

Specific Sequences of the Sm and Sm-like (Lsm) Proteins Mediate Their Interaction with the Spinal Muscular Atrophy Disease Gene Product (SMN)*

Received for publication, April 18, 2000, and in revised form, May 25, 2000
Published, JBC Papers in Press, June 12, 2000, DOI 10.1074/jbc.M003299200

Westley John Friesen and Gideon Dreyfuss‡

From the Howard Hughes Medical Institute and Department of Biochemistry & Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6148

The spinal muscular atrophy disease gene product (SMN) is crucial for small nuclear ribonuclear protein (snRNP) biogenesis in the cytoplasm and plays a role in pre-mRNA splicing in the nucleus. SMN oligomers interact avidly with the snRNP core proteins SmB, -D1, and -D3. We have delineated the specific sequences in the Sm proteins that mediate their interaction with SMN. We show that unique carboxyl-terminal arginine- and glycine-rich domains comprising the last 29 amino acids of SmD1 and the last 32 amino acids of SmD3 are necessary and sufficient for SMN binding. Interestingly, SMN also interacts with at least two of the U6-associated Sm-like (Lsm) proteins, Lsm4 and Lsm6. Furthermore, the carboxyl-terminal arginine- and glycine-rich domain of Lsm4 directly interacts with SMN. This suggests that SMN also functions in the assembly of the U6 snRNP in the nucleus and in the assembly of other Lsm-containing complexes. These findings demonstrate that arginine- and glycine-rich domains are necessary and sufficient for SMN interaction, and they expand further the range of targets of the SMN protein.

Spinal muscular atrophy (SMA)¹ is a neuromuscular disease that results in muscular weakness and atrophy due to degeneration of motor neurons of the spinal cord (1–4). Over 98% of SMA patients have mutations or deletions of the Survival Motor Neuron (*SMN1*) gene, which is present as an inverted repeat on chromosome 5 at 5q13 (5–7). Only deletions or mutations in the telomeric copy of *SMN* (*SMN1*) lead to SMA (7–14). The centromeric copy of the gene (*SMN2*) produces mostly an alternatively spliced form of SMN deleted of amino acids encoded by exon 7 (*SMN Δ Ex7*), and *SMN2* can only partially compensate for mutations or deletions in *SMN1* (7, 15, 16). Indeed, some patients, instead of complete deletion of *SMN1*, have shorter deletions of at least exon 7 or single point mutations within the conserved YG domain (17, 18).

As measured by coimmunoprecipitation from cytoplasmic extracts, SMN exists in a complex with Sm proteins (19), Gemin2 (formerly SIP1) (19), Gemin4 (20), and the DEAD box RNA helicase Gemin3 (21). The amino terminus of SMN tightly associates with Gemin2, whereas the carboxyl-terminal con-

served YG domain is necessary for self-association and interaction with the core snRNP Sm proteins (19, 22). We have recently demonstrated that oligomerization of SMN through the conserved YG domain is required for efficient interaction with SmB, -D1, and -D3 and suggested that upon oligomerization a high affinity Sm protein binding site is formed (23).

SnRNPs are formed in the cytoplasm where the core Sm proteins bind to the Sm site on U snRNA, and following hypermethylation of the m⁷G cap to a 2,2,7-trimethyl-gaunosine (m³G) cap the snRNP is transported to the nucleus (24–30). Using oocyte injections, it was demonstrated that the SMN-Gemin2 complex plays a role in spliceosomal snRNP assembly that takes place in the cytoplasm. Injection of Gemin2-specific monoclonal antibodies inhibited, whereas injection of SMN-specific monoclonal antibodies stimulated, snRNP biogenesis (31). In HeLa cells SMN deleted of the amino-terminal 27 amino acids (*SMN Δ N27*) functions as a dominant negative mutant sequestering the SMN complex in large cytoplasmic and nuclear bodies (32). *SMN Δ N27* causes accumulation of Sm proteins and U snRNA in the cytoplasm, suggesting that this mutant blocks snRNP assembly prior to cap hypermethylation and nuclear import. *In vitro* SMN stimulates, whereas *SMN Δ N27* and anti-SMN antibodies inhibit, splicing following preincubation in nuclear extracts, demonstrating that SMN may have a role in nuclear regeneration of snRNPs (32). SnRNP assembly and nuclear splicing are basic cellular functions required for cellular viability. Indeed, mice (33), *Caenorhabditis elegans* (34), and *Schizosaccharomyces pombe* (35, 36), all lacking SMN, are not viable, which confirms that SMN plays an essential role. Taken together these data suggest that the SMA phenotype is caused by a motor neuron-specific response to defects in the cytoplasmic assembly and/or nuclear cycling of the splicing machinery.

The biogenesis of U6 snRNP is different from that of the other snRNPs. Unlike other snRNAs, which are transcribed by RNA polymerase II, U6 is transcribed by RNA polymerase III (37). In addition, U6 does not leave the nucleus (38–40), has a γ -monomethyl triphosphate cap structure (41), and has a protein core consisting of eight Sm-like (Lsm) proteins (42–44). In yeast it has been shown that Lsm4p, Lsm6p, and Lsm7p directly contact the U6 snRNA as measured by ultraviolet light cross-linking experiments (45). The Lsm-U6 RNA complex is relatively salt-labile and dissociates in the presence of competitor RNA, suggesting that the interaction between Lsm and U6 snRNA is more dynamic in nature than that of the Sm core snRNPs, which are extremely stable (44).

Here we have mapped the regions on the Sm proteins that are responsible for SMN binding. We show that oligomerized SMN interacts with the arginine- and glycine-containing carboxyl-terminal regions of SmB, SmD1, and SmD3. Surpris-

* This work was supported by a grant from the National Institutes of Health. 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.

‡ An 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; snRNP, small nuclear ribonuclear protein; GST, glutathione S-transferase.

ingly, we found that SMN also interacts directly with several of the Lsm proteins. We suggest that the SMN complex has a role in the assembly and/or regeneration of Lsm-containing complexes, including the U6 snRNP.

MATERIALS AND METHODS

Plasmid Construction and Protein Production—All plasmids for expression of Sm and Lsm proteins and deletions or fragments thereof were constructed by polymerase chain reaction amplification of cDNA with specific primers containing restriction sites for subcloning. Expressed sequence tag cDNA clones encoding Lsm2, -3, -4, -5, -6, -7, and -8 (IMAGE clone 1168053, IMAGE clone 278511, ATCC clone 160595, IMAGE clone 2368661 (a kind gift from Bertrand Seraphin), IMAGE clone 925894, and IMAGE clone 2325387, respectively) were obtained and subcloned into pET vectors (Novagen). All constructs used were confirmed by automated DNA cycle sequencing.

All [³⁵S]methionine-labeled proteins were produced *in vitro* using the TNT T7-coupled reticulocyte system (Promega) in the presence of [³⁵S]methionine (Amersham Pharmacia Biotech) according to the manufacturer's protocol. In Fig. 3, shown below, all Sm and Sm deletion peptides were expressed from pcDNA3 with an in-frame myc tag (*myc*-pcDNA3) (46). All Sm and Sm peptide fragments were produced from pET vectors with (His)₆-T7 tag (see Fig. 4A, below). SMN, SMNY272C (tyrosine at position 272 changed to cysteine), SMNΔYG (deleted of the carboxyl-terminal 26 amino acids), and SMNΔEx7 (deleted of amino acids encoded by exon 7) were produced from *myc*-pcDNA3. These *myc* tag constructs produce two bands, because translation starts at both the *myc* tag ATG and the native SMN ATG.

Recombinant GST (Amersham Pharmacia Biotech) and (His)₆-T7 (Novagen)-tagged fusion proteins were produced according to the manufacturers' protocols. Following elution from Ni²⁺ columns all (His)₆-T7-tagged proteins were extensively dialyzed in 20 mM Tris, pH 7.5, 300 mM NaCl, 2 mM EDTA, 40% glycerol, and 0.5 mM phenylmethylsulfonyl fluoride and stored at -20 °C. For oligomerized GST-SMN/SMN production, two 500-ml cultures, one expressing GST-SMN and one expressing (His)₆-T7-SMN, were induced and incubated at 17 °C with shaking for 16 h. The cultures were centrifuged, and each cell pellet was recovered in 20 ml of resuspension buffer (20 mM Tris, pH 7.5, 400 mM NaCl, 1 mM dithiothreitol, and one tablet per 50 ml of Complete EDTA-free protease inhibitor mixture tablets (Roche Molecular Biochemicals)). Resuspended cells were mixed and lysed by sonication. Triton X-100 was added to 1%, and the lysate was incubated with mild agitation at 4 °C for 1 h. The lysate was clarified by centrifugation, and oligomers of SMN were purified on glutathione-Sepharose (Amersham Pharmacia Biotech). Purification of GST-SMN/SMN in this manner produces immobilized oligomerized SMN, which binds to SmB, -D1, and -D3 similarly to the previously reported oligomerized SMN (23). Prior to use, all recombinant proteins were visualized on Coomassie-stained SDS-polyacrylamide gels to confirm purity and lack of degradation.

In Vitro Binding—All binding reactions were carried out in 1 ml of binding buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 0.2 mM EDTA, 0.05% Nonidet P-40, 2 mM dithiothreitol, and one tablet per 50 ml of Complete EDTA-free protease inhibitor mixture) and incubated 1–2 h at 4 °C followed by five washes with 1 ml of binding buffer. All binding assays employed glutathione-Sepharose-immobilized GST fusion proteins (1–2 μg). (His)₆-T7-tagged fusion proteins (0.5–0.7 μg) were used for direct binding assays, and for binding of *in vitro* translated [³⁵S]methionine-labeled proteins, 8 μl of each translation mixture was used. Bound proteins were separated by SDS-polyacrylamide gel electrophoresis and detected by Western blotting with the T7 tag antibody (Novagen) or autoradiography.

RESULTS

SMN, but Not SMN Mutants Found in SMA Patients, Directly Interacts with Smd1 and Smd3—Among the eight core Sm proteins, three of them, Smd1, -D3 and -B, have been shown to avidly interact with oligomerized SMN, and only SmB has been shown to interact directly with SMN (23). To determine if Smd1 and -D3 can also directly interact with SMN, we produced recombinant (His)₆-T7-tagged Smd1 and -D3. As shown in Fig. 1A, these proteins bound directly to immobilized GST-SMN/SMN but not to GST alone. In the reverse experiment, (His)₆-T7-tagged SMN bound GST-Smd1 and GST-Smd3 but not GST alone (Fig. 1B) demonstrating that Smd1 and -D3 interact directly with SMN. To determine if, as is the

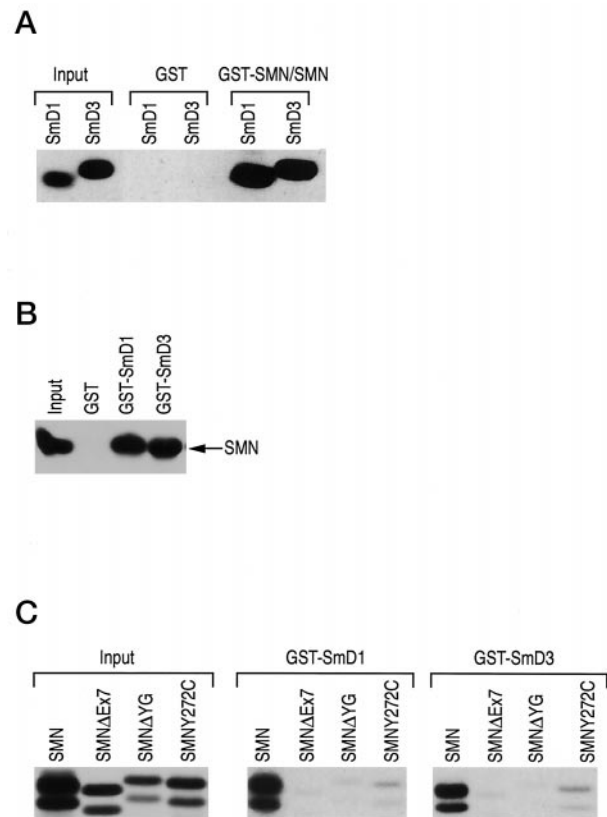


FIG. 1. SMN directly interacts with Smd1 and Smd3. A, (His)₆-T7-tagged Smd1 and Smd3 (as indicated) were incubated with immobilized GST-SMN/SMN or GST followed by washing, boiling in SDS sample buffer, and Western analysis with anti-T7-tag antibody. The *input* lanes show 10% of the total used in each binding reaction. B, (His)₆-T7-tagged SMN was incubated with the indicated immobilized fusion proteins, and binding was measured as in A. The *input* lane shows 20% of the total used in each binding reaction. C, the indicated *in vitro* translated, [³⁵S]methionine-labeled, SMN proteins were incubated with immobilized GST-Smd1 and Smd3 as indicated followed by SDS-polyacrylamide gel electrophoresis and fluorography. The *input* panel shows 10% of each protein used in the binding experiments.

case with SmB, SMN mutations found in some SMA patients are deficient in Smd1 and Smd3 interactions, we tested the ability of *in vitro* produced, [³⁵S]methionine-labeled, SMN and SMN mutants to interact with GST, GST-Smd1, and GST-Smd3. Fig. 1C shows that SMN but not SMN deleted of amino acids encoded by exon 7 (SMNΔEx7), SMN deleted of the YG box (SMNΔYG), or SMN with tyrosine at position 272 changed to cysteine (SMNY272C) bound GST-Smd1 and GST-Smd3. No binding with GST alone was observed (data not shown). Thus, as is the case with SmB, SMN binds directly to Smd1 and -D3, and SMN mutants reduce these interactions.

Sm Protein Carboxyl-terminal Arginine- and Glycine-rich Domains Are Required for SMN Binding—The Sm protein sequence requirements for SMN interaction have not been elucidated. From visual inspection of the primary sequence of the human Sm proteins we observed that those proteins, which interact most strongly with SMN (SmB, -D1, and -D3), contain extended carboxyl termini. Further examination revealed that the carboxyl termini of these proteins contain clusters of glycine and arginine residues (Fig. 2). Most strikingly, Smd1 contains nine contiguous glycine-arginine (GR) repeats. Smd3 contains fewer arginine and glycine residues but is markedly arginine- and glycine-rich. SmB contains a very extended carboxyl terminus, which is most noticeably proline-rich but also contains five GRG and three GMR repeats. Because the Sm proteins that interact weakly (or not at all) with SMN do not



FIG. 2. **SmB, -D1, -D3, and Lsm4 contain arginine- and glycine-rich regions in their carboxyl termini.** An alignment of Sm and Lsm proteins is shown. *Dark shading* indicates identities, and *light shading* indicates similarities. Sm motifs 1 and 2 are indicated. Adjacent arginine and glycine residues are *underlined*, and *arrowheads with numbers* indicate carboxyl-terminal deletions and fragments used in this study.

contain carboxyl-terminal arginine- and glycine-rich domains, we tested whether these domains of SmD1, -D3, and -B were responsible for SMN interaction. We deleted the carboxyl-terminal arginine- and glycine-rich regions and analyzed the ability of these deletions to interact with recombinant GST-SMN/SMN. These deletions did not eliminate any portion of Sm motifs 1 or 2 common to all Sm proteins (Figs. 2 and 3A). Deletion of the carboxyl-terminal 99 or 139 amino acids of SmB (SmB Δ c99 and SmB Δ c139, respectively) abolished SMN binding. Deletion of 29 amino acids from the carboxyl terminus of SmD1 (SmD1 Δ c29) and 32 amino acids from the carboxyl terminus of SmD3 (SmD3 Δ c32) also abolished SMN interaction (Fig. 3B). Thus, the arginine- and glycine-rich domains of SmB, -D1, and -D3 are necessary for interaction with SMN.

The Carboxyl-terminal Arginine- and Glycine-rich Domains of Sm Proteins Are Sufficient for SMN Binding—To determine whether the arginine- and glycine-rich Sm protein sequences are sufficient for interaction with SMN, carboxyl-terminal fragments were produced separately. In the case of SmB, the carboxyl-terminal 99 (SmBc99) and 139 (SmBc139) amino acids were translated *in vitro* and SMN binding was assayed. Fig. 4A shows that these peptides did not bind SMN (see "Discussion"). SmD1 carboxyl-terminal 79 (SmD1c79) and 54 (SmD1c54) amino acids and SmD3 carboxyl-terminal 83 (SmD3c83) and 58 (SmD3c58) amino acids were also translated *in vitro* and tested for binding to immobilized SMN. All of the SmD1 and SmD3 carboxyl-terminal fragments bound SMN with avidities com-

parable to that of the full-length proteins (Fig. 4B) and did not bind GST alone (data not shown). In the reverse experiment we fused GST to the carboxyl-terminal 29 and 32 amino acids of SmD1 (GST-SmD1c29) and SmD3 (GST-SmD3c32), respectively. These peptides were produced as recombinant proteins and assayed for interaction with wild type SMN and SMN mutants (SMN Δ Ex7, SMNY272C, and SMN Δ YG). Wild type SMN bound the GST fusion proteins, whereas two of the SMN mutants (SMN Δ Ex7 and SMN Δ YG) showed greatly reduced or undetectable binding (Fig. 4C). In contrast, under these conditions SMNY272C has only slightly reduced affinity for the carboxyl termini of SmD1 and -D3. GST-SmD1c29 and GST-SmD3c32 interact directly with (His)₆-T7-tagged SMN (data not shown). Thus, the carboxyl-terminal arginine- and glycine-rich domains of SmD1 and -D3 are sufficient for SMN interaction and, like the full-length SmD1 and -D3 proteins, have markedly reduced affinity for SMN deleted of Ex7 or the YG box.

SMN Interacts with the Carboxyl-terminal Arginine- and Glycine-rich Domain of Lsm4—It has recently been shown that the human U6 snRNP contains Sm-like proteins (Lsm) instead of the Sm proteins found in the U1, U2, U4, and U5 snRNPs (42–44). Because the Lsm proteins may function in a similar manner to the core Sm proteins, we were interested to determine if the Lsm proteins can also interact with SMN. To facilitate these experiments we obtained cDNA clones corresponding to the human Lsm2, -3, -4, -5, -6, -7, and -8 proteins.

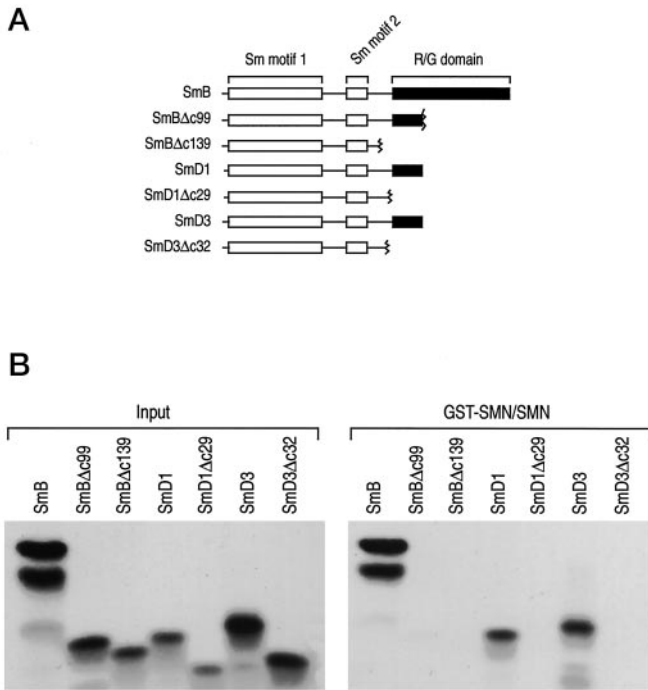


FIG. 3. Carboxyl-terminal arginine- and glycine-rich domains of SmB, -D1, and -D3 are required for interaction with SMN. *A*, schematic representation of full-length and deletion mutants used for binding in *B*. Open boxes represent Sm motifs as indicated. Closed boxes represent carboxyl-terminal arginine- and glycine-rich domains (R/G domain). Jagged lines indicate sites of carboxyl-terminal deletions. *B*, the indicated *in vitro* translated, [³⁵S]methionine-labeled, Sm and Sm deletion mutants were tested for interaction with immobilized SMN as in Fig. 1C. The *input* panel shows 20% of each protein used in the binding experiments.

These proteins were then produced by *in vitro* transcription and translation in the presence of [³⁵S]methionine and tested for binding to SMN. Lsm2, -4, -6, and -7 bound immobilized SMN, with Lsm4 exhibiting higher affinity than the other binders under the assay conditions used (Fig. 5A). None of these Lsm proteins bound GST alone (data not shown). Interestingly, only one of the Lsm proteins, Lsm4, has an arginine- and glycine-rich carboxyl domain (Fig. 2). To confirm a direct interaction between Lsm4 and SMN and to determine if, like Sm proteins, the carboxyl-terminal arginine- and glycine-rich domain of Lsm4 interacts with SMN, recombinant GST fusions of full-length Lsm4 (GST-Lsm4), Lsm4 deleted of the carboxyl-terminal 54 amino acids (GST-Lsm4Δc54), and the carboxyl-terminal 54 amino acids of Lsm4 (GST-Lsm4c54) were produced separately. As shown in Fig. 5B, GST-Lsm4 and GST-Lsm4c54 bound SMN with similar avidities, whereas deletion of the carboxyl-terminal 54 amino acids greatly reduced SMN binding. Thus, the carboxyl-terminal arginine- and glycine-rich domain of Lsm4 is necessary and sufficient for direct interaction with SMN. Because SMN mutants found in SMA patients affect interaction with Sm proteins, we tested their ability to interact with Lsm4. Fig. 5C show that these SMN mutations abolished the binding to Lsm4.

DISCUSSION

We demonstrate that SMN directly interacts with SmD1 and -D3, and mutants that alter the YG domain of SMN greatly reduce these interactions (Fig. 1). We have previously shown that peptides encompassing the YG box specifically compete for SmB interaction (19). Recently, it has been suggested that the tudor domain of SMN is necessary and sufficient for interaction with the Sm proteins. It has also been shown that a point

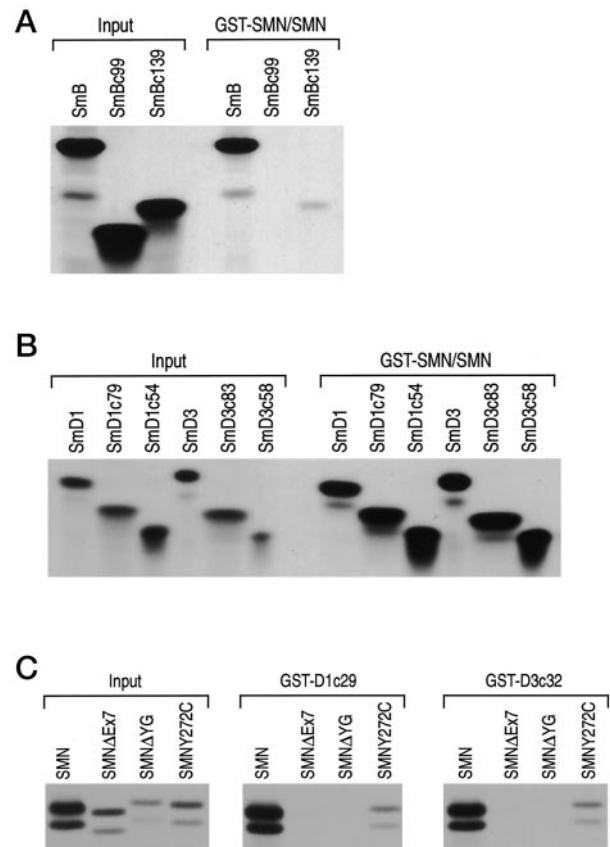


FIG. 4. Carboxyl-terminal arginine- and glycine-rich domains of SmD1 and -D3, but not that of SmB, are sufficient for SMN interaction. *A*, *in vitro* [³⁵S]methionine-labeled SmB and the indicated SmB fragments were tested for immobilized SMN binding as in Fig. 1C. *Input* lanes show 10% of the input used in each binding reaction. *B*, as indicated, *in vitro* [³⁵S]methionine-labeled SmD1, SmD3, and fragments were tested for interaction with immobilized SMN as in Fig. 1C. *Input* lanes show 10% of the input used in each binding reaction. *C*, the indicated [³⁵S]methionine-labeled SMN and mutant SMN proteins were bound to GST-SmD1c29 or GST-SmD3c32, as indicated, and processed as in Fig. 1C. The *input* panel shows 10% of each protein used in the binding experiments.

mutation in the tudor domain, which changes a glutamate residue at position 134 to lysine (E134K), reduces Sm interaction (47). These inconsistent data can be explained if SMN self-association through the YG domain results in the formation of a high affinity Sm binding site, which includes amino acids found in the tudor domain, the YG domain, and possibly other regions of the protein (23). Regardless of deletion mapping results designed to find the Sm binding site on SMN, it remains clear that SMN mutations found in SMA patients are deficient in Sm protein interaction.

We show that arginine- and glycine-rich domains found at the carboxyl termini of SmB, -D1, and -D3 are necessary for SMN interaction. We further demonstrate that SmD1 and -D3 carboxyl termini with relatively uninterrupted arginine-glycine repeats are also sufficient for SMN interaction. In contrast, SmB has GRG and GMR repeats spread throughout the 151 amino acids from the end of Sm motif 2 to the carboxyl terminus (Fig. 2). This region, when expressed without the Sm domain, is not sufficient for SMN interaction. This result can be explained if the three-dimensional structure of SmB brings the GRG/GMR repeats into close proximity to form a surface for SMN interaction, and deletion of the Sm domain disrupts this fold. It is possible that this region forms an arginine- and glycine-rich binding surface that is structurally similar to the correspondingly shorter regions of SmD1 and -D3. Crystal

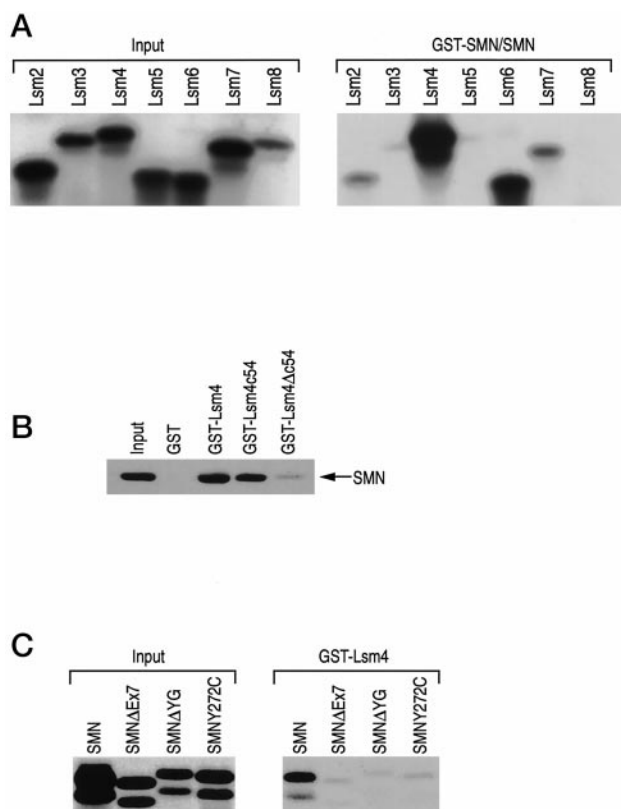


FIG. 5. SMN interacts with Lsm proteins *in vitro*. *A*, the indicated *in vitro* translated, [³⁵S]methionine-labeled Lsm proteins were incubated with immobilized GST-SMN/SMN and binding was assayed as in Fig. 1C. *Input* lanes show 10% of each protein used in the binding reactions. *B*, the indicated GST fusion proteins were incubated with (His)₆-T7-tagged SMN and processed as in Fig. 1A. The *input* lane shows 10% of the total used in each binding reaction. *C*, the indicated *in vitro* produced SMN and SMN mutant proteins were incubated with immobilized GST-Lsm4, and binding was assayed as in Fig. 1C. The *input* panel shows 10% of each protein used in the binding experiments.

structures of heterodimers of SmD3-SmB and SmD1-SmD2 show that the tertiary structure of the Sm domains contains an amino-terminal α -helix followed by a bent five-stranded anti-parallel β -sheet (48). Unfortunately, the SmB-SmD3 structure did not include amino acids carboxyl-terminal to Sm motif 2, and the carboxyl terminus of SmD1 was not ordered in the crystal structure of SmD1-SmD2. In light of the present work it will be interesting to determine the structure of the carboxyl-terminal regions of SmB, -D1, and -D3, and it is possible that these regions remain unstructured until bound by SMN.

We show that the carboxyl termini of SmD1 and -D3 are sufficient for interaction with SMN. It is striking that such a small region that, in the case of SmD1, is comprised almost exclusively of arginine-glycine repeats (Fig. 2) is responsible for SMN interaction. This suggests that SMN interacts with any arginine-glycine repeat sequence of sufficient length. However, this is not the case, because heteronuclear RNP A1 and A2, which have regions containing RGG repeats, do not interact with SMN (data not shown).

Electron microscopy (49), Sm-Sm interaction studies (50–52), x-ray crystallography, and structural modeling suggest that the assembled snRNP Sm core is a seven-membered ring consisting of one copy of each of the Sm proteins, with the snRNA bound in the central hole of the ring (48). Prior to snRNP assembly, cytoplasmic Sm proteins appear to form SmD1-SmD2 and SmB-SmD3 dimers and SmE-SmF-SmG trimers (50, 53, 54). The capacity of SMN to interact with at least one component of each of the pre-snRNP Sm complexes (or

particularly efficiently with at least the two SmB and SmD complexes) suggests that SMN serves as a scaffold on which the seven-membered snRNP ring forms. SMN readily oligomerizes *in vitro* (23) and is found in large oligomers in cellular extracts (19, 21). Oligomeric SMN interaction with non-Sm motif carboxyl-terminal regions of SmB, -D1, and -D3 will likely not interfere with Sm-Sm and Sm-snRNA interactions required for proper snRNP assembly, thus supporting our speculation that SMN can serve as a scaffold for proper snRNP assembly.

Gemin3 and Gemin4, two components of the SMN complex, bind subsets of the Sm proteins. Gemin3 binds SmB, -D2, and -D3, whereas Gemin4, similar to SMN, binds SmB, -D1, -D2, -D3, and -E (20, 21). The series of SmB, -D1, and -D3 deletion mutants used to map the SMN binding site on these Sm proteins was used in interaction studies with GST-Gemin3 and GST-Gemin4. Although these studies did not exhaustively map the Sm protein sequences required for these interactions, they did demonstrate that the carboxyl-terminal arginine- and glycine-rich sequences of SmB, -D1, and -D3 are dispensable for interaction with Gemin3 and Gemin4 (data not shown). Thus, SMN interacts with the same Sm proteins bound by Gemin3 and Gemin4 but on different regions of the Sm proteins. The SMN complex is, therefore, likely to have a firm multivalent grip on the Sm proteins during snRNP assembly.

We also report here a novel set of interactions between SMN and Lsm proteins (Fig. 5). The carboxyl terminus of Lsm4 contains an arginine- and glycine-rich domain that, when expressed on its own, is sufficient for direct interaction with SMN (Fig. 5B). We further show that SMN interacts with Lsm2, -6, and -7. These proteins have no carboxyl-terminal arginine- and glycine-rich domain and, thus, are likely have a different mode of SMN binding. By analogy with the crucial role of the SMN complex in cytoplasmic snRNP assembly and SMN interaction with Sm proteins, the interaction with Lsm proteins suggests that SMN likewise functions in the assembly of U6 snRNP. The fact that SMN interacts with at least two of the Lsm proteins and the similarities in the size and overall shape of Lsm and Sm cores (44, 49) suggest that SMN also serves as a scaffold for Lsm core assembly. This also brings up the possibility that SMN plays a more direct and more extensive role in splicing. Because the Lsm-U6 snRNA interaction is less stable than that of the Sm core with U1, U2, U4, and U5 snRNAs, it has been suggested that the Lsm proteins play a more dynamic role in splicing (44). Given SMN's role in splicing (32), it is possible that SMN functions in the rearrangement of the Lsm proteins during splicing.

In yeast, Lsm proteins have recently been shown to have a role in mRNA decay. Specifically, there appears to be a complex of Lsm1–7 involved in mRNA degradation in the cytoplasm and a complex of Lsm2–8 involved in splicing in the nucleus (55). Thus, it is possible that SMN also has a role in RNA degradation in the cytoplasm. The role of SMN in Lsm function is currently being investigated.

We originally cloned SMN by interaction with the RGG box of heteronuclear RNP U, and from this work it now emerges to be a general theme that SMN interacts with arginine- and glycine-rich domains. RGG boxes are often found in RNA binding proteins (56, 57), which is consistent with the role of SMN in assembly of RNPs. Finally, we note that arginines in RGG boxes are often subject to dimethyl arginine modification (58–61). It is possible that these arginine- and glycine-rich domains in Sm proteins are similarly modified, and this could serve to modulate their interactions with SMN and to regulate snRNP assembly.

Acknowledgments—We thank Drs. Bernard Charroux, Sergey Paushkin, and Zissimos Mourelatos for discussion and insight. In par-

ticular, we thank Drs. Livio Pellizzoni and Sara Nakielny for helpful suggestions and critical reading of the manuscript.

REFERENCES

1. Roberts, D. F., Chavez, J., and Court, S. D. (1970) *Arch. Dis. Child.* **45**, 33–38
2. Pearn, J. H. (1973) *J. Med. Genet.* **10**, 260–265
3. Pearn, J. (1978) *J. Med. Genet.* **15**, 409–413
4. Czeizel, A., and Hamula, J. (1989) *J. Med. Genet.* **26**, 761–763
5. Brzustowicz, L. M., Lehner, T., Castilla, L. H., Penchaszadeh, G. K., Wilhelmsen, K. C., Daniels, R., Davies, K. E., Leppert, M., Ziter, F., Wood, D., Dubowitz, V., Zerres, K., Hausmanowa-Petrusewicz, I., Ott, J., Munsat, T. L., and Gilliam, T. C. (1990) *Nature* **344**, 540–541
6. Melki, J., Sheth, P., Abdelhak, S., Burlet, P., Bachelot, M. F., Lathrop, M. G., Frezal, J., and Munnich, A. (1990) *Lancet* **336**, 271–273
7. Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., Le Paslier, D., Frézal, J., Cohen, D., Weissenbach, J., Munnich, A., and Melki, J. (1995) *Cell* **80**, 155–165
8. Bussaglia, E., Clermont, O., Tizzano, E., Lefebvre, S., Bürglen, L., Cruaud, C., Urtizberea, J. A., Colomer, J., Munnich, A., Baiget, M., and Melki, J. (1995) *Nat. Genet.* **11**, 335–337
9. Chang, J. G., Jong, Y. J., Huang, J. M., Wang, W. S., Yang, T. Y., Chang, C. P., Chen, Y. J., and Lin, S. P. (1995) *Am. J. Hum. Genet.* **57**, 1503–1505
10. Cobben, J. M., van der Steege, G., Grootsholten, P., de Visser, M., Scheffer, H., and Buys, C. H. (1995) *Am. J. Hum. Genet.* **57**, 805–808
11. Hahnen, E., Forkert, R., Marke, C., Rudnik-Schoneborn, S., Schonling, J., Zerres, K., and Wirth, B. (1995) *Hum. Mol. Genet.* **4**, 1927–1933
12. Hahnen, E. T., and Wirth, B. (1996) *Hum. Genet.* **98**, 122–123
13. Rodrigues, N. R., Owen, N., Talbot, K., Ignatius, J., Dubowitz, V., and Davies, K. E. (1995) *Hum. Mol. Genet.* **4**, 631–634
14. Velasco, E., Valero, C., Valero, A., Moreno, F., and Hernandez-Chico, C. (1996) *Hum. Mol. Genet.* **5**, 257–263
15. Gennarelli, M., Lucarelli, M., Capon, F., Pizzuti, A., Merlini, L., Angelini, C., Novelli, G., and Dallapiccola, B. (1995) *Biochem. Biophys. Res. Commun.* **213**, 342–348
16. Gavrillov, D. K., Shi, X., Das, K., Gilliam, T. C., and Wang, C. H. (1998) *Nat. Genet.* **20**, 230–231
17. Burghes, A. H. (1997) *Am. J. Hum. Genet.* **61**, 9–15
18. Talbot, K., Ponting, C. P., Theodosiou, A. M., Rodrigues, N. R., Surtees, R., Mountford, R., and Davies, K. E. (1997) *Hum. Mol. Genet.* **6**, 497–500
19. Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997) *Cell* **90**, 1013–1021
20. Charroux, B., Pellizzoni, L., Perkinson, R. A., Yong, J., Shevchenko, A., Mann, M., and Dreyfuss, G. (2000) *J. Cell Biol.* **148**, 1177–1186
21. Charroux, B., Pellizzoni, L., Perkinson, R. A., Shevchenko, A., Mann, M., and Dreyfuss, G. (1999) *J. Cell Biol.* **147**, 1181–1194
22. Lorson, C. L., Strasswimmer, J., Yao, J. M., Baleja, J. D., Hahnen, E., Wirth, B., Le, T., Burghes, A. H., and Androphy, E. J. (1998) *Nat. Genet.* **19**, 63–66
23. Pellizzoni, L., Charroux, B., and Dreyfuss, G. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11167–11172
24. DeRobertis, E. M. (1983) *Cell* **32**, 1021–1025
25. Mattaj, I. W., and De Robertis, E. M. (1985) *Cell* **40**, 111–118
26. Mattaj, I. W. (1986) *Cell* **46**, 905–911
27. Neuman de Vegvar, H. E., and Dahlberg, J. E. (1990) *Mol. Cell. Biol.* **10**, 3365–3375
28. Fischer, U., and Luhrmann, R. (1990) *Science* **249**, 786–790
29. Hamm, J., Darzynkiewicz, E., Tahara, S. M., and Mattaj, I. W. (1990) *Cell* **62**, 569–577
30. Fischer, U., Sumpster, V., Sekine, M., Satoh, T., and Luhrmann, R. (1993) *EMBO J.* **12**, 573–583
31. Fischer, U., Liu, Q., and Dreyfuss, G. (1997) *Cell* **90**, 1023–1029
32. Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998) *Cell* **95**, 615–624
33. 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
34. 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
35. Owen, N., Doe, C. L., Mellor, J., and Davies, K. E. (2000) *Hum. Mol. Genet.* **9**, 675–684
36. Hannus, S., Buhler, D., Romano, M., Seraphin, B., and Fischer, U. (2000) *Hum. Mol. Genet.* **9**, 663–674
37. Dahlberg, J. E., and Lund, E. (1991) *Science* **254**, 1462–1463
38. Vankan, P., McGuigan, C., and Mattaj, I. W. (1990) *EMBO J.* **9**, 3397–3404
39. Terns, M. P., Dahlberg, J. E., and Lund, E. (1993) *Genes Dev.* **7**, 1898–1908
40. Boelens, W. C., Palacios, I., and Mattaj, I. W. (1995) *RNA (N. Y.)* **1**, 273–283
41. Singh, R., and Reddy, R. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 8280–8283
42. Seraphin, B. (1995) *EMBO J.* **14**, 2089–2098
43. Salgado-Garrido, J., Bragado-Nilsson, E., Kandels-Lewis, S., and Seraphin, B. (1999) *EMBO J.* **18**, 3451–3462
44. Achsel, T., Brahm, H., Kastner, B., Bach, A., Wilm, M., and Luhrmann, R. (1999) *EMBO J.* **18**, 5789–5802
45. Mayes, A. E., Verdone, L., Legrain, P., and Beggs, J. D. (1999) *EMBO J.* **18**, 4321–4331
46. Siomi, H., and Dreyfuss, G. (1995) *J. Cell Biol.* **129**, 551–560
47. Buhler, D., Raker, V., Luhrmann, R., and Fischer, U. (1999) *Hum. Mol. Genet.* **8**, 2351–2357
48. Kambach, C., Walke, S., Young, R., Avis, J. M., de la Fortelle, E., Raker, V. A., Luhrmann, R., Li, J., and Nagai, K. (1999) *Cell* **96**, 375–387
49. Kastner, B., Bach, M., and Luhrmann, R. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 1710–1714
50. Raker, V. A., Plessel, G., and Luhrmann, R. (1996) *EMBO J.* **15**, 2256–2269
51. Fury, M. G., Zhang, W., Christodouloupolous, I., and Zieve, G. W. (1997) *Exp. Cell Res.* **237**, 63–69
52. Camasses, A., Bragado-Nilsson, E., Martin, R., Seraphin, B., and Bordonne, R. (1998) *Mol. Cell. Biol.* **18**, 1956–1966
53. Lehmeier, T., Raker, V., Hermann, H., and Luhrmann, R. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **1**, 12317–12321
54. Hermann, H., Fabrizio, P., Raker, V. A., Foulaki, K., Hornig, H., Brahm, H., and Luhrmann, R. (1995) *EMBO J.* **14**, 2076–2088
55. Tharun, S., He, W., Mayes, A. E., Lennertz, P., Beggs, J. D., and Parker, R. (2000) *Nature* **404**, 515–518
56. Kiledjian, M., and Dreyfuss, G. (1992) *EMBO J.* **11**, 2655–2664
57. Burd, C. G., and Dreyfuss, G. (1994) *Science* **265**, 615–621
58. Beyer, A. L., Christensen, M. E., Walker, B. W., and LeStourgeon, W. M. (1977) *Cell* **11**, 127–138
59. Boffa, L. C., Karn, J., Vidali, G., and Allfrey, V. G. (1977) *Biochem. Biophys. Res. Commun.* **74**, 969–976
60. Liu, Q., and Dreyfuss, G. (1995) *Mol. Cell. Biol.* **15**, 2800–2808
61. Henry, M. F., and Silver, P. A. (1996) *Mol. Cell. Biol.* **16**, 3668–3678