Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization

Nils Neuenkirchen^{1,†,‡}, Clemens Englbrecht^{1,‡}, Jürgen Ohmer¹, Thomas Ziegenhals¹, Ashwin Chari^{2,**} & Utz Fischer^{1,3,*}

Abstract

The assembly of spliceosomal U snRNPs depends on the coordinated action of PRMT5 and SMN complexes *in vivo*. These *trans*-acting factors enable the faithful delivery of seven Sm proteins onto snRNA and the formation of the common core of snRNPs. To gain mechanistic insight into their mode of action, we reconstituted the assembly machinery from recombinant sources. We uncover a stepwise and ordered formation of distinct Sm protein complexes on the PRMT5 complex, which is facilitated by the assembly chaperone pICln. Upon completion, the formed pICln-Sm units are displaced by new pICln-Sm protein substrates and transferred onto the SMN complex. The latter acts as a Brownian machine that couples spontaneous conformational changes driven by thermal energy to prevent mis-assembly and to ensure the transfer of Sm proteins to cognate RNA. Investigation of mutant SMN complexes provided insight into the contribution of individual proteins to these activities. The biochemical reconstitution presented here provides a basis for a detailed molecular dissection of the U snRNP assembly reaction.

Keywords assembly; pICln; PRMT5; SMN; snRNP

Subject Category RNA Biology

DOI 10.15252/embj.201490350 | Received 20 October 2014 | Revised 11 May 2015 | Accepted 12 May 2015 | Published online 11 June 2015

The EMBO Journal (2015) 34: 1925–1941

Introduction

Macromolecular complexes perform vital activities in virtually all cells. The cellular environment in which these complexes are assembled is highly crowded, and thus not only substantially perturbs diffusion-driven self-assembly, but also increases the likelihood of non-productive interactions. Assembly processes therefore require

trans-acting factors, which sequester subunits of macromolecular complexes and safeguard them from engaging in unwanted interactions (Ellis, 2006; Chari & Fischer, 2010). Compelling evidence has emerged in recent years that the formation of even comparatively simple structures often requires a plethora of assembly factors (Zemp *et al*, 2009; Chari & Fischer, 2010; Liu *et al*, 2010). An extreme example is the assembly of the common Sm core structure of spliceosomal snRNPs (U1, U2, U4/U6 and U5), which utilizes more assembly factors than parts to be assembled (Fischer *et al*, 2011).

Even though self-recognition of RNA and protein counterparts is sufficient for Sm core assembly *in vitro* (Raker *et al*, 1996), at least 12 *trans*-acting factors, united in PRMT5 and SMN complexes, participate in this process *in vivo* in animals (Meister *et al*, 2002; Paushkin *et al*, 2002). Other eukaryotes, such as *Drosophila melanogaster* and *Trypanosoma brucei*, only contain SMN and Gemin2 genes in their genome, and sometimes also genes for additional SMN complex subunits (Kroiss *et al*, 2008; Palfi *et al*, 2009). However, it remains to be elucidated whether and how these additional SMN complex subunits contribute to snRNP biogenesis (Shpargel *et al*, 2009). The cellular snRNP assembly pathway can be divided into two distinct temporal phases (see Supplementary Fig S1 for a schematic overview of the snRNP assembly pathway). The early phase is dominated by the assembly chaperone pICln (Friesen *et al*, 2001; Meister *et al*, 2001b). pICln accepts newly synthesized Sm proteins and delivers them to the PRMT5 complex, which consists of PRMT5 and WD45 (also referred to as MEP50) (Friesen *et al*, 2001; Meister *et al*, 2001b). The PRMT5 subunit catalyzes symmetrical dimethylation of arginine residues (sDMA) within the Sm proteins B-B', D1 and D3 (Friesen & Dreyfuss, 2000; Brahms *et al*, 2001). Recently, a second methyltransferase was identified in humans and *Drosophila*, termed PRMT7 that is capable of arginine methylation and contributes to the snRNP assembly process (Gonsalvez *et al*, 2007, 2008). Additionally, pulse-chase experiments in HeLa cells have revealed that Sm protein methylation is

¹ Department of Biochemistry, Biocenter, University of Würzburg, Würzburg, Germany

² Research Group of 3D Electron Cryomicroscopy, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

³ Department of Radiation Medicine and Applied Sciences, University of California, San Diego, San Diego, CA, USA

*Corresponding author. Tel: +49 931 318 4029; E-mail: utz.fischer@biozentrum.uni-wuerzburg.de

**Corresponding author. Tel: +49 551 201 1654; E-mail: ashwin.chari@mpibpc.mpg.de

[‡]These authors contributed equally to this work

[†]Present address: Yale Stem Cell Center and Department of Cell Biology, Yale University School of Medicine, New Haven, CT, 06520, USA

required for proper snRNP biogenesis (Gonsalvez *et al*, 2007). It is believed from experiments in cell extracts that sDMA modification enhances the affinity of Sm proteins for the SMN complex *in vivo* (Brahms *et al*, 2001; Friesen *et al*, 2001; Meister *et al*, 2001b; Matera & Wang, 2014). However, in purified, reconstituted systems the effect is less apparent (Chari *et al*, 2008; Kroiss *et al*, 2008; Zhang *et al*, 2011). The early assembly phase eventually segregates into two lines. In one assembly line, a stable hexameric ring intermediate is formed termed the 6S complex, which is composed of pICln and the Sm proteins D1, D2, F, E and G, and pre-organizes these Sm proteins into spatial positions adopted in the assembled Sm core domain (Fisher *et al*, 1985; Chari *et al*, 2008; Grimm *et al*, 2013). The other assembly line consists of pICln-D3/B, which may likely not dissociate from the PRMT5 complex (Chari *et al*, 2008). The result of both assembly lines is that Sm proteins are kinetically trapped and fail to proceed in the assembly pathway. The late phase of snRNP formation is dominated by the SMN complex, which resolves this kinetic trap by accepting the pre-organized Sm proteins from pICln and catalyzes the subsequent joining of Sm proteins with snRNA (Chari *et al*, 2008; Zhang *et al*, 2011; Grimm *et al*, 2013).

While the basic principles in the cellular formation of snRNPs have been understood in great detail, several mechanistic aspects await their elucidation. For example, in the early assembly phase, the PRMT5 complex appears to contain all seven Sm proteins. Nevertheless, the understanding of how the two lines of assembly entailing 6S- and pICln-D3/B complexes emerge remains entirely elusive. Likewise, apart from SMN and Gemin2, which mediate Sm protein binding and assembly (Fischer *et al*, 1997; Meister *et al*, 2001a; Kroiss *et al*, 2008; Zhang *et al*, 2011; Grimm *et al*, 2013), the functions of other SMN complex subunits (Gemin3-8 and unrip) are less clear (Chari & Fischer, 2010). RNAi depletion of Gemin3-8 and unrip has been shown to perturb snRNP assembly in cell culture (Feng *et al*, 2005; Grimmer *et al*, 2005; Shpargel & Matera, 2005). However, it remains to be seen whether this requirement for Gemin3-8 and unrip stems from a direct function of these subunits in snRNP assembly or whether this is due to indirect effects such as the destabilization of the SMN complex or other possible roles of these subunits in other steps of snRNP biogenesis. Additionally, several, in part contradictory, functions have been ascribed to each of these subunits in snRNP assembly by a multitude of experimental approaches (Charroux *et al*, 1999, 2000; Baccon *et al*, 2002; Gubitza *et al*, 2002; Pellizzoni *et al*, 2002a; Carissimi *et al*, 2005, 2006a,b; Battle *et al*, 2006, 2007; Otter *et al*, 2007). As an example, snRNP assembly was shown to be ATP dependent when analyzed in extracts (Meister *et al*, 2001a) and with purified SMN complexes (Pellizzoni *et al*, 2002b). By examining the domain structure of proteins in the SMN complex, an ATP requirement has been attributed to the Gemin3 subunit, a DEAD-box helicase. Nevertheless, subsequent studies using purified systems showed that the assembly reaction could occur in the absence of ATP (Chari *et al*, 2008; Kroiss *et al*, 2008). Therefore, a deep mechanistic and eventually structure-based analysis of the reaction mechanism of the assisted snRNP assembly process is necessary.

The mechanistic dissection of many macromolecular machines, foremost the translation machinery and the RNA polymerase II transcription cycle, has greatly benefited over the last decades from the availability of totally reconstituted systems (Rodnina & Wintermeyer, 1995; Myers *et al*, 1997; Pisarev *et al*, 2007). To enable such studies

for the assisted formation of U snRNPs, we present here the first total *in vitro* reconstitution of the entire assembly machinery (i.e., PRMT5 and SMN complexes) from recombinant sources. Using this system, we obtained mechanistic insight into the formation of the pICln-Sm assembly intermediates on the PRMT5 complex and their subsequent transfer onto the SMN complex. The investigation of the late phase allowed us to identify subunits of the SMN complex that are important for snRNP assembly and RNA proofreading. Moreover, reduced levels of SMN or expression of mutant versions thereof elicits the neuromuscular disorder spinal muscular atrophy (SMA) (Lunn & Wang, 2008; Burghes & Beattie, 2009; Kolb & Kissel, 2011). The reconstituted system presented here enables us to dissect the contribution of disease-causing mutations and deletions and undoubtedly will be a valuable tool for future mechanistic and eventually structure-based studies of the snRNP assembly reaction mechanism.

Results

Reconstitution of recombinant PRMT5 complex

The cytosolic assembly phase of U snRNPs starts with the sequestration of newly synthesized Sm proteins onto the PRMT5 complex and their subsequent release as distinct pICln-Sm assembly intermediates (i.e., 6S and pICln-D3/B). To analyze in a reconstituted system how the PRMT5 complex generates these units, we initially established protocols for the production of the recombinant human PRMT5, WD45 and pICln proteins (i.e., the components of the PRMT5 complex) and the seven Sm proteins (B, D1, D2, D3, E, F and G). PRMT5 and WD45 were obtained by co-expression in the MultiBac expression system (Berger *et al*, 2004) and were purified (Fig 1A, lane 1, Fig 1C and Supplementary Fig S2). pICln and the Sm protein heterooligomers D1/D2, F/E/G and D3/B were expressed in bacteria and purified as described previously (Fig 1A, lanes 2–5) (Chari *et al*, 2008). Subsequently, pICln was incubated with D1/D2 and D3/B, or a mixture of D1/D2 and F/E/G. From these proteins, we reconstituted the pICln-Sm assembly intermediates, pICln-D1/D2, pICln-D3/B and the 6S complex as described (see Fig 1A, lanes 6–8 and Supplementary Fig S3) (Chari *et al*, 2008). One crucial step in the initial phase of snRNP assembly is the recruitment of all Sm proteins to the PRMT5 complex and the methylation of a subset of them (i.e., B, D1 and D3) (Brahms *et al*, 2001; Friesen *et al*, 2001; Meister *et al*, 2001b). To test whether the recombinant PRMT5 complex is enzymatically active in our *in vitro* system, PRMT5/WD45 was incubated with either individual Sm protein heterooligomers or their corresponding pICln complexes. Methylation was measured using an established *in vitro* methylation assay with [³H]-labeled S-adenosylmethionine (SAM) as a co-factor (Frankel *et al*, 2002). Recombinant PRMT5/WD45 efficiently symmetrically di-methylated the guanidine group of arginine residues of free and pICln-bound Sm proteins B, D1 and D3 (Fig 1B and Supplementary Fig S2D), as evident by the incorporation of [³H]-labeled methyl groups (see also Supplementary Information and Supplementary Fig S4 for a detailed characterization of the modification). Thus, our recombinant system of the early assembly phase is functional with respect to binding and methylation of individual Sm protein heterooligomers as well as pICln-Sm complexes.

Figure 1. *In vitro* reconstitution of the human PRMT5 complex.

- A Purified proteins and reconstituted protein complexes of the early phase of cytoplasmic snRNP biogenesis. PRMT5 and WD45 were co-expressed in Sf21 insect cells (lane 1). PICln (lane 2) and Sm protein complexes D1/D2 (lane 3), D3/B (lane 4) and F/E/G (lane 5) were produced in bacterial cells. After the individual expression and purification of PICln and Sm protein complexes, pICln-D1/D2 (lane 6), pICln-D1/D2/F/E/G (= 6S; lane 7) and pICln-D3/B (lane 8) were reconstituted *in vitro* and resolved by gel filtration chromatography. Purified proteins were separated by SDS-PAGE and visualized by Coomassie staining.
- B Autoradiography of *in vitro* methylated Sm and pICln-Sm protein complexes using radioactively labeled [³H]-SAM as the methyl group donor.
- C–J Reconstitution of PRMT5-Sm protein complexes. PRMT5/WD45 alone (C) or in combination with D1/D2 (D), F/E/G (E), D3/B (F), pICln (G), pICln-D1/D2 (H), 6S (I) or pICln-D3/B (J) was incubated at 4°C overnight and resolved by gel filtration chromatography. Separated protein complexes were subsequently analyzed by SDS-PAGE and silver staining. Left panels: schematic of formed protein complexes; right panels: SDS-PAGE.

Data information: Asterisks indicate protein degradation products.

Stepwise formation of the 6S assembly intermediate on the PRMT5 complex

The availability of the recombinant PRMT5 complex allowed us to start a series of experiments addressing the order of events leading to the formation of the 6S complex, that is, the most prominent assembly intermediate arising from the PRMT5 complex. For this, we incubated PRMT5/WD45 with Sm and pICln-Sm heterooligomers and analyzed the resulting complexes by gel filtration chromatography (Fig 1C–J). Consistent with a previous study, PRMT5/WD45 alone migrated as a stoichiometric heterooctamer in gel filtration chromatography (Antonyasamy *et al*, 2012; Fig 1C). We found that pICln, D1/D2 and D3/B alone bound inefficiently to PRMT5/WD45, while F/E/G did not bind at all (Fig 1D–G, note that unbound D1/D2 elutes in a volume not analyzed on the gels, see also Supplementary Fig S3). The weak binding of D1/D2 and D3/B could be inferred from our finding that both are substrates for methylation by PRMT5/WD45 even in the absence of pICln (see Fig 1B, lanes 2 and 4). However, when equimolar amounts of Sm proteins were provided in a pICln-bound form, they were present in higher abundance in the PRMT5 complex (Fig 1H–J). As pICln is able to bind independently to PRMT5/WD45 (Fig 1G), the most likely explanation is that pICln serves as the bridging factor between PRMT5/WD45 and Sm proteins.

In vivo, pICln is part of both 20S and 6S complexes (Chari *et al*, 2008). The 20S complex consists of PRMT5, WD45, pICln and all Sm heterooligomers. Of note, the relative stoichiometry of the Sm proteins within the 20S complex differs strongly *in vivo*. Whereas D1/D2 and D3/B are often stoichiometric to each other and to the PRMT5/WD45 dimer, F/E/G is underrepresented and in some cases even absent (Friesen *et al*, 2001; Chari *et al*, 2008). In contrast, the 6S complex is a stoichiometric hexamer consisting of pICln-D1/D2/F/E/G. Noteworthy, this unit or parts thereof must have been in contact with the 20S complex as D1 is symmetrically dimethylated. Hence, the 6S complex represents an assembly intermediate bridging the early and late phases of snRNP biogenesis (Chari *et al*, 2008). These findings, along with the observation that stable incorporation of F/E/G into the 6S complex requires cooperative binding via pICln and D2, suggest a stepwise assembly of 6S on PRMT5/WD45.

To investigate this possibility, we initially incubated recombinant PRMT5/WD45 with the pICln-D1/D2 trimer, which led to the formation of a defined complex (Fig 2A and B). Fractions containing the PRMT5/WD45/pICln-D1/D2 complex (Fig 2B) were then pooled, incubated with F/E/G and re-analyzed by gel filtration chromatography. This resulted in the stable recruitment of the F/E/G trimer to the PRMT5 complex (Fig 2C). As the F/E/G trimer does not interact

with any of the PRMT5 complexes independently (Fig 1E), this recruitment is likely to reflect the formation of the 6S complex. To proceed in the assembly reaction, the 6S complex must leave the PRMT5/WD45 unit (Chari *et al*, 2008). In order to elucidate the basis for this reaction, we asked whether the release is diffusion driven, dependent on D1 methylation or caused by another protein factor involved in the snRNP assembly pathway. Neither the prolonged incubation of the PRMT5 complex nor methylation of D1 resulted in 6S complex release, as evident by gel filtration analysis (Fig 2E, lanes 1–4). However, addition of pICln, pICln-D1/D2 or pICln-D3/B quantitatively expelled the 6S complex (Fig 2E, lanes 5–10). To evaluate whether pICln-mediated dissociation of the 6S complex is a consequence of its reduced affinity to PRMT5/WD45, we reconstituted a PRMT5/WD45/pICln-D1/D2 complex (Fig 2F and G). The reconstituted complex was stable over time and was also not dissociated by the addition of 6S, indicating a higher affinity of pICln-D1/D2 toward PRMT5/WD45 and proving directionality in the assembly and release of the 6S complex (Fig 2H). These data suggest a sequence of events, in which initially pICln-D1/D2 binds to PRMT5/WD45 and the methylation reaction on D1 can occur. The complex then accepts F/E/G, which upon joining pICln-D1/D2 forms 6S that remains associated with the PRMT5 complex. PICln alone or containing either D1/D2 or D3/B expels the 6S complex due to their higher affinity toward PRMT5/WD45 and drives the assembly reaction in a forward direction.

Reconstitution of recombinant SMN complex

The results above illustrate how Sm building blocks are generated on the PRMT5 complex and are made available for the SMN complex-dominated late phase of assembly. The core machinery of the SMN complex consists of eight core proteins (SMN and Gemin2–8) and the seven Sm protein “substrates” that are transferred onto snRNA. The SMN-mediated assembly process can be subdivided into: (i) the binding of Sm proteins to the SMN complex, (ii) the concomitant release of the assembly chaperone pICln, and finally (iii) the assembly of the Sm proteins and snRNA to mature snRNP cores. To establish an assay that allowed us to recapitulate these events and to gain insight into its mechanism, we generated all components of the human SMN complex in a recombinant form. Based on a recently published interaction map of the SMN complex (Otter *et al*, 2007), the components SMN, Gemin2, Gemin6, Gemin7 and Gemin8 form a central protein network, which allowed their co-expression in bacteria and purification as a stable unit as reported (termed SMNΔGemin3–5, Fig 3A, lane 1) (Chari *et al*, 2008). The three remaining and peripheral components Gemin3, Gemin4 and Gemin5 were produced in insect cells, as their size prevented

Figure 2.

Figure 2. The 6S assembly intermediate is formed in a stepwise manner on PRMT5/WD45.

- A Schematic experimental outline of 6S formation.
- B PRMT5/WD45 and a 2-fold molar excess of pICln-D1/D2 were incubated at 4°C overnight and resolved by gel filtration chromatography.
- C Fractions containing PRMT5/WD45/pICln-D1/D2 were pooled (B, enclosed by dashed line), incubated with a 5-fold molar excess of F/E/G and separated by size fractionation.
- D Experimental outline of 6S release: gel filtration fractions containing PRMT5/WD45/6S (C, enclosed by dashed line) were pooled and incubated with buffer only, methylated for 1 h at 37°C with 1 mM SAM, or incubated with pICln, pICln-D1/D2 or pICln-D3/B overnight at 4°C.
- E SDS-PAGE of the gel filtration elution peak fractions containing the PRMT5 complex (I) or pICln-Sm protein complexes (II). Protein complexes were detected by silver staining.
- F–H 6S is unable to replace pICln-D1/D2 from the PRMT5 complex. (F) Schematic outline of 6S competition. The PRMT5/WD45/pICln-D1/D2 complex was formed as described above (G). Respective elution fractions were pooled, incubated overnight at 4°C in the absence or presence of 6S and separated by size. (H) SDS-PAGE of the concentrated elution peak fractions containing the PRMT5 complex (I) and pICln-Sm protein complexes (II).
- Data information: Asterisks indicate protein degradation products.
Source data are available online for this figure.

expression in *E. coli*, and were purified to homogeneity (Fig 3A, lane 2). As determined by gel filtration analysis, Gemin3/Gemin4 formed a heterohexamer, whereas Gemin5 formed a tetramer. Upon incubation of these Gemin3 with SMNΔGemin3-5, the complete reconstitution of the SMN complex was accomplished (Fig 3B, lane 2). The recombinant SMN complex was first characterized by its ability to bind Sm proteins and to release pICln (Fig 3B). For this, we incubated immobilized SMN complex with pICln-D3/B and 6S, that is, those pICln-Sm complexes released from the PRMT5 complex. Analysis by gel electrophoresis revealed Sm protein transfer from both pICln-Sm intermediates (Fig 3B, lanes 3 and 4, see also Supplementary Fig S5B). Of note, even though the Sm proteins bound efficiently to the SMN complex, pICln was absent as determined by Western blotting (Fig 3B, lanes 3 and 4 lower panel; lanes 5 and 6 show input 6S and pICln-D3/B complexes. See also Supplementary Fig S5B, lower panel). Thus, similar to its endogenous counterpart (Meister *et al*, 2001a; Pellizzoni *et al*, 2002b; Chari *et al*, 2008), the recombinant SMN complex accepts pre-organized Sm proteins and simultaneously expels pICln.

Next, we tested whether the reconstituted SMN complex also promotes the assembly of bound Sm proteins onto snRNA. For this purpose, the SMN complex was initially incubated with either the 6S complex or a mixture of 6S and pICln-D3/B, resulting in loading with Sm proteins of the subcore assembly intermediate and the mature core, respectively (Chari *et al*, 2008). Hereafter, both complexes were incubated with either [³²P]-labeled wild-type U1 snRNA or a mutant thereof lacking a functional Sm site (U1ΔSm snRNA) and RNP formation was visualized by native gel electrophoresis (Fig 3C). Whereas both complexes failed to assemble Sm proteins onto U1ΔSm snRNA (Fig 3C, lanes 7 and 8), efficient formation of the subcore and core snRNP on the wild-type U1 snRNA could be observed (Fig 3C, lanes 3 and 4). We conclude that the recombinant SMN complex is active in snRNP core assembly onto RNAs containing a functional Sm site and, regarding this activity, is indistinguishable from the endogenous complex.

The SMN complex acts as a Brownian machine in U snRNP assembly

The availability of functional recombinant SMN complex enabled us to investigate individual components in the assembly reaction. We initially focused on the function of the peripheral SMN complex components Gemin3, 4 and 5. The Gemin3-4 heterodimer has been implied to hydrolyze ATP in the course of snRNP assembly, while

Gemin5 was shown to bind specifically to precursor and mature snRNAs and thus identifies RNA targets for the assembly reaction (Yong *et al*, 2010). Our recombinant system allowed us to directly test whether these proteins indeed perform the above-mentioned activities in the assembly reaction.

To analyze the energy requirement of the assembly reaction, we initially incubated Sm protein heterooligomers with U1 snRNA in the absence of SMN complex (Fig 3D, lane 1). In accordance with previous reports (Raker *et al*, 1996; Meister *et al*, 2001a; Pellizzoni *et al*, 2002b; Wan *et al*, 2005), assembly occurred in a spontaneous reaction as core formation was observed not only at 37°C but also at 4°C (Fig 3D, lanes 1 and 6). In striking contrast, snRNP core formation was entirely blocked at 4°C when Sm proteins were loaded onto the recombinant SMN complex prior to incubation with U1 snRNA (Fig 3D, lanes 9 and 10). When the reaction was carried out at 37°C, however, the assembly was efficient (Fig 3D, lanes 4 and 5). These results, in conjunction with our observation that ATP has no measurable effect on the assembly reaction (Fig 3E), suggest that the SMN complex could be considered a Brownian machine that couples spontaneous conformational changes driven by thermal energy to the directed delivery of Sm proteins onto snRNA.

Gemin5 is dispensable for snRNA identification and snRNP assembly *in vitro*

Next, we investigated whether Gemin5 was required for snRNA identification in our *in vitro* system as reported previously (Battle *et al*, 2006). For this, we generated [³²P]-labeled U1 snRNA and pre-U1 snRNA as target RNAs. The 3'-extension of the latter has been shown to form a stem-loop, which constitutes the snRNP code recognized by Gemin5 (Yong *et al*, 2010). Both snRNAs were subsequently incubated with either the complete recombinant SMN complex or SMNΔGemin3-5, and the assembly reaction was analyzed over time by native gel electrophoresis (Fig 4A). Surprisingly, both SMN complexes were active in snRNP assembly and also displayed very similar assembly kinetics for wild-type U1 snRNA (Fig 4A, upper panel and Fig 4B) and its precursor [Fig 4A, lower panel and Fig 4B; note that pre-U1 snRNA was assembled faster than the wild-type RNA as reported earlier (Yong *et al*, 2010)]. These data indicate that Gemin3, 4 and 5, while being an integral part of the SMN complex, are not essential for the identification of target RNA and their faithful assembly into the Sm core domain *in vitro*.

The SMN complex also prevents mis-assembly of Sm proteins onto non-target snRNAs (Pellizzoni *et al*, 2002b). We hence asked

Figure 3. *In vitro* reconstitution of assembly-active SMN complexes.

A SDS-PAGE of the purified SMN complex lacking Gemin3-5 and baculovirus-expressed and purified Gemin3-5 (lanes 1 and 2, respectively).
 B SMNΔGemin3-5 complex combined with Gemin3-5 was immunoprecipitated with anti-SMN-coupled beads. Immobilized entire complex was incubated with pICln-Sm complexes: 6S complex (lane 3), 6S and pICln-D3/B (lane 4), or treated with buffer only (lane 2). LC indicates the light chain of the antibody. Asterisks mark unspecific bands. The lower panel shows a Western blot against pICln.
 C *In vitro* assembly reactions on $[\alpha\text{-}^{32}\text{P}]$ -labeled U1 snRNA with SMN complexes depicted in (B) (lanes 2–4). Sm site dependency was controlled via a mutant version U1ΔSm (lanes 6–8). Lanes 1 and 5 show RNAs without proteins.
 D *In vitro* assembly reactions on $[\alpha\text{-}^{32}\text{P}]$ -labeled U1 snRNA with Sm core proteins alone (lanes 1 and 6) or pre-bound to the SMN complex (lanes 5 and 10) at 37°C or 4°C, respectively.
 E SMN complex loaded with Sm core proteins was incubated with $[\alpha\text{-}^{32}\text{P}]$ -labeled U1 snRNA in the absence or with increasing amounts of ATP (lanes 2 and 3–5, respectively). Lane 1 shows free RNA.

whether the recombinant SMN complex also fulfills this unique proofreading activity. Additionally, encouraged by the results above, we asked whether this proposed proofreading activity is dependent on Gemin3-5. For this purpose, we incubated a mixture of different cellular RNAs (3'-end-labeled with ^{32}P) with

Sm proteins that were either bound to SMNΔGemin3-5 or present as free heterooligomers under physiological conditions. The reaction mixture was subsequently immunoprecipitated with antibodies against the Sm proteins, and the co-precipitated RNAs were detected by electrophoresis. Under these conditions, the SMN

Figure 4. Gemins3, 4 and 5 are dispensable for snRNA identification and snRNP assembly *in vitro*.

A Assembly activity for SMN(WT)ΔGemin3-5 (left panel) and the entire SMN complex (right panel) over time with human U1 snRNA (upper panel) and pre-snRNA (lower panel). Lanes 1 and 9 show RNA only.

B Densitometric quantification of the experiment shown in (A). Measurements were performed with the software Image Lab (Bio-Rad) and fitted to saturation kinetics by Solver in Microsoft Excel ($n = 2$). Data points indicate the average value; bars represent the data range of all measurements.

C Autoradiography of the specificity assay. Sm core proteins alone or pre-bound to SMN(WT)ΔGemin3-5 were incubated with [α - 32 P]-labeled total cell RNA. The mixture was subsequently immunoprecipitated with an antibody against Sm proteins (Y12), and bound fractions were resolved on a urea gel after phenol extraction. Lane 1 shows the input.

Source data are available online for this figure.

complex assembled Sm cores only onto snRNAs, whereas other RNAs of the mixture (i.e., tRNAs) were no substrates (Fig 4C, lane 3). In striking contrast, incubation of free Sm protein heterooligomers leads to their non-specific association with any RNA present in the mixture (Fig 4C, lane 2).

We conclude that the recombinant SMN complex assembles Sm core domains in an ATP-independent manner and proofreads the assembly onto cognate snRNA. However, neither the putative ATPase/helicase Gemin3 nor the snRNA binding protein Gemin5 is required for this reaction *in vitro*.

SMA-causing missense mutations or deletions in SMN affect either the activity of the SMN complex or its assembly

Finally, we investigated the function of the SMN protein in U snRNP assembly. Structural, biochemical and bioinformatic studies have revealed that SMN together with Gemin2 forms the core of the SMN complex that mediates the binding of Sm proteins and their assembly onto U snRNA. SMN is essential for the formation of the SMN complex and hence for the assembly reaction per se. Mutations within the SMN gene cause the devastating neuromuscular disorder

spinal muscular atrophy (SMA). How SMN acts mechanistically, however, is still unclear. While it is impossible to assemble an SMN complex without the SMN protein itself, pathogenic missense

mutations have been described in SMA patients that allow their functional characterization in our reconstituted system. We initially focused our studies on a missense mutation in SMN that is located in

Figure 5. Recombinant SMN(E134K) complex shows defects in snRNP assembly.

A Transfer of Sm proteins from pICln-Sm complexes to SMN complexes containing either wild-type or mutant SMN protein. Immobilized wild-type and mutant complexes were incubated with the 6S complex (lanes 2 and 6), 6S and pICln-D3/B (lanes 3 and 7), or treated with buffer only (lanes 1 and 5). Analysis of retained proteins was achieved by a SDS-PAGE.

B Immobilized wild-type and mutant complexes shown in (A) were subsequently used for *in vitro* assembly reactions with [α - 32 P]-U1 snRNA. snRNP formation was monitored by native gel electrophoresis (lanes 1–3 and 5–7, respectively). Lane 4 shows RNA only.

C *In vitro* snRNP assembly reactions were performed as described above with increasing concentrations of the SMN(WT) Δ Gemin3-5 or SMN(E134K) Δ Gemin3-5 complexes, respectively.

D Quantification of the experiments shown in (C) with Image Lab (Bio-Rad). The data were fitted to the Boltzmann equation by Solver in Microsoft Excel ($n = 2$). Rep1 and Rep2 indicate the datasets of independent experiments.

Source data are available online for this figure.

Figure 6. SMN complex formation and assembly activity with mutant SMN proteins.

A Recombinant wild-type SMN protein was used to assemble the SMNΔGemin3-5 complex. After incubation with 6S complex, the mixture was separated by gradient centrifugation and the peak fraction (dashed box) was used for snRNP assembly assays (right panel) with radiolabeled U1 snRNA.

B–D The reconstitution and assembly activity of SMN complexes containing either SMN(D44V), SMN(ΔExon7) or SMN(ΔYG) is shown in (B), (C) and (D), respectively.

its central Tudor domain. The mutation alters a glutamate to a lysine in position 134 [termed SMN(E134K)] and has been shown to disrupt an aromatic cage designed to bind dimethylarginines of Sm proteins (Tripsianes *et al*, 2011). To functionally characterize this mutant in the context of the SMN complex, we co-expressed and purified the SMN(E134K) Δ Gemin3-5 complex as described above for its wild-type counterpart (see Fig 5A and Materials and Methods for details). A stable mutant complex was obtained whose biochemical composition was indistinguishable from the wild-type complex. This illustrated that the mutant protein had folded properly and did not induce an assembly defect on the SMN complex itself (Fig 5A, compare lanes 1 and 5). Furthermore, the mutant complex bound Sm proteins and expelled pICln as described for its wild-type counterpart (Fig 5A, lanes 2–3 and 6–7, Supplementary Fig S5C, lanes 15–16). To test whether the SMN(E134K) mutation caused a defect in U snRNP assembly, wild-type and mutant complexes were loaded with Sm proteins of the subcore and the core proteins, respectively, and were subsequently incubated with [32 P]-labeled wild-type U1 snRNA. U snRNP assembly was then analyzed by native gel electrophoresis as described above (Fig 5B). Strikingly, the mutant complex, although not entirely inactive, displayed a markedly reduced assembly activity (Fig 5B, lanes 6 and 7). A very similar picture was observed when wild-type and mutant SMN complexes also containing Gemin3-5 were analyzed (Supplementary Fig S5). To gain insight into the molecular basis of this defect, the dissociation constant (K_d) of the assisted assembly process was determined. Whereas the assembly reaction mediated by the wild-type complex had a K_d of approximately 126 nM, this value doubled with the mutant complex (Fig 5C and D). These data suggest that the Sm proteins are less efficiently transferred onto the U snRNA because they are either more tightly bound or mis-arranged on the mutant complex.

To extend our studies further, we analyzed three additional mutant SMN Δ Gemin3-5 complexes containing SMN with either a pathogenic missense mutation (SMN(D44V)), a C-terminal deletion (SMN(Δ Exon7)) that is the major product of the second human SMN gene (SMN2), or an artificial deletion of the highly conserved C-terminal YG box (SMN(Δ YG box)). The complexes were generated as described for SMN(E134K) Δ Gemin3-5 and analyzed by gradient centrifugation and *in vitro* assembly assays. As shown in Fig 6B, SMN(D44V) readily formed the SMN Δ Gemin3-5 complex and also displayed assembly activity comparable to the wild-type complex (compare Fig 6A with B). We note, however, that the SMN(D44V) Δ Gemin3-5 complex appeared to sediment in a lighter molecular weight range than the wild-type complex on density gradients, indicating that the stoichiometry of the complex had changed. Likewise, SMN(Δ Exon7) formed SMN complexes, which sedimented in a similar molecular weight range as the SMN(D44V) Δ Gemin3-5 complex. The snRNP assembly activity of the SMN(Δ Exon7) Δ Gemin3-5 complex, however, was entirely abrogated (Fig 6C, right panel), even though no gross defects in the SMN Δ Gemin3-5 complex were appreciable and Sm proteins were bound (Fig 6C, left panel). On the other hand, the artificial SMN(Δ YG box) Δ Gemin3-5 complex had a severe propensity to aggregate in our hands (Fig 6D, left panel). Concomitantly, this mutant SMN Δ Gemin3-5 complex was inactive in snRNP assembly (Fig 6D, right panel). Therefore, we conclude that SMN mutations/deletions affect the SMN complex differentially. Whereas some may affect the assembly activity directly (SMN

(E134K)), others may affect the integrity/stoichiometry of the complex (i.e., SMN(Δ Exon7) and SMN(Δ YG box)) or even processes yet to be identified (SMN(D44V)). The detailed structural, biochemical and functional analysis of these and potential other mutant complexes will hence enable the dissection of the assembly pathway and provide insight into the etiology of SMA.

Discussion

The detailed mechanistic analysis of the mode of action of cellular macromolecular complexes greatly benefits from the availability of reconstituted systems. Keeping with this notion, a series of *in vivo* experiments, extract systems and partial purification from cell lysates have defined the basic machinery and principles of cellular snRNP assembly (Friesen & Dreyfuss, 2000; Meister *et al*, 2001a,b; Pellizzoni *et al*, 2002b; Battle *et al*, 2006; Yong *et al*, 2010). Moreover, most recently, partial reconstitution (Chari *et al*, 2008) and structural characterization of key intermediates (Zhang *et al*, 2011; Sarachan *et al*, 2012; Grimm *et al*, 2013) have provided mechanistic insight at near atomic resolution. Despite these achievements, several aspects remain unanswered and a matter of debate. To establish experimental systems that may help to resolve these issues, we report here the total reconstitution of the human snRNP assembly machinery from recombinant proteins.

Employing this system, we show that the 6S complex, a key assembly intermediate, which constitutes a kinetic trap in snRNP assembly (Chari *et al*, 2008), is formed in a stepwise manner on the PRMT5 complex (Fig 7A). The heterotrimer pICln-D1/D2 binds to the PRMT5/WD45 heterodimer, which we confirm to form a quaternary structure, composed of 4 PRMT5/WD45 heterodimers (Antonyamy *et al*, 2012). Note that this quaternary structure enables the simultaneous binding of pICln-D1/D2 and pICln-D3/B to the PRMT5/WD45 heterooctamer. In this complex, mono- and symmetrical dimethylation of B, D1 and D3 can occur. *In vivo*, the PRMT5 complex contains only substoichiometric amounts of F/E/G compared to D1/D2 and D3/B (Chari *et al*, 2008). Using an *in vitro* competition assay, we are able to provide a basis for this finding and show that the 6S complex is formed in a sequential manner on the PRMT5 complex with pICln-D1/D2 binding initially and F/E/G joining subsequently (Fig 2). Surprisingly, neither methylation nor F/E/G binding is sufficient for 6S complex dissociation from PRMT5/WD45. This process appears to be feed-forward driven, where the addition of a new pICln-D1/D2 or pICln-D3/B heterotrimer displaces the 6S complex (Fig 2). Therefore, it might be a combination of two effects that cause the newly formed 6S complex from interacting with the PRMT5 complex and support it to pass on to the SMN complex. First, pICln-D1/D2 or pICln-D3/B expels 6S from the PRMT5 complex. Second, the methylated D1 protein in 6S might make it more susceptible to bind to the SMN complex.

In the late phase of snRNP biogenesis, the SMN complex sequesters the Sm proteins from the assembly intermediates 6S and pICln-D3/B and expels pICln. Subsequently, the snRNA interacts with the SMN complex-bound Sm proteins to form snRNPs. In order to address specific steps of the late phase, we initially verified whether the SMN complex, reconstituted from recombinant proteins, indeed had the same activities as its cellular counterpart.

Figure 7. Model of early and late phases of cytoplasmic snRNP assembly.

- A** Early phase: Sm protein heterooligomers D1/D2 and D3/B interact with pICln to form pICln-Sm protein complexes (1). These are recruited to PRMT5/WD45 forming a 20S complex and obtain symmetrical dimethylations in arginine side chains (sDMA, indicated by a star symbol) in their C-terminal domains (2). The addition of F/E/G to this results in the formation of a ring-shaped 6S assembly intermediate on PRMT5/WD45 (3). Further supplementation of pICln-Sm protein complexes releases 6S and again establishes a 20S complex (4). The pICln-D3/B moiety presumably remains attached to PRMT5/WD45 as experimental evidence of cytoplasmic localization of pICln-D3/B alone is lacking (5). As this complex can be detected *in vivo* only to a very small extent, it is likely to be a short-lived species (†). In either case, pICln serves as a molecular chaperone and prevents the interaction of the associated Sm proteins with snRNA (6).
- B** Late phase: Previously formed and methylated 6S and pICln-D3/B interact with the SMN complex. Whereas pICln is expelled, the Sm proteins are transferred onto the SMN complex (1). It has recently been shown that Gemin2 coordinates the binding of five Sm proteins (D1/D2/F/E/G) (Zhang *et al*, 2011), while D3/B is likely bound to SMN, Gemin2 and/or Gemin8 (Chari *et al*, 2008). Then, the snRNA associates with the loaded SMN complex and the Sm proteins are assembled around the Sm site (2). Finally, this complex is translocated to the nucleus where the mature snRNP dissociates from the SMN complex (3). All steps shown in B as well as the RNA proofreading activity occur without measurable participation of Gemin3, 4 and 5 *in vitro*. The entire reaction is driven by thermal energy only and hence identifies the SMN complex as a Brownian machine.

Quite unexpectedly, the central SMN complex components are both necessary and sufficient for all aspects of snRNP core formation *in vitro*. This finding is in contrast to former reports, which have emphasized the function of Gemin5 in binding to snRNAs and pre-snRNAs in a sequence-specific manner (Battle *et al*, 2006; Yong *et al*, 2010). While we were able to recapitulate that pre-snRNAs enhanced snRNP biogenesis, we found Gemin3-5 to be dispensable for assembly and proofreading of snRNAs (Figs 3 and 4). However,

this result confirms our former findings that a minimal subcomplex of SMN/Gemin2 is sufficient to carry out the assembly reaction *in vitro* (Chari *et al*, 2008). Notably, *Drosophila* and *Trypanosome* SMN complexes consist of only these two subunits *in vivo*, enforcing this notion (Kroiss *et al*, 2008; Palfi *et al*, 2009). Accordingly, as previously reported by us (Kroiss *et al*, 2008), SMN and Gemin2 were alone sufficient to proofread snRNAs. The view that emerges from the experiments presented here in Fig 4C and Kroiss *et al*

(2008) is that SMN and Gemin2 subunits provide a platform, where Sm proteins themselves are able to proofread the presence of cognate snRNAs. Additionally, in our reconstituted system we find no dependence of SMN complex-mediated snRNP assembly on ATP. We thus put forward the hypothesis that the SMN complex is a Brownian machine driven by thermal energy (Fig 7B). Interestingly, several laboratories including ours have shown that assembly depends on metabolic energy in cellular extracts (Meister *et al*, 2001a; Pellizzoni *et al*, 2002b). This raises the possibility that steps in the biogenesis pathway not covered in our *in vitro* system are driven by the hydrolysis of ATP.

Finally, we focused on the SMN protein itself. Since decreased levels of functional SMN protein lead to spinal muscular atrophy, we hypothesized that distinct patient mutations in the SMN protein give rise to defects in snRNP assembly. Consequently, we generated a recombinant SMN complex with a common patient mutation (E134K). We demonstrate that this mutation decreases the rate of snRNP formation, while it does not affect the binding of Sm proteins of assembly intermediates (Fig 5). This suggests a crucial role of the SMN protein in the transfer of Sm proteins from the SMN complex onto snRNA. It is interesting to note that the E134K mutation inhibits SMN's ability to interact with sDMA-methylated Sm proteins when analyzed in the absence of other Gemins (Bühler *et al*, 1999; Tripsianes *et al*, 2011). This finding may hint to an interaction of SMN's Tudor domain with Sm proteins at stages of the assembly pathway that are not covered in our *in vitro* assay. Furthermore, we assessed the SMN(D44V), SMN(Δ Exon7) and SMN(Δ YG box) mutations in the SMN gene for their capacity to form complexes of similar quaternary structure as the wild-type SMN gene and interrogated the snRNP assembly activity of these mutant complexes (Fig 6). The SMN(D44V) complex appears to form complexes that are somewhat smaller than the wild type and is not impaired in snRNP assembly. It might well be true that this SMN mutation exhibits a defect in snRNP biogenesis not recapitulated in our *in vitro* assay. The SMN(Δ Exon7) complex appeared to have a similar quaternary structure as the SMN(D44V) complex. However, this complex was entirely abrogated in its snRNP assembly activity. Of note, the SMN(Δ Exon7) protein is the predominant product expressed from the SMN2 gene and has also been shown to be unstable in human cells (Cho & Dreyfuss, 2010). Our data indicate that the lack of exon 7 not only interferes with the stability of SMN but also with its function in the assembly reaction. In striking contrast to the aforementioned two mutant SMN proteins, we found the SMN(Δ YG box) protein to impart a strong aggregation behavior for the SMN complex. This finding was surprising, considering that the YG box is believed to induce oligomerization of the SMN complex (Martin *et al*, 2012; Praveen *et al*, 2014) and is a common source of patient mutations (Wirth, 2000). In keeping with the aggregation propensity of this mutant SMN complex, there was no snRNP assembly activity.

Cellular and organismic approaches have thus far been utilized to study the effects of disease-causing mutations and/or deletions of the SMN gene (Lefebvre *et al*, 1995; Wang & Dreyfuss, 2001; Shpargel & Matera, 2005; Praveen *et al*, 2014). The reconstitution system we describe here enables mechanistic analyses of snRNP assembly and the significance of disease-causing mutations in detail. Our *in vitro* reconstitution scheme will therefore be essential for the understanding of the molecular etiology of the devastating disease spinal muscular atrophy.

Materials and Methods

Expression and purification of PRMT5/WD45

The genes encoding human His₆-tagged PRMT5 isoform 1, WD45 and enhanced green fluorescent protein (EGFP) were cloned between the *EcoRI* and *XhoI* sites (PRMT5, WD45) and *NdeI* and *StuI* sites (EGFP) of a modified pFBDM vector (pFBDM4) of the MultiBac system (Berger *et al*, 2004 and Supplementary Methods). This plasmid was transformed into *E. coli* DH10MultiBac cells, and positive colonies were identified by blue/white screening. Twelve micrograms of isolated bacmid DNA was transfected into 3×10^6 Sf21 insect cells using Cellfectin II (Invitrogen) and incubated for 10 days at 27°C in 10 ml EX-CELL TiterHigh™ medium (Sigma-Aldrich). The P1 virus titer of this was applied to 250 ml of Sf21 insect cell suspension culture (2×10^6 cells/ml) at 100 rpm and 27°C for 72–79 h. Baculovirus titer concentrations were determined by end-point dilution identifying EGFP-expressing cells (O'Reilly *et al*, 1993). Sf21 insect cells (2×10^6 cells/ml) were infected with recombinant baculovirus titer at 3 MOI (PRMT5, WD45, EGFP) and incubated for 72–79 h. The overexpressed protein was purified using Ni-NTA chromatography (Qiagen) with a binding buffer of 20 mM HEPES-NaOH (pH 7.5), 1 M NaCl, 10% (v/v) glycerol, 10 mM imidazole, 1 mM PMSF, 20 mg/l aprotinin, 20 mg/l leupeptin/pepstatin, 0.2 mM AEBSF and 5 mM β -mercaptoethanol, and an elution buffer lacking protease inhibitors but containing 250 mM imidazole. Elution fractions were pooled, dialyzed twice against 20 mM HEPES-NaOH (pH 7.5), 90 mM NaCl and 5 mM DTT at 4°C for 3 h and applied to anion exchange chromatography (HiTrap Q 1 ml, GE Healthcare). Proteins were eluted by increasing the NaCl concentration. Finally, recombinant proteins were separated by gel filtration chromatography (Superose 6 10/300 GL, GE Healthcare) in 20 mM HEPES-NaOH (pH 7.5), 200 mM NaCl and 5 mM DTT.

Expression and purification of Gemin3, Gemin4 and Gemin5

Genes encoding human His₆- and GST-tagged Gemin3 and Gemin5 were introduced into MCS2 of individual pFBDM4 transfer vectors using the *NcoI* and *NotI* sites and EGFP into MCS1 applying the *NdeI* and *StuI* sites of the same constructs. His₆-tagged Gemin4 was inserted into MCS1 of pFBDM4 using the *EcoRI* and *XhoI* sites, whereas EGFP was introduced into MCS2 via the *NcoI* and *NotI* sites. All coding sequences of protein affinity tags were followed by a tobacco etch virus (TEV) cleavage site. Resulting transfer vectors were transfected into Sf21 insect cells as described above and used to generate P2 baculovirus titers. Sf21 insect cells (2×10^6 cells/ml) were infected at 3 MOI with combinations of recombinant baculoviruses (GST-Gemin3/EGFP + His₆-Gemin4/EGFP, His₆-Gemin3/EGFP + His₆-Gemin4/EGFP +/- GST-Gemin5/EGFP or His₆-Gemin5 alone) and incubated for 72–79 h at 27°C.

Cells were harvested by centrifugation, resuspended in buffer C [50 mM sodium phosphate (pH 7.5), 500 mM NaCl, 10 mM EDTA, 1 mM spermidine, 1 mM TCEP] containing protease inhibitors and broken by sonication. A cleared lysate was prepared by ultracentrifugation in a 45Ti rotor (Beckman) for 1 h at 185,500 g and 4°C. The cleared lysate was incubated with Glutathione Sepharose® 4B (GE Healthcare) in batch for 2 h at 4°C. After extensive washing with

buffer C, the protein complex was eluted with buffer C containing 40 mM glutathione. Fractions were identified by SDS–PAGE and stored at 4°C until further use.

Expression and purification of the SMNΔGemin3-5 complex

For the expression and purification of the SMNΔGemin3-5 complex, *E. coli* BL21 Star™ (DE3) cells (Novagen) were co-transformed with both Gemin7-Gemin6:pET21a and Gemin2-Gemin8- His₆-GST-TEV-SMN:pET28b* plasmids (for cloning strategy see Chari *et al*, 2008). Additionally, the plasmid pRARE (Novagen) was co-transformed. Cells were cultured in Super Broth medium containing 2% glucose, 500 mM sorbitol, 1 mM betaine, 100 µg/ml ampicillin, 25 µg/ml kanamycin and 34 µg/ml chloramphenicol. Cultures were incubated at 37°C to an OD₆₀₀ of 0.2 and cooled to 15°C, and protein expression was induced with 0.5 mM IPTG at an OD₆₀₀ of 0.4 for additional 20 h. Cells were harvested by centrifugation, resuspended in buffer A [50 mM imidazole (pH 6.8), 300 mM Na₂SO₄, 10 mM EDTA, 10% (w/v) galactose, 1 mM spermidine and 1 mM TCEP] containing protease inhibitors and broken by sonication. A cleared lysate was prepared by ultracentrifugation in a 45Ti rotor (Beckman) for 1 h at 72,400 g and 4°C. This was then incubated with Glutathione Sepharose® 4B (GE Healthcare) in batch for 2 h at 4°C. After extensive washing with buffer A and buffer B [50 mM imidazole (pH 6.8), 150 mM Na₂SO₄, 10% (w/v) galactose, 1 mM spermidine and 1 mM TCEP], the Sepharose beads were supplemented with a 1:50 ratio (protease:protein) of TEV protease and incubated for 14 h at 4°C. The supernatant and four wash fractions containing the cleaved SMNΔGemin3-5 complex were identified by SDS–PAGE and stored at 4°C until further use.

Expression and purification of pICln and Sm protein heterooligomers D1/D2, D3/B and F/E/G

pICln, the heterodimers D1/D2 and D3/B as well as the heterotrimer F/E/G were expressed in *E. coli* and purified as described (Kambach *et al*, 1999; Chari *et al*, 2008).

Reconstitution of pICln-Sm protein complexes and PRMT5-Sm protein complexes

pICln-Sm protein complexes were reconstituted as described previously (Chari *et al*, 2008). To reconstitute PRMT5 complexes *in vitro*, recombinant PRMT5/WD45 was incubated with a 2- to 5-fold excess of Sm protein heterooligomers (D1/D2, F/E/G or D3/B), pICln or pICln-Sm protein complexes (pICln-D1/D2, 6S or pICln-D3/B) in 20 mM HEPES-NaOH (pH 7.5), 1 M NaCl, 10% (v/v) glycerol and 5 mM β-mercaptoethanol. These were then dialyzed overnight at 4°C against the same buffer containing 200 mM NaCl. Finally, protein complexes were resolved by gel filtration chromatography (Superose 6 10/300 GL, GE Healthcare). The formation of protein complexes was verified by SDS–PAGE and subsequent silver staining. Protein standards of known molecular weight (dextran blue: 2,000 kDa, thyroglobulin: 669 kDa, ferritin: 440 kDa and BSA: 67 kDa) were applied separately to the same gel filtration column. The resulting elution profiles were plotted against the elution volume and are indicated above the SDS gels.

Sequential formation of 6S on PRMT5/WD45 and subsequent release

Recombinant PRMT5/WD45 was incubated with a 2-fold excess of pICln-D1/D2 in 20 mM HEPES-NaOH (pH 7.5), 1 M NaCl and 5 mM DTT and dialyzed against 20 mM HEPES-NaOH (pH 7.5), 200 mM NaCl and 5 mM DTT overnight at 4°C. Complex formation was monitored by gel filtration chromatography using identical buffer conditions on a Superose 6 10/300 GL column (GE Healthcare) and subsequent SDS–PAGE. Elution fractions containing the protein complex of PRMT5/WD45/pICln-D1/D2 were pooled and incubated with a 5-fold excess of F/E/G, and reconstitution was analyzed as before. Finally, fractions containing PRMT5/WD45/6S were treated with a 3.5-fold excess of pICln-D1/D2 and assayed for complex formation.

In vitro methylation of Sm proteins by recombinant PRMT5/WD45

One picomole of recombinant His₆-PRMT5/WD45 was incubated with a hundredfold molar excess of D1/D2, pICln-D1/D2, 6S, D3/B or pICln-D3/B and 219 pmol radioactively labeled co-factor S-adenosylmethionine [SAM; mixture of 50% [³H]-SAM (Perkin Elmer: 10 Ci/mmol) and 50% SAM (Sigma-Aldrich)] and a reaction buffer containing 100 mM Hepes-NaOH (pH 8.2), 200 mM NaCl and 5 mM DTT in a total volume of 20 µl at 37°C for 60 min. Reactions were stopped by the addition of 6× SDS–PAGE loading buffer and subsequent incubation at 95°C for 5 min. Proteins were separated on a 13% SDS gel and fixed with 30% (v/v) methanol and 10% (v/v) acetic acid. The radioactive signal was amplified by incubation with NAMP-100 amplifying reagent (GE Healthcare) for 45 min. Subsequently, gels were dried and exposed to Amersham Hyperfilm™ MP (GE Healthcare) at –80°C for 18 h.

In vitro reconstitution of recombinant SMN complexes

In order to reconstitute the entire human SMN complex from recombinant sources, 40 pmol of bacterially expressed SMNΔGemin3-5 (containing either the wild-type SMN protein or the mutants E134K, D44V, ΔExon7 or ΔYG box) was combined with: (i) buffer only, (ii) 40 pmol of insect cell-expressed Gemin3/Gemin4, (iii) Gemin5, or (iv) Gemin3/Gemin4 + Gemin5. Each of these reactions was then supplemented with: (i) buffer only, (ii) 200 pmol of 6S, (iii) 200 pmol pICln-D3/B, or (iv) 200 pmol of 6S and pICln-D3/B. Finally, BSA was added to a final concentration of 1 mg/ml to prevent protein precipitation. Reaction samples were dialyzed overnight against 20 mM HEPES-NaOH (pH 7.5), 200 mM NaCl and 5 mM DTT at 4°C followed by a dialysis against the same buffer lacking DTT for 3 h at 4°C. Samples were centrifuged at 13,000 g for 30 min at 4°C, and the supernatant was incubated with 40 µl Protein G-Sepharose™ beads (GE Healthcare) coupled with 2.5 mg/ml 7B10 antibody (anti-SMN) at 600 rpm and 4°C for 90 min. Beads were washed twice with 1.2 ml 1× PBS, 300 mM NaCl and 0.01% NP-40, and twice with 1× PBS and 300 mM NaCl. Finally, the beads were resuspended in 20 µl 1× PBS and 300 mM NaCl, supplemented with 5 µl 6× SDS–PAGE loading buffer and analyzed by 12.5% SDS–PAGE and Coomassie staining. For subsequent band shifts, larger protein amounts were deployed in complex reconstitutions.

Native gel electrophoresis of RNA–protein complexes

Band shift assays were performed essentially as described (Meister *et al*, 2001a; Chari *et al*, 2008). In brief, 5 pmol proteins were added to 0.025 pmol radiolabeled RNA in 16 µl reactions with 0.4 U/µl RNasin and 0.1 µg/µl tRNA and BSA. The mixtures were incubated for 60 min at 37°C (or respective time). After incubation, the mixtures were briefly centrifuged, supplemented with heparin to a final concentration of 0.5 µg/µl and separated on 6% native polyacrylamide gels (acrylamide/bisacrylamide ratio 80:1) containing 4% (v/v) glycerol and 1× TBE buffer. Gels were pre-run for 1 h at 4°C at 100 V in 1× TBE and run with samples for 2 h at 4°C at 300 V. Gels were exposed wet to Amersham Hyperfilm™ MP (GE Healthcare) at –80°C. Densitometric measurements were quantified with the Image Lab software integrated into the Gel Doc™ XR+ system (Bio-Rad). Free RNA was taken as a reference (100%). Values were fitted to saturation kinetics via the Solver add-in of Microsoft Excel and plotted. All values shown are averages of two independent experiments.

In vitro transcription of U snRNAs

[³²P]-labeled *X. laevis* U1 snRNA and U1ΔSm snRNA (Hamm *et al*, 1987) were obtained by an *in vitro* run-off transcription. pUC9 vectors containing the coding sequences were linearized with *Bam*HI and purified by phenol–chloroform extraction. The transcription was carried out at 37°C for 4 h. Transcripts were separated by electrophoresis on a 5% polyacrylamide gel under denaturing conditions. Respective bands were cut out, purified via ethanol precipitation, resuspended in water and stored at –20°C until further use.

Cloning of the human U1 snRNA and pre-snRNA

Primers hU1-for-*Eco*RI-T7 and Pre-hU1-rev-*Bam*HI were used to amplify human U1 pre-snRNA from a genomic DNA preparation. After *Eco*RI and *Bam*HI cleavage, the fragment was ligated into an analogously cleaved pUC19 vector. An altered Sm site was introduced by site-specific mutations with primers hU1-ΔSm-upr and hU1-ΔSm-lwr.

In vitro transcription was carried out with hU1-for-*Eco*RI-T7 and hU1-rev (human U1 snRNA) and hU1-for-*Eco*RI-T7 and pre-hU1-rev (human U1 pre-snRNA).

hU1-D_{Sm}-upr: 5'-GGAACTCGACTGCATACGGACTCGTAGTGGGG GACTG-3'

hU1-D_{Sm}-lwr: 5'-CAGTCCCCACTACGAGTCCGTATGCAGTCGAGT TTCC-3'

hU1-for-*Eco*RI-T7: 5'-GGAATTCCTAATACGACTCACTATAGGATACT TACCTGGCAGGGGAGATACC-3'

Pre-hU1-rev-*Bam*HI: 5'-CGGGATCCCGAAAAGATATGACCTTGGC GTACAGTCTG-3'

hU1-rev: 5'-CAGGGGAAAGCGGAACGCAGTCCCCACTACCACAA ATTATGC-3'

prehU1-rev: 5'-AAAAGATATGACCTTGGCGTACAGTCTG-3'

The RNA concentration was measured via a spectrophotometer at 260 nm.

Assembly reactions in vitro

RNA was prepared from HeLa nuclear cell extract via TRIzol® (Life Tech) treatment according to the manufacturer's specification.

Additionally, snRNAs were isolated from snRNPs (Sumpter *et al*, 1992) and mixed with nuclear RNA (10 mg each). The RNA mixture was 3'-end-labeled with pCp overnight and subsequently phenolized. A total of 80 pmol of Sm core proteins loaded onto SMNΔGemin3-5, Sm core proteins alone or no protein was incubated with 500 ng of labeled RNA at 37°C or 4°C for 1 h in the presence of 1 mg/ml BSA in 1× PBS. Notably, this reaction was performed in the absence of non-specific competitors such as heparin. Sm proteins and bound RNA were immunoprecipitated via Y12 antibody coupled to protein G-Sepharose (1.5 h, 4°C on a head-over-tail rotor). The beads were washed three times with 1× PBS and phenolized. The resulting RNA pellet was resuspended in 20 µl ddH₂O. Samples were separated on an 8% denaturing urea gel (30 W constantly).

Supplementary information for this article is available online:

<http://emboj.emboipress.org>

Acknowledgements

We thank I. Mattaj for providing plasmids and members of the laboratory for help and criticism. This work was supported by grants of the DFG (FI573-8/1 to UF and CH1098-1/1 to AC), and the Rudolf-Virchow-Centre of Experimental Medicine, Würzburg, to UF.

Author contributions

NN and CE contributed equally to this work. NN, CE, AC and UF designed the experiments and analyzed the results. NN, CE, AC, JO and TZ performed the experiments. NN, CE, AC and UF wrote the manuscript. UF and AC supervised the research.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Antonyamy S, Bonday Z, Campbell RM, Doyle B, Druzina Z, Gheyi T, Han B, Jungheim LN, Qian Y, Rauch C, Russell M, Sauder JM, Wasserman SR, Weichert K, Willard FS, Zhang A, Emtage S (2012) Crystal structure of the human PRMT5:MEP50 complex. *Proc Natl Acad Sci USA* 109: 17960–17965
- Baccon J, Pellizzoni L, Rappsilber J, Mann M, Dreyfuss G (2002) Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex. *J Biol Chem* 277: 31957–31962
- Battle DJ, Lau CK, Wan L, Deng H, Lotti F, Dreyfuss G (2006) The Gemin5 protein of the SMN complex identifies snRNAs. *Mol Cell* 23: 273–279
- Battle DJ, Kasim M, Wang J, Dreyfuss G (2007) SMN-independent subunits of the SMN complex. Identification of a small nuclear ribonucleoprotein assembly intermediate. *J Biol Chem* 282: 27953–27959
- Berger I, Fitzgerald DJ, Richmond TJ (2004) Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol* 22: 1583–1587

- Brahms H, Meheus L, de Brabandere V, Fischer U, Lührmann R (2001) Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* 7: 1531–1542
- Bühler D, Raker V, Lührmann R, Fischer U (1999) Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. *Hum Mol Genet* 8: 2351–2357
- 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
- Carissimi C, Baccon J, Straccia M, Chiarella P, Maiolica A, Sawyer A, Rappsilber J, Pellizzoni L (2005) Unrip is a component of SMN complexes active in snRNP assembly. *FEBS Lett* 579: 2348–2354
- Carissimi C, Saieva L, Baccon J, Chiarella P, Maiolica A, Sawyer A, Rappsilber J, Pellizzoni L (2006a) Gemin8 is a novel component of the survival motor neuron complex and functions in small nuclear ribonucleoprotein assembly. *J Biol Chem* 281: 8126–8134
- Carissimi C, Saieva L, Gabanella F, Pellizzoni L (2006b) Gemin8 is required for the architecture and function of the survival motor neuron complex. *J Biol Chem* 281: 37009–37016
- Chari A, Golas MM, Klingenhäger 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
- Chari A, Fischer U (2010) Cellular strategies for the assembly of molecular machines. *Trends Biochem Sci* 35: 676–683
- Charroux B, Pellizzoni L, Perkinson RA, Shevchenko A, Mann M, Dreyfuss G (1999) Gemin3: a novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. *J Cell Biol* 147: 1181–1194
- Charroux B, Pellizzoni L, Perkinson RA, Yong J, Shevchenko A, Mann M, Dreyfuss G (2000) Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli. *J Cell Biol* 148: 1177–1186
- Cho S, Dreyfuss G (2010) A degon created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes Dev* 24: 438–442
- Ellis RJ (2006) Molecular chaperones: assisting assembly in addition to folding. *Trends Biochem Sci* 31: 395–401
- Feng W, Gubitz AK, Wan L, Battle DJ, Dostie J, Golembe TJ, Dreyfuss G (2005) Gemin3 modulate the expression and activity of the SMN complex. *Hum Mol Genet* 14: 1605–1611
- Fischer U, Liu Q, Dreyfuss G (1997) The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* 90: 1023–1029
- Fischer U, Englbrecht C, Chari A (2011) Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdiscip Rev RNA* 2: 718–731
- Fisher DE, Conner GE, Reeves WH, Wisniewolski R, Blobel G (1985) Small nuclear ribonucleoprotein particle assembly in vivo: demonstration of a 6S RNA-free core precursor and posttranslational modification. *Cell* 42: 751–758
- Frankel A, Yadav N, Lee J, Branscombe TL, Clarke S, Bedford MT (2002) The novel human protein arginine N-methyltransferase PRMT6 is a nuclear enzyme displaying unique substrate specificity. *J Biol Chem* 277: 3537–3543
- Friesen WJ, Dreyfuss G (2000) Specific sequences of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene product (SMN). *J Biol Chem* 275: 26370–26375
- Friesen WJ, Paushkin S, Wyce A, Massenet S, Pesiridis GS, Van Duyne G, Rappsilber J, Mann M, Dreyfuss G (2001) The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol* 21: 8289–8300
- Gonsalvez GB, Tian L, Ospina JK, Boisvert FM, Lamond AI, Matera AG (2007) Two distinct arginine methyltransferases are required for biogenesis of Sm-class ribonucleoproteins. *J Cell Biol* 178: 733–740
- Gonsalvez GB, Praveen K, Hicks AJ, Tian L, Matera AG (2008) Sm protein methylation is dispensable for snRNP assembly in *Drosophila melanogaster*. *RNA* 14: 878–887
- 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
- Grimmler M, Otter S, Peter C, Muller F, Chari A, Fischer U (2005) Unrip, a factor implicated in cap-independent translation, associates with the cytosolic SMN complex and influences its intracellular localization. *Hum Mol Genet* 14: 3099–3111
- Gubitz AK, Mourelatos Z, Abel L, Rappsilber J, Mann M, Dreyfuss G (2002) Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J Biol Chem* 277: 5631–5636
- Hamm J, Kazmaier M, Mattaj JW (1987) In vitro assembly of U1 snRNPs. *EMBO J* 6: 3479–3485
- 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
- Kolb SJ, Kissel JT (2011) Spinal muscular atrophy: a timely review. *Arch Neurol* 68: 979–984
- Kroiss M, Schultz J, Wiesner J, Chari A, Sickmann A, Fischer U (2008) Evolution of an RNP assembly system: a minimal SMN complex facilitates formation of UsnRNPs in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 105: 10045–10050
- Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, Le Paslier D, Frézal J, Cohen D, Weissenbach J, Munnich A, Melki J (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80: 155–165
- Liu C, Young AL, Starling-Windhof A, Bracher A, Saschenbrecker S, Rao BV, Rao KV, Berninghausen O, Mielke T, Hartl FU, Beckmann R, Hayer-Hartl M (2010) Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature* 463: 197–202
- Lunn MR, Wang CH (2008) Spinal muscular atrophy. *Lancet* 371: 2120–2133
- Martin R, Gupta K, Ninan NS, Perry K, Van Duyne GD (2012) The survival motor neuron protein forms soluble glycine zipper oligomers. *Structure* 20: 1929–1939
- Matera AG, Wang Z (2014) A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* 15: 108–121
- Meister G, Bühler D, Pillai R, Lottspeich F, Fischer U (2001a) A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs. *Nat Cell Biol* 3: 945–949
- Meister G, Eggert C, Bühler D, Brahms H, Kambach C, Fischer U (2001b) Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr Biol* 11: 1990–1994
- Meister G, Eggert C, Fischer U (2002) SMN-mediated assembly of RNPs: a complex story. *Trends Cell Biol* 12: 472–478
- Myers LC, Leuther K, Bushnell DA, Gustafsson CM, Kornberg RD (1997) Yeast RNA polymerase II transcription reconstituted with purified proteins. *Methods* 12: 212–216
- O'Reilly DR, Miller LK, Luckow VA (1993) *Baculovirus Expression Vectors: A Laboratory Manual*. Oxford: Oxford University Press

- Otter S, Grimm M, Neuenkirchen N, Chari A, Sickmann A, Fischer U (2007) A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *J Biol Chem* 282: 5825–5833
- Palfi Z, Jae N, Preusser C, Kaminska KH, Bujnicki JM, Lee JH, Gunzl A, Kambach C, Urlaub H, Bindereif A (2009) SMN-assisted assembly of snRNP-specific Sm cores in trypanosomes. *Genes Dev* 23: 1650–1664
- Paushkin S, Gubitz AK, Massenet S, Dreyfuss G (2002) The SMN complex, an assemblyosome of ribonucleoproteins. *Curr Opin Cell Biol* 14: 305–312
- Pellizzoni L, Baccon J, Rappsilber J, Mann M, Dreyfuss G (2002a) Purification of native survival of motor neurons complexes and identification of Gemin6 as a novel component. *J Biol Chem* 277: 7540–7545
- Pellizzoni L, Yong J, Dreyfuss G (2002b) Essential role for the SMN complex in the specificity of snRNP assembly. *Science* 298: 1775–1779
- Pisarev AV, Unbehaun A, Hellen CU, Pestova TV (2007) Assembly and analysis of eukaryotic translation initiation complexes. *Methods Enzymol* 430: 147–177
- Praveen K, Wen Y, Gray KM, Noto JJ, Patlolla AR, Van Duyne GD, Matera AG (2014) SMA-causing missense mutations in survival motor neuron (Smn) display a wide range of phenotypes when modeled in *Drosophila*. *PLoS Genet* 10: e1004489
- 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
- Rodnina MV, Wintermeyer W (1995) GTP consumption of elongation factor Tu during translation of heteropolymeric mRNAs. *Proc Natl Acad Sci USA* 92: 1945–1949
- Sarachan KL, Valentine KG, Gupta K, Moorman VR, Gledhill JM Jr, Bernens M, Tommos C, Wand AJ, Van Duyne GD (2012) Solution structure of the core SMN-Gemin2 complex. *Biochem J* 445: 361–370
- Shpargel KB, Matera AG (2005) Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins. *Proc Natl Acad Sci USA* 102: 17372–17377
- Shpargel KB, Praveen K, Rajendra TK, Matera AG (2009) Gemin3 is an essential gene required for larval motor function and pupation in *Drosophila*. *Mol Biol Cell* 20: 90–101
- Sumpter V, Kahrs A, Fischer U, Kornstadt U, Lührmann R (1992) In vitro reconstitution of U1 and U2 snRNPs from isolated proteins and snRNA. *Mol Biol Rep* 16: 229–240
- 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
- Wan L, Battle DJ, Yong J, Gubitz AK, Kolb SJ, Wang J, Dreyfuss G (2005) The survival of motor neurons protein determines the capacity for snRNP assembly: biochemical deficiency in spinal muscular atrophy. *Mol Cell Biol* 25: 5543–5551
- Wang J, Dreyfuss G (2001) Characterization of functional domains of the SMN protein in vivo. *J Biol Chem* 276: 45387–45393
- Wirth B (2000) An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum Mutat* 15: 228–237
- 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
- Zemp I, Wild T, O'Donohue MF, Wandrey F, Widmann B, Gleizes PE, Kutay U (2009) Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. *J Cell Biol* 185: 1167–1180
- 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