SMN-assisted assembly of snRNP-specific Sm cores in trypanosomes

Zsofia Palfi, Nicolas Jaé, Christian Preußer, et al.

Genes Dev. 2009 23: 1650-1664
Access the most recent version at doi:10.1101/gad.526109

SMN-assisted assembly of snRNP-specific Sm cores in trypanosomes

Zsofia Palfi,[1,7] Nicolas Jaé,[1,7] Christian Preußer,[1,7] Katarzyna H. Kaminska,[2,3] Janusz M. Bujnicki,[2,3] Ju Huck Lee,[4] Arthur Günzl,[4] Christian Kambach,[5] Henning Urlaub,[6] and Albrecht Bindereif[1,8]

[1]Institute of Biochemistry, Justus Liebig University of Giessen, D-35392 Giessen, Germany; [2]Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, PL-02-109 Warsaw, Poland; [3]Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, PL-61-614 Poznan, Poland; [4]Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06030, USA; [5]Structural Biology, Paul Scherrer Institute, CH-5232 Villigen, Switzerland; [6]Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

Spliceosomal small nuclear ribonucleoproteins (snRNPs) in trypanosomes contain either the canonical heptameric Sm ring (U1, U5, spliced leader snRNPs), or variant Sm cores with snRNA-specific Sm subunits (U2, U4 snRNPs). Searching for specificity factors, we identified SMN and Gemin2 proteins that are highly divergent from known orthologs. SMN is splicing-essential in trypanosomes and nuclear-localized, suggesting that Sm core assembly in trypanosomes is nuclear. We demonstrate in vitro that SMN is sufficient to confer specificity of canonical Sm core assembly and to discriminate against binding to nonspecific RNA and to U2 and U4 snRNAs. SMN interacts transiently with the SmD3B subcomplex, contacting specifically SmB. SMN remains associated throughout the assembly of the Sm heteroheptamer and dissociates only when a functional Sm site is incorporated. These data establish a novel role of SMN, mediating snRNP specificity in Sm core assembly, and yield new biochemical insight into the mechanism of SMN activity.

[*Keywords*: SMN; snRNA; snRNP; splicing; trypanosomes]

Supplemental material is available at http://www.genesdev.org.

Received February 10, 2009; revised version accepted June 2, 2009.

Before pre-mRNA splicing catalysis can occur, each intron–exon unit has to assemble into a large spliceosome complex, composed of small nuclear ribonucleoproteins (snRNPs) and many additional non-snRNP proteins. This occurs in an ordered multistep process with the spliceosome undergoing several conformational transitions (Brow 2002; Will and Lührmann 2006). Each of the snRNPs is assembled from snRNA and proteins that are characterized very well in their biochemical composition for the mammalian system (Will and Lührmann 2001). The protein components can be divided into the Sm core; a heteroheptameric, ring-like complex of seven Sm proteins (SmD3B, D1D2, EFG in the U1, U2, U4, and U5 snRNPs; LSm2-8 in the U6 snRNP); and additional, snRNA-specific proteins (Kambach et al. 1999; for review, see Khusial et al. 2005).

In contrast to the mammalian and yeast systems, only a few protein components of the splicing machinery have been identified in trypanosomes. Trypanosomes are particularly interesting since their expression of protein-coding genes requires *trans* splicing, and in addition—at least for a small number of genes—*cis* splicing (for review, see Liang et al. 2003): The U2, U4/U6, and U5 snRNPs are considered to be general and essential splicing factors, whereas the SL RNP and the U1 snRNP represent *trans*- and *cis*-splicing-specific components, respectively.

Recently, first examples of what we called "Sm core variation" were described in spliceosomal snRNPs from trypanosomes: First, in the U2 snRNP, two of the canonical Sm polypeptides, SmD3B, are replaced by two U2-specific Sm proteins, Sm16.5K/15K (Wang et al. 2006). Second, at least in a fraction of the U4 snRNP, the U4-specific Ssm4/LSm2 substitutes for the canonical SmD3 (Wang et al. 2006; Tkacz et al. 2007; N. Jaé and A. Bindereif, unpubl.). In sum, the U2 and U4 Sm core variants differ in one or two of the canonical Sm subunits. As shown so far for U2, this Sm core variation mediates RNA-binding specificity for the unusual Sm site in the U2 snRNA (Wang et al. 2006). The phenomenon raises new questions on the specificity of biogenesis and assembly of Sm cores, questions that are relevant also for the different mammalian Sm cores found in the spliceosomal snRNPs versus the U7 snRNP (Pillai et al. 2003; Schümperli and Pillai 2004).

Our classical concept of snRNP biogenesis relies primarily on studies in the *Xenopus* oocyte and mammalian

[7]**These authors contributed equally to this work.**
[8]**Corresponding author.**
E-MAIL albrecht.bindereif@chemie.bio.uni-giessen.de; FAX 49-641-9935 419.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.526109.

cells: In the cytoplasmic phase, the U1, U2, U4, and U5 snRNAs assemble to Sm cores and are trimethylated in their cap structure, a requirement for their subsequent nuclear reimport (for review, see Will and Lührmann 2001); in contrast, the U6 snRNA carries a γ-monomethyl cap structure and does not leave the nucleus for its biogenesis, including its association with the LSm2-8 proteins (for review, see Khusial et al. 2005). Progress over the last 10 years has focused on the detailed assembly mechanism of the classical Sm core: Three multicomponent protein complexes were identified that assist in the Sm site-specific assembly of Sm cores on the spliceosomal snRNAs (for review, see Meister et al. 2002; Paushkin et al. 2002; Yong et al. 2004; Eggert et al. 2006).

In the center of this machinery there is the SMN (survival of motor neurons) complex, consisting of the SMN protein and at least seven associated proteins, called Gemins 2-8 (see Otter et al. 2007, and references therein). Importantly, mutations in the human SMN gene cause a common neuromuscular disease, spinal muscular atrophy (SMA), and defective snRNP biogenesis (Lefebvre et al. 1995; Winkler et al. 2005, and references therein). In mammalian cells, SMN localizes primarily to the cytoplasm (Liu et al. 1997), but there is also a minor fraction of SMN concentrated in nuclear foci called gems, often associated with Cajal bodies and of unclear significance (Liu and Dreyfuss 1996).

Before mammalian Sm proteins integrate into the ring structure and associate tightly with the Sm site on the snRNA, SmB, D1, and D3 undergo a characteristic modification: symmetric arginine dimethylation (sDMA) at multiple RG dipeptides within their C-terminal tail regions. This takes place in the cytoplasm and creates an epitope recognized by the monoclonal Y12 antibody (Brahms et al. 2000, 2001). sDMA modification strongly increases the affinity of Sm proteins to SMN (Brahms et al. 2001; Friesen et al. 2001a) and is catalyzed by the so-called methylosome; this 20S complex contains the methyltransferase PRMT5, the WD40 protein MEP50, both of which bind subsets of several Sm proteins, as well as the pICln protein (Friesen et al. 2001b, 2002; for review, see Yong et al. 2004). pICln also exists in a smaller, snRNA-free complex of 6S, in which it interacts with the Sm domains of a subset of Sm proteins (Friesen et al. 2001b), specifically the SmD1D2-EFG pentamer (Chari et al. 2008). The current model implies that the Sm proteins are handed over from this latter complex, which is not competent to bind snRNA, to the SMN complex, which mediates closure of the Sm ring over the Sm site on the snRNA (Chari et al. 2008).

In sum, Sm proteins can interact with multiple components of the snRNP assembly machinery, either through their Sm domains or their C-terminal tails carrying the sDMA-modified RG residues with (1) pICln in the 6S complex via Sm domains (Pu et al. 1999; Friesen et al. 2001b); (2) Gemins 6/7 of the SMN complex that themselves contain an Sm fold structure (Ma et al. 2005); or (3) the SMN protein itself, which contacts—through its internal Tudor domain—the methylated RG dipeptides in the C-terminal tails of SmD1, SmB, and SmD3 (Selenko et al. 2001; Sprangers et al. 2003). (4) In addition, as proposed by those studies, the Tudor domain of SMN resembles the classical Sm fold and may interact directly with the Sm core domain. For each of these interactions, certain specificities or preferences for different subsets of Sm proteins have been characterized. However, it has remained unclear in the mammalian system how these overlapping specificities integrate into an organized pathway of Sm core assembly.

Recent studies have clearly demonstrated that the SMN complex is crucial in facilitating specific Sm core assembly, directing it to authentic Sm sites in snRNAs and discriminating against nonspecific RNAs (Pellizzoni et al. 2002). However, many questions remained: What is the biochemical basis of specificity for the Sm sites? How can SMN interact with all seven Sm subunits in the last stage of Sm core assembly? In sum, in the mammalian system, it is not only the multitude of factors involved, but also post-translational protein modifications that pose additional difficulties in a detailed biochemical dissection of the Sm core assembly pathway.

Here we report the identification of the trypanosomal SMN protein, which is highly divergent in sequence and domain structure from other known orthologs. The identification of SMN was based on a mass-spectrometric protein analysis of SmB-associated components of the splicing machinery of *Trypanosoma brucei*, which in addition yielded >30 components of the trypanosomal spliceosome. In contrast to its mammalian counterpart, the trypanosome SMN is nuclear-localized and specifically and stably interacts with the SmD3B subcomplex. In vitro reconstitution assays revealed that SMN by itself is sufficient to confer specificity in Sm core assembly; i.e., to discriminate against binding of the canonical Sm core to nonspecific RNA and to snRNAs, which assemble specific variant cores not containing the SmD3B heterodimer. SMN stays associated throughout the assembly of the Sm heteroheptamer and dissociates only when a functional Sm site RNA is incorporated. In sum, these data establish a novel role of SMN, mediating snRNP specificity in Sm core assembly, and yield new biochemical insight into the mechanism of SMN activity.

Results

Identification of spliceosomal protein components in T. brucei*: TAP tag purification of SmB, SMN, and Gemin2*

Spliceosomal snRNPs in trypanosomes contain either the canonical heptameric Sm ring, as in the U1, U5, and spliced leader (SL) snRNPs, or variant Sm cores, where single canonical Sm subunits are exchanged by snRNA-specific components, as in the U2 and U4 snRNPs. To search for specificity factors, we applied tandem affinity purification (TAP) (Schimanski et al. 2005), based on C-terminally PTP-tagged canonical SmB. Proteins copurifying with SmB (Fig. 1A) were identified by mass spectrometry (MS) and are listed in Table 1: Twenty-four factors can be grouped into proteins that are bona fide

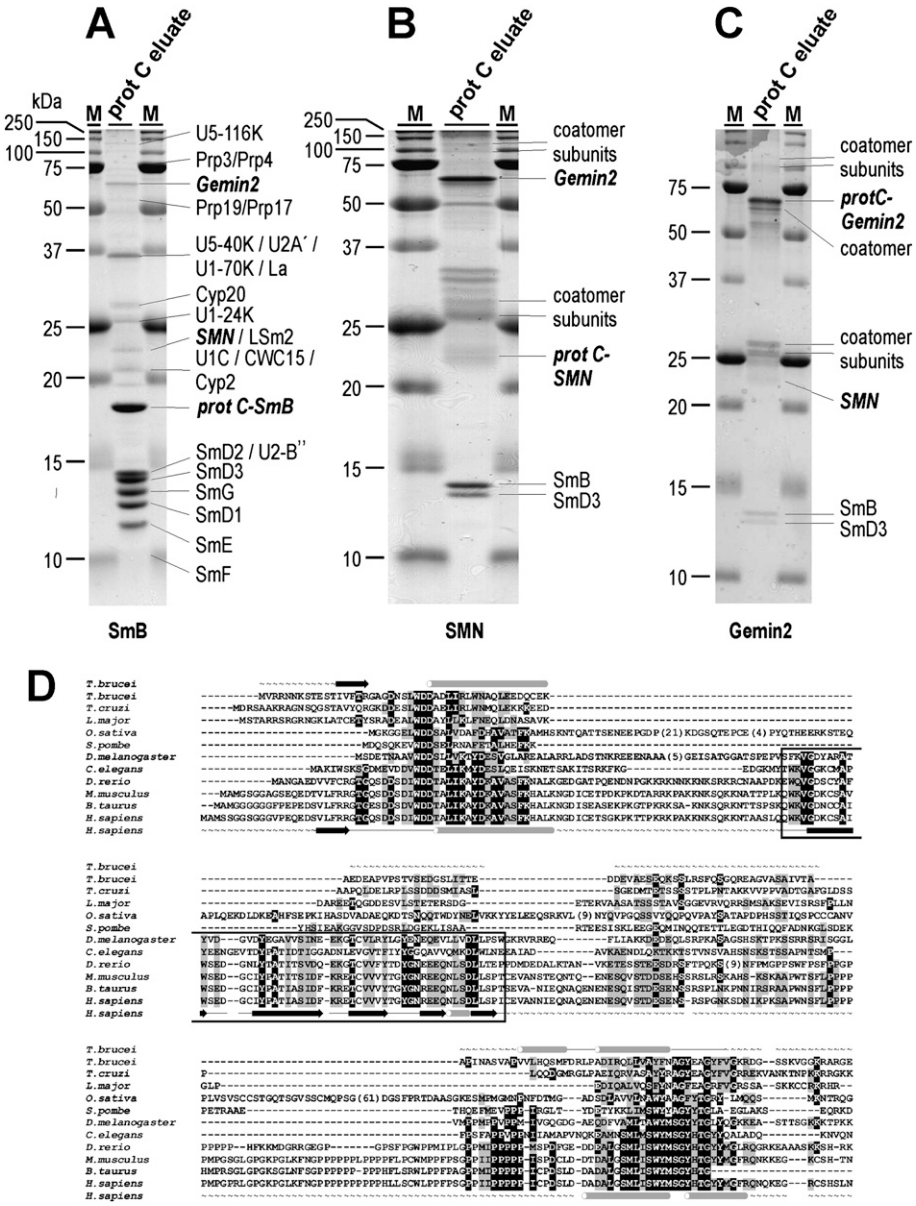

Figure 1. TAP tag purification of SmB, SMN, and Gemin2 complexes: SMN protein identification and multiple alignment. (*A–C*) TAP tag purifications of SmB, SMN, and Gemin2 complexes: Extracts were prepared from *T. brucei* cell lines that stably express tagged proteins SmB (*A*), SMN (*B*), or Gemin2 (*C*), and used to affinity-purify tagged complexes. Protein analysis by SDS-PAGE and Coomassie staining is shown only for the second purification step (lanes "prot C eluate"). Some identities of purified proteins are indicated (for a more complete list, see Table 1). (M) Protein marker (in kilodaltons). (*D*) Multiple alignment of SMN protein sequences. SMN sequences are denoted by their abbreviated genus and species names (see below). Residues that are conserved or whose physicochemical character is conserved in >40% of sequences are highlighted in black and gray, respectively. The secondary structures of SMN from *T. brucei* and *H. sapiens* are shown above and below the alignment, respectively. α-Helices are represented as tubes, and β-strands as arrows, while disordered regions are shown as "∼." The number of omitted residues is indicated in parentheses. The Tudor domains of the metazoan sequences are boxed. SMN from the following species are listed (NCBI gene identification [GI] numbers are in parentheses after the species name; for the three trypanosomatid species, also their GeneDB numbers): (*T. brucei*) *Trypanosoma brucei* TREU927 (74025984; GeneDB: Tb11.01.6640); (*T. cruzi*) *Trypanosoma cruzi* strain CL Brener (71414378; GeneDB:00.1047 05350 8445.40); (*L. major*) *Leishmania major* (157874007; GeneDB: F32.1802); (*O. sativa*) *Oryza sativa* (115448075); (*S. pombe*) *Schizosaccharomyces pombe* (19114000); (*D. melanogaster*) *Drosophila melanogaster* (17647939); (*C. elegans*) *Caenorhabditis elegans* (71981679); (*D. rerio*) *Danio rerio* (18859385); (*M. musculus*) *Mus musculus* (74194985); (*B. taurus*) *Bos taurus* (27462896); (*H. sapiens*) *Homo sapiens* (1737214).

Table 1. *Summary of SmB-, SMN-, and Gemin2-associated proteins*

Factor group/name	Mol. mass (kDa)	GeneDB #	SmB	SMN	Gemin2	Reference
Sm proteins						
SmB	12.3	Tb927.2.4540	+	+	+	Palfi et al. 2000
SmD3	12.3	Tb927.4.890	+	+	+	Z. Palfi, unpubl.
SmD1	11.7	Tb927.7.3120	+			Palfi et al. 2000
SmD2	12.5	Tb927.2.5850	+			Palfi et al. 2000
SmE	9.6	Tb927.6.2700	+			Palfi et al. 2000
SmF	8.3	Tb09.211.1695	+			Palfi et al. 2000
SmG	8.9	Tb11.01.5915	+			Palfi et al. 2000
LSm proteins (U6-specific)						
LSm2	13.2	Tb927.8.5180	+			Tkacz et al. 2008
LSm4	14.1	Tb11.01.5535	+			Liu et al. 2004
LSm7	10.2	Tb927.5.4030	+			Liu et al. 2004
LSm8	14.0	Tb927.3.1780	+			Liu et al. 2004
U1 snRNP						
U1-70K	31.7	Tb927.8.4830	+			Palfi et al. 2005
U1C	21.7	Tb10.70.5640	+			Palfi et al. 2005
U1-24K	24.1	Tb927.3.1090	+			Palfi et al. 2005
U2 snRNP						
U2 A′ (U2-40K)	36.4	Tb10.61.2410	+			Cross et al. 1993
U2 B″	13.6	*Tb927.3.3480	+			Preußer et al. 2009
Sm16.5K (also called Ssm2-2)	14.7	Tb10.70.2250	+			Wang et al. 2006; Tkacz et al. 2007
Sm15K (also called Ssm2-1)	12.7	Tb927.6.4340	+			Wang et al. 2006; Tkacz et al. 2007
U4/U6 snRNP						
PRP3 (90K)	63.2	*Tb09.160.2900	+			
PRP4 (60K)	65.4	*Tb10.70.7190	+			
Ssm4/LSm2[a] (U4-spec.)	23.2	Tb927.7.6380	+			Tkacz et al. 2007; N. Jaé and A. Bindereif, in prep.
U5 snRNP						
PRP8	277.0	Tb09.211.2420	+		+	Lücke et al. 1997
116K	105.4	*Tb11.01.7080	+			
40K WD repeat protein	34.9	*Tb11.01.2940	+			
SMN	17.0	*Tb11.01.6640	+	+	+	
Gemin2	55.4	*Tb10.70.1350	+	+	+	
Prp19/CDC5/CWC						
CWC21	16.1	*Tb09.160.2110	+			
CWC22	66.8	*Tb11.01.2520	+			
CWC15	22.3	*Tb10.389.1830	+			
PRP19	54.2	*Tb927.2.5240	+			
Coatomer (α)	132.0	Tb927.4.450		+	+	Maier et al. 2001
Coatomer (β)	110.0	Tb927.1.2570		+	+	Maier et al. 2001
Coatomer (γ)	97.5	Tb11.01.3740		+		Maier et al. 2001
Coatomer (δ)	57.3	Tb927.8.5250		+	+	Maier et al. 2001
Coatomer (ζ)	20.5	Tb10.70.2980		+		Maier et al. 2001
Coatomer (ϵ)	34.8	Tb11.01.6530		+	+	Maier et al. 2001
β′ COP protein	93.9	Tb927.2.6050		+	+	Maier et al. 2001
Poly(A)-binding protein 1	62.1	Tb09.211.2150	+			
PRP17	52.8	*Tb927.3.1930	+			
Importin α subunit	58.0	Tb927.6.2640	+			
Cyclophilin (cyp2)	21.6	Tb927.8.2090	+			
Cyclophilin (cyp-20)	27.1	*Tb927.8.6280	+			
HSP70	75.3	Tb11.01.3110	+	+		
La protein	37.6	Tb10.70.5360	+	+		
Translation elongation factor eEF-1 α	49.1	Tb10.70.5670		+		

T. brucei proteins were identified by MS after separation by SDS-PAGE. Each protein is described by name, its molecular mass (in kilodaltons), by the systematic GeneDB name (http://www.genedb.org/genedb/tryp/), and if applicable, by literature reference. An asterisk before the GeneDB number indicates that this protein is annotated in GeneDB only as a hypothetical protein; therefore, note that some of our assignments based on sequence similarity may be putative. Proteins were grouped into Sm proteins; LSm proteins; U1, U2, U4/U6, and U5 snRNP proteins; and others. In addition, the table lists in which of the affinity-purified complexes (SmB, SMN, or Gemin2) each protein is represented by a significant number of peptides (+).

[a]Initially identified as U6-specific LSm2 (Liu et al. 2004), later renamed "Ssm4" by Tkacz et al. (2007) and independently confirmed as a U4-specific Sm protein (Tkacz et al. 2007; N. Jaé and A. Bindereif, in prep.).

snRNP proteins, including the complete set of seven canonical Sm proteins, four LSm proteins, the three known U1-specific proteins, four U2-specific proteins, three U4-specific proteins, and three U5-specific proteins (see Table 1 for references).

In addition, we identified several proteins likely to represent snRNP proteins or spliceosome-associated factors. When systematically searching with these MS-derived candidates for homologous proteins, we discovered that one of the hypothetical proteins, Tb11.01.6640 (predicted molecular mass of 17.0 kDa), showed a very low level of sequence similarity to SMN from other species (e.g., 12% identity to human SMN). Similarly, we discovered among the SmB-selected proteins a putative Gemin2 homolog (Tb10.70.1350; predicted molecular mass of 55.4 kDa). In addition to SMN and Gemin2 candidates, this group of splicing-related factors includes four components of the Prp19/CDC5 complex. As studied primarily in yeast and mammals, there are at least seven proteins associated with Prp19 in this complex, which is recruited to the spliceosome before its activation (Deckert et al. 2006, and references therein). We also found PRP17, two cyclophilins (Cyp2 and Cyp20), and La. Finally, our proteomics approach revealed 14 proteins that are listed in the *T. brucei* genome database (http://www.GeneDB.org) as hypothetical proteins and that exhibited no clear sequence similarity to known splicing factors (data not shown).

Based on extensive work primarily in the human system, SMN—as part of a large protein complex, which includes at least seven Gemin proteins (Gemin2-8)—plays a crucial role in cytoplasmic Sm core assembly (see above for references). Therefore, it may also participate in snRNP biogenesis in trypanosomes, possibly specifying Sm core variation. To obtain more evidence for this novel trypanosome protein being a functional counterpart of SMN, we performed a second round of affinity purification, using PTP-tagged SMN. A *T. brucei* cell line was created that stably expresses a PTP-tagged version of the putative SMN protein. A two-step affinity purification and MS protein identification resulted in a limited set of polypeptides (Fig. 1B; protein identities summarized in Table 1). As the most prominent proteins copurifying with SMN in trypanosomes, we found the following: first, SmB and SmD3, but not the other core Sm proteins; second, all seven coatomer (COP1) subunits (α, β, γ, δ, ϵ, ζ, and β′ COP protein) (see Discussion); third, the same putative Gemin2 homolog that had appeared in the SmB list (see above), in addition to the tagged SMN protein itself (prot C-SMN). We searched—without success—among the SMN- and Gemin2-copurifying proteins for potential homologs of other Gemins (Gemin3 through Gemin8), suggesting that in trypanosomes, there may be only a simple core complex of SMN–Gemin2.

To obtain further confirmation on the putative Gemin2 ortholog, a third cell line was generated, stably expressing PTP-tagged Gemin2, and TAP and MS protein analysis were carried out again (Fig. 1C; Table 1). As observed for SMN, SmD3B and five out of the seven known coatomer (COP1) subunits (α, β, δ, ϵ, and β′ COP protein) represented the most prominent bands; in addition, SMN was present as a minor band, confirming the copurification of Gemin2 through TAP-tagged SMN and vice versa. In sum, these results strongly indicate that one or several protein complexes exist containing SMN, Gemin2, and SmD3B, and thereby strongly support the identities of the trypanosome SMN and Gemin2 proteins.

Comparing the three data sets, we found only small overlaps of proteins (Table 1): SmB clearly coselected most spliceosomal proteins, most likely as snRNP and/or larger spliceosomal complexes. In contrast, both SMN and Gemin2 complexes contain only the SmD3B heterodimer out of the Sm protein family, a surprising specificity, since human SMN is known to interact with all Sm proteins (see the Discussion).

Trypanosome SMN and Gemin2 are highly divergent from known orthologs

Figure 1D shows a multiple sequence alignment of the human SMN (294 amino acids), the structurally best-characterized ortholog, with SMN sequences from the three trypanosomatid species *T. brucei*, *Trypanosoma cruzi*, and *Leishmania major*. Clearly the trypanosome SMN proteins group together with *Schizosaccharomyces pombe* (152 amino acids) (Hannus et al. 2000; Owen et al. 2000; Paushkin et al. 2000)—the shortest known SMN proteins (153–159 amino acids). Although the overall sequence similarity is very low, the multiple alignment reveals two regions of significant local similarity: first, the N-terminal region, through which the human protein interacts with Gemin2, corresponding to amino acids 1–51 in the *T. brucei* SMN; second, the C-terminal region (amino acids 91–157), containing two GY dipeptides (in *L. major* also GF) that are implicated as the so-called YG-box in SMN oligomerization (Pellizzoni et al. 1999). Surprisingly, the central Tudor domain appears to be conserved only in animal metazoan SMN orthologs and is absent from sequences from other species, including trypanosomes, fungi, and plants. The few nonmetazoan SMN-like proteins that do include the Tudor domain are limited to *Alveolata*; e.g., *Plasmodium falciparum* (Hossain et al. 2008). Since they have diverged significantly from both trypanosomal and human SMNs, we omitted them in our alignment.

Structure prediction of the metazoan SMN sequences reveals that the N- and C- terminal conserved regions have helical conformation, while the extended stretches of insertions between those two regions and the Tudor domain are likely to be intrinsically disordered. On the other hand, in the nonmetazoan SMN orthologs (including the trypanosome proteins), the entire central part is predicted to be disordered, without any trace of stable structure in the region corresponding to the metazoan Tudor domain. Thus, trypanosomal proteins appear to represent minimal versions of SMN, devoid of the central Tudor domain. A phylogenetic tree calculated from the N- and C-terminal regions, which are conserved among all SMN orthologs, confirmed the division into Tudor-containing and Tudor-missing variants (Supplemental Fig. 1).

Finally, the putative Gemin2 ortholog from *T. brucei*, which we identified here and which had not been annotated

previously in GeneDB, shows only very limited sequence similarity to Gemin2 proteins from other species and is much larger than its mammalian counterpart (495 vs. 280 amino acids) (see Supplemental Fig. 2). To our knowledge, the trypanosomatid Gemin2 proteins are the largest Gemin2 orthologs known. Comparison of the putative trypanosomatid Gemin2 homologs with a large collection of protein families (Pfam) using a rigorous method for the detection of remote homology (HHsearch) revealed a close relationship of their C termini (residues 195–495 in the *T. brucei* Gemin2) to the PF04938 family (SMN-interacting protein 1 [SIP1] family [Gemin2]), as represented by statistically significant scores: probability 97.9, *E*-value 0.00049, and *P*-value 5.3e-09. No functional relevance of this region is documented in the literature. To confirm this assignment, we demonstrated both in vitro and in vivo the direct interaction of Gemin2 and SMN (Supplemental Fig. 3A,B).

Trypanosome SMN protein is essential and predominantly nuclear

The expression of SMN in procyclic *T. brucei* cells was silenced by doxycycline-inducible RNAi (Wang et al. 2000). The stably transfected, clonal SMN-RNAi cell line showed normal growth, when RNAi was not induced (Fig. 2A, −Dox). However, after doxycycline-induced SMN silencing, cell growth was rapidly affected (+Dox). Growth inhibition was detected already 1–2 d after induction, and the cell number dropped significantly until day 6, indicating that SMN is essential for cellular functions and growth. Both real-time and semiquantitative RT–PCR confirmed that the level of SMN mRNA was specifically reduced already 1 d after RNAi induction to levels ~30% (Fig. 2B).

To investigate whether SMN is required for splicing in vivo, the abundance of α-tubulin mRNA and α-tubulin

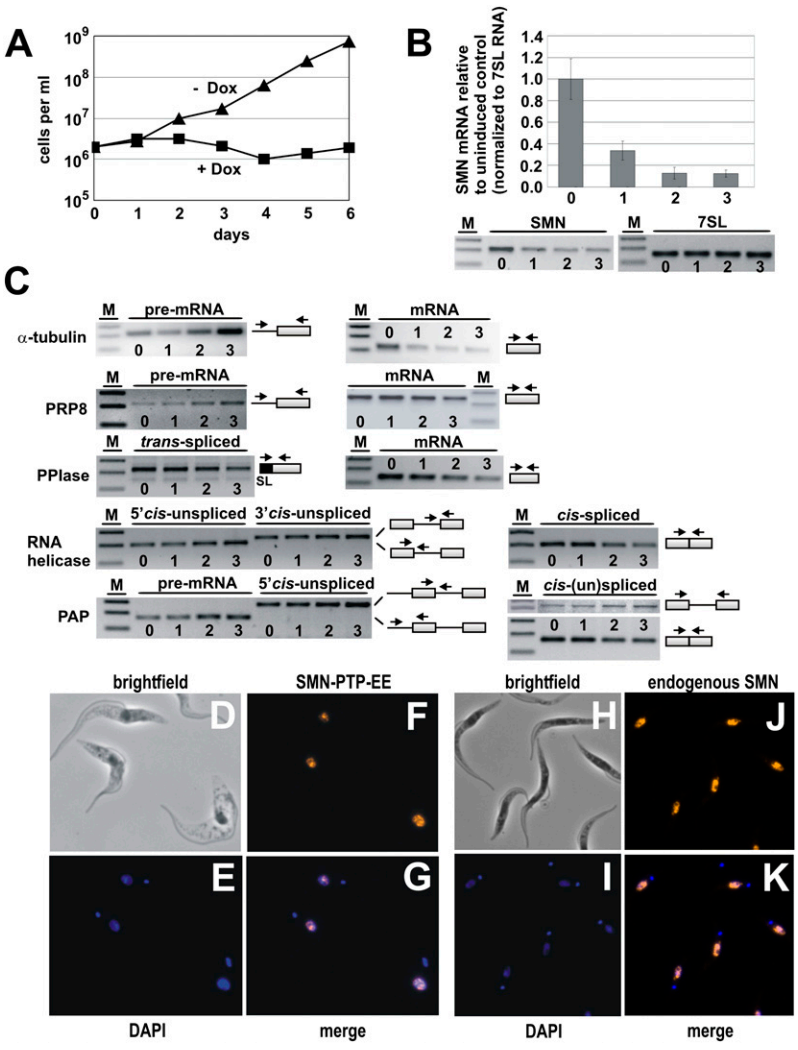

Figure 2. SMN is splicing-essential and predominantly nuclear in trypanosomes. (*A*) SMN is essential for growth. Growth curve of a representative clonal procyclic *T. brucei* cell line, in which RNAi was induced by the addition of doxycycline. Cells were grown in the absence (−Dox; line with triangles) or in the presence of doxycycline (+Dox; line with squares). Cells were counted every day and diluted back to 2×10^6 cells/mL. (*B*) RNAi down-regulation of SMN expression. SMN mRNA was measured by real-time (*top*) or semiquantitative RT–PCR (*bottom*) 1, 2, or 3 d after RNAi induction (as indicated) or from uninduced cells (time point 0). As a control, 7SL RNA was analyzed from the same RNA samples. (M) Markers (100, 200, and 300 bp). (*C*) Inhibition of *cis*- and *trans*-splicing by SMN down-regulation. The following *T. brucei* genes (GeneDB numbers in parentheses) were analyzed by RT–PCR for splicing defects: α-tubulin (Tb927.1.2400), *PRP8* (Tb09.211.2420), a putative RNA helicase (Tb927.8.1510), *PAP* [poly(A) polymerase; Tb927.3.3160], *PPIase* (peptidyl-prolyl *cis-trans* isomerase; Tb927.8.690). To detect pre-mRNA and mRNA, gene-specific primers (mRNA) or combinations of exon-, intergenic-region-, intron-, or SL-specific primers were used, and RNA samples corresponding to those described for *B*. Primer positions and orientations are schematically represented on the side. *Cis*-spliced introns were assayed for the RNA helicase and *PAP* genes: For *PAP*, both spliced and unspliced intron could be detected by the same primer pair (longer exposure shown for the *top* panel); for the RNA helicase, the unspliced intron was too long to be detected by the two exon primers. (M) Markers (100, 200, and 300 bp; for panel *PAP*, *cis*-unspliced: 800 and 900 bp). (*D–K*) Localization of SMN in *T. brucei*. Cells expressing exclusively PTP-tagged SMN were fixed (*D*, bright field) and stained with DAPI (*E*). (*F*) In parallel, SMN-PTP was detected by the anti-protA primary and the Alexa 594 secondary antibody. (*G*) Merge of DAPI and SMN-PTP staining. (*H–K*) In addition, the endogenous SMN protein was localized by a polyclonal antibody specific for the *T. brucei* SMN protein.

pre-mRNA was determined by semiquantitative RT–PCR (Fig. 2C). Clearly, the α-tubulin mRNA signal, which represents mostly spliced mRNA, was strongly reduced upon SMN RNAi induction. At the same time, RT–PCR signals for the α-tubulin pre-mRNA strongly increased, reflecting the accumulation of unspliced α-tubulin pre-mRNA. In addition, splicing of four other genes was assayed during the time course of SMN knockdown: *PRP8* (testing pre-mRNA and total mRNA); *PPIase* (*trans*-spliced mRNA and total mRNA); an RNA helicase and *PAP* (for those two genes testing *cis*-splicing of their internal intron). In each case, we observed splicing defects after SMN knockdown; splicing defects differed in their severity, most likely reflecting differences in RNA stability and expression levels. For two cases (*PRP8* pre-mRNA and RNA helicase *cis*-spliced mRNA), these effects were also confirmed by real-time RT–PCR assays (data not shown). In sum, these assays indicate that SMN is essential for in vivo splicing activity in trypanosomes. We note that a recent genome-wide study in the mammalian system identified differential effects of SMN deficiency on splicing of individual pre-mRNAs (Zhang et al. 2008).

Finally, since SMN in metazoan organisms is compartmentalized between the cytoplasm and nuclear gems, we examined its cellular localization in *T. brucei*. Cells expressing exclusively PTP-tagged SMN were subjected to indirect immunofluorescence, using the protein A epitope of the PTP tag (Fig. 2D–G). In addition, to rule out potential artifacts due to the PTP tag, we localized the endogenous SMN in wild-type *T. brucei* cells, using our polyclonal anti-SMN antibody (Fig. 2H–K). Based on both assays, we conclude that—in contrast to mammalian cells—SMN occurs in trypanosomes predominantly in the nucleoplasm.

Trypanosome SMN stably interacts with SmB

Since mass-spectrometric analyses of SMN-associated proteins had yielded SmD3B as a major interaction partner in trypanosomes, we wanted to confirm this in vitro (Fig. 3A). Each of the three canonical Sm subcomplexes (His-SmD1D2, His-Flag-SmD3B, His-SmEFG) and the U2-specific, His-tagged Sm16.5K/15K heterodimer were incubated with GST-SMN, or as a control, with GST alone. After GST pull-down, copurified proteins were analyzed by Coomassie staining. Clearly, only the SmD3B subcomplex efficiently interacted with GST-SMN. The same result was obtained with His-SmD3B (see Fig. 3B,

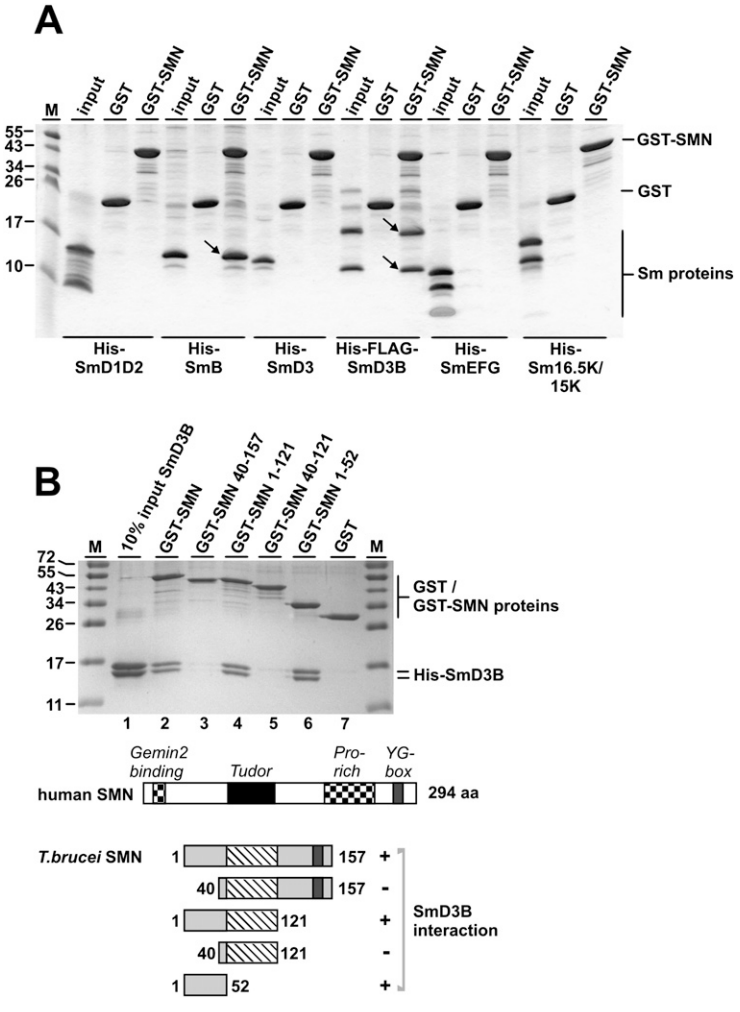

Figure 3. The N-terminal region of trypanosome SMN specifically interacts with subcomplex SmD3B. (*A*) GST-*T. brucei* SMN, or GST alone as a control (as indicated *above* the lanes) was immobilized on glutathione-Sepharose and incubated with Sm subcomplexes His-SmD1D2, His-Flag-SmD3B, His-SmEFG, His-Sm16.5K/15K, or with His-tagged Sm proteins His-SmB or His-SmD3 (see *below* the lanes). For comparison, 10% of each input is shown (lanes "input"). Bound protein was analyzed by Tricine-SDS-PAGE and Coomassie staining. The arrows point to His-Flag-SmD3B and His-SmB copurifying with GST-SMN. (M) Protein marker (in kilodaltons). (*B*) GST-*T. brucei* SMN protein or deletion derivatives (as indicated *above* lanes *2–6*), or GST alone as a control (lane *7*) were bound to glutathione-Sepharose and incubated with SmD3B (10% of input shown in lane *1*). Bound protein was analyzed by SDS-PAGE and Coomassie staining. (M) Protein markers (in kilodaltons). (*Below*) The domain structure of human SMN and the *T. brucei* SMN derivatives are schematically represented (with their activities in SmD3B interaction summarized [+/−]).

lane 2). This interaction is highly specific, since none of the other subcomplexes nor GST resulted in significant interactions.

Does SMN interact directly with SmB or SmD3? To answer this question, we tested in the same manner the separate His-tagged SmB and SmD3 proteins for SMN interaction (Fig. 3A). This revealed that SMN contacts specifically and directly SmB, not SmD3. The efficiency of interaction with the single SmB subunit was comparable with that with the SmD3B subcomplex.

We conclude that the trypanosome SMN protein stably and specifically interacts with the SmD3B heterodimer, directly contacting SmB. What region in the trypanosomal SMN protein mediates the SmB interaction? In the human system, structural and mutational evidence points to both the central Tudor and the C-terminal regions, perhaps involving a conformational switch between these two regions (see above for references); in contrast, the trypanosome SMN appears to lack a Tudor domain (see above). Therefore, we next tested by GST pull-down assays the following derivatives of the *T. brucei* SMN for interaction with the recombinant SmD3B heterodimer (Fig. 3B): GST-SMN containing the full-length *T. brucei* protein (amino acids 1–157); GST-SMN 40–157, GST-SMN 1–121, GST-SMN 40–121, and GST-SMN 1–52; GST alone served as a control. As a result, only full-length SMN, SMN 1–121, and SMN 1–52 were positive in SmD3B interaction; the two others with the N-terminal part deleted were negative. We conclude that the N-terminal 52 amino acids of the trypanosome SMN are necessary and sufficient to interact with SmD3B.

Trypanosome SMN protein confers specificity to canonical Sm core assembly in vitro: evidence for a minimal SMN system

To assay for activity of the trypanosome SMN protein, we reconstituted canonical Sm cores in vitro: The three canonical Sm subcomplexes—SmD1D2, SmEFG, and SmD3B, each of them His-tagged in one component—were incubated with total RNA from *T. brucei*, followed by His tag pull-down and analysis of bound RNAs by either Northern blot hybridization for U2, SL, and U5 snRNAs (Fig. 4A,B), or by silver staining to detect all RNAs (Fig. 4C). The reconstitution of canonical Sm cores required that all three Sm subcomplexes were present during the incubation with total RNA (Supplemental Fig. 4). Reconstitutions were done either in the absence of SMN (mock reactions) (Fig 4A–C, lanes 2) or in the presence of increasing amounts of SMN protein (100, 200, and 400 pmol) (Fig 4A–C, lanes 3–5; schematic outline of in vitro Sm core assembly given below; Fig. 4A,B). Initially we compared two reconstitution protocols: adding SMN together with the Sm proteins and total RNA (Fig. 4A), or preincubating SMN and Sm proteins before the addition of total RNA (Fig. 4B). Without SMN/Sm preincubation, SMN did not change significantly the pattern of Sm-bound RNAs (Fig. 4A, lanes 2–5). In contrast, when SMN was preincubated with Sm proteins before RNA addition, the "illegitimate" assembly of U2 snRNA with the canonical Sm core was strongly reduced (Fig. 4B, lanes 2–5); Sm core assembly on SL RNA and U5 snRNA, on the other hand, was not altered significantly. SMN by itself does not bind RNA (Supplemental Fig. 4, lane 2). In sum, this experiment gave a first indication that the trypanosome SMN by itself is able to suppress—within the complex mixture of total RNA–canonical Sm core assembly on U2 snRNA. Initial in vitro assembly assays had also been performed with both recombinant SMN and Gemin2 proteins combined, as well as with both factors alone; surprisingly, the same effect was found, whether the SMN–Gemin2 combination or SMN alone was used; in contrast, we observed no significant effect with Gemin2 alone (data not shown). Therefore, we focused our detailed functional study on SMN, which appears to be sufficient for conferring specificity to canonical Sm core assembly.

Next, we repeated these in vitro assembly reactions, preincubating SMN with canonical Sm subcomplexes prior to the addition of total RNA, but detecting RNA by silver staining (Fig. 4C). After reconstitution in the absence of SMN, some of the abundant RNAs, in particular high-molecular-weight RNA, sRNA1 (one of the fragmented ribosomal RNAs), and 5.8S RNA, were coprecipitated by Sm subcomplexes, reflecting nonspecific RNA binding of Sm cores (Fig. 4C, lanes 1,2). Adding increasing amounts of SMN strongly reduced this nonspecific RNA binding, in a dose-dependent manner (Fig. 4C, lanes 3–5), confirming the activity of SMN to suppress nonspecific RNA binding of canonical Sm cores.

To characterize this activity of SMN in more detail, we then did reconstitutions with the canonical Sm subcomplexes (Fig. 4D), using the following ^{32}P-labeled RNA transcripts: full-length *T. brucei* U2 snRNA (U2 WT); full-length *T. brucei* U5 snRNA (U5 WT); full-length Sm-mutant U5 snRNA (U5 mut, Sm site ACUUUG changed to ACAAAG); U4 snRNA 3′ half including the Sm site (U4 WT, nucleotides 69–110); U4 snRNA 3′ half with mutant Sm site (U4 mut, AGUUUG changed to AGAAAG); and an unrelated RNA of 101 nucleotides (nt) derived from the human *SLC2A2* gene (control). Note that, in vivo, only the U5 snRNP contains the canonical Sm core, both U2 and U4 snRNPs carry their snRNA-specific Sm cores without the canonical SmD3B unit.

Reconstitution reactions were carried out with two different mixtures of ^{32}P-labeled RNA transcripts: control, U5 WT, U4 WT (Fig. 4D, input lane 1); U2 WT, U5 mut, U4 mut (Fig. 4D, input lane 5), each reaction in the absence and presence of recombinant SMN (Fig. 4D, cf. lanes 2 and 3, and lanes 6 and 7). A control reaction included only His-tagged SMN (Fig. 4D, lanes 4,8). The following effects of SMN addition were observed: The control RNA associated to a low extent with the canonical Sm core; however, this was completely abolished by SMN. Both U2 WT and U4 WT snRNAs (which in vivo are present in noncanonical Sm cores) efficiently assembled with the canonical Sm proteins, but SMN addition reduced this to insignificant levels (Fig. 4D, cf. lanes 2 and 3, and lanes 6 and 7). In contrast, binding of canonical Sm proteins to the U5 WT snRNA was not altered by

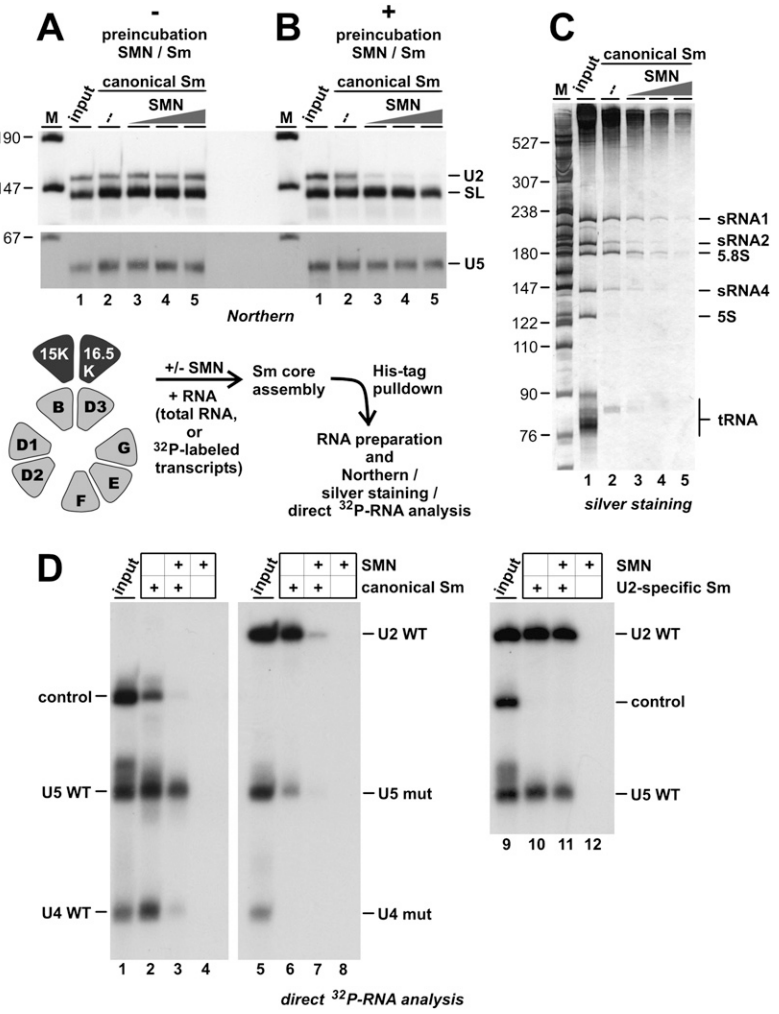

Figure 4. Trypanosome SMN specifies canonical Sm core assembly in vitro. (*A*,*B*) Canonical Sm core assembly with total RNA from *T. brucei*. His-tagged subcomplexes of the canonical Sm core (SmD1D2, SmEFG, and SmD3B) were incubated with total *T. brucei* RNA, in the absence of SMN (lanes *2*) or in the presence of increasing amounts of SMN protein (100, 200, and 400 pmol; lanes *3–5*). (Lanes *1*) In addition, 20% of RNA input is shown. (*A*) Sm subcomplexes and SMN were incubated with total RNA simultaneously (*A*), or the Sm subcomplexes were first preincubated with SMN, before addition of total RNA (*B*). The experimental procedure is schematically shown *below*. RNA assembled in vitro into Sm cores was recovered by His tag pull-down and analyzed by denaturing PAGE and Northern blotting with snRNA-specific probes, detecting U2, SL, and U5 snRNAs (shown from two different exposures). RNA positions are marked on the *right*. (M) Markers (in nucleotides). (*C*) Canonical Sm core assembly with total RNA from *T. brucei*, carried out as in *B* (with preincubation of SMN and Sm subcomplexes), and detecting Sm-bound RNAs by silver staining. Positions of major RNAs (ribosomal RNAs and tRNA) are marked on the *right*. (M) Markers (in nucleotides). (*D*) His-tagged subcomplexes of the canonical Sm core (His-SmD1D2, His-SmEFG, His-SmD3B; *left* and *middle* panels) or of the U2-specific Sm core (His-SmD1D2, His-SmEFG, His-Sm16.5K/15K; *right* panel) were incubated with ^{32}P-labeled *T. brucei* snRNA transcripts (marked on the sides), in the absence (lanes *2*,*6*,*10*) or presence of SMN (lanes *3*,*7*,*11*). (Lanes *4*,*8*,*12*) As a control, RNA binding to His-tagged SMN alone was tested. (Lanes *1*,*5*,*9*) In addition, 20% of input RNAs is shown. Sm-bound RNAs were recovered by His tag pull-down and detected by direct ^{32}P-RNA analysis. The following ^{32}P-labeled RNA transcripts were used: an unrelated RNA of 101 nt (control), (U2 WT) U2 snRNA; (U4 WT) U4 snRNA; (U4 mut) Sm-mutant U4 snRNA; (U5 WT) U5 snRNA; (U5 mut) Sm-mutant U5 snRNA.

SMN addition (Fig. 4D, lanes 2,3). When testing Sm site mutant derivatives of U4 and U5, we found that association with the canonical Sm core was undetectable (U4 mutant) or very low (U5 mutant) already without SMN (Fig. 4D, lane 6); SMN addition completely abolished the low level seen for U5 mutant snRNA (Fig. 4D, lanes 6,7).

To test whether this striking discrimination activity of SMN operates only for the canonical Sm core, reconstitutions were carried out with the U2-specific Sm subcomplex (His-Sm16.5K/15K) and His-SmD1D2, SmEFG, and a mix of three ^{32}P-labeled RNA transcripts (U2 WT, control, U5 WT) (see Fig. 4D, input lane 9), each reaction in the absence and presence of SMN (Fig. 4D, cf. lanes 10 and 11). A control reaction included only His-tagged SMN (Fig. 4D, lane 12). SMN addition did not affect association of the U2-specific Sm core with U2 snRNA, nor did it significantly change binding to U5 snRNA (Fig. 4D, cf. lanes 10 and 11). The control RNA did not interact with the U2-specific Sm core. In summary, we demonstrated by in vitro Sm core assembly that the trypanosome SMN by itself has strong discriminatory activities: SMN is able to abolish or strongly reduce nonspecific associations of the canonical Sm core with an unrelated RNA, with snRNAs that in vivo do not bind the canonical core (U2 and U4), or with snRNAs containing Sm site mutations. The discrimination activity of SMN is apparently restricted to the Sm core containing the canonical SmD3B subcomplex (see also the Discussion). Additional evidence for this conclusion comes from assembly reactions under protein-competitive conditions with both canonical and U2-specific Sm subcomplexes present, using direct protein analysis (see Supplemental Fig. 5). Another characteristic of the trypanosomal Sm core assembly is that all three canonical subcomplexes—SmD3B, SmD1D2, and SmEFG—can assemble in the absence of RNA (Supplemental Fig. 6A); this contrasts with the mammalian system, where SmD3B and SmEFG in the presence of an Sm site-containing RNA form an Sm subcore that—after SmD3B addition—can be chased into the complete Sm core complex (Raker et al. 1996). On the other hand, stable U2-specific Sm cores

assemble only in the presence of U2 snRNA (Supplemental Fig. 6B–D).

Trypanosome SMN protein functions as an RNA-dependent chaperone in canonical Sm core assembly: a novel SMN/SmD3B intermediate

Since the trypanosome SMN stably interacts with SmD3B, we next asked whether this trimeric complex functions as an intermediate in canonical Sm core assembly. We compared in vitro assembly reactions containing—in addition to SmD1D2 and SmEFG—either the free SmD3B heterodimer or the SMN/SmD3B complex. The trimeric SMN/SmD3B complex was purified by a two-step affinity protocol, based on the GST part on SMN and the His tag on SmD3 (Fig. 5A). Canonical Sm cores were assembled with wild-type U1 snRNA (73 nt) or a U1 mutant RNA, in which the Sm site was inactivated by a triple mutation (ACUUUG to ACAAAG). For these assembly assays, we used Flag pull-down, relying on the single Flag tag on SmD3. For each of the four reconstitutions, an aliquot of the total reaction (Fig. 5B, lanes 5–8), the Flag peptide-eluted material (Fig. 5B, lanes 9–12), and the supernatant (Fig. 5B, lanes 13–16) were analyzed.

Comparing reconstitutions in the absence of SMN revealed that only wild-type U1 RNA, but not Sm-mutant U1 RNA, could be assembled (Fig. 5B, cf. lanes 5 and 6, lanes 9 and 10, and lanes 13 and 14). However, when SMN was added in the form of the SMN/SmD3B complex, reconstitution was slightly more efficient than in the absence of SMN, and strictly depended on an intact Sm site (Fig. 5B, cf. lanes 7 and 8, lanes 11 and 12, and lanes 15 and 16). Importantly, when wild-type U1 RNA was assembled, we detected SMN in the Flag-pull-down material only at very low levels; in contrast, SMN was present at much higher levels when U1 mutant RNA was used (Fig. 5B, cf. lanes 11 and 12). Accordingly, the reverse distribution was found in the supernatant fractions (Fig. 5B, cf. lanes 15 and 16). In conclusion, this provides evidence for a stable complex containing all seven Sm proteins and SMN, which can be detected in the presence of U1 mutant-Sm RNA (Fig. 5B, lane 12). In contrast, in the presence of wild-type U1 RNA with a functional Sm site, SMN dissociates (Fig. 5B, lane 11) and is released into the supernatant (Fig. 5B, lane 15). This provides direct biochemical evidence for a transient chaperone function of SMN, based on its direct interaction with the SmB component.

Discussion

The starting point of this study was a systematic protein analysis in *T. brucei* of factors associated with the SmB protein, based on TAP tag affinity purification and MS.

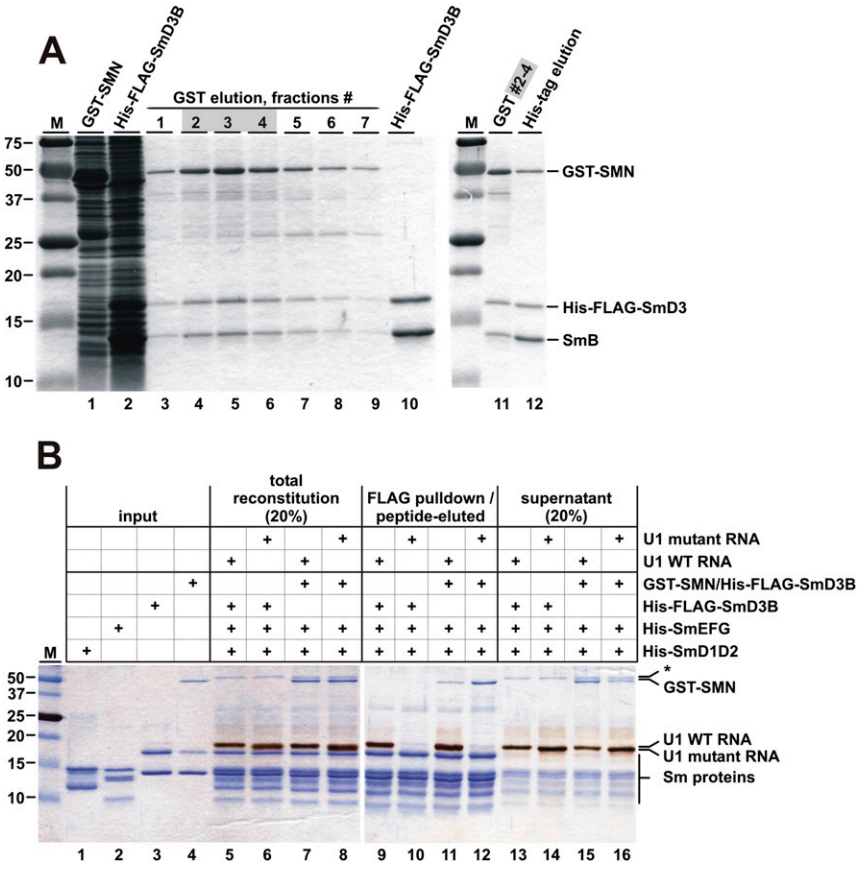

Figure 5. Evidence for chaperone activity of trypanosome SMN during canonical Sm core assembly. (*A*) Two-step purification of SMN/SmD3B complex. Bacterial lysates containing GST-SMN (lane *1*) or His-Flag-tagged SmD3B (lane *2*) were combined (for comparison, see purified His-Flag-SmD3B in lane *10*). GST-SMN/His-Flag-SmD3B complex formed was purified first through its GST portion (see eluted fractions in lanes *3–9*, and pooled fractions #2–#4 in lane *11*), and second, through its His tag on SmD3 (eluted pooled complex in lane *12*). (*B*) Reconstitution of canonical Sm core results in release of SMN, depending on a functional Sm site RNA. Canonical Sm cores were reconstituted from these subcomplexes: His-tagged SmD1D2, His-tagged SmEFG, and His-Flag double-tagged SmD3B (alternatively, purified GST-SMN/His-Flag-SmD3B complex; see input lanes *1–4*). Reconstitutions were carried out in the presence of wild-type or Sm-mutant U1 snRNA. For each of the four reconstitutions (combination of components indicated *above* the lanes), a 20% aliquot of the total reaction (lanes *5–8*), the Flag-pull-down material (lanes *9–12*), and a 20% aliquot of the supernatant (lanes *13–16*) were analyzed for protein and RNA by SDS-PAGE and sequential Coomassie and silver staining. (Note that proteins and RNAs stain differentially; [*] nonspecific protein band; the diffuse bands in the 25- to 35-kDa range are most likely due to partially denatured Sm proteins.)

This approach revealed a large set of proteins, including 24 snRNP protein components and several putative non-snRNP factors.

Unexpectedly, we found among the newly identified proteins SMN and Gemin2 candidates, initially based only on low sequence similarities with the known SMN and Gemin2 protein orthologs. Nevertheless, we are confident of their assignment: First, as known from other species (e.g., Liu et al. 1997), the SMN and Gemin2 candidates interact directly with each other, as demonstrated by GST interaction assays and consistent with that both SMN- or Gemin2-based TAP tag purifications yielded the respective partner protein. Second, SMN interacts specifically and efficiently in vitro with SmB and the SmD3B heterodimer. Third, SMN is splicing-essential, and recombinant SMN was functional in Sm core assembly in vitro (see below). Fourth, both TAP-tagged SMN and Gemin2 coselect the same coatomer subunits, probably reflecting a coatomer interaction with the SMN–Gemin2 complex. The coatomer complex functions in retrograde vesicular transport between Golgi and ER compartments (Maier et al. 2001). We do not know the relevance of this coatomer interaction, but the result raises the interesting possibility that a minor, cytoplasmic fraction of SMN may assist in the assembly of other, non-splicing-related protein complexes.

The trypanosome SMN strongly deviates from the mammalian counterpart in sequence and domain organization, and only two sequence blocks in the N- and C-terminal regions are conserved: In the N terminus, where the human SMN interacts with Gemin2 (Liu et al. 1997), there is a short cluster of a few conserved positions around WDD (amino acids 25–27 in *T. brucei*); in the C terminus, some aromatic positions are conserved (called the GY region in the human SMN and implicated in SMN oligomerization) (Pellizzoni et al. 1999, and references therein). On the other hand, our alignment revealed a striking difference in the central region of SMN orthologs. While the human SMN and its metazoan orthologs possess a well-structured and conserved Tudor domain engaged in Sm protein binding, the corresponding sequence in the trypanosomal proteins and their nonmetazoan orthologs lacks any conservation of Tudor-specific residues and is predicted to be completely devoid of tertiary structure. Thus, bioinformatics predictions indicate with confidence that nonmetazoan SMN orthologs lack the Tudor domain.

The trypanosome SMN also lacks the hydrophobic aromatic residues that in the human SMN (W102, Y109, Y127, Y130) stack on the methyl groups of sDMA groups in the C-terminal tails of the Sm proteins (Selenko et al. 2001). Consistent with this difference, the trypanosome Sm proteins appear to lack DMA: First, there is no long C-terminal extension with RG dipeptides beyond the Sm domain, and second, *T. brucei* Sm proteins do not react with the monoclonal Y12 antibody (Palfi et al. 1991), which recognizes the sDMA epitope (Brahms et al. 2001). Interestingly, Gonsalvez et al. (2008) found that Sm protein methylation is dispensable for snRNP assembly in *Drosophila*.

How does the trypanosome SMN system work biochemically? As we demonstrated here, trypanosome SMN by itself is functional in suppressing illegitimate assembly of the canonical Sm core on RNAs not carrying a bona fide canonical Sm site. Such "wrong" RNAs include unrelated RNAs as well as snRNAs with mutated canonical Sm sites. Our results indicate that in the absence of SMN, most RNAs bind canonical Sm proteins to some extent, but in the presence of SMN, this is reduced to insignificant or undetectable levels. Note that this discrimination activity requires preincubation of SMN with the canonical Sm subcomplexes, suggesting that SMN/SmD3B complex formation is the important principle underlying Sm site-dependent and correct Sm core assembly. This also explains our result that association of the U2-specific Sm core with the Sm site in the U2 snRNA is not affected by SMN addition. Thus, SMN action appears to be restricted to canonical Sm core assembly (see Fig. 6 for our model).

We demonstrated directly by in vitro Sm core assembly that the preformed SMN/SmD3B complex represents an intermediate and associates with the remaining Sm subcomplexes (Fig. 5). Importantly, SMN was released after assembly of a canonical Sm core, depending on a correct Sm site. For example, U1 snRNA with a mutated Sm site was not stably bound, and SMN was not released. We do not know whether the complex of SMN and all seven Sm proteins, which we detect in the absence of a functional Sm site, is an open structure or a closed Sm

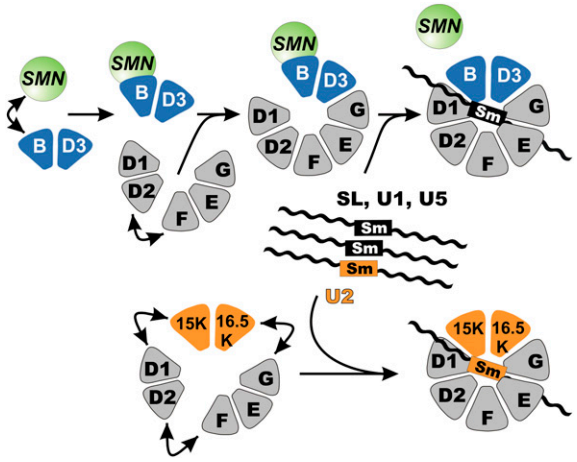

Figure 6. SnRNP-specific Sm core assembly mediated by SMN chaperone: Model canonical Sm cores (*top*; SL, U1, and U5 snRNPs), which contain the SmD3B subunit (in blue), are assembled with SMN (in green) acting as a chaperone. Initially, SMN interacts specifically with SmD3B, forming the SMN/SmD3B complex, which then accepts the other, common Sm subcomplexes (SmD1D2 and SmEFG; in gray). The resulting complex of SMN and all seven Sm proteins may be an open linear (as shown here) or a closed circular structure. Concomitantly with the stable integration of the snRNA, SMN is released, depending on a functional Sm site. (*Bottom*) In contrast, U2-specific variant Sm cores, which do not contain SmD3B, but instead the Sm16.5K/15K subcomplex (in orange), assemble in an SMN-independent manner.

ring-like complex. In sum, we provided direct evidence that SMN acts as a chaperone in Sm core assembly, based on a specific transient intermediate, the SMN/SmD3B complex, and that SMN is involved in recognizing a functional Sm site.

How does this assembly pathway for canonical Sm cores in trypanosomes compare with the model proposed for mammalian Sm core assembly (Chari et al. 2008, and references therein)? The most important difference appears to be that mammalian Sm cores require at least the SMN–Gemin2 complex and pICln, the latter component acting as an assembly chaperone and forming a closed ring structure together with SmD1D2 and SmEFG. This pICln–Sm complex, which is blocked in snRNA binding, cooperates with a second pICln complex, carrying SmD3B, in transferring all seven Sm proteins onto SMN. Only then, in a less well-defined step, can ring closure around the Sm site take place (Chari et al. 2008). In comparison, trypanosome SMN recruits specifically SmD3B; SMN accompanies the SmD3B unit until all seven Sm subunits are bound, and it leaves this complex only after a correct Sm site has been incorporated. Therefore, the trypanosome SMN appears to combine functional roles of both the mammalian pICln and SMN, acting as a chaperone and specificity factor in the assembly of canonical Sm cores (Fig. 6). Despite these apparent differences, we may have uncovered in the trypanosome system an intrinsic and essential activity of the SMN protein—transient SmD3B interaction and Sm site proofreading—that in mammalian Sm core assembly cannot be easily studied in the same manner, due to its many accessory components, additional protein modifications, and higher complexity of multiple assembly stages.

Since our in vitro assays indicate that SMN acts specifically during canonical Sm core assembly in trypanosomes, how are the variant Sm cores of U2 and U4 built up? They may follow a default pathway, or there might be additional, snRNP-specific assembly factors with SMN-like function. Gemin2, which we identified from *T. brucei*, may play such a role. In sum, although trypanosomes entail the additional complication of parallel assembly of different Sm cores, our study suggests that—compared with mammals—Sm core assembly in trypanosomes is achieved through a much less complex, SMN-based system: First, at least in vitro, SMN proved to be sufficient for activity in trypanosomes. Second, we did not find any obvious orthologs of most additional components known from vertebrates, such as Gemins3 to 8 (data not shown). As shown recently by Kroiss et al. (2008), in some metazoans, fungi, and plants only subsets of the Gemin3-8 proteins appear to operate, and in *Drosophila*, the SMN–Gemin2 dimer is sufficient for Sm core assembly in vitro. Third, arginine methylations in the Sm proteins that may regulate SMN interaction in humans appear to be missing in trypanosomes (see above for references) and dispensable in *Drosophila* (Gonsalvez et al. 2008).

Finally, the result that the trypanosome SMN is nuclear-localized suggests that the entire snRNP biogenesis pathway in trypanosomes proceeds in the nucleus, different from the situation in higher eukaryotes, but consistent with previous SL RNA and U2 snRNA localizations by RNA-FISH (Biton et al. 2006; Tkacz et al. 2007). The possibility that the localization of SMN is influenced by the PTP tag is very unlikely since we found that silencing of SMN expression is lethal and cells that exclusively expressed PTP-tagged SMN grew normally; in addition, endogenous SMN localized to the nucleoplasm as the PTP-tagged version did (Fig. 2). In trypanosomes, biogenesis pathways of SL RNP and the spliceosomal snRNPs are not well characterized. Whether the SL RNP assembly pathway includes a cytoplasmic phase remains a controversial issue (Mandelboim et al. 2003; Zeiner et al. 2003, 2004; Biton et al. 2006; Zamudio et al. 2009). An interesting open question in this context is what prevents the Sm proteins after their synthesis in the cytoplasm from binding to noncognate RNAs: We assume they rapidly assemble to Sm subcomplexes, which do not bind snRNAs (see Supplemental Fig. 4) and which are imported into the nucleus, due to their own nuclear localization signals, as shown for *T. brucei* SmB (Girard et al. 2004). The assembly to heptamer complexes may then proceed only in the nuclear compartment, with SMN ensuring specificity of canonical Sm core assembly. The identification of the trypanosomal SMN as a key player in snRNP biogenesis has opened up new possibilities to address these long-standing issues, including the question whether the assembly of the SL RNP and the spliceosomal snRNPs follows a common route in the trypanosome cell.

Materials and methods

Culture conditions for the procyclic form of *T. brucei* strain 427, extract preparation, TAP of protein complexes, MS and database analysis, the construction and expression of recombinant proteins, Western blotting, protein–protein interaction assays by GST pull-down, in vitro transcription of snRNAs, and oligonucleotide sequences are described in the Supplemental Material.

Purification of the trimeric SMN–SmD3B protein complex

Bacterial cell lysates from *Escherichia coli* BL-21 cells expressing GST-tagged *T. brucei* SMN (extract 1) and His-Flag-tagged SmD3B proteins (extract 2) were combined and incubated for 45 min at 30°C. The complex formed of the three recombinant proteins was purified through a two-step affinity procedure based on the GST tag on SMN and the His tag on SmD3. First, the trimeric complex was purified by HPLC on a 1-mL GSTrap FF column (ÄKTA Purifier; Amersham-GE Healthcare) and eluted in 50 mM Tris-HCl (pH 8.0) containing 10 mM reduced glutathione. Second, peak fractions were pooled, diluted in 5 vol of His-binding buffer (50 mM Na-phosphate at pH 8.0, 200 mM NaCl, 0.02% NP-40), and the trimeric complex was further purified on Ni-NTA agarose beads. Bound proteins were eluted by 300 mM imidazole-containing His-binding buffer, dialyzed against 1× Sm storage buffer (20 mM Tris-HCl at pH 7.5, 200 mM NaCl, 10% glycerol, 5 mM β-mercaptoethanol) and concentrated by ultrafiltration on Microcon Ultracel YM-3 spin columns (Millipore).

Reconstitution of recombinant Sm cores

Reconstitution of canonical Sm cores with total RNA from *T. brucei* or with transcribed RNAs was followed by His tag

pull-down of reconstituted complexes (Fig. 4). For the assays, 150 pmol of purified His-tagged Sm subcomplexes (for canonical Sm core: His-SmD3B, His-SmD1D2, and His-SmEFG; for U2 Sm core: His-Sm16.5K/15K, His-SmD1D2, and His-SmEFG) were mixed in 10 μL of 5× reconstitution buffer (100 mM Tris-HCl at pH 7.5, 1 M NaCl, 25 mM $MgCl_2$, 5 mM DTT, without NP-40), with or without the addition of His-tagged SMN protein. The volume of the samples was adjusted to 40 μL, and the reactions were preincubated for 15 min at 30°C. Then 10 μg of *T. brucei* total RNA (prepared with TRIzol reagent; Invitrogen) or ~50 ng (2×10^7 cpm) of full-length snRNA transcripts (TbU2-WT, TbU5-WT, TbU5-mutSm, or a 101-nt control RNA), and ~10 ng (2×10^6 cpm) of short RNA transcripts (TbU4-3′ half WT or mutSm) were added to the reactions, to give a final volume of 50 μL. The reconstitution reactions were further incubated for 30 min at 30°C and then for 15 min at 37°C. Control reconstitution reactions were done similarly by using only two His-tagged Sm subcomplexes in different combinations (SmD1D2 + SmD3B, SmD1D2 + SmEFG, SmEFG + SmD3B; 150 pmol of each subcomplex) with 10 μg of total RNA.

For His tag pull-down, the reconstituted Sm complexes were incubated with 25 μL of packed Ni-NTA agarose beads in 0.5 mL of 1× reconstitution buffer (20 mM Tris-HCl at pH 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 0.02% NP-40, 0.5 mM DTT), for 2 h at 4°C. After washing with the same buffer (three times with 1 mL), the beads were resuspended in 1× proteinase K buffer (100 mM Tris-HCl at pH 7.5, 150 mM NaCl, 1% SDS, 12.5 mM EDTA) and the RNAs were recovered by phenol-chloroform extraction and ethanol precipitation. The coselected RNAs were analyzed by 10% denaturing polyacrylamide gel electrophoresis and autoradiography or Northern blotting with a mixture of snRNA-specific DIG-labeled probes (Bell and Bindereif 1999).

For reconstitution of canonical Sm cores with the trimeric GST-SMN/His-Flag-SmD3B complex and Flag-pull-down assays (Fig. 5B), either the trimeric protein complex or His-Flag-SmD3B subcomplex was used for reconstitution together with His-SmD1D2 and His-SmEFG subcomplexes. One-hundred-fifty picomoles of each complex were combined in 1× reconstitution buffer (as above, without NP-40; 40-μL reaction) with 5 μg (200 pmol) of full-length TbU1-WT or TbU1-mutSm RNA. The samples were incubated for 30 min at 30°C, then for 15 min at 37°C. Reconstituted Sm complexes were purified by Flag pull-down on anti-Flag beads (25 μL of packed M2 agarose beads per reaction; Sigma), and washed three times in 1 mL of 1× reconstitution buffer (as above, containing 0.02% NP-40). Bound proteins and RNAs were eluted by two sequential 30-min incubations at room temperature with 3× Flag peptide (Sigma; 200 ng/μL in 1× reconstitution buffer + 0.02% NP-40). Eluted proteins and RNAs were recovered by precipitation and analyzed by 15% SDS-PAGE and sequential Coomassie and silver staining.

RNAi silencing of SMN expression and RT–PCR

The *T. brucei* SMN coding region (Tb11.01.6640) from positions 191 to 463 was PCR-amplified and cloned as a stem–loop module into pLEW100 vector (Wirtz et al. 1999) according to a well-established cloning strategy (Wang et al. 2000). *T. brucei* 29-13 cells were transfected with 15 μg of the SacII-linearized SMN-pLEW100 construct and cloned by limiting dilution in the presence of G418 (15 μg/mL), hygromycin (50 μg/mL), and phleomycin (2.5 μg/mL) (Schimanski et al. 2006). Silencing was induced with doxycycline (1 μg/mL). Cells were counted every day and diluted to 2×10^6 cells per milliliter.

For semiquantitative and real-time RT–PCR analysis (iCycler, Bio-Rad; iQ5 software), RNA was prepared from cells after 1, 2, or 3 d of RNAi induction or from uninduced cells, using TRizol reagent (Invitrogen) and further purified by the RNeasy Mini Kit (Qiagen). cDNA was generated by SuperScript II RT (Invitrogen) combined with gene-specific reverse primers, or by using the iScript cDNA synthesis kit (Bio-Rad). Amplification products were analyzed by agarose gel electrophoresis. 7SL RNA specific primers were used in control reactions for the semiquantitative RT–PCR as well as for normalization of the real-time RT–PCR results.

Immunofluorescence

T. brucei cells (2×10^7 to 4×10^7) expressing SMN-PTP-EE (exclusively expressing PTP-tagged SMN) or wild-type *T. brucei* cells were harvested, washed with PBS, and resuspended in 1 mL of PBS. Twenty-five microliters of the suspension were fixed on coverslips with 4% paraformaldehyde in PBS for 30 min at 4°C. The cells were blocked with 1% cold-water fish gelatine (Sigma-Aldrich) in PBS and incubated with a 1:40,000 dilution of rabbit anti-protein A primary antibody (Sigma-Aldrich), or rabbit anti-SMN antibody developed against the *T. brucei* SMN protein and affinity-purified (see in the Supplemental Material under "Western Blotting"), at a dilution of 1:200. Following washing with PBS containing 0.05% Tween 20, 4,6-diamidino-2-phenylindol (1 μg/mL DAPI), and goat anti-rabbit Alexa Fluor 594 secondary antibody (1:400 dilution; Invitrogen) were added. After washing, coverslips were mounted in DABCO-containing Mowiol embedding medium. For imaging, a Zeiss Axioskop 20 microscope and the Axio Vision software were used.

Acknowledgments

We thank Utz Fischer, Ashwin Chari, and Michael Sattler for discussions; Tu N. Nguyen and Tianpeng Gu for essential help; as well as Monika Raabe for excellent technical assistance in MS. This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG Sonderforschungsbereich 535, DIP German-Israel Project Cooperation and GK1384; to A.B.), the European-Commission-funded Network of Excellence EURASNET (to A.B., J.B., and H.U.), and the Fonds der Chemischen Industrie (to A.B.), and by NIH grant AI059377 (to A.G.).

References

Bell M, Bindereif A. 1999. Cloning and mutational analysis of the *Leptomonas seymouri* U5 snRNA gene: Function of the Sm site in core RNP formation and nuclear localization. *Nucleic Acids Res* **27:** 3986–3994.

Biton M, Mandelboim M, Arvatz G, Michaeli S. 2006. RNAi interference of XPO1 and Sm genes and their effect on the spliced leader RNA in *Trypanosoma brucei*. *Mol Biochem Parasitol* **150:** 132–143.

Brahms H, Raymackers J, Union A, de Keyser F, Meheus L, Lührmann R. 2000. The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies. *J Biol Chem* **275:** 17122–17129.

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.

Brow DA. 2002. Allosteric cascade of spliceosome activation. *Annu Rev Genet* **36:** 333–360.

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.

Cross M, Wieland B, Palfi Z, Günzl A, Röthlisberger U, Lahm HW, Bindereif A. 1993. The *trans*-spliceosomal U2 snRNP protein 40K of *Trypanosoma brucei*: Cloning and analysis of functional domains reveals homology to a mammalian snRNP protein. *EMBO J* **12:** 1239–1248.

Deckert J, Hartmuth K, Boehringer D, Behzadnia N, Will CL, Kastner B, Stark H, Urlaub H, Lührmann R. 2006. Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Mol Cell Biol* **26:** 5528–5543.

Eggert C, Chari A, Laggerbauer B, Fischer U. 2006. Spinal muscular atrophy: The RNP connection. *Trends Mol Med* **12:** 113–121.

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.

Friesen WJ, Paushkin S, Wyce A, Massenet S, Pesiridis GS, Van Duyne G, Rappsilber J, Mann M, Dreyfuss G. 2001b. The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol* **21:** 8289–8300.

Friesen WJ, Wyce A, Paushkin S, Abel L, Rappsilber J, Mann M, Dreyfuss G. 2002. A novel WD repeat protein component of the methylosome binds Sm proteins. *J Biol Chem* **277:** 8243–8247.

Girard C, Mouaikel J, Neel H, Bertrand E, Bordonné R. 2004. Nuclear localization properties of a conserved protuberance in the Sm core complex. *Exp Cell Res* **299:** 199–208.

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.

Hannus S, Bühler D, Romano M, Séraphin B, Fischer U. 2000. The *Schizosaccharomyces pombe* protein Yab8p and a novel factor, Yip1p, share structural and functional similarity with the spinal muscular atrophy-associated proteins SMN and SIP1. *Hum Mol Genet* **9:** 663–674.

Hossain MJ, Korde R, Singh S, Mohmmed A, Dasaradhi PV, Chauhan VS, Malhotra P. 2008. Tudor domain proteins in protozoan parasites and characterization of *Plasmodium falciparum* tudor staphylococcal nuclease. *Int J Parasitol* **38:** 513–526.

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.

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.

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* **105:** 10045–10050.

Lefebvre S, Bürglen 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.

Liang XH, Haritan A, Uliel S, Michaeli S. 2003. *Trans* and *cis* splicing in trypanosomatids: Mechanism, factors, and regulation. *Eukaryot Cell* **2:** 830–840.

Liu Q, Dreyfuss G. 1996. A novel nuclear structure containing the survival of motor neurons protein. *EMBO J* **15:** 3555–3565.

Liu Q, Fischer U, Wang F, Dreyfuss G. 1997. The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* **90:** 1013–1021.

Liu Q, Liang XH, Uliel S, Belahcen M, Unger R, Michaeli S. 2004. Identification and functional characterization of Lsm proteins in *Trypanosoma brucei*. *J Biol Chem* **279:** 18210–18219.

Lücke S, Klöckner T, Palfi Z, Boshart M, Bindereif A. 1997. *Trans* mRNA splicing in trypanosomes: Cloning and analysis of a PRP8-homologous gene from *Trypanosoma brucei* provides evidence for a U5-analogous RNP. *EMBO J* **16:** 4433–4440.

Ma Y, Dostie J, Dreyfuss G, Van Duyne GD. 2005. The Gemin6-Gemin7 heterodimer from the survival of motor neurons complex has an Sm protein-like structure. *Structure* **13:** 883–892.

Maier AG, Webb H, Ding M, Bremser M, Carrington M, Clayton C. 2001. The coatomer of *Trypanosoma brucei*. *Mol Biochem Parasitol* **115:** 55–61.

Mandelboim M, Barth S, Biton M, Liang XH, Michaeli S. 2003. Silencing of Sm proteins in *Trypanosoma brucei* by RNA interference captured a novel cytoplasmic intermediate in spliced leader RNA biogenesis. *J Biol Chem* **278:** 51469–51478.

Meister G, Eggert C, Fischer U. 2002. SMN-mediated assembly of RNPs: A complex story. *Trends Cell Biol* **12:** 472–478.

Otter S, Grimmler 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.

Owen N, Doe CL, Mellor J, Davies KE. 2000. Characterization of the *Schizosaccharomyces pombe* orthologue of the human survival motor neuron (SMN) protein. *Hum Mol Genet* **9:** 675–684.

Palfi Z, Günzl A, Cross M, Bindereif A. 1991. Affinity purification of *Trypanosoma brucei* small nuclear ribonucleoproteins reveals common and specific protein components. *Proc Natl Acad Sci* **88:** 9097–9101.

Palfi Z, Lücke S, Lahm HW, Lane WS, Kruft V, Bragado-Nilsson E, Séraphin B, Bindereif A. 2000. The spliceosomal snRNP core complex of *Trypanosoma brucei*: Cloning and functional analysis reveals seven Sm protein constituents. *Proc Natl Acad Sci* **97:** 8967–8972.

Palfi Z, Schimanski B, Günzl A, Lücke S, Bindereif A. 2005. U1 small nuclear RNP from *Trypanosoma brucei*: A minimal U1 snRNA with unusual protein components. *Nucleic Acids Res* **33:** 2493–2503.

Paushkin S, Charroux B, Abel L, Perkinson RA, Pellizzoni L, Dreyfuss G. 2000. The survival motor neuron protein of *Schizosacharomyces pombe*. Conservation of survival motor neuron interaction domains in divergent organisms. *J Biol Chem* **275:** 23841–23846.

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, Charroux B, Dreyfuss G. 1999. SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. *Proc Natl Acad Sci* **96:** 11167–11172.

Pellizzoni L, Yong J, Dreyfuss G. 2002. Essential role for the SMN complex in the specificity of snRNP assembly. *Science* **298:** 1775–1779.

Pillai RS, Grimmler M, Meister G, Will CL, Lührmann R, Fischer U, Schümperli 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.

Preußer C, Palfi Z, Bindereif A. 2009. Special Sm core complex functions in assembly of the U2 small nuclear ribonucleoprotein of *Trypanosoma brucei*. *Euk. Cell* doi: 10.1128/EC.00090-09.

Pu WT, Krapivinsky GB, Krapivinsky L, Clapham DE. 1999. pICln inhibits snRNP biogenesis by binding core spliceosomal proteins. *Mol Cell Biol* **19:** 4113–4120.

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.

Schimanski B, Nguyen TN, Günzl A. 2005. Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell* **4:** 1942–1950.

Schimanski B, Brandenburg J, Nguyen TN, Caimano MJ, Günzl A. 2006. A TFIIB-like protein is indispensable for spliced leader RNA gene transcription in *Trypanosoma brucei*. *Nucleic Acids Res* **34:** 1676–1684.

Schümperli 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.

Selenko P, Sprangers R, Stier G, Bühler D, Fischer U, Sattler M. 2001. SMN tudor domain structure and its interaction with the Sm proteins. *Nat Struct Biol* **8:** 27–31.

Sprangers R, Groves MR, Sinning I, Sattler M. 2003. High-resolution X-ray and NMR structures of the SMN Tudor domain: Conformational variation in the binding site for symmetrically dimethylated arginine residues. *J Mol Biol* **327:** 507–520.

Tkacz ID, Lustig Y, Stern MZ, Biton M, Salmon-Divon M, Das A, Bellofatto V, Michaeli S. 2007. Identification of novel snRNA-specific Sm proteins that bind selectively to U2 and U4 snRNAs in *Trypanosoma brucei*. *RNA* **13:** 30–43.

Tkacz ID, Cohen S, Salmon-Divon M, Michaeli S. 2008. Identification of the heptameric Lsm complex that binds U6 snRNA in *Trypanosoma brucei*. *Mol Biochem Parasitol* **160:** 22–31.

Wang Z, Morris JC, Drew ME, Englund PT. 2000. Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. *J Biol Chem* **275:** 40174–40179.

Wang P, Palfi Z, Preusser C, Lücke S, Lane WS, Kambach C, Bindereif A. 2006. Sm core variation in spliceosomal small nuclear ribonucleoproteins from *Trypanosoma brucei*. *EMBO J* **25:** 4513–4523.

Will CL, Lührmann R. 2001. Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13:** 290–301.

Will CL, Lührmann R. 2006. Spliceosome structure and function. In *The RNA world* (eds. TR Cech and JF Atkins), pp. 369–400. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Winkler C, Eggert C, Gradl D, Meister G, Giegerich M, Wedlich D, Laggerbauer 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.

Wirtz E, Leal S, Ochatt C, Cross GA. 1999. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol* **99:** 89–101.

Yong J, Wan L, Dreyfuss G. 2004. Why do cells need an assembly machine for RNA–protein complexes? *Trends Cell Biol* **14:** 226–232.

Zamudio JR, Mittra B, Chattopadhyay A, Wohlschlegel JA, Sturm NR, Campbell DA. 2009. *Trypanosoma brucei* spliced leader RNA maturation by the cap 1 2′-O-ribose methyltransferase and SLA1 H/ACA snoRNA pseudouridine synthase complex. *Mol Cell Biol* **29:** 1202–1211.

Zeiner GM, Sturm NR, Campbell DA. 2003. Exportin 1 mediates nuclear export of the kinetoplastid spliced leader RNA. *Eukaryot Cell* **2:** 222–230.

Zeiner GM, Foldynová S, Sturm NR, Lukes J, Campbell DA. 2004. SmD1 is required for spliced leader RNA biogenesis. *Eukaryot Cell* **3:** 241–244.

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.