biochem Sci* **46**: 772–784. doi:10.1016/j.tibs.2021.04.003
- Lotti F, Imlach WL, Saieva L, Beck ES, Hao le T, Li DK, Jiao W, Mentis GZ, Beattie CE, McCabe BD, et al. 2012. An SMN-dependent U12 splicing event essential for motor circuit function. *Cell* **151**: 440–454. doi:10.1016/j.cell.2012.09.012
- Lührmann R, Kastner B, Bach M. 1990. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochim Biophys Acta Gene Struct Expr* **1087**: 265–292. doi:10.1016/0167-4781(90)90001-I
- Mandel CR, Bai Y, Tong L. 2008. Protein factors in pre-mRNA 3'-end processing. *Cell Mol Life Sci* **65**: 1099–1122. doi:10.1007/s00018-007-7474-3
- Matera AG, Terns RM, Terns MP. 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* **8**: 209–220. doi:10.1038/nrm2124
- Mayeda A, Krainer AR. 1999. Preparation of HeLa cell nuclear and cytosolic S100 extracts for in vitro splicing. *Methods Mol Biol* **118**: 309–314. doi:10.1385/1-59259-676-2:309
- Meister G, Eggert C, Buhler D, Brahms H, Kambach C, Fischer U. 2001. Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr Biol* **11**: 1990–1994. doi:10.1016/S0960-9822(01)00592-9
- Monani UR. 2005. Spinal muscular atrophy: a deficiency in a ubiquitous protein; a motor neuron-specific disease. *Neuron* **48**: 885–896. doi:10.1016/j.neuron.2005.12.001
- Mulvaney KM, Blomquist C, Acharya N, Li R, Ranaghan MJ, O'Keefe M, Rodriguez DJ, Young MJ, Kesar D, Pal D, et al. 2021. Molecular basis for substrate recruitment to the PRMT5 methylosome. *Mol Cell* **81**: 3481–3495.e7. doi:10.1016/j.molcel.2021.07.019
- Musiani D, Bok J, Massignani E, Wu L, Tabaglio T, Ippolito MR, Cuomo A, Ozbek U, Zorgati H, Ghoshdastider U, et al. 2019. Proteomics profiling of arginine methylation defines PRMT5 substrate specificity. *Sci Signal* **12**: eaat8388. doi:10.1126/scisignal.aat8388
- Najbauer J, Johnson BA, Young AL, Aswad DW. 1993. Peptides with sequences similar to glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins. *J Biol Chem* **268**: 10501–10509. doi:10.1016/S0021-9258(18)82227-X
- Neuenkirchen N, Chari A, Fischer U. 2008. Deciphering the assembly pathway of Sm-class U snRNPs. *FEBS Lett* **582**: 1997–2003. doi:10.1016/j.febslet.2008.03.009
- Neuenkirchen N, Englbrecht C, Ohmer J, Ziegenhals T, Chari A, Fischer U. 2015. Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization. *EMBO J* **34**: 1925–1941. doi:10.15252/embj.201490350
- Paknia E, Chari A, Stark H, Fischer U. 2016. The ribosome cooperates with the assembly chaperone pICln to initiate formation of snRNPs. *Cell Rep* **16**: 3103–3112. doi:10.1016/j.celrep.2016.08.047
- Pellizzoni L. 2007. Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep* **8**: 340–345. doi:10.1038/sj.embor.7400941
- Pelz JP, Schindelin H, van Pee K, Kuper J, Kisker C, Diederichs K, Fischer U, Grimm C. 2015. Crystallizing the 6S and 8S spliceosomal assembly intermediates: a complex project. *Acta Crystallogr D Biol Crystallogr* **71**: 2040–2053. doi:10.1107/S1399004715014832
- Pesiridis GS, Diamond E, Van Duyn GD. 2009. Role of pICln in methylation of Sm proteins by PRMT5. *J Biol Chem* **284**: 21347–21359. doi:10.1074/jbc.M109.015578
- Petterson EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605–1612. doi:10.1002/jcc.20084
- Pillai RS, Will CL, Lührmann R, Schumperli D, Muller B. 2001. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J* **20**: 5470–5479. doi:10.1093/emboj/20.19.5470
- Pillai RS, Grimm M, Meister G, Will CL, Lührmann R, Fischer U, Schumperli D. 2003. Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev* **17**: 2321–2333. doi:10.1101/gad.274403
- Pu WT, Krapivinsky GB, Krapivinsky L, Clapham DE. 1999. pICln inhibits snRNP biogenesis by binding core spliceosomal proteins. *Mol Cell Biol* **19**: 4113–4120. doi:10.1128/MCB.19.6.4113
- Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**: 290–296. doi:10.1038/nmeth.4169
- Raker VA, Plessel G, Lührmann R. 1996. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*. *EMBO J* **15**: 2256–2269. doi:10.1002/j.1460-2075.1996.tb00579.x
- Raker VA, Hartmuth K, Kastner B, Lührmann R. 1999. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol Cell Biol* **19**: 6554–6565. doi:10.1128/MCB.19.10.6554
- Schumperli D, Pillai RS. 2004. The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell Mol Life Sci* **61**: 2560–2570. doi:10.1007/s00018-004-4190-0
- Selenko P, Sprangers R, Stier G, Buhler D, Fischer U, Sattler M. 2001. SMN Tudor domain structure and its interaction with the Sm proteins. *Nat Struct Biol* **8**: 27–31. doi:10.1038/83014
- Skrajna A, Yang XC, Dadlez M, Marzluff WF, Dominski Z. 2018. Protein composition of catalytically active U7-dependent processing complexes assembled on histone pre-mRNA containing biotin and a photo-cleavable linker. *Nucleic Acids Res* **46**: 4752–4770. doi:10.1093/nar/gky133
- Smith HO, Tabiti K, Schaffner G, Soldati D, Albrecht U, Birnstiel ML. 1991. Two-step affinity purification of U7 small nuclear ribonucleoprotein particles using complementary biotinylated 2'-O-methyl oligoribonucleotides. *Proc Natl Acad Sci* **88**: 9784–9788. doi:10.1073/pnas.88.21.9784
- Stefanovic B, Hackl W, Lührmann R, Schumperli D. 1995. Assembly, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into *Xenopus* oocytes. *Nucleic Acids Res* **23**: 3141–3151. doi:10.1093/nar/23.16.3141
- Stopa N, Krebs JE, Shechter D. 2015. The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. *Cell Mol Life Sci* **72**: 2041–2059. doi:10.1007/s00018-015-1847-9
- Sun Y, Zhang Y, Aik WS, Yang XC, Marzluff WF, Walz T, Dominski Z, Tong L. 2020. Structure of an active human histone pre-mRNA 3'-end processing machinery. *Science* **367**: 700–703. doi:10.1126/science.aaz7758
- Sun Y, Aik WS, Yang XC, Marzluff WF, Dominski Z, Tong L. 2021. Reconstitution and biochemical assays of an active human histone pre-mRNA 3'-end processing machinery. *Methods Enzymol* **655**: 291–324. doi:10.1016/bs.mie.2021.03.021
- Timm DE, Bowman V, Madsen R, Rauch C. 2018. Cryo-electron microscopy structure of a human PRMT5:MEP50 complex. *PLoS One* **13**: e0193205. doi:10.1371/journal.pone.0193205
- Tisdale S, Lotti F, Saieva L, Van Meerbeke JP, Crawford TO, Sumner CJ, Mentis GZ, Pellizzoni L. 2013. SMN is essential for

- the biogenesis of U7 small nuclear ribonucleoprotein and 3'-end formation of histone mRNAs. *Cell Rep* **5**: 1187–1195. doi:10.1016/j.celrep.2013.11.012
- Tisdale S, Van Alstyne M, Simon CM, Mentis GZ, Pellizzoni L. 2022. SMN controls neuromuscular junction integrity through U7 snRNP. *Cell Rep* **40**: 111393. doi:10.1016/j.celrep.2022.111393
- Tripsianes K, Madl T, Machyna M, Fessas D, Englbrecht C, Fischer U, Neugebauer KM, Sattler M. 2011. Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat Struct Mol Biol* **18**: 1414–1420. doi:10.1038/nsmb.2185
- Wahl MC, Fischer U. 2016. The right pick: structural basis of snRNA selection by Gemin5. *Genes Dev* **30**: 2341–2344. doi:10.1101/gad.293084.116
- Wilkinson ME, Charenton C, Nagai K. 2020. RNA splicing by the spliceosome. *Annu Rev Biochem* **89**: 359–388. doi:10.1146/annurev-biochem-091719-064225
- Will CL, Luhrmann R. 2001. Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13**: 290–301. doi:10.1016/S0955-0674(00)00211-8
- Winkler C, Eggert C, Gradl D, Meister G, Giegerich M, Wedlich D, Lagerbauer B, Fischer U. 2005. Reduced U snRNP assembly causes motor axon degeneration in an animal model for spinal muscular atrophy. *Genes Dev* **19**: 2320–2330. doi:10.1101/gad.342005
- Yang XC, Burch BD, Yan Y, Marzluff WF, Dominski Z. 2009. FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3' end processing of histone pre-mRNAs. *Mol Cell* **36**: 267–278. doi:10.1016/j.molcel.2009.08.016
- Yang XC, Sabath I, Debski J, Kaus-Drobek M, Dadlez M, Marzluff WF, Dominski Z. 2013. A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3'-end processing. *Mol Cell Biol* **33**: 28–37. doi:10.1128/MCB.00653-12
- Yang XC, Sun Y, Aik WS, Marzluff WF, Tong L, Dominski Z. 2020. Studies with recombinant U7 snRNP demonstrate that CPSF73 is both an endonuclease and a 5'–3' exonuclease. *RNA* **26**: 1345–1359. doi:10.1261/ma.076273.120
- Yi H, Mu L, Shen C, Kong X, Wang Y, Hou Y, Zhang R. 2020. Negative cooperativity between Gemin2 and RNA provides insights into RNA selection and the SMN complex's release in snRNP assembly. *Nucleic Acids Res* **48**: 895–911. doi:10.1093/nar/gkz1135
- Yong J, Kasim M, Bachorik JL, Wan L, Dreyfuss G. 2010. Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol Cell* **38**: 551–562. doi:10.1016/j.molcel.2010.03.014
- Zhang Z, Lotti F, Dittmar K, Younis I, Wan L, Kasim M, Dreyfuss G. 2008. SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* **133**: 585–600. doi:10.1016/j.cell.2008.03.031
- Zhang R, So BR, Li P, Yong J, Glisovic T, Wan L, Dreyfuss G. 2011. Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146**: 384–395. doi:10.1016/j.cell.2011.06.043
- Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**: e42166. doi:10.7554/eLife.42166