

Identification and Characterization of Gemin7, a Novel Component of the Survival of Motor Neuron Complex*

Received for publication, April 10, 2002, and in revised form, June 10, 2002
Published, JBC Papers in Press, June 13, 2002, DOI 10.1074/jbc.M203478200

Jennifer Baccon‡, Livio Pellizzoni‡, Juri Rappsilber§¶, Matthias Mann§, and Gideon Dreyfuss‡||

From the ‡Howard Hughes Medical Institute and Department of Biochemistry & Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6148 and §Protein Interaction Laboratory, Center for Experimental Bioinformatics and Department of Biochemistry & Molecular Biology, University of Southern Denmark, DK-5230 Odense M, Denmark

The survival of motor neurons (SMN) protein is the product of the gene mutated or deleted in the neurodegenerative disease, spinal muscular atrophy. SMN is part of a large macromolecular complex that also contains Gemin2, Gemin3, Gemin4, Gemin5, and Gemin6. The SMN complex functions in the assembly of spliceosomal small nuclear ribonucleoproteins and probably other ribonucleoprotein particles. We have identified a novel protein component of the SMN complex termed Gemin7 using native purified SMN complexes and peptide sequencing by mass spectrometry. Coimmunoprecipitation and immunolocalization experiments demonstrate that Gemin7 is a component of the SMN complex and colocalizes with SMN in the cytoplasm and in gems. Binding experiments show that Gemin7 interacts directly with SMN and Gemin6 and mediates the association of Gemin6 with the SMN complex. The amino acid sequence of Gemin7 does not contain any recognizable motifs with the exception of several arginine and glycine repeats that are necessary for its interaction with SMN. Moreover, Gemin7 interacts with several Sm proteins of spliceosomal small nuclear ribonucleoproteins, in particular, with SmE. With the identification of Gemin7, the inventory of the core components of the SMN complex appears essentially complete.

mice, chicken DT40 cells, *Drosophila melanogaster*, *Caenorhabditis elegans*, and *Schizosaccharomyces pombe* (3–9).

SMN associates with Gemin2 (10), Gemin3 (11), Gemin4 (12), Gemin5 (13), Gemin6, and at least one hitherto unidentified component to form a large multi-protein complex (14). The SMN complex is found both in the cytoplasm and in the nucleus where it is concentrated in gems, nuclear bodies similar in size and number to Cajal bodies (coiled bodies (CBs)) and often associated with them (15). A number of additional proteins have been identified that interact directly with SMN and/or the Gemins but are not integral components of the SMN complex described above. These proteins, referred to as substrates in contrast to the more tenaciously associated integral components of the SMN complex, include the Sm and Lsm proteins of the spliceosomal snRNPs (10, 16–18), the small nucleolar RNP proteins fibrillarin and GAR1 (19, 20), RNA helicase A (21), and the human RNP proteins hnRNPQ (22, 23), hnRNPU (15), and hnRNPR (23). Other proteins including coilin (24) and p53 (25) have also been reported to interact with SMN. Importantly, SMN, but not SMN mutants found in SMA patients, interacts directly with its substrates (16–24), strongly suggesting a relationship between the SMA phenotype and a defective SMN interaction with these proteins. Most of the characterized substrates contain arginine- and glycine-rich (RG) domains that are necessary for their direct interaction with SMN (18–22). Specific arginine residues of these domains are dimethylated *in vivo* (26, 27), and this modification dramatically enhances the affinity for SMN, indicating that SMN preferentially interacts with posttranslationally modified RG-rich protein targets (26, 28). This modification is carried out by a macromolecular complex containing the arginine methyltransferase JBP1 (PRMT5), pICln, and MEP50 (29–31). Previous studies have provided evidence that the SMN complex functions in the assembly/restructuring of spliceosomal snRNPs and in pre-mRNA splicing and transcription (17, 21, 32–36). The SMN complex may also play a role in the metabolism of small nucleolar RNPs involved in the processing and modification of ribosomal RNA (19, 20).

The determination of the full complement of components of the SMN complex and its substrates will facilitate the mechanistic analysis of the activities of the SMN complex. A thorough knowledge of how this macromolecular machine participates in the assembly and function of several RNP complexes should lead to an understanding of the underlying pathogenesis of SMA. To identify all of the components of the SMN complex and to characterize its biochemical, structural, and functional properties, we developed a system for the purification of native SMN complexes from HeLa cells (14). From these purified SMN complexes, novel proteins were identified by peptide microsequencing by mass spectrometry. Here, we characterize what is apparently the last of the components of the SMN complex,

Spinal muscular atrophy (SMA),¹ an autosomal recessive disease characterized by degeneration of motor neurons of the spinal cord resulting in muscular atrophy, is the most common genetic cause of infant mortality, affecting ~1 in 6000 newborns (1). Over 95% of SMA patients harbor deletions or mutations in the survival of motor neurons (SMN) gene, the SMA-determining gene (2). SMN is essential for viability in all eukaryotic organisms and cell types tested including humans,

* This work was supported in part by grants from the National Institutes of Health (to G. D.), the Association Française contre les Myopathies (to G. D.), and the Danish National Research Foundation to the Center for Experimental Bioinformatics. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY114106.

¶ Marie Curie Fellow.

|| Investigator of the Howard Hughes Medical Institute. To whom correspondence should be addressed. Tel.: 215-898-0398; Fax: 215-573-2000; E-mail: gdreyfuss@hhmi.upenn.edu.

¹ The abbreviations used are: SMA, spinal muscular atrophy; SMN, survival of motor neurons; RNP, ribonucleoprotein; snRNP, small nuclear RNP; hnRNP, human RNP; RG, arginine-glycine; YG, tyrosine-glycine; GST, glutathione S-transferase; TRITC, tetramethylrhodamine isothiocyanate.

termed Gemin7. We show that Gemin7 associates *in vivo* with the SMN complex and is present in gems. These features are characteristic of the SMN complex components. Gemin7 interacts directly with Gemin6 and SMN *in vitro* and bridges the interaction between Gemin6 and SMN. The evolutionarily conserved YG box of SMN is important for efficient interaction with Gemin7. Additionally, Gemin7 associates with a unique subset of Sm proteins. Interestingly, Gemin7 is the first integral component of the SMN complex found to contain RG repeats. The amino terminus of Gemin7 containing these RG repeats is necessary but not sufficient for the interaction between Gemin7 and SMN.

EXPERIMENTAL PROCEDURES

Affinity Purification of the SMN Complex—The SMN complex was purified from HeLa Tet-ON cells stably expressing FLAG-tagged Gemin2 under an inducible promoter as described previously (14). Washes in RSB-500 (10 mM Tris-HCl, pH 7.4, 2.5 mM MgCl₂, 500 mM NaCl) containing 0.02% Nonidet P-40 were performed prior to elution from the anti-FLAG affinity resin using a low pH elution buffer (Sigma). The purified complex was analyzed by SDS-PAGE and silver staining.

Protein Sequencing by Mass Spectrometry—Bands were excised from a single one-dimensional or two-dimensional silver-stained polyacrylamide gel and in gel digested with trypsin as described previously (37). Tryptic peptides in the supernatant were analyzed by nanoelectrospray tandem mass spectrometry as described previously (38) using a QSTAR (MDS Sciex, Toronto, Canada) equipped with a nanoelectrospray ion source (MDS Proteomics, Odense, Denmark). Comprehensive protein and expressed sequence tag databases were searched using the protein and peptide software suite (MDS Proteomics). No limitations on protein molecular weight and species of origin were imposed.

DNA Constructs—The cDNA corresponding to Gemin7 was generated by PCR amplification from the Marathon-Ready cDNA library (CLONTECH). The insert was cloned downstream of the cytomegalovirus promoter into a modified pcDNA3 vector (Invitrogen) containing either the *myc*-tag or the FLAG-tag sequences at the amino terminus. Gemin7 wild-type (wt) and the deletion of the amino-terminal 32 amino acids (Δ N32) were cloned into the pet28 vector (Novagen) to obtain the amino-terminal His-tagged fusion. Gemin6 and Gemin7 were cloned into a petGST-TEV vector to obtain the amino-terminal GST fusion containing a TEV protease site upstream of the cDNA insert. The constructs for *myc*-tagged SMN, Gemin2, Gemin3, Gemin4, Gemin6, his-tagged SMN wild type and mutants, FLAG-tagged Gemin2, V5-tagged Gemin5, GST-tagged Gemin6, his-tagged Gemin6, and *myc*-tagged Sm proteins were described previously (10–14). All of the constructs were analyzed by automated DNA sequencing.

Antibodies—Antibodies used in these experiments were as follows: 2B1, mouse IgG1 monoclonal anti-SMN (15); 2E17, mouse IgG1 monoclonal anti-Gemin2 (10); 12H12, mouse IgG1 monoclonal anti-Gemin3 (11); 17D10, mouse IgG1 monoclonal anti-Gemin4; 10G11, mouse IgG1 monoclonal anti-Gemin5 (13); rG6, affinity-purified rabbit polyclonal anti-Gemin6 (36); 4B10, mouse IgG2a monoclonal anti-hnRNP1 (39); R288, rabbit polyserum anti-p80 coilin (40); 9E10, mouse IgG1 monoclonal anti-Myc; A-14, affinity-purified rabbit polyserum anti-Myc 9E10 epitope (Santa Cruz Biotechnology); and M2, mouse IgG1 monoclonal anti-FLAG (Sigma).

Cell Culture and Immunoprecipitation Experiments—HeLa PV or 293T cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen). Cells

were transfected by the standard calcium phosphate method. Following overnight incubation with DNA, cells were washed, and fresh medium was added. Immunoprecipitation experiments were performed from 293T total cell extracts followed by Western blotting as described previously (39).

Immunofluorescence Microscopy—HeLa PV cells plated on glass coverslips were briefly washed with PBS, fixed in 2% formaldehyde-PBS for 30 min at room temperature and permeabilized in 0.5% Triton X-100-PBS for 5 min at room temperature. Cells were blocked in 3% bovine serum albumin in PBS for 1 h at room temperature. Double-label immunofluorescence experiments were performed by separate sequential incubations of each primary antibody diluted in PBS containing 3% bovine serum albumin followed by the specific secondary coupled to fluorescein isothiocyanate or TRITC. All of the incubations were at room temperature for 1 h. Indirect epifluorescence microscopy was performed with a Nikon Eclipse E800 microscope. Digital images were collected with a Cook Sencis high performance camera and processed with the IP lab software.

In Vitro Binding Experiments—Recombinant GST, GST-Gemin6, and GST-Gemin7 proteins were expressed in *Escherichia coli* BL21 (DE3) and purified by affinity chromatography on glutathione-Sepha-

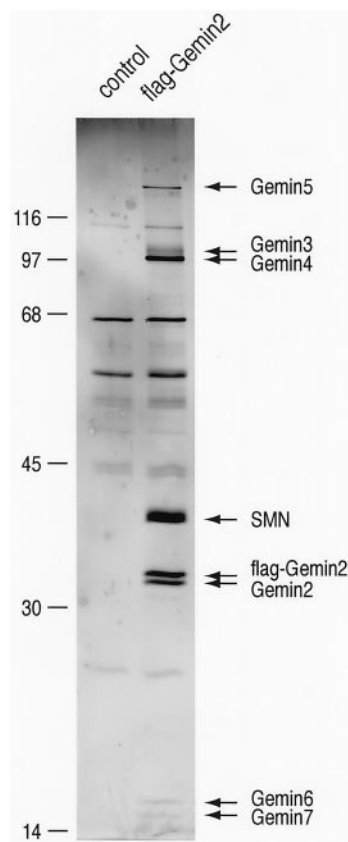


FIG. 1. **Identification of Gemin7.** SDS-PAGE and silver staining of the SMN complex purified from cells stably expressing FLAG-Gemin2 (*flag-Gemin2*) and control purification from the parental Tet-ON cell line (*control*). The integral SMN complex components are indicated, and molecular mass standards are shown on the left.

FIG. 2. **Alignment of Gemin7 protein sequences.** Protein sequences corresponding to human Gemin7 (GenBankTM accession number AY114106) and its putative mouse ortholog (GenBankTM accession number BAB26836) were aligned using ClustalW alignment of the MacVector 6.5 program. Dark gray boxes indicate identical amino acids; light gray boxes indicate similar amino acids.

human	1	M Q T P V N I P V P V - - L R L P R G P D G F S R G F A P D G R R A P	33
mouse	1	M Q S P L T I P V P V P V L R L P R G P D G F S R G F A S D G R R T I	35
human	34	L R P E V P E I Q E C P I A Q E S L E S Q E Q R A R A A L R R E Y L R	68
mouse	36	L R P E V G E G - - - H I Q D P P E S Q E Q R A R A T L R R E Y L R	66
human	69	S L L A M V G H Q V S F T L H E G V R V A A H F G A T D L D V A N F Y	103
mouse	67	S L L A M V G H P V S F T L H E G V H V T A Q F G A T D L D V A N F Y	101
human	104	V S Q L Q T P I G V Q A E A L L R C S D I I S Y T F K P	131
mouse	102	V S Q L Q T P I G V Q A E A L L R C S D I I S Y S F K L	129

rose beads (Amersham Biosciences) according to the manufacturer's instructions. GST-TEV-Gemin6 was coexpressed with either His-tagged Gemin7 wt or Δ N32 and purified in the same manner. *In vitro* binding experiments were carried out using purified recombinant proteins (5 μ g) bound to glutathione-Sepharose beads, and the indicated proteins translated *in vitro* in the presence of [³⁵S]methionine (Amersham Biosciences) as described previously (16). His-Gemin7 was translated *in vitro* in the presence of [³⁵S]methionine + [³⁵S]cysteine. Direct *in vitro* binding experiments were carried out using 2 μ g of purified

recombinant His-tagged SMN wild-type and mutants or Gemin6 proteins.

RESULTS AND DISCUSSION

Identification of Gemin7 by Mass Spectrometry Analysis of Purified SMN Complexes—We have previously described a system that is suitable for the purification of native SMN complexes containing all of the core components from cell lines expressing FLAG-tagged SMN or Gemin2 (14). Here, we purified the SMN complex from FLAG-Gemin2 extracts using anti-FLAG affinity resin and high stringency salt washes. The purified SMN complex was analyzed by SDS-PAGE and silver staining along with nonspecific proteins purified from extracts of the parental HeLa Tet-ON cell line as a comparison (Fig. 1). Using this purification scheme, the integral SMN complex components remain associated, whereas other SMN interactors including the Sm proteins are either barely detectable or completely absent. An analysis by peptide sequencing of these purified SMN complexes has led to the identification of novel proteins including Gemin6 and a 15-kDa protein (14). Here, we present the characterization of the 15-kDa protein component of the SMN complex, which we term Gemin7, previously indicated as SIP3 (Fig. 1) (10, 14). The band corresponding to this protein was excised from the gel and digested with trypsin, and the resulting peptides were sequenced by nano-electrospray mass spectrometry as described under "Experimental Procedures." A peptide sequence (RRAPLRPEVPEIQECPIAQESLESQEQR) was obtained that exactly matched several expressed sequence tag sequences in the data base, and the corresponding cDNA was cloned by PCR amplification from a cDNA library. Gemin7 cDNA contains both start and stop codons, and an in-frame stop codon is present upstream of the initial methionine indicating that the cDNA represents the full-length transcript. Gemin7 cDNA encodes a protein of 131 amino acids of predicted molecular mass (14.5 kDa) and isoelectric point (pI = 7.27) that is consistent with the size and the pI of the protein eluted and analyzed by mass spectrometry (Fig. 2) (see also Ref. 14). Data base searches identified a putative murine ortholog of Gemin7 (identity 82%, BAB26836). The alignment of the human and mouse Gemin7 is presented in Fig. 2. No obvious candidate orthologs were found in other organisms where SMN is present such as *C. elegans*, *S. pombe*, and *D. melanogaster* (4–8), suggesting that the sequences of the homologs may be evolutionarily divergent, or alternatively, Gemin7 is not present in these organisms. Gemin7 does not contain any known motifs that may suggest possible functions;

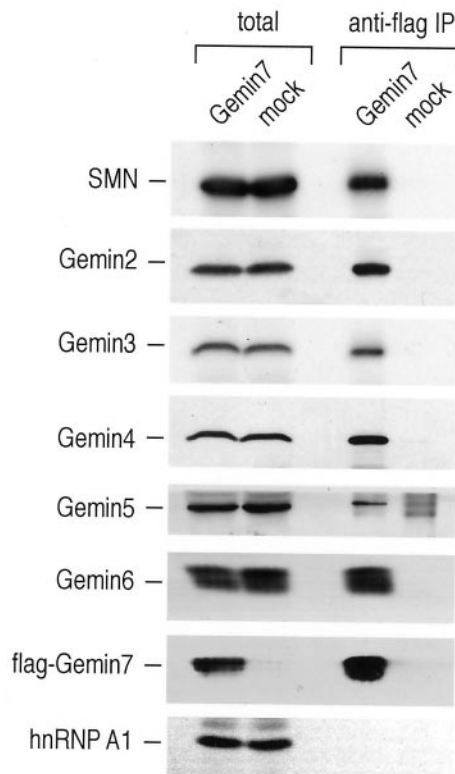
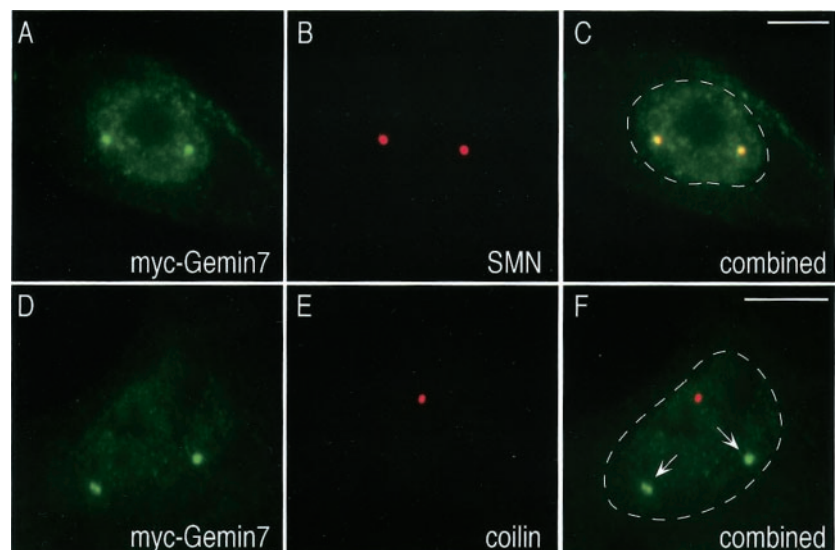


FIG. 3. Gemin7 is a component of the SMN complex. 293T cells were transiently transfected with a pcDNA3 vector encoding FLAG-tagged Gemin7 (*Gemin7*) or empty pcDNA3 vector (*mock*). Total extracts were prepared from these cells and analyzed by immunoprecipitation using anti-FLAG antibodies (*anti-flag IP*). Bound proteins were analyzed by SDS-PAGE, and Western blotting with antibodies against the proteins is indicated on the left. The total represents 5% of the extract used for the immunoprecipitation. The *Gemin5* panel shows some background bands, because a long exposure is usually required using this antibody to detect immunoprecipitated Gemin5 by Western blotting.

FIG. 4. Gemin7 localizes to gems. HeLa PV cells were transiently transfected with a pcDNA3 vector encoding *myc*-tagged Gemin7 (*myc-Gemin7*) and analyzed by indirect immunofluorescence and epifluorescence microscopy. *A–C*, double labeling experiments of cells expressing *myc-Gemin7* using anti-Myc (*A*) or anti-SMN (*B*) antibodies. *D–F*, double labeling experiments of cells expressing *myc-Gemin6* using anti-Myc (*D*) or anti-coilin (*E*) antibodies. Combined images are shown in *C* and *F*. Arrows indicate gems that are distinct from CBs. The intensity of the signal of the cytoplasm is underestimated compared with that of the nucleoplasm because of the focal plane that was chosen to show gems and CBs. Dotted lines demarcate the nucleus. Scale bar represents 5 μ m.



however, it does contain several arginine and glycine residues, two RG pairs and one Gly-Arg-Arg, at the amino terminus.

Gemin7 Is a Novel Component of the SMN Complex—The observation that Gemin7 copurifies with SMN under highly stringent conditions indicates that it is a *bona fide* component of the SMN complex. To confirm this observation, we transiently transfected 293T cells with a vector encoding a FLAG-tagged Gemin7 fusion protein (Gemin7) or with an empty vector as a control (mock). Immunoprecipitation using anti-FLAG antibodies was performed on total extracts, and bound proteins were analyzed by SDS-PAGE and Western blotting. Fig. 3 shows that FLAG-Gemin7 is expressed and immunoprecipitated efficiently by anti-FLAG antibodies. All of the components of the SMN complex, SMN, Gemin2, Gemin3, Gemin4, Gemin5, and Gemin6, coimmunoprecipitated with FLAG-Gemin7. In contrast, the abundant protein hnRNPA1 is not immunoprecipitated with FLAG-Gemin7, indicating that the immunoprecipitation is specific. None of the SMN complex components is immunoprecipitated from a control extract that does not express Gemin7. These results demonstrate that Gemin7 is a novel component of the SMN complex.

Gemin7 and SMN Colocalize in Gems—SMN localizes in both the cytoplasm and the nucleus where it is highly concentrated in gems (15). Gems are nuclear bodies that are similar in size and number to CBs and can be found, depending on the cell type or tissue analyzed, either separated or associated with them (15, 41–43). All of the components of the SMN complex including Gemin2, Gemin3, Gemin4, Gemin5, and Gemin6 have an identical subcellular localization to that of SMN (10–15). In contrast, most of the substrates of the SMN complex such as snRNP and small nucleolar RNP proteins localize elsewhere in the cells including CBs but not gems when these structures are found separated (15). We examined the subcellular localization of Gemin7 by indirect immunofluorescence microscopy on HeLa PV cells transiently transfected with Myc-tagged Gemin7. In this cell line, gems and CBs are found mostly separated (15). Fig. 4A shows that in addition to diffuse cytoplasmic and nucleoplasmic distribution, Gemin7 localizes to discrete nuclear domains. To distinguish between gems and CBs, we carried out double labeling experiments with anti-SMN and anti-coilin antibodies. The Gemin7 colocalizes perfectly with SMN in gems (Fig. 4, A–C) but not with coilin (Fig. 4, D–F), the marker of CBs (40). Importantly, Gemin7 staining without coilin can be observed in transfected cells. We conclude that Gemin7 is a novel component of the SMN complex and colocalizes with other components of the SMN complex in gems.

Gemin7 Interacts with SMN and Gemin6—We further studied the interactions of Gemin7 *in vitro* with the other known components of the SMN complex using SMN, Gemin2, Gemin3, Gemin4, Gemin5, Gemin6, and Gemin7 *in vitro* translated proteins (Fig. 5A). GST-Gemin7 binds efficiently to *in vitro* translated Gemin6 but not to any of the other complex components. No binding to GST was observed (data not shown). The interaction of Gemin7 with Gemin6 is direct as demonstrated by *in vitro* binding between purified recombinant proteins. His-tagged Gemin6 (His-Gemin6) binds efficiently to GST-Gemin7 but does not bind to GST-Gemin6 or to GST alone (Fig. 5B). Upon longer exposure of the gel in Fig. 5A, a weak interaction between Gemin7 and SMN could be detected. Therefore, we tested this interaction using purified recombinant proteins. Purified His-tagged SMN binds directly and efficiently to GST-Gemin7 but not to GST-Gemin6 or GST (Fig. 5C). These experiments demonstrate that Gemin7 binds directly to both SMN and Gemin6.

Gemin7 Is Required for the Interaction between Gemin6 and SMN—Next, we tested whether Gemin7 can interact simulta-

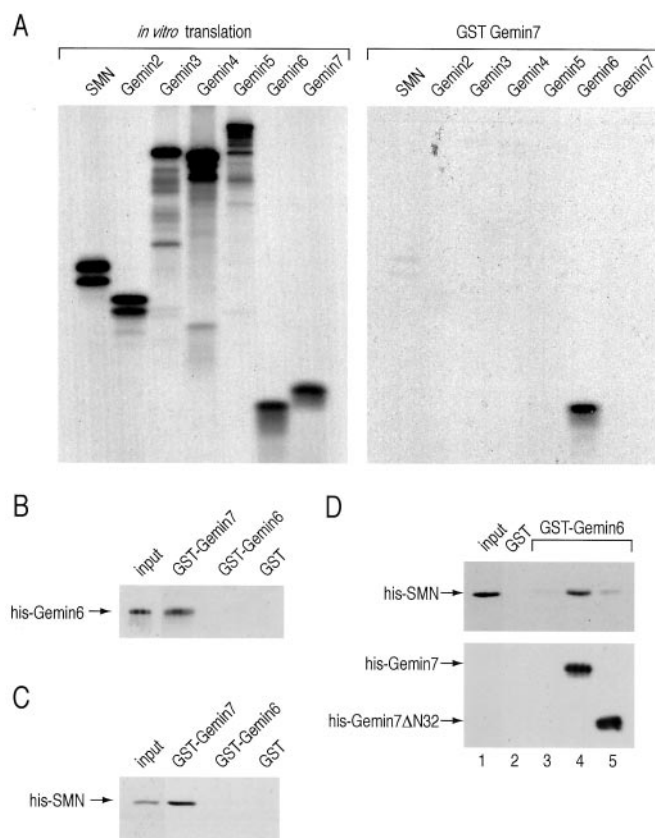


FIG. 5. Gemin7 interacts *in vitro* with Gemin6 and SMN. A, SMN complex components were *in vitro* translated in the presence of [³⁵S]methionine or [³⁵S]methionine + cysteine (Gemin7) and incubated with either GST-Gemin7 or GST (data not shown) recombinant proteins as described under “Experimental Procedures.” Following extensive washes, bound proteins were analyzed by SDS-PAGE and autoradiography. The *in vitro* translation panel shows 5% input. B, recombinantly produced His-Gemin6 was incubated with GST, GST-Gemin6, or GST-Gemin7, and following extensive washes, bound proteins were analyzed by SDS-PAGE and Western blotting. The input shows 20% input. C, recombinantly produced His-SMN was bound to GST-Gemin7 or GST alone and analyzed by SDS-PAGE and Western blotting. The input lane shows 5% input. D, recombinantly produced GST (lane 2), GST-Gemin6 alone (lane 3), and GST-Gemin6 copurified with His-Gemin7 (lane 4) or His-Gemin7ΔN32 (lane 5) were incubated with His-SMN, and following extensive washes, bound proteins were analyzed by SDS-PAGE and Western blotting. The input (lane 1) shows 20% input of His-SMN.

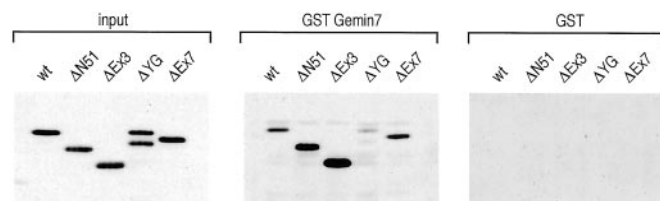


FIG. 6. The YG box of SMN is necessary for interaction with Gemin7. Bacterially produced His-tagged SMN wild-type (wt) and mutants were incubated with bacterially produced GST-Gemin7 or GST alone. Following extensive washes, bound proteins were analyzed by SDS-PAGE and Western blotting with anti-T7 antibodies. The input panel shows 5% input.

neously with both Gemin6 and SMN. For these direct binding experiments, GST-Gemin6 was expressed alone or coexpressed with His-tagged Gemin7 wt or Gemin7 ΔN32. Both the wt and ΔN32 proteins copurified efficiently with Gemin6, indicating that the RG domain of Gemin7 is not necessary for the interaction with Gemin6 (Fig. 5D, lanes 4 and 5, lower panel). GST alone (lane 2) and GST-Gemin6 (lane 3) do not interact with His-tagged SMN, whereas the GST-Gemin6/His-Gemin7 pro-

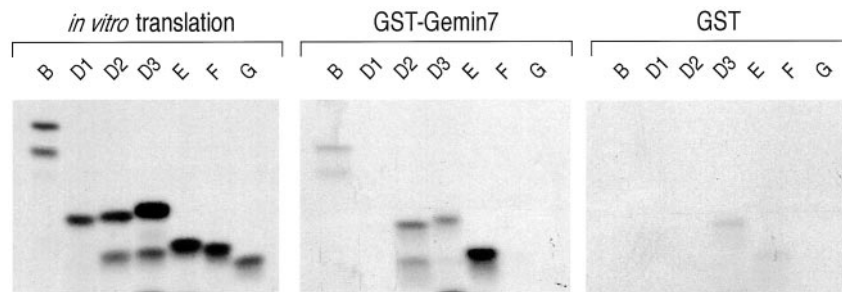


FIG. 7. **Gemin7 interacts *in vitro* with several Sm proteins of spliceosomal snRNPs.** Sm proteins were translated *in vitro* in the presence of [³⁵S]methionine and incubated either with GST-Gemin7 or with GST purified from bacteria as described under "Experimental Procedures." Following extensive washes, bound proteins were analyzed by SDS-PAGE and autoradiography. The *in vitro* translation panel shows 5% input.

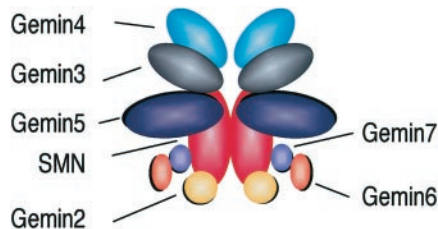


FIG. 8. **The updated SMN complex.** The model depicts the seven known integral components of the SMN complex and their reciprocal interactions in dimeric form.

tein complex interacts efficiently and directly with purified recombinant His-SMN (*lane 4*). This efficient interaction with SMN is impaired when Gemin7 Δ N32 is used for binding (*lane 5*). These results demonstrate that Gemin7 is required for the interaction of Gemin6 with the SMN complex, serving as a bridge to SMN, and that the RG-containing region of Gemin7 is necessary for efficient interaction with SMN. We then tested the RG domain alone of Gemin7 comprising amino acids 1–32 for binding to SMN and found that this region is not sufficient for the Gemin7 interaction with SMN (data not shown). Thus, the RG-rich region of Gemin7 is necessary but not sufficient for direct interaction between Gemin7 and SMN. SmB, SmD1, and SmD3 also require their RG domains for interaction with SMN (18). Furthermore, arginine residues within the RG domains of these Sm proteins are symmetrically dimethylated *in vivo*, and this modification enhances the binding to SMN (26–28). It will be interesting to determine whether Gemin7 receives a similar modification *in vivo*. To identify the region of SMN necessary for its interaction with Gemin7, we tested a panel of recombinantly produced His-tagged SMN wild-type and deletion mutant proteins in direct *in vitro* binding experiments. Recombinant GST-Gemin7 or GST as a control was incubated with SMN wild-type or SMN deleted of the amino-terminal 51 amino acids (Δ N51), the Tudor domain, encoded by the sequence of Exon3 (Δ Ex3), the YG box (Δ YG), or the amino acids encoded by Exon7 (Δ Ex7) (Fig. 6). The deletion of the evolutionarily conserved YG box of SMN impaired Gemin7 binding, whereas all of the other deletion mutants efficiently interacted with Gemin7. Therefore, both the Gemin2 binding domain encoded by Exon1/2A (10) and the Tudor domain are dispensable for SMN interaction with Gemin7. In contrast, the evolutionarily conserved YG box, which is important for SMN oligomerization and interaction with RG-containing protein targets (6, 18), is necessary for Gemin7 interaction.

Gemin7 Interacts with a Subset of Sm Proteins—We next investigated the interactions of Gemin7 with the Sm proteins of spliceosomal snRNPs, the best characterized substrates of the SMN complex (10, 16–18, 22, 24, 26, 32). *In vitro* translated [³⁵S]methionine-labeled Sm proteins were incubated with purified recombinant GST-Gemin7 or GST as a control. As

shown in Fig. 7, Gemin7 interacts strongly with SmE and to a lesser extent with SmB, SmD2, and SmD3. None of the Sm proteins interacts with GST alone. These experiments demonstrate that Gemin7 interacts with a specific subset of Sm proteins. Gemin7 interacts most avidly with SmE, demonstrating a different Sm binding pattern as compared with the other components of the SMN complex. Specifically, SMN interacts avidly with the symmetrically dimethylated arginines of the carboxyl-terminal RG tails of SmD1 and SmD3 and with SmB (16, 18, 28). The DEAD box RNA helicase Gemin3 interacts with SmB and SmD3, (11) and Gemin4 and Gemin5 both interact with SmB, SmD1, SmD2, SmD3, and SmE (12, 13), Gemin6 interacts most avidly with SmD2 and SmE but also with SmD1 and SmD3 (14). These interactions provide further evidence for multiple contacts between individual Sm proteins and components of the SMN complex that are likely to be important for the function of the SMN complex in snRNP core assembly.

CONCLUSIONS

From our analysis of the purified SMN complex, it appears that with Gemin7 the inventory of the integral constituents of the SMN complex is or is close to being complete. In Fig. 8 we present an overall scheme depicting the SMN complex based on the known composition and interactions among components. For simplicity of presentation, the complex is presented as a dimer, although it forms higher order oligomers *in vivo* (10–12, 16, 34, 44). The stoichiometry of the individual components is unknown; however, the individual interactions within the complex have been delineated. SMN interacts directly with Gemin2, Gemin3, Gemin5, and Gemin7 (Fig. 5C) (10, 11, 13), whereas Gemin4 associates with the complex via Gemin3 (12). We now have demonstrated that the interaction of Gemin6 with the SMN complex requires Gemin7 (Fig. 5D). Further studies on how all of the SMN complex constituents work together to facilitate RNP biogenesis will provide insights into the pathogenesis of spinal muscular atrophy.

Acknowledgments—We thank Dr. Eng Tan for anti-p80 coilin antibodies. We are grateful to members of our laboratory for stimulating discussions and, in particular, Drs. Zissimos Mourelatos, Amelie Gubitz, and Sergey Paushkin for critical comments on this paper.

REFERENCES

- Melki, J. (1997) *Curr. Opin. Neurol.* **10**, 381–385
- Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., and Zeviani, M. (1995) *Cell* **80**, 155–165
- Schrank, B., Gotz, R., Gunnarsen, J. M., Ure, J. M., Toyka, K. V., Smith, A. G., and Sendtner, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 9920–9925
- Miguel-Aliaga, I., Culetto, E., Walker, D. S., Baylis, H. A., Sattelle, D. B., and Davies, K. E. (1999) *Hum. Mol. Genet.* **8**, 2133–2143
- Owen, N., Doe, C. L., Mellor, J., and Davies, K. E. (2000) *Hum. Mol. Genet.* **9**, 675–684
- Paushkin, S., Charroux, B., Abel, L., Perkinson, R. A., Pellizzoni, L., and Dreyfuss, G. (2000) *J. Biol. Chem.* **275**, 23841–23846
- Hannus, S., Buhler, D., Romano, M., Seraphin, B., and Fischer, U. (2000) *Hum. Mol. Genet.* **9**, 663–674

8. Miguel-Aliaga, I., Chan, Y. B., Davies, K. E., and van den Heuvel, M. (2000) *FEBS Lett.* **486**, 99–102
9. Wang, J., and Dreyfuss, G. (2001) *J. Biol. Chem.* **276**, 9599–9605
10. Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997) *Cell* **90**, 1013–1021
11. Charroux, B., Pellizzoni, L., Perkinson, R. A., Shevchenko, A., Mann, M., and Dreyfuss, G. (1999) *J. Cell Biol.* **147**, 1181–1194
12. Charroux, B., Pellizzoni, L., Perkinson, R. A., Yong, J., Shevchenko, A., Mann, M., and Dreyfuss, G. (2000) *J. Cell Biol.* **148**, 1177–1186
13. Gubitza, A. K., Mourelatos, Z., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) *J. Biol. Chem.* **277**, 5631–5636
14. Pellizzoni, L., Baccon, J., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) *J. Biol. Chem.* **277**, 7540–7545
15. Liu, Q., and Dreyfuss, G. (1996) *EMBO J.* **15**, 3555–3565
16. Pellizzoni, L., Charroux, B., and Dreyfuss, G. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11167–11172
17. Buhler, D., Raker, V., Luhrmann, R., and Fischer, U. (1999) *Hum. Mol. Genet.* **8**, 2351–2357
18. Friesen, W. J., and Dreyfuss, G. (2000) *J. Biol. Chem.* **275**, 26370–26375
19. Pellizzoni, L., Baccon, J., Charroux, B., and Dreyfuss, G. (2001) *Curr. Biol.* **11**, 1079–1088
20. Jones, K. W., Gorzynski, K., Hales, C. M., Fischer, U., Badbanchi, F., Terns, R. M., and Terns, M. P. (2001) *J. Biol. Chem.* **276**, 38645–38651
21. Pellizzoni, L., Charroux, B., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001) *J. Cell Biol.* **152**, 75–85
22. Mourelatos, Z., Abel, L., Yong, J., Kataoka, N., and Dreyfuss, G. (2001) *EMBO J.* **20**, 5443–5452
23. Rossoll, W., Kroning, A. K., Ohndorf, U. M., Steegborn, C., Jablonka, S., and Sendtner, M. (2002) *Hum. Mol. Genet.* **11**, 93–105
24. Hebert, M. D., Szymczyk, P. W., Shpargel, K. B., and Matera, A. G. (2001) *Genes Dev.* **15**, 2720–2729
25. Young, P. J., Day, P. M., Zhou, J., Androphy, E. J., Morris, G. E., and Lorson, C. L. (2002) *J. Biol. Chem.* **277**, 2852–2859
26. Brahms, H., Meheus, L., de Brabandere, V., Fischer, U., and Luhrmann, R. (2001) *RNA* **7**, 1531–1542
27. Brahms, H., Raymackers, J., Union, A., de Keyser, F., Meheus, L., and Luhrmann, R. (2000) *J. Biol. Chem.* **275**, 17122–17129
28. Friesen, W. J., Massenet, S., Paushkin, S., Wyce, A., and Dreyfuss, G. (2001) *Mol. Cell* **7**, 1111–1117
29. Friesen, W. J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G. S., Van Duynne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001) *Mol. Cell Biol.* **21**, 8289–8300
30. Friesen, W. J., Wyce, A., Paushkin, S., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) *J. Biol. Chem.* **277**, 8243–8247
31. Meister, G., Eggert, C., Buhler, D., Brahms, H., Kambach, C., and Fischer, U. (2001) *Curr. Biol.* **11**, 1990–1994
32. Fischer, U., Liu, Q., and Dreyfuss, G. (1997) *Cell* **90**, 1023–1029
33. Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998) *Cell* **95**, 615–624
34. Meister, G., Buhler, D., Laggerbauer, B., Zobawa, M., Lottspeich, F., and Fischer, U. (2000) *Hum. Mol. Genet.* **9**, 1977–1986
35. Meister, G., Buhler, D., Pillai, R., Lottspeich, F., and Fischer, U. (2001) *Nat. Cell Biol.* **3**, 945–949
36. Yong, J., Pellizzoni, L., and Dreyfuss, G. (2002) *EMBO J.* **21**, 1188–1196
37. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.* **68**, 850–858
38. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) *Nature* **379**, 466–469
39. Pinol-Roma, S., Choi, Y. D., Matunis, M. J., and Dreyfuss, G. (1988) *Genes Dev.* **2**, 215–227
40. Andrade, L. E., Tan, E. M., and Chan, E. K. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 1947–1951
41. Matera, A. G., and Frey, M. R. (1998) *Am. J. Hum. Genet.* **63**, 317–321
42. Carvalho, T., Almeida, F., Calapez, A., Lafarga, M., Berciano, M. T., and Carmo-Fonseca, M. (1999) *J. Cell Biol.* **147**, 715–728
43. Young, P. J., Le, T. T., Dunckley, M., Nguyen, T. M., Burghes, A. H., and Morris, G. E. (2001) *Exp. Cell Res.* **265**, 252–261
44. Paushkin, S., Gubitza, A. K., Massenet, S., and Dreyfuss, G. (2002) *Curr. Opin. Cell Biol.* **14**, 305–312