

The Gemin6-Gemin7 Heterodimer from the Survival of Motor Neurons Complex Has an Sm Protein-like Structure

Yingli Ma, Josée Dostie, Gideon Dreyfuss,
and Gregory D. Van Duyne*

Department of Biochemistry & Biophysics
Howard Hughes Medical Institute
University of Pennsylvania School of Medicine
Philadelphia, Pennsylvania 19104

Summary

The survival of motor neurons (SMN) protein, product of the disease gene of the common neurodegenerative disease spinal muscular atrophy, is part of the large multiprotein “SMN complex.” The SMN complex functions as an assembly machine for small nuclear ribonucleoproteins (snRNPs)—the major components of the spliceosome. Here, we report the crystal structure of two components of the human SMN complex, Gemin6 and Gemin7. Although Gemin6 and Gemin7 have no significant sequence similarity with Sm proteins, both adopt canonical Sm folds. Moreover, Gemin6 and Gemin7 exist as a heterodimer, and interact with each other via an interface similar to that which mediates interactions among the Sm proteins. Together with binding experiments that show that the Gemin6/Gemin7 complex binds to Sm proteins, these findings provide a framework for considering how the SMN complex, with Gemin6 and Gemin7 as tools, might organize Sm proteins for formation of Sm rings on snRNA targets.

Introduction

The survival of motor neurons (SMN) complex is a multiprotein machine that plays a central role in the biogenesis of ribonucleoprotein complexes (RNPs), including spliceosomal small nuclear RNPs (snRNPs) (Fischer et al., 1997; Meister et al., 2001; Pellizzoni et al., 1998) and, possibly, small nucleolar RNPs (snoRNPs) (Jones et al., 2001; Pellizzoni et al., 2001a), heterogeneous nuclear RNPs (hnRNPs) (Mourelatos et al., 2001), and transcriptosomes (Pellizzoni et al., 2001b). The SMN complex is located both in the cytoplasm of cells, where immunofluorescence studies indicate a diffuse distribution, and in the nucleus, where it is highly concentrated in nuclear bodies termed gems (Liu and Dreyfuss, 1996; Liu et al., 1997). Consistent with its critical role in the biogenesis of snRNPs, the SMN protein is essential for viability in all eukaryotic organisms tested thus far, including mouse, chicken, fly, worm, and fission yeast (Hannus et al., 2000; Miguel-Aliaga et al., 1999, 2000; Owen et al., 2000; Paushkin et al., 2000; Schrank et al., 1997; Wang and Dreyfuss, 2001). Reduced levels of SMN as a result of deletions or loss-of-function mutations in the *SMN* gene cause spinal muscular atrophy (SMA), a common neurodegenerative disease characterized by the degeneration of motor

neurons of the spinal cord (Lefebvre et al., 1995). SMA is the leading genetic cause of infant mortality, affecting ~1 in 6000 newborns (Melki, 1997).

In addition to SMN, the SMN complex contains several proteins called Gemin2–7 (Gubitz et al., 2004; Paushkin et al., 2002). Of these, Gemin3 contains a DEAD-box RNA helicase motif (Charroux et al., 1999) and Gemin5 contains multiple WD repeats (Gubitz et al., 2002), but the remaining components of the SMN complex contain no recognizable domains that could provide insight into their functional roles. Gemin2, Gemin3, and Gemin5 interact directly with SMN, whereas Gemin4 is recruited to the complex through its interaction with Gemin3 (Charroux et al., 1999, 2000; Gubitz et al., 2002; Liu et al., 1997). Similarly, Gemin6 associates with the SMN complex indirectly through Gemin7, which binds to SMN through its N-terminal 30 amino acid residues (Baccon et al., 2002; Pellizzoni et al., 2002a).

Although several studies suggest a function for the SMN complex in the nucleus (Pellizzoni et al., 2001b), its role in the biogenesis of snRNPs in the cytoplasm is best understood. The major snRNPs (named by their snRNA components U1, U2, U4/U6, and U5) are essential components of the spliceosome, a dynamic protein/RNA machine that mediates the nuclear excision of introns from pre-mRNA. Each snRNP consists of one U snRNA molecule, a core composed of seven highly conserved Sm proteins, and several snRNP-specific proteins (Will and Luhrmann, 2001). With the exception of U6, the major U snRNAs are exported to the cytoplasm where the Sm protein core is assembled on the Sm site of the U snRNA and the 5'-cap is hypermethylated (Mattaj and De Robertis, 1985). The partially assembled snRNP is believed to contain a heptameric ring of Sm proteins encircling the Sm site, which is a short, single-stranded, uridine-rich sequence flanked by two RNA stem-loop structures (Will and Luhrmann, 2001). The assembly of the Sm core is required for hypermethylation of the cap structure, and provides a nuclear localization signal for import of mature snRNP back into the nucleus (Kleinschmidt et al., 1989; Raker et al., 1996, 1999; Sumpter et al., 1992).

The Sm core readily assembles on U snRNAs *in vitro* in the absence of ATP when purified RNAs and total snRNP proteins are mixed (Raker et al., 1996, 1999; Sumpter et al., 1992). However, this process requires ATP hydrolysis in HeLa cell and *Xenopus laevis* egg extracts (Kleinschmidt et al., 1989; Meister et al., 2001). *In vitro*, the SMN complex has been shown to mediate the assembly of Sm proteins on the Sm site of U snRNAs by a mechanism that requires ATP binding, but not ATP hydrolysis (Pellizzoni et al., 2002b). In this process, the SMN complex is thought to function as an assembly chaperone that prevents illicit assembly of the Sm proteins on irrelevant uridine-rich RNAs (Pellizzoni et al., 2002b).

As part of an overall effort to study the structure and function of the SMN complex, we are investigating the structures and biochemical properties of individual pro-

*Correspondence: vanduyne@mail.med.upenn.edu

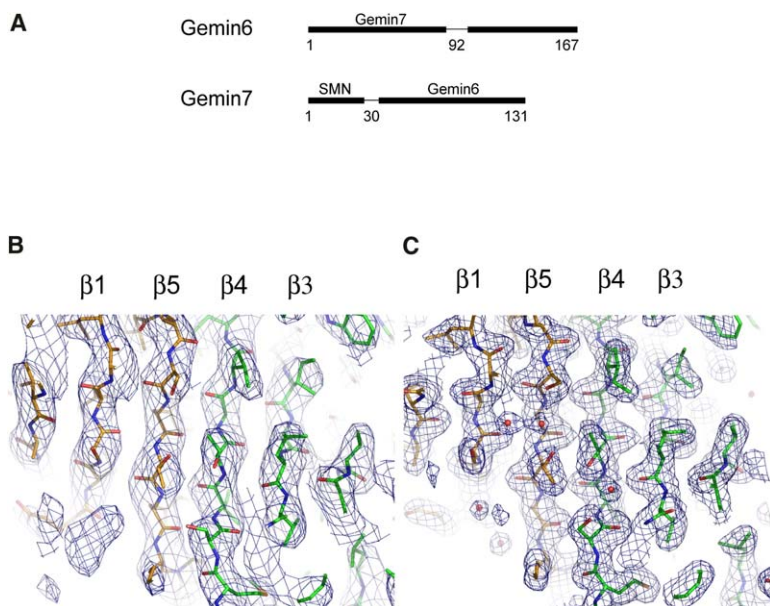


Figure 1. Domain Structure and Electron Density for the Gemin6/7 Complex at the β 4- β 5 Interface

(A) Schematic of domain organization in Gemin6 and Gemin7. The region on Gemin7 required for interaction with SMN and the domains that interact with one another in Gemin6 and Gemin7 are indicated.

(B) Experimental electron density map for crystal form II from Solve/Resolve phasing at 2.8 Å resolution, contoured at 1.5 σ . Gemin6 is colored green (right half) and Gemin7 is orange (left half).

(C) σ_A -weighted $2F_o - F_c$ electron density for the final refined structure in crystal form I at 2 Å resolution, contoured at 1.5 σ . The orientation and coloring of the indicated β strands is the same as that shown in (A).

teins and stable subcomplexes in this system. Here we report the crystal structure of the human Gemin6/Gemin7 (hereafter referred to as Gemin6/7) complex. Although neither protein displays any significant sequence homology to Sm protein family members, the structure reveals that Gemin6 and Gemin7 both contain Sm-like folds and that the Gemin6/7 interface is similar to the interfaces observed between Sm proteins. We also show that the Gemin6/7 heterodimer is competent to bind and perhaps to organize Sm proteins. Together, the structural and biochemical data indicate that Gemin6/7 may participate in formation of higher order complexes with Sm or Sm-like proteins and could function as an Sm-surrogate during snRNP assembly.

Results

Structure Determination

Secondary structure prediction methods indicated that the N-terminal 30 residues of Gemin7 that have been shown to be required for interaction with SMN are likely to be unstructured in the free Gemin6/7 complex (the domain structures of Gemin6 and Gemin7 are shown schematically in Figure 1A). These residues were therefore removed from the constructs used for crystallographic studies. Limited trypsin proteolysis of the Gemin6/7 complex, coupled with MALDI-TOF mass spectral analysis and N-terminus sequencing of tryptic fragments also identified a domain boundary in Gemin6, near residue 92 (data not shown). Coexpression and copurification of Gemin6 (residues 1–92) with Gemin7 (residues 31–131) confirmed that the N-terminal domain of Gemin6 was both necessary and sufficient for tight association with Gemin7. However, only the full length Gemin6/Gemin7 (residues 31–131) complex formed diffraction quality crystals.

Crystals of the Gemin6/7 complex that diffracted to 2.8 Å resolution were produced by vapor diffusion, and

experimental phases were determined from multiwavelength anomalous scattering data using SeMet-substituted Gemin6 and Gemin7. A second crystal form diffracted to higher resolution, but was not amenable to multiwavelength phasing using SeMet-substituted proteins. We therefore used molecular replacement to position Gemin6/7 in the higher resolution crystal form and refined the resulting structure to 2.0 Å resolution. Crystallographic data and the results of phasing and refinement are summarized in Table 1, and representative electron density of the Gemin6/7 structure is shown in Figure 1B. There are two Gemin6/7 heterodimers in the asymmetric unit of both crystal forms, and all four of the independent heterodimer structures are similar, with pairwise rms deviations ranging from 0.3–1.1 Å for superposition of C α atoms, excluding the variable β 3- β 4 loop region (discussed below).

Overall Structure of the Gemin6/Gemin7 Complex

Gemin6 and Gemin7 have similar folds, with a five-stranded bent β sheet flanked by α helices (Figures 2A and 2B). The two β sheets in the heterodimer are connected via β 4 of Gemin6 and β 5 of Gemin7, resulting in the formation of a continuous 10-stranded β sheet (Figure 2C). The longer N-terminal helix of Gemin7 packs tightly into a hydrophobic pocket formed by α 1, β 2- β 4, and α 2 of Gemin6. This hydrophobic interface is flanked by a network of hydrogen bonds formed between α 1 side chains in Gemin7 and pocket residues of Gemin6. Overall, the Gemin6/7 interface buries a total of 2577 Å² solvent-accessible surface area with a shape complementarity index of 0.77, which is consistent with a highly specific interaction (Lawrence and Colman, 1993).

The two independent Gemin6/7 heterodimers interact in the crystal to form a two-fold-symmetric (Gemin6/7)₂ heterotetramer via the β 3- β 4 loop of Gemin6 (Figure 2D). The β 3- β 4 hairpin extends from the

Table 1. Crystallographic Data

Diffraction Data				
Crystal form	I	II		
Space group	P4 ₂ 2 ₁ 2	P6 ₁ 22		
Cell constants (Å)	a = 136.0, c = 81.5	a = 107.1, c = 289.1		
	Native	SeMet λ1	SeMet λ2	SeMet λ3
Resolution (Å)	50–2.0	50–2.8	50–2.8	50–2.8
Completeness (%)	95.6 (66.0)	95.4 (65.2)	95.7 (67.4)	95.2 (64.3)
Redundancy ^a	32.4 (4.7)	28.6 (3.9)	28.5 (3.9)	28.3 (4.0)
<I/σ>	36.8 (3.7)	24.0 (3.5)	23.7 (3.5)	23.9 (3.4)
R _{sym} (%) ^b	7.2 (31)	7.8 (26.3)	7.7 (26.4)	7.8 (27.6)
SOLVE Phasing				
Resolution				2.8 Å
Overall figure of merit				0.55
Overall Z score				36.3
Refinement				
Resolution	2.0 Å			
R _{working} (%)	21.7 (27)			
R _{free} (%)	25.2 (33)			
Average B (Å ²)	32.5			
Rms bonds (Å)	0.019			
Rms angles (°)	1.283			

Values in parentheses are for the highest resolution shell. Wavelengths are λ₁ = 0.97974Å (inflection), λ₂ = 0.97961Å (peak), and λ₃ = 0.97190Å (remote).

^aRedundancy = the number of observations/the number of unique reflections.

^bR_{sym} = $\sum_n |I_n - \langle I \rangle_n| / \sum I_n$, where $\langle I \rangle_n$ is the average intensity over symmetry equivalents.

main body of each heterodimer, exposing side chains that can self-associate primarily via polar and electrostatic interactions. This interface buries a total of 861 Å² solvent-accessible surface area and the same heterotetramer is observed in both crystal forms, despite the fact that crystal packing is different in the two lattices. To establish whether the higher order oligomer

observed in the crystals is also present in solution, we analyzed the Gemin6/7 complex at three different concentrations using sedimentation equilibrium ultracentrifugation. At complex concentrations ranging from 9 to 123 μM, radial distributions measured at equilibrium for three speeds were fit well by a single-species heterodimer model (that does not dissociate), with no indi-

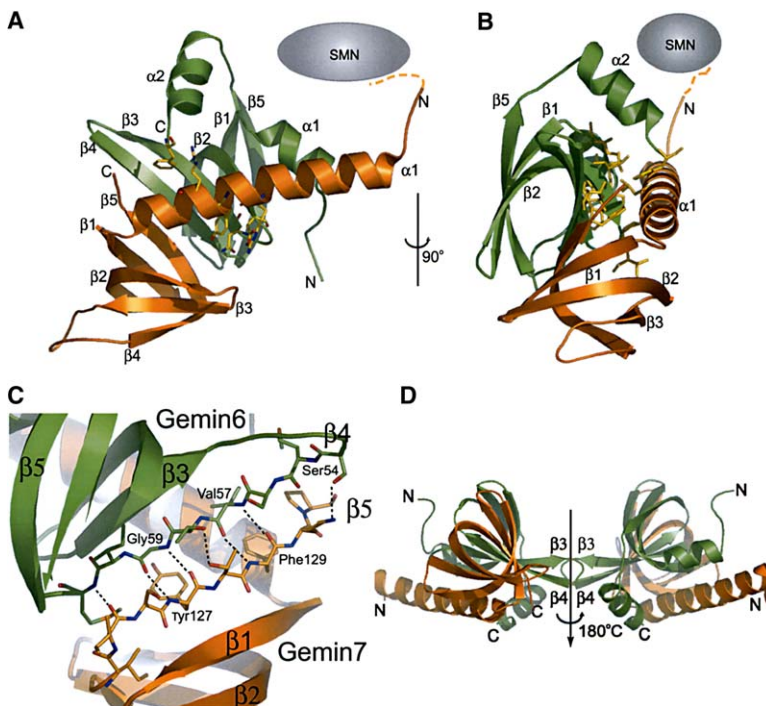


Figure 2. Structure of the Gemin6/7 Complex (A) Ribbon diagram of the Gemin6/7 complex, with Gemin6 and Gemin7 colored green and orange, respectively. Residues forming polar interactions between Gemin7-α1 and Gemin6 are drawn as sticks. The N-terminal 30 residues of Gemin7 are shown schematically as a dashed orange line. (B) Orthogonal view to that shown in (A), but with residues forming hydrophobic interactions between Gemin7-α1 and Gemin6 drawn as sticks. (C) Closeup of the β4-β5 interface between Gemin6 and Gemin7 that forms an extended β sheet. (D) The Gemin6/7 heterotetramer present in both crystal forms of the Gemin6/7 complex. The dimerization interface involves the β3-β4 region and the symmetry dyad is indicated.

cation of heterotetramer formation (see [Experimental Procedures](#)). Although the centrifugation data indicates that the Gemin6/7 heterodimer does not self-associate to any significant extent in the absence of other proteins, we cannot rule out a role for this interaction in the context of the SMN complex, where higher order oligomerization via SMN is known to play an important functional role ([Pellizzoni et al., 1999](#)). In this case, Gemin6/7 may be present in multiple copies, where the weak affinity interaction observed here in the Gemin6/7 crystals could be functionally relevant.

Gel electrophoresis analysis of dissolved crystals confirmed that full-length Gemin6 was present in both Gemin6/7 complex crystal forms, but electron density for the C-terminal domain of Gemin6 (residues 87–167) is not sufficiently well defined in either crystal form to allow tracing of the polypeptide chain. Small helical segments could be identified in difference electron density maps of crystal form II, but the sequence could not be assigned, and connectivity to the corresponding N-terminal domain could not be established. We believe that the missing C-terminal domain of Gemin6 is folded in the Gemin6/7 crystals because the isolated domain can be overexpressed and purified and is resistant to mild trypsin proteolysis (data not shown). Straightforward crystallographic refinement to the residuals shown in [Table 1](#) supports the conclusion that the Gemin6 C-terminal domain is disordered and, therefore, does not contribute to the high-resolution diffraction data. Although unexpected, the disordering of an entire domain in a high-resolution crystal structure has been observed (for example, see [Rice and Steitz, 1994](#)). Interestingly, N-terminal sequencing of limited proteolysis products of the Gemin6/7 complex showed that trypsin cleaves Gemin6 after Lys92 to produce N-terminal and C-terminal domains with a boundary that closely matches that observed in the crystals.

Gemin6 and Gemin7 Resemble Sm Proteins

Primary amino acid sequence analyses of Gemin6 and Gemin7 did not indicate significant homology with known proteins in the databases ([Baccon et al., 2002](#); [Pellizzoni et al., 2002a](#)). However, the Gemin6/7 structure reveals that both Gemin6 and Gemin7 contain folds that are similar to those observed in the spliceosomal Sm proteins. The closest matches to proteins with known three-dimensional structure using DALI ([Holm and Sander, 1993](#)) are the bacterial Sm-like Hfq ([Schumacher et al., 2002](#)), human SmB, SmD3, SmD1, and SmD2 ([Kambach et al., 1999](#)), and archaeal Sm ([Mura et al., 2001](#); [Toro et al., 2001](#)) proteins. The Sm proteins are evolutionarily conserved, with homologs from bacteria to humans, and, like Gemin6 and Gemin7, they contain an N-terminal helix followed by a 5-stranded, highly bent antiparallel β sheet. This protein family is also characterized by two sequence motifs, Sm1 and Sm2, which are separated by a sequence of variable length ([Hermann et al., 1995](#); [Seraphin, 1995](#)). An alignment of Gemin6 and Gemin7 with the human Sm proteins is shown in [Figure 3](#).

Among the human Sm proteins, Gemin6 is most similar in structure to SmD1 (PDB entry 1B34), with an rms deviation of 1.7 Å over 58 α -carbons, and Gemin7 most

closely superimposes with SmB (PDB entry 1D3B), with an rms deviation of 1.9 Å over 60 α -carbons. A representative superposition of Gemin6, Gemin7, and human SmD3 is shown in [Figure 4A](#). The primary source of variation among the human Sm protein structures and between Gemin6/7 and the Sm proteins exists in the two terminal segments and in the variable β 3– β 4 region. The α 1-helix in Gemin7 is also substantially longer than in either the Sm proteins or in Gemin6. Of the four human Sm proteins with experimental structures, only SmD1 has a C-terminal helix corresponding to that observed in Gemin6, but the helices are oriented differently in the two cases (data not shown).

Structure-based sequence alignment of the seven human Sm proteins with Gemin6 and Gemin7 indicates that strongly conserved amino acids in the Sm1 and Sm2 motifs, which are important for the Sm fold, are conserved in Gemin6 and Gemin7 ([Figure 3](#)). For example, in human SmB/B', Asn39 hydrogen bonds with Gly74 and Asp35, linking the C-terminal segments of strands β 2 and β 4 with the N terminus of β 3. In Gemin6, Asn43, Gly62, and Asp38 form an identical hydrogen bonding network. Similarly, SmD3 residues Asp37, Asn40, and Tyr62 and Gemin7 residues Asp96, Asn101, and Tyr103 form the same hydrogen bonding patterns. As expected, Gemin6 and Gemin7 also share the characteristic pattern of hydrophobic residues found in the Sm1 and Sm2 motifs, although the identities of these core packing residues are not strongly conserved.

The structural similarity between Gemin6 and Gemin7 and the Sm proteins also extends to the interface formed in the Gemin6/7 heterodimer ([Figure 2C](#)). Sm-like proteins display a propensity to oligomerize with themselves or with other Sm proteins by formation of a head-to-tail continuous β sheet via interactions between edge strands (β 4 and β 5) of the interacting domains ([Collins et al., 2003](#); [Kambach et al., 1999](#); [Mura et al., 2001](#); [Schumacher et al., 2002](#); [Toro et al., 2001](#)). In the absence of U snRNAs, the seven Sm proteins have been shown to preferentially form three stable heteromeric subcomplexes, SmD3/B, SmD1/D2, and SmF/E/G ([Lehmeier et al., 1990](#); [Raker et al., 1996](#)). Crystal structures for the D3/B and D1/D2 Sm domain cocomplexes have been reported ([Kambach et al., 1999](#)), and Gemin6/7 superimposes on these structures as a heterodimer reasonably well, with rms deviations of 2.5 and 3.3 Å, respectively, over C α atoms ([Figure 4B](#)). The largest difference between the Gemin6/7 and the SmB/D3 or SmD1/D2 interfaces involves the longer N-terminal helix found in Gemin7, which allows formation of a more substantial interaction with Gemin6 than is observed between the human Sm proteins. This more intimate Gemin6/7 interface is reflected in the higher shape complementarity index ([Lawrence and Colman, 1993](#)) for Gemin6/7 (0.77) compared to those for SmB/D3 (0.72) and SmD1/D2 (0.73).

A structural similarity to Sm proteins has also been noted for the SMN tudor domain, which lacks the N-terminal helix and fifth β strand found in the Sm folds ([Selenko et al., 2001](#)). Unlike the Sm proteins, however, the SMN tudor domain is monomeric in solution, which may reflect its inability to form a homomeric β 4– β 5 interface. In principle, the tudor domain could bind to proteins with Sm folds via the β 4 edge of its β sheet to

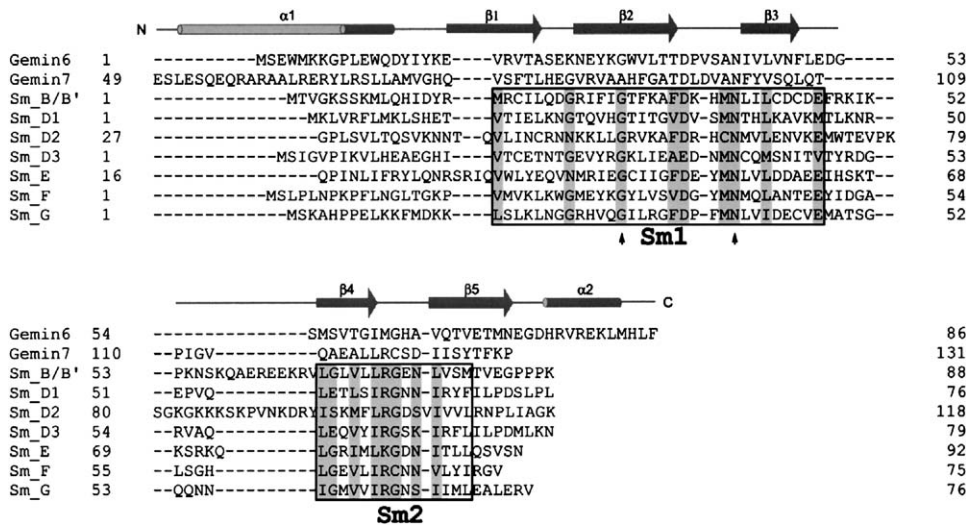


Figure 3. Sequence Alignment of Gemin6, Gemin7, and the Human Sm Proteins

The seven human Sm proteins are aligned based on sequence homology, with shaded regions indicating $\geq 57\%$ identity among the sequences shown. Strictly conserved Gly and Asn are indicated with arrowheads. Sm1 and Sm2 motifs are boxed. Gemin6 and Gemin7 sequences are aligned with the human Sm protein sequences based on structural similarity. Secondary structure elements of Gemin6 and Gemin7 are shown above the sequences.

form interfaces that resemble those discussed above for Sm proteins and for Gemin6/7. So far, there is no evidence for this type of “ $\beta 4$ - $\beta 5$ ” interaction between the SMN tudor domain and Sm proteins. Instead, the SMN tudor domain has been shown to interact with the RG-rich tails in SmB, SmD1, and SmD3, and symmetrical dimethylation of arginine residues in these tails is important for high affinity binding (Brahms et al., 2001; Selenko et al., 2001).

The Gemin6/Gemin7 Complex Binds Sm Proteins

Both Gemin6 and Gemin7 have been shown previously to interact in vitro with the human Sm proteins, and the similarity between Gemin6/7 and the Sm proteins immediately suggests how they are likely to interact. In principle, however, there are two distinct binding surfaces on both Gemin6 and Gemin7 that could be used

to bind Sm proteins (Figure 5). The $\beta 5$ surface of Gemin6 and the $\beta 4$ surface of Gemin7 are both exposed and available for interaction with Sm proteins in the Gemin6/7 complex, whereas the $\beta 4$ surface of Gemin6 and $\beta 5$ surface of Gemin7 would only be available to interact with Sm proteins if the Gemin6/7 complex were disrupted. If Sm proteins bind only to the $\beta 4$ surface of Gemin6 and $\beta 5$ surface of Gemin7 when the proteins are assayed independently, then we would expect Sm binding to Gemin6 and Gemin7 to be abolished upon Gemin6/7 heterodimer formation.

To determine if the Gemin6/7 complex is capable of binding Sm proteins, we carried out in vitro binding experiments comparing Sm proteins binding to Gemin6 alone with Sm proteins binding to the Gemin6/7 complex. In these experiments, in vitro translated [^{35}S]-Met labeled Sm proteins were incubated with Gemin6/7

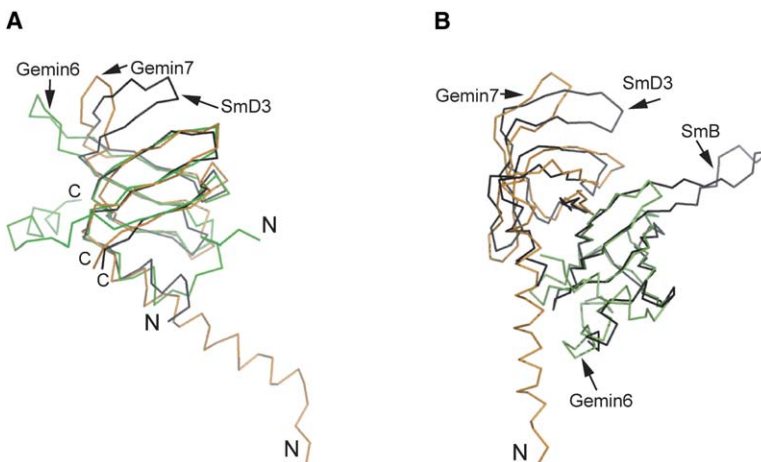


Figure 4. Superposition of Gemin6, Gemin7, and the Human Sm Proteins

(A) Superposition of α backbones of human SmD3 (black), Gemin6 (green), and Gemin7 (orange). The variable $\beta 3$ - $\beta 4$ regions are indicated.

(B) Superposition of the Gemin6/7 complex (green and orange) with the human SmD3/B complex (black).

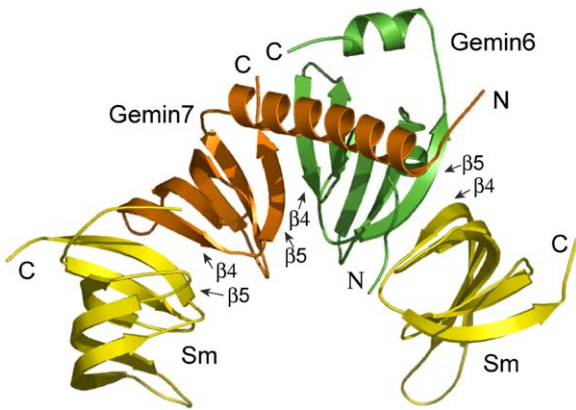


Figure 5. Speculative Model for Gemin6/7 Complex Interactions with Sm Proteins

SmD3 (yellow) was docked to either end of the Gemin6/7 complex (green and gold), assuming that the SmD3/Gemin6 and SmD3/Gemin7 interfaces would be similar to those observed in SmD3/SmB and SmD1/SmD2. No significant steric clashes are present in the modeled interfaces. Three additional Sm proteins can be readily added to the model to form a closed heptameric ring, as previously shown for the Sm proteins (Kambach et al., 1999).

complex immobilized on glutathione agarose beads via the GST-Gemin6 fusion. After extensive washing, bound proteins were stripped from the beads and analyzed by SDS-PAGE. As shown in Figure 6, the Gemin6/7 complex binds with highest affinity to Sm B, SmD2, SmD3, and SmE. Only weak binding was observed for SmD1, SmF, and SmG, and no binding was observed for immobilized GST alone (data not shown). Because the tandem affinity procedure used here to purify the Gemin6/7 heterodimer provides a stoichiometric 1:1 complex, and because we have so far not observed any dissociation of the Gemin6/7 complex under stringent washing conditions, we are confident that the interactions observed with Sm proteins are with the intact Gemin6/7 complex and not with immobilized Gemin6 that has lost Gemin7. These results are consistent with a model in which the Gemin6/7 complex can bind Sm proteins via the exposed $\beta 5$ and $\beta 4$ strands of Gemin6 and Gemin7, respectively (Figure 5).

The binding of Sm proteins to the Gemin6/7 complex

shows a similar pattern of relative affinities as that observed for Gemin6 alone, although the binding is somewhat weaker for the Gemin6/7 complex. This difference in binding affinities suggests that the $\beta 4$ surface of Gemin6 may bind efficiently to Sm proteins in the absence of Gemin7, leading to a higher overall affinity than for the Gemin6/7 complex. It is also possible that the N terminus of Gemin7 modulates the binding to Sm proteins through the $\beta 5$ interface of Gemin6 (see Figure 5). Interestingly, even the weakest binding observed between Sm proteins and Gemin6 and Gemin6/7 complex in this experiment may be relevant in vivo. Coexpression experiments in bacteria have revealed that GST-Gemin6 forms a complex with hexahistidine-tagged SmD1 (which binds weakest among the Sm proteins to Gemin6 and Gemin6/7) when the two proteins are coexpressed, and the resulting Gemin6/SmD1 complex is stable through both glutathione agarose and Ni-NTA agarose purification steps (data not shown).

Discussion

Based on the crystal structures of the human Sm D3/B and Sm D1/D2 complexes, Nagai and colleagues proposed that human Sm proteins form heptameric rings when assembled on the Sm sites of U snRNAs (Kambach et al., 1999). This model is supported by negative-stain electron microscopy data of U1 snRNP, which reveals a doughnut-shaped core that could readily accommodate the seven Sm proteins (Stark et al., 2001). Further support comes from recently described crystal structures of an archaeal Sm-like protein that forms a homoheptameric ring (Mura et al., 2001; Toro et al., 2001) and the bacterial Sm-like protein Hfq, which forms a homohexameric ring (Schumacher et al., 2002). In this context, the role of the SMN complex in mediating spliceosomal snRNPs assembly is to promote Sm ring formation on Sm sites of the appropriate snRNAs, and to disallow assembly on irrelevant RNA molecules containing uridine-rich stretches that would otherwise resemble Sm sites.

Two important features of the SMN-mediated snRNP assembly process in vivo are the sequestration of Sm proteins by components of the methylosome (Friesen et al., 2001; Meister and Fischer, 2002) and specific re-

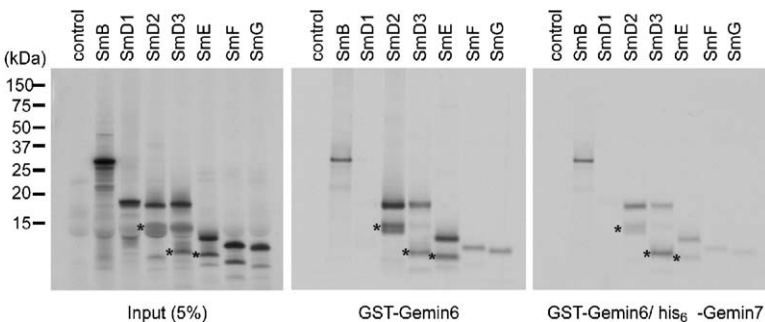


Figure 6. The Gemin6/7 Complex Binds Sm Proteins

The indicated Sm proteins were produced individually by in vitro transcription and translation in rabbit reticulocyte lysates in the presence of [³⁵S]Methionine. In vitro-translated proteins were incubated with 2 μ g of purified recombinant GST-Gemin6 (middle panel) or with GST-Gemin6/Gemin7 complex (right panel) coupled to glutathione beads. After extensive washing, bound proteins were resolved by SDS-PAGE and visualized by fluorography. Input (left panel) represents 5% of the labeled protein used for binding.

Positions of molecular weight markers are indicated and proteolysis products that we believe to be Sm core domains are indicated by asterisks. The fragments of SmD2, SmD3, and SmE also bind to Gemin6 and Gemin6/7 and may be present at higher concentrations in the middle and right panels relative to the input panel because of continued proteolysis during binding and washing. In some cases, the fragments may also bind with enhanced affinity to Gemin6 or Gemin6/7.

cruitment of snRNA molecules by the SMN complex (Yong et al., 2004a). The current view of this process is that Sm proteins are symmetrically dimethylated in their RG-rich repeat regions by the methylosome, which in turn allows them to be recruited to the SMN complex and assembled specifically onto recruited snRNA Sm sites (Meister et al., 2002; Yong et al., 2004b). The roles of individual protein components of the SMN complex have not yet been identified in terms of this assembly process, but Sm protein binding and organization, snRNA recruitment, and Sm loading onto the RNA clearly must be accommodated in any model for understanding the mechanism of ATP-dependent assembly.

An intriguing possibility is that Gemin6 and Gemin7 participate in binding and organizing Sm proteins in preparation for snRNP assembly. The model shown in Figure 5 implies that Gemin6/7 could serve as an Sm dimer surrogate, binding to individual Sm proteins or to Sm subcomplexes to form a partial ring in preparation for snRNA loading. This Gemin6/7/Sm assembly could be an open ring (as shown in Figure 5) or a closed heptameric ring, with Gemin6/7 acting as a temporary Sm dimer. Interestingly, previous work has shown that snRNP core assembly occurs in a distinct, stepwise manner, with SmD1, D2, E, F, and G assembling on the U snRNA Sm site first to form a stable subcore, followed by addition of SmD3/B (Andersen and Zieve, 1991; Raker et al., 1996). The structural comparison of Gemin6/7 with Sm protein heterodimers revealed closest similarity with SmD3/B, supporting the idea that the Sm domains of Gemin6/7 could function as an SmD3/B surrogate during snRNP assembly. Other components of the SMN complex clearly must also contribute to the process of Sm recruitment from the methylosome, which involves recognition of symmetrically dimethylated arginine residues in the RG-rich tails of a subset of the Sm proteins.

An alternative interpretation for the role of Gemin6 and 7 in the SMN complex is that Gemin6 and 7 are Sm-like proteins that are components of a yet-to-be-identified RNP complex. In this case, one might conclude that Gemin6/7 copurifies with the SMN complex because it is an RNP assembly substrate, rather than an assembly factor. Although this possibility cannot be entirely ruled out, we do not favor this viewpoint because Gemin6 and 7 remain associated with the SMN complex at very high stringency—conditions that completely remove the Sm proteins. Furthermore, immunofluorescence microscopy of Gemin6 and 7 in HeLa cells indicates that both proteins colocalize with SMN, with a characteristic diffuse pattern in the cytosol and discreet localization to gems in the nucleus (Baccon et al., 2002; Pellizzoni et al., 2002a). Other components of the SMN complex (SMN and Gemins 2–5) show a similar colocalization pattern (Liu et al., 1997; Charroux et al., 1999; Gubitzi et al., 2002), but Sm proteins and other likely assembly substrates known to interact with SMN, such as GAR1 and fibrillarin, do not show this characteristic cellular colocalization.

The model for Gemin6/7-Sm protein interactions shown in Figure 5 is consistent with what is known about the oligomeric properties of Sm-like proteins, with the types of interactions shown to exist between human Sm proteins, and with the ability of Gemin6 and

7 to form Sm-like interfaces. However, a functional role for Gemin6/7 participating in formation of a ring (or partial ring) of Sm proteins is at this point still speculative. An alternative or additional mode of interaction between Gemin6 and 7 and Sm proteins could involve the RG-rich C-terminal tails of SmB, SmD3, and SmD1, which are known to be important in SMN-Sm protein interactions (Friesen and Dreyfuss, 2000; Brahm et al., 2001). Given the structural similarity between the SMN tudor domain, the Sm proteins, and Gemin6/7, it is plausible that such an interaction could also play a role in Gemin6/7 interacting with Sm proteins. However, we believe it unlikely that the RG-rich tails of Sm proteins play a major role in binding to Gemin6/7 for several reasons. First, neither Gemin6 nor Gemin7 contain an obvious aromatic binding pocket or acidic surface that could correspond to that observed in the SMN tudor domain and which is implicated in interacting with sDMA residues (Sprangers et al., 2003). Second, the binding preferences of Sm proteins to both Gemin6 and Gemin6/7 show no correlation with the presence or absence of tails that would presumably be methylated to some extent in the eukaryotic extracts used in producing Sm proteins by *in vitro* translation. Indeed, SmD2 and SmE bind to Gemin6 and to Gemin6/7 with the highest affinity, yet neither has an RG-rich tail and neither is known to be methylated *in vivo*. Alternatively, SmD1 does have an RG-rich C-terminal tail, yet binds most weakly to both Gemin6 and Gemin6/7. As described above, bacterially expressed SmD1 does form a stable complex with Gemin6, but SmD1 is not methylated in bacteria (Brahm et al., 2000). Although we cannot rule out a possible role for the RG-rich tails of Sm proteins in interacting with Gemin6/7, our current data indicates that this is not likely to be the primary binding mechanism.

There is still a considerable gap in our understanding of how the SMN complex assembles snRNPs on a detailed molecular level. Work toward achieving this goal would be greatly aided by structural models of the intact SMN complex or of each individual component of the complex, along with *in vitro* assays of snRNP assembly that would facilitate testing of mechanistic hypotheses. The crystal structure of the Gemin6/7 complex reported here is an important step toward building a structural model for this system, and together with NMR and crystallographic data for the SMN tudor domain (Selenko et al., 2001; Sprangers et al., 2003), represents the only structural data available at this point for the SMN complex. Major progress has also been made recently in the development of *in vitro* assays for snRNP assembly, where the SMN complex is either derived from cellular extracts or is purified from HeLa cells (Meister et al., 2001; Pellizzoni et al., 2002b). Indeed, preliminary experiments indicate that extracts from cells in which Gemin6 protein levels have been reduced by small interfering RNA molecules are deficient in snRNP assembly (G.D., unpublished data).

The discovery that Gemin6 and 7 have Sm fold structures indicates that this protein superfamily is larger than previously thought and includes proteins that do not contain canonical Sm motifs (Seraphin, 1995; Hermann et al., 1995), including the SMN tudor domain. It is intriguing that the SMN complex might take advan-

tage of the structures and oligomeric properties of Sm superfamily proteins by using an Sm-like assembly factor to interact with and organize Sm protein substrates. With continuing work toward developing a structural model for the SMN complex and improved assays for snRNP assembly, a molecular understanding of how this assembly machine functions may soon be within reach.

Experimental Procedures

Expression and Purification of the Gemin6/7 Complex

A cleavable glutathione-S-transferase (GST) fusion of full length Gemin6 was coexpressed with hexahistidine-tagged Gemin7 (residues 31–131) using compatible T7 expression plasmids in strain BL21(DE3) at 17°C. In both cases, the coding regions were subcloned into vectors containing the respective fusion tags followed by cleavage sites for tobacco etch virus (TEV) protease (Parks et al., 1994). After purification of the complex on glutathione-agarose resin (Sigma) followed by TALON resin (BD Biosciences), both the GST and hexahistidine tags were removed by overnight dialysis at 4°C in the presence of TEV protease. The Gemin6/7 complex was further purified by MonoQ anion-exchange chromatography (Pharmacia), followed by size-exclusion chromatography on a Bio-Rad SEC-250 column. The SEC-250 elution time indicated an apparent molecular weight of 30 kDa relative to calibration with globular protein standards. The Gemin6/7 complex was concentrated to 60 mg/ml in 10 mM sodium HEPES, pH 7.0, 5 mM DTT, using Centricon YM-10 devices (Millipore). Selenomethionine was incorporated into the Gemin6/7 complex by expression in strain B834(DE3) using minimal growth medium containing 100 mg/L D,L-selenomethionine (Sigma).

Analytical Ultracentrifugation

Sedimentation equilibrium experiments were performed at 20°C with an XL-A analytical ultracentrifuge (Beckman) and a Ti An60 rotor with six-channel charcoal-filled epon centerpieces and quartz windows. The Gemin6/7 complex (9.1 μ M, 24.8 μ M, and 122.6 μ M) was centrifuged at 20,000, 25,000 and 30,000 rpm, respectively, for 14 hr, at which time equilibrium was established. Radial absorption scan data at 281 nm, 291 nm, and 297 nm, respectively, for each of the three concentrations could be fit well by a single species model, with an experimental molecular weight of 29.3 kDa (the purified Gemin6/7 complex calculated molecular weight is 30.2 kDa) using ULTRASCAN data analysis software (Demeler, 2003).

Crystallization, Data Collection, and Structure Determination

Crystals of the Gemin6/7 complex were grown at 18°C from hanging drops at a concentration of 20 mg/ml in the presence of 100 mM sodium acetate, pH 4.7, and 12% 2-methyl-2,4-pentanediol (MPD). Crystals were transferred into a cryoprotectant solution containing 100 mM sodium acetate, pH 4.7, 25% MPD before being flash-frozen in liquid nitrogen prior to data collection at 100 K. Two crystal forms were obtained from this condition: form I diffracted to higher resolution, but only form II yielded diffraction-quality crystals using SeMet-containing proteins. The structure was therefore determined initially in crystal form II using multiwavelength phasing from SeMet proteins, and was later refined to high resolution using data from crystal form I. Both forms contain two Gemin6/7 heterodimers in the crystallographic asymmetric unit.

Native data for crystal form I and multiwavelength anomalous scattering data for crystal form II were measured at the Advanced Light Source Beamline 8.2.2, using an ADSC Quantum 315 CCD detector. Diffraction data were processed using the HKL package (Otwinowski and Minor, 1997). Multiwavelength phasing (based on 11 ordered Se sites) and phase improvement were performed using SOLVE/RESOLVE (Terwilliger, 2000). The experimental and solvent-flattened electron density maps were of high quality, and all four protein chains were readily traced, except that the C-terminal 83 residues of Gemin6 did not have interpretable electron density. Iterative rounds of model building with O (Jones et al., 1991) and refinement with REFMAC5 (Murshudov et al., 1999) resulted in a well-

defined structure that was used as a molecular replacement model to compute starting phases for crystal form I. Both heterodimers in the asymmetric unit were readily positioned using AMORE (Navaza, 1994), and REFMAC5 refinement combined with model adjustments rapidly converged to yield the statistics shown in Table 1. The final refined model includes residues 1–86 for both Gemin6 chains, residues 47–131 for one Gemin7 chain, residues 49–131 for the second Gemin7 chain, and 108 solvent molecules modeled as water. All residues have allowed values of (ϕ , ψ) backbone dihedral angles. Structural illustrations and electron density (Figures 1, 2, 4, and 5) were prepared with Pymol (Delano, 2002).

In Vitro Protein Binding Assays

His₆-tagged Gemin7 (residues 31–131) and GST-Gemin6 (full length) were coexpressed as described above. The complex was first purified on TALON resin to remove any excess Gemin6 before it was incubated with glutathione-agarose beads (Sigma) and washed according to the manufacturer's protocol. Protein complexes bound to glutathione-agarose beads were quantified by Coomassie staining of SDS-PAGE gels relative to protein standards. Human Sm proteins were produced by in vitro transcription/translation using rabbit reticulocyte lysate (Promega) in the presence of [³⁵S]methionine according to the vendor's instructions. For binding experiments, 10 μ l of in vitro-translated products were incubated with 2 μ g of either GST, GST-Gemin6, or GST-Gemin6/His₆-Gemin7 (residues 31–131) complex bound to glutathione beads in binding buffer (50 mM Tris-HCl, pH 7.4, 2 mM EDTA, 200 mM NaCl, 0.1% NP40, Roche protease inhibitor cocktail), and mixed for 1 hr at 4°C. Beads were then washed five times with 0.5 ml binding buffer. Bound proteins were eluted in SDS-PAGE sample buffer, resolved by SDS-PAGE, and detected by fluorography.

Acknowledgments

We thank Dr. Mark Lemmon and members of the Dreyfuss and Van Duyne laboratories for stimulating discussions, and Dr. Steve Stayrook for expert crystallographic assistance. Use of the Advanced Light Source Beamlines 8.2.1 and 8.2.2 and assistance from Beamline personnel are gratefully acknowledged. The 8.2.1 and 8.2.2 Beamlines are supported by grants from the Howard Hughes Medical Institute. This work was supported by funding from the Association Française Contre les Myopathies (AFM). G.D. is an Investigator and G.V. an Assistant Investigator of the Howard Hughes Medical Institute.

Received: February 3, 2005

Revised: March 25, 2005

Accepted: March 28, 2005

Published: June 7, 2005

References

- Andersen, J., and Zieve, G.W. (1991). Assembly and intracellular transport of snRNP particles. *Bioessays* 13, 57–64.
- Baccon, J., Pellizzoni, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex. *J. Biol. Chem.* 277, 31957–31962.
- Brahms, H., Raymackers, J., Union, A., de Keyser, F., Meheus, L., and Luhrmann, 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., and Luhrmann, 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.
- Charroux, B., Pellizzoni, L., Perkinson, R.A., Shevchenko, A., Mann, M., and 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, R.A., Yong, J., Shevchenko, A., Mann, M., and 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.
- Collins, B.M., Cubeddu, L., Naidoo, N., Harrop, S.J., Kornfeld, G.D., Dawes, I.W., Curmi, P.M., and Mabbutt, B.C. (2003). Homomeric ring assemblies of eukaryotic Sm proteins have affinity for both RNA and DNA. Crystal structure of an oligomeric complex of yeast SmF. *J. Biol. Chem.* **278**, 17291–17298.
- Delano, W. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).
- Demeler, B. (2003). UltraScan (<http://www.ultrascan.uthscsa.edu>).
- Fischer, U., Liu, Q., and Dreyfuss, G. (1997). The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* **90**, 1023–1029.
- Friesen, W.J., and 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, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001). The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol. Cell Biol.* **21**, 8289–8300.
- Gubitz, A.K., Mourelatos, Z., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J. Biol. Chem.* **277**, 5631–5636.
- Gubitz, A.K., Feng, W., and Dreyfuss, G. (2004). The SMN complex. *Exp. Cell Res.* **296**, 51–56.
- Hannus, S., Buhler, D., Romano, M., Seraphin, B., and 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.
- Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahm, H., and Luhrmann, R. (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J.* **14**, 2076–2088.
- Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
- Jones, T.A., Zou, J.Y., and Cowan, S.W. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A.* **47**, 110–119.
- Jones, K.W., Gorzyski, K., Hales, C.M., Fischer, U., Badbanchi, F., Terns, R.M., and Terns, M.P. (2001). Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *J. Biol. Chem.* **276**, 38645–38651.
- Kambach, C., Walke, S., Young, R., Avis, J.M., de la Fortelle, E., Raker, V.A., Luhrmann, R., Li, J., and Nagai, K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**, 375–387.
- Kleinschmidt, A.M., Patton, J.R., and Pederson, T. (1989). U2 small nuclear RNP assembly in vitro. *Nucleic Acids Res.* **17**, 4817–4828.
- Lawrence, M.C., and Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950.
- Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., and Zeviani, M. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**, 155–165.
- Lehmeier, T., Foulaki, K., and Luhrmann, R. (1990). Evidence for three distinct D proteins, which react differentially with anti-Sm autoantibodies, in the cores of the major snRNPs U1, U2, U4/U6 and U5. *Nucleic Acids Res.* **18**, 6475–6484.
- Liu, Q., and 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., and 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.
- Mattaj, I.W., and De Robertis, E.M. (1985). Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. *Cell* **40**, 111–118.
- Meister, G., and Fischer, U. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J.* **21**, 5853–5863.
- Meister, G., Buhler, D., Pillai, R., Lottspeich, F., and Fischer, U. (2001). A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs. *Nat. Cell Biol.* **3**, 945–949.
- Meister, G., Eggert, C., and Fischer, U. (2002). SMN-mediated assembly of RNPs: a complex story. *Trends Cell Biol.* **12**, 472–478.
- Melki, J. (1997). Spinal muscular atrophy. *Curr. Opin. Neurol.* **10**, 381–385.
- Miguel-Aliaga, I., Culetto, E., Walker, D.S., Baylis, H.A., Sattelle, D.B., and Davies, K.E. (1999). The *Caenorhabditis elegans* orthologue of the human gene responsible for spinal muscular atrophy is a maternal product critical for germline maturation and embryonic viability. *Hum. Mol. Genet.* **8**, 2133–2143.
- Miguel-Aliaga, I., Chan, Y.B., Davies, K.E., and van den Heuvel, M. (2000). Disruption of SMN function by ectopic expression of the human SMN gene in *Drosophila*. *FEBS Lett.* **486**, 99–102.
- Mourelatos, Z., Abel, L., Yong, J., Kataoka, N., and Dreyfuss, G. (2001). SMN interacts with a novel family of hnRNP and spliceosomal proteins. *EMBO J.* **20**, 5443–5452.
- Mura, C., Cascio, D., Sawaya, M.R., and Eisenberg, D.S. (2001). The crystal structure of a heptameric archaeal Sm protein: implications for the eukaryotic snRNP core. *Proc. Natl. Acad. Sci. USA* **98**, 5532–5537.
- Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 247–255.
- Navaza, J. (1994). AmoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
- Owen, N., Doe, C.L., Mellor, J., and Davies, K.E. (2000). Characterization of the *Schizosaccharomyces pombe* orthologue of the human survival motor neuron (SMN) protein. *Hum. Mol. Genet.* **9**, 675–684.
- Parks, T.D., Leuther, K.K., Howard, E.D., Johnston, S.A., and Dougherty, W.G. (1994). Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal. Biochem.* **216**, 413–417.
- Paushkin, S., Charroux, B., Abel, L., Perkinson, R.A., Pellizzoni, L., and Dreyfuss, G. (2000). The survival motor neuron protein of *Schizosaccharomyces pombe*: conservation of survival motor neuron interaction domains in divergent organisms. *J. Biol. Chem.* **275**, 23841–23846.
- Paushkin, S., Gubitz, A.K., Massenet, S., and Dreyfuss, G. (2002). The SMN complex, an assemblysome of ribonucleoproteins. *Curr. Opin. Cell Biol.* **14**, 305–312.
- Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* **95**, 615–624.
- Pellizzoni, L., Charroux, B., and Dreyfuss, G. (1999). SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. *Proc. Natl. Acad. Sci. USA* **96**, 11167–11172.
- Pellizzoni, L., Baccon, J., Charroux, B., and Dreyfuss, G. (2001a). The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Curr. Biol.* **11**, 1079–1088.
- Pellizzoni, L., Charroux, B., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001b). A functional interaction between the survival motor neuron complex and RNA polymerase II. *J. Cell Biol.* **152**, 75–85.
- Pellizzoni, L., Baccon, J., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002a). Purification of native survival of motor neurons com-

plexes and identification of Gemin6 as a novel component. *J. Biol. Chem.* **277**, 7540–7545.

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

Raker, V.A., Plessel, G., and Luhrmann, 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.

Raker, V.A., Hartmuth, K., Kastner, B., and Luhrmann, R. (1999). Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol. Cell. Biol.* **19**, 6554–6565.

Rice, P.A., and Steitz, T.A. (1994). Model for a DNA-mediated synaptic complex suggested by crystal packing of gamma delta resolvase subunits. *EMBO J.* **13**, 1514–1524.

Schrank, B., Gotz, R., Gunnensen, J.M., Ure, J.M., Toyka, K.V., Smith, A.G., and Sendtner, M. (1997). Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc. Natl. Acad. Sci. USA* **94**, 9920–9925.

Schumacher, M.A., Pearson, R.F., Moller, T., Valentin-Hansen, P., and Brennan, R.G. (2002). Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J.* **21**, 3546–3556.

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

Seraphin, B. (1995). Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *EMBO J.* **14**, 2089–2098.

Sprangers, R., Groves, M.R., Sinning, I., and 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.

Stark, H., Dube, P., Luhrmann, R., and Kastner, B. (2001). Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* **409**, 539–542.

Sumpter, V., Kahrs, A., Fischer, U., Kornstadt, U., and Luhrmann, R. (1992). In vitro reconstitution of U1 and U2 snRNPs from isolated proteins and snRNA. *Mol. Biol. Rep.* **16**, 229–240.

Terwilliger, T.C. (2000). Maximum likelihood density modification. *Acta Crystallogr. D55*, 965–972.

Toro, I., Thore, S., Mayer, C., Basquin, J., Seraphin, B., and Suck, D. (2001). RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J.* **20**, 2293–2303.

Wang, J., and Dreyfuss, G. (2001). A cell system with targeted disruption of the SMN gene: functional conservation of the SMN protein and dependence of Gemin2 on SMN. *J. Biol. Chem.* **276**, 9599–9605.

Will, C.L., and Luhrmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.* **13**, 290–301.

Yong, J., Golembe, T.J., Battle, D.J., Pellizzoni, L., and Dreyfuss, G. (2004a). snRNAs contain specific SMN-binding domains that are essential for snRNP assembly. *Mol. Cell. Biol.* **24**, 2747–2756.

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

Accession Codes

The refined atomic coordinates for the Gemin6/Gemin7 complex have been deposited in the Protein Data Bank, with accession code 1Y96.