

Characterization of Functional Domains of the SMN Protein *in Vivo**

Received for publication, June 1, 2001, and in revised form, September 20, 2001
Published, JBC Papers in Press, September 25, 2001, DOI 10.1074/jbc.M105059200

Jin Wang and Gideon Dreyfuss‡

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

The *Survival of Motor Neurons (SMN)* is the disease gene of spinal muscular atrophy. We have previously established a genetic system based on the chicken pre-B cell line DT40, in which expression of SMN protein is regulated by tetracycline, to study the function of SMN *in vivo*. Depletion of SMN protein is lethal to these cells. Here we tested the functionality of mutant SMN proteins by determining their capacity to rescue the cells after depletion of wild-type SMN. Surprisingly, all of the spinal muscular atrophy-associated missense mutations tested were able to support cell viability and proliferation. Deletion of the amino acids encoded by exon 7 of the *SMN* gene resulted in a partial loss of function. A mutant SMN protein lacking both the tyrosine/glycine repeat (in exon 6) and exon 7 failed to sustain viability, indicating that the C terminus of the protein is critical for SMN activity. Interestingly, the Tudor domain of SMN, encoded by exon 3, does not appear to be essential for SMN function since a mutant deleted of this domain restored cell viability. Unexpectedly, a chicken SMN mutant (Δ N39) lacking the N-terminal 39 amino acids that encompass the Gemin2-binding domain also rescued the lethal phenotype. Moreover, the level of Gemin2 in Δ N39-rescued cells was significantly reduced, indicating that Gemin2 is not required for Δ N39 to perform the essential function of SMN in DT40 cells. These findings suggest that SMN may perform a novel function in DT40 cells.

SMN1 and *SMN2* genes are different (2). Although both pre-mRNAs are spliced into full-length SMN mRNAs, most of the *SMN2*-derived pre-mRNA is alternatively spliced to generate an SMN mRNA lacking exon 7 (Ex7) that encodes the last C-terminal 16 amino acids. In addition to the deletions, missense point mutations have been found in SMA patients who still retain one *SMN1* allele (12).

The mechanism by which a reduction of the SMN protein causes motor neuron degeneration remains unclear. SMN does not contain any domain with a known function.

However, alignment of SMN proteins of several species ranging from *Schizosaccharomyces pombe* to human reveals two highly conserved regions located at the N and C termini, which suggests functional importance of the two regions (13–15). A number of proteins have been shown to interact with SMN directly and therefore represent potential functional partners of SMN (16–23). Of these, Gemin2 (formerly known as SIP1 for SMN-interacting protein 1) has been well characterized (16, 24). Gemin2 interacts directly with SMN *in vitro* and can be co-immunoprecipitated with SMN from cell extracts. SMN and Gemin2 are co-localized in cells. Fractionation of cell extracts by chromatography showed that the SMN and Gemin2 proteins coeluted in a high molecular mass peak, indicating that SMN and Gemin2 exist in a macromolecular complex *in vivo*. Moreover, depletion of the SMN protein in cells results in a significant reduction of the level of Gemin2, suggesting that Gemin2 is stabilized by SMN binding (25). The interaction of SMN and Gemin2 is mediated by the phylogenetically conserved N terminus of SMN (16). Recently, a protein in the fission yeast *S. pombe* that binds to the N terminus of the yeast ortholog of SMN has been identified and shown to share low sequence homology with Gemin2 (26).

In addition to Gemin2, several other proteins are co-immunoprecipitated with SMN from cell extracts, including Gemin3, Gemin4, and the core Sm proteins of spliceosomal small nuclear ribonucleoproteins (snRNPs). Gemin3 is a DEAD box-containing putative RNA helicase (18), whereas Gemin4 does not share significant homology with any known protein (22). The C terminus of SMN, including the sequence encoded by Ex7 and the highly conserved tyrosine/glycine-rich sequence (Y/G box), is required for SMN to interact directly with Gemin3 and the Sm proteins as well as for its oligomerization (18, 27, 28). Deletion of Ex7 and, to a lesser extent, a point mutation found in SMA patients that changes a tyrosine residue in the Y/G box to cysteine (Y272C) significantly reduces these interactions. The central region of SMN shares homology with the Tudor domain of the *Drosophila* Tudor protein (29–31). Recently, the NMR structure of the SMN Tudor domain has been determined (32). The structure and *in vitro* binding assays suggest that the Tudor domain of SMN mediates the interaction with Sm proteins (32, 33). Several lines of evidence have suggested that SMN plays a critical role in snRNP metabolism, pre-mRNA splicing, and gene transcription (24, 34, 35).

Spinal muscular atrophy (SMA),¹ a motor neuron degenerative disease, is an autosomal recessive genetic disorder (1). The disease gene of SMA has been determined as the *Survival of Motor Neurons (SMN)* gene located at human chromosome 5q13 (2–9). Due to a duplication at 5q13, the human genome contains two copies of the *SMN* gene. Deletions of the telomeric *SMN* gene (*SMN1*) have been found in a vast majority of SMA patients. As a result, the SMN protein is solely expressed from the centromeric *SMN* gene (*SMN2*), which often leads to a lower SMN level in SMA patients compared with unaffected individuals. The extent of reduction of the SMN level appears to closely correlate with the severity of the disease (10, 11). Although the two *SMN* genes encode an identical protein, the splicing patterns of the pre-mRNAs transcribed from the

* This work was supported by Grant NIH/R01-NS40575-02 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.

‡ 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; Ex7, exon 7; snRNP, small nuclear ribonucleoprotein; tet, tetracycline; mAb, monoclonal antibody; cSMN, chicken SMN; huSMN, human SMN.

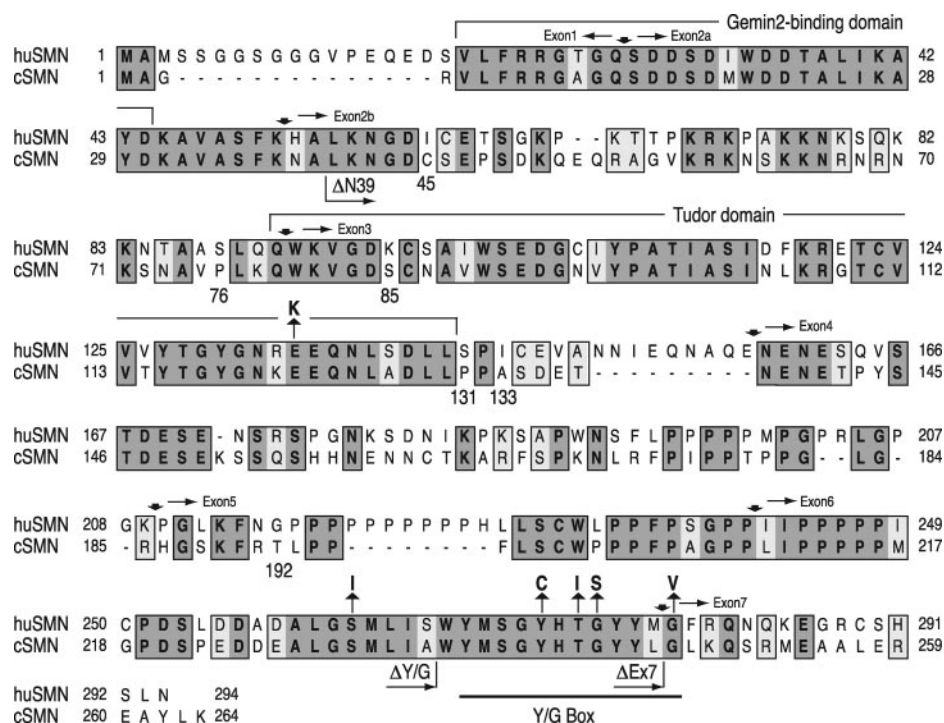


FIG. 1. Alignment of human and chicken SMN protein sequences. The identical and similar residues shared by huSMN and cSMN are shown in dark- and light-gray boxes, respectively. The highly conserved Y/G box is underlined. The Gemin2-binding domain and the Tudor domain are also indicated. Numbers indicate the positions of amino acids. Exon junctions of mRNAs encoding the huSMN and cSMN proteins are marked with small downward-pointing arrows. The missense point mutations found in SMA patients are indicated above each residue that is affected. The first or last residue of the terminal deletion mutants $\Delta N39$, $\Delta Ex7$, and $\Delta Y/G$ is marked. The amino acids of cSMN removed in the internal deletion mutants ΔK , $\Delta Tudor$, and ΔP are 45–76, 85–133, and 131–192, respectively.

Genetic studies have shown that *SMN* is an essential gene in various organisms, indicating that the SMN protein is required for fundamental cellular process(es) (13–15, 25, 26, 36); yet partial loss of SMN function leads to selective motor neuron degeneration. Mice that are null in the *smn* gene and carry a human *SMN2* transgene are viable, but develop symptoms that bear resemblance to human SMA (37, 38). Histological studies in these mice have confirmed the presence of defects and massive loss of motor neurons in the spinal cord. These findings further demonstrate that *SMN* is the disease gene of SMA, but do not reveal the specific functions of SMN.

To gain further insight into the function of SMN, we employed the chicken pre-B cell line DT40 to develop a cell-based genetic system (25). In the conditional *SMN* knockout cell line S5, the endogenous *SMN* gene is disrupted by homologous recombination, and the SMN protein is expressed from a cDNA under the control of a tetracycline (tet)-repressible promoter. Depletion of SMN by tet repression results in cell death, indicating that the function of SMN is required for viability of DT40 cells (25). Using cell viability and proliferation as an indicator, we have set up a functional assay to test whether a given protein can functionally substitute for SMN. Here we determined the role of each domain of SMN in its essential function.

MATERIALS AND METHODS

Plasmid Constructs—To generate missense and deletion mutations, human and chicken SMN cDNAs were used as templates for polymerase chain reaction-based mutagenesis. Appropriate primers were purchased from Life Technologies, Inc. Polymerase chain reaction products were digested with the appropriate restriction enzymes and ligated into the retroviral vector pMX (39). The resulting clones were fully sequenced to confirm each mutation.

Cell Culture and Retroviral Infections—S5 cells were routinely maintained in RPMI 1640 medium (Life Technologies, Inc.) containing 10% fetal bovine serum (Hyclone Laboratories), 1% chicken serum (Sigma), 2 mM L-glutamine, 50 μ M 2-mercaptoethanol, and 10 ng/ml tetracycline (Sigma). To deplete SMN, S5 cells were split into medium containing 1 μ g/ml tetracycline. Retroviral production and infection were performed essentially as described (25). S5 cells rescued by retroviral infections were maintained in medium containing 1 μ g/ml tetracycline.

Antibodies, Immunoblotting, and Immunoprecipitation—Anti- β -tubulin antibody was purchased from Sigma. A rabbit anti-SMN poly-

clonal antibody and anti-Gemin2 monoclonal antibody (mAb) 2E17 have been described (16, 25). Anti-SMN mAb 62E7 was generated in our laboratory. The number of cells in each sample was counted prior to protein analysis. A pellet of 1×10^6 cells was lysed in 40 μ l of 2 \times SDS sample buffer. 20 μ l of the lysate was loaded on a 12.5% SDS-polyacrylamide gel. Immunoblotting was performed as described previously (18). The dilutions of antibodies were 1:3000 for anti- β -tubulin antibody, 1:500 for rabbit anti-SMN polyclonal antibody, and 1:1000 for mAb 62E7.

To radioactively label cellular proteins, 5×10^6 cells were incubated for 6 h in 2 ml of methionine-free RPMI 1640 medium containing 5% fetal bovine serum, 1% chicken serum, and 40 μ Ci/ml [35 S]methionine (Amersham Pharmacia Biotech). 1 μ g/ml tetracycline was added to the labeling medium for cells rescued by retroviral infections. Immunoprecipitation of the SMN complex was performed essentially as described (18).

RESULTS

Analysis of the SMA-associated Point Mutations—About 4% of SMA patients retain at least one allele of the *SMN1* gene (12). Intragenic mutations have been found in a few of them. Several missense mutations cluster in and around the highly conserved Y/G box, including S262I, Y272C, T274I, G275S, and G279V (Fig. 1). Another missense mutation converts glutamic acid to lysine (E134K) in the Tudor domain of SMN. All these amino acids are conserved between chicken SMN (cSMN) and human SMN (huSMN) (Fig. 1). Among them, Y272C and G279V have been reported to be associated with the most severe form of the disease, type I SMA (12).

In the S5 cell line, the SMN protein is expressed exclusively from a cSMN cDNA driven by a tet-repressible promoter. Addition of 1 μ g/ml tet to the culture medium shuts off the expression and results in depletion of the cSMN protein and consequent cell death. Previously, we have shown that expression of the huSMN protein by retroviral infection completely rescued S5 cells treated with 1 μ g/ml tet (25). Using the same assay, we tested whether the missense mutations mentioned above have deleterious effects on the function of SMN that is required for DT40 cell viability. S5 cells were infected with a retrovirus containing wild-type huSMN cDNA or an huSMN cDNA harboring one missense mutation. Following infection, the cells were incubated with 1 μ g/ml tet. The majority of cells from all infections survived and continued to proliferate. Ly-

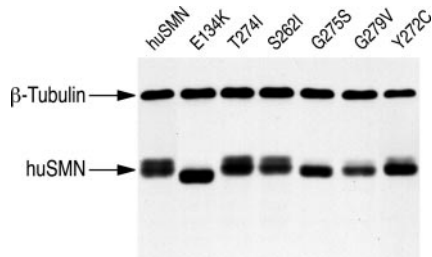


FIG. 2. SMN carrying SMA-associated point mutations rescues cell death following depletion of wild-type SMN. S5 cells were infected with a retrovirus containing wild-type huSMN cDNA or an huSMN cDNA with one missense mutation as indicated. Following infection, cells were incubated in medium containing 1 μ g/ml tet. Lysates of cells that survived the tet treatment were analyzed by Western blotting using rabbit anti-SMN polysera. Anti- β -tubulin antibody was used as a control for equal protein loading. The positions of β -tubulin and huSMN are indicated.

sates of the rescued cells were resolved by SDS-polyacrylamide gel electrophoresis and analyzed by Western blotting using an anti-SMN antibody (Fig. 2). As expected, the cSMN protein, which is smaller than huSMN, was completely depleted in the presence of 1 μ g/ml tet and was thus undetectable on the Western blot (also see Fig. 3). The level of each mutant huSMN expressed in the cells was comparable to that of wild-type huSMN. For an unknown reason, huSMN E134K showed a slightly faster migration upon SDS-polyacrylamide gel electrophoresis. The same blot was probed with anti- β -tubulin antibody as a control for equal loading of total protein (Fig. 2). The rescued cells were maintained in medium containing 1 μ g/ml tet, and the growth rates of the various cell lines were compared. No significant difference was detected between the cells expressing wild-type and mutant huSMN proteins (data not shown). As a negative control, none of the S5 cells infected with a retrovirus containing an enhanced green fluorescence protein cDNA remained viable after a 1-week incubation with 1 μ g/ml tet. We also mutated tyrosine 240 of cSMN to cysteine, which is equivalent to huSMN Y272C (Fig. 1). The mutant cSMN protein rescued the lethal phenotype as efficiently as did wild-type cSMN (data not shown). Taken together, these data indicate that the SMA-associated missense mutations tested here do not have significant effects on the function of SMN, at least as assessed by cell viability.

The C Terminus of SMN Is Essential for SMN Function—Next we asked whether the highly conserved C terminus of SMN is required for SMN activity. To this end, we generated deletion mutants of cSMN and huSMN and tested whether expression of these mutants could rescue the tet-induced lethality as described above. S5 cells were infected with a retrovirus containing a cSMN or an huSMN cDNA deleted of Ex7 (Δ Ex7) and subsequently cultured in medium containing 1 μ g/ml tet. Viable cells were readily seen after 1 week, indicating that both cSMN Δ Ex7 and huSMN Δ Ex7 are functional. However, clear differences were noted when compared with S5 cells infected with a virus containing a full-length SMN cDNA. A large percentage of S5 cells infected with the Δ Ex7 virus died during tet treatment, and the remaining viable cells grew poorly. However, the growth rates of Δ Ex7-rescued cells gradually increased when maintained in medium containing 1 μ g/ml tet. Eventually, cSMN Δ Ex7-rescued cells reached a steady doubling time of 22 h, whereas the growth pattern of huSMN Δ Ex7-rescued cells was nearly identical to that of either cSMN- or huSMN-rescued cells, with a doubling time of 16 h. At this point, expression of the full-length protein or Δ Ex7 was analyzed by Western blotting using an anti-SMN antibody (Fig. 3). As expected, in cSMN Δ Ex7-rescued cells, only Δ Ex7 (but no cSMN) was detected (Fig. 3, lane 3), and the level of

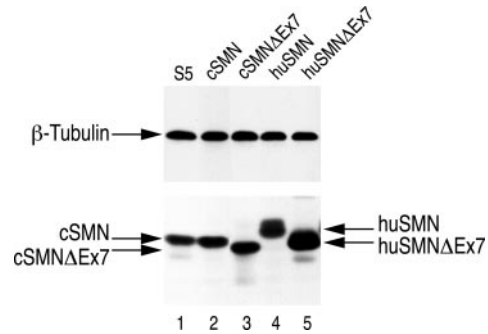


FIG. 3. Deletion of Ex7 results in a partial loss of function. S5 cells were infected with a retrovirus containing a cDNA encoding cSMN, cSMN Δ Ex7, huSMN, or huSMN Δ Ex7 and selected in the presence of 1 μ g/ml tet. Viable cells from tet selection were maintained in medium containing 1 μ g/ml tet. After cells expressing cSMN Δ Ex7 or huSMN Δ Ex7 reached a steady growth rate, lysates of the rescued cells (lanes 2–5) and control S5 cells (lane 1) were analyzed by Western blotting using rabbit anti-SMN polysera. The positions of β -tubulin, cSMN, cSMN Δ Ex7, huSMN, and huSMN Δ Ex7 are indicated.

Δ Ex7 was comparable to that of cSMN in cSMN-rescued cells (compare lanes 2 and 3). Expression of huSMN Δ Ex7 was more efficient than that of huSMN in the rescued cells (Fig. 3, compare lanes 4 and 5). We also noticed that the gain-of-growth rate of Δ Ex7-rescued cells was accompanied by an increase in the Δ Ex7 level. Apparently, cells with high expression of the Δ Ex7 protein were selected during the course of this experiment. Taken together, these results indicate that deletion of Ex7 results in a partial loss of SMN function and that the partial loss of function is compensated by higher expression of the mutant protein. Deletion of the Y/G box in addition to Ex7 (Δ Y/G) (Fig. 1) abolished SMN function, as neither cSMN Δ Y/G nor huSMN Δ Y/G restored cell viability in the presence of 1 μ g/ml tet. Expression of Δ Y/G proteins was confirmed by Western blotting shortly after retroviral infection and tet treatment, when cell death had not yet occurred (data not shown). Thus, the C terminus of SMN including the Y/G box and Ex7 is essential for the function of SMN.

Functional Analysis of the Interaction between SMN and Gemin2—Sequences near the N terminus of SMN mediate its interaction with Gemin2 (16). Most, if not all, of the SMN and Gemin2 proteins in cells are bound to each other and exist in a high molecular mass complex, suggesting that Gemin2 is a critical functional partner of SMN. We therefore anticipated that deletion of the Gemin2-binding domain would disrupt the function of SMN. However, S5 cells treated with 1 μ g/ml tet were rescued by expression of a mutant cSMN protein (Δ N39) with the first 39 amino acids truncated, including the Gemin2-binding domain and the entire phylogenetically conserved N terminus (Fig. 1). Δ N39-rescued cells showed only a slightly lower growth rate compared with full-length cSMN-rescued cells, but no other obvious abnormality (data not shown). Lysates of S5, cSMN-depleted S5, and cSMN- and Δ N39-rescued cells were analyzed by Western blotting using an anti-SMN antibody (Fig. 4, upper panel). In Δ N39-rescued cells, there was no detectable full-length cSMN, whereas Δ N39 was efficiently expressed (upper panel, lane 4). As we reported previously (25), the Gemin2 level was greatly reduced following depletion of SMN (lower panel, lane 2), suggesting that Gemin2 is unstable when not bound with SMN. If Δ N39 indeed does not interact with Gemin2, a similar reduction of Gemin2 would be expected in Δ N39-rescued cells. The same lysates were therefore probed with an anti-Gemin2 antibody (lower panel). As predicted, a significant decrease in the amount of Gemin2 was observed in Δ N39-rescued cells compared with S5 or cSMN-rescued cells (lower panel, lane 4). The extent of this decrease was identical

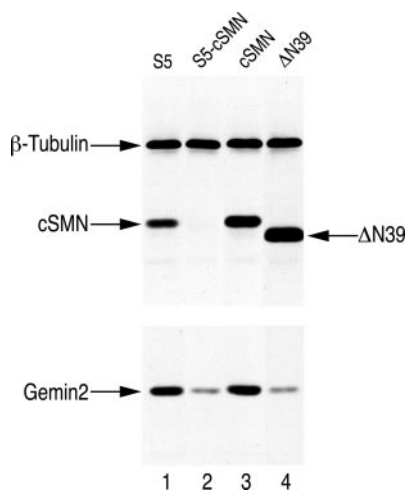


FIG. 4. The N-terminal Gemin2-binding domain of SMN is functionally dispensable for DT40 cell viability. S5 cells were infected with a retrovirus containing full-length cSMN or Δ N39 cDNA and cultured in the presence of 1 μ g/ml tet. Lysates of control S5 cells (lane 1), S5 cells depleted of cSMN (*S5-cSMN*; lane 2), and S5 cells rescued by expression of full-length cSMN (lane 3) or Δ N39 (lane 4) were analyzed by Western blotting using either rabbit anti-SMN polyclonal (upper panel) or anti-Gemin2 mAb 2E17 (lower panel). The positions of β -tubulin, cSMN, Δ N39, and Gemin2 are indicated.

to that found following SMN depletion (lower panel, compare lanes 2 and 4). We also confirmed that Δ N39 did not interact with Gemin2 in an *in vitro* binding assay (data not shown). Immunoprecipitation of extracts of Δ N39-rescued cells using an anti-SMN antibody failed to copurify Gemin2 and *vice versa* (see Fig. 7). Therefore, an interaction with Gemin2 is not required for Δ N39 to perform the function of SMN that is essential for viability of DT40 cells.

The Tudor Domain of SMN Is Dispensable for Cell Viability—The unexpected finding regarding Δ N39 prompted us to investigate whether any other regions of SMN, besides the C terminus, are essential. Internal deletions were generated to remove amino acids 45–76 (Δ K), 85–133 (Δ Tudor), or 131–192 (Δ P) of cSMN (Fig. 1). S5 cells were transduced with a retrovirus containing Δ K, Δ P, or Δ Tudor cDNA and incubated with 1 μ g/ml tet. The majority of cells remained viable after 1 week. The growth rate of the mutant-rescued cells was nearly identical to that of the full-length cSMN-rescued cells. Expression of each mutant cSMN was confirmed by Western blot analysis of lysates of the rescued cells (Fig. 5). Thus, despite the fact that the SMN Tudor domain is conserved among vertebrates (although we note that it is entirely absent in the *S. pombe* Smn protein), our data indicate that it is functionally dispensable. On the other hand, it is not surprising that the poorly conserved region covering either amino acids 45–76 or 131–192 is not essential for SMN function.

The SMN Complex in Mutant SMN-rescued Cells—Biochemical studies have demonstrated that huSMN is a component of a large multimeric complex *in vivo*, which also contains Gemin2, Gemin3, and Gemin4 as well as associated proteins including the Sm proteins (16, 18, 22). We wondered whether the same complex exists also in DT40 cells. Total proteins of DT40 cells were radioactively labeled with [³⁵S]methionine and subjected to immunoprecipitation using anti-SMN mAb 62E7 or anti-Gemin2 mAb 2E17 (Fig. 6, lanes 2 and 5). Several proteins were specifically co-immunoprecipitated with cSMN, including Gemin2 and a set of proteins whose molecular masses are very similar to those of Gemin3, Gemin4, and Sm proteins. Therefore, the composition of the SMN complex is similar in chicken and human.

Next, we analyzed the SMN complex in mutant cSMN-res-

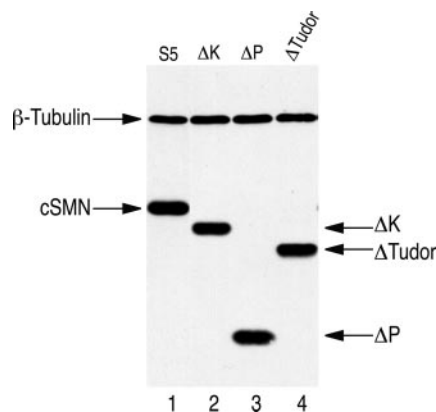


FIG. 5. Functional analysis of the central region of the SMN protein. S5 cells were infected with a retrovirus containing a cDNA encoding the internal deletion mutant Δ K, Δ P, or Δ Tudor and incubated in medium containing 1 μ g/ml tet. Lysates of control S5 cells (lane 1) and S5 cells rescued by expression of Δ K (lane 2), Δ P (lane 3), or Δ Tudor (lane 4) were analyzed by Western blotting using anti-SMN mAb 62E7. The positions of β -tubulin, cSMN, Δ K, Δ P, and Δ Tudor are indicated.

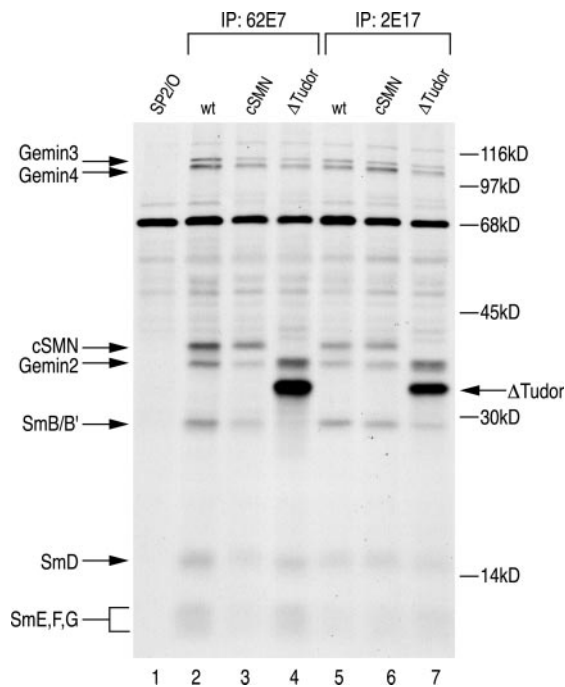


FIG. 6. Sm proteins are associated with SMN Δ Tudor *in vivo*. Total proteins of DT40 cells (*wt*; lanes 2 and 5) and S5 cells rescued by expression of full-length cSMN (lanes 3 and 6) or Δ Tudor (lanes 4 and 7) were radioactively labeled with [³⁵S]methionine and subjected to immunoprecipitation (IP) using either anti-SMN mAb 62E7 (lanes 2–4) or anti-Gemin2 mAb 2E17 (lanes 5–7). Immunoprecipitation of DT40 cell extracts by control antibody SP2/O was used as a control for antibody specificity (lane 1). The positions of protein bands of interest and protein molecular mass markers are indicated.

cued cells. Since these mutants support cell viability, the changes in the SMN complex caused by the mutations may not be essential for the function of the SMN complex. S5 cells rescued by expression of cSMN, cSMN Δ Ex7, Δ N39, or Δ Tudor were incubated in medium containing [³⁵S]methionine. Total cell lysate from equal numbers of these cells was subjected to immunoprecipitation using an anti-SMN or anti-Gemin2 antibody. Immunoprecipitation from DT40 cell lysates was used as a control for the SMN complex. As in DT40 cells, all the components of the SMN complex were copurified by either mAb 62E7 or 2E17 from cSMN-rescued cells (Fig. 6, lanes 3 and 6) and from cells expressing only Δ Tudor (lanes 4 and 7). The

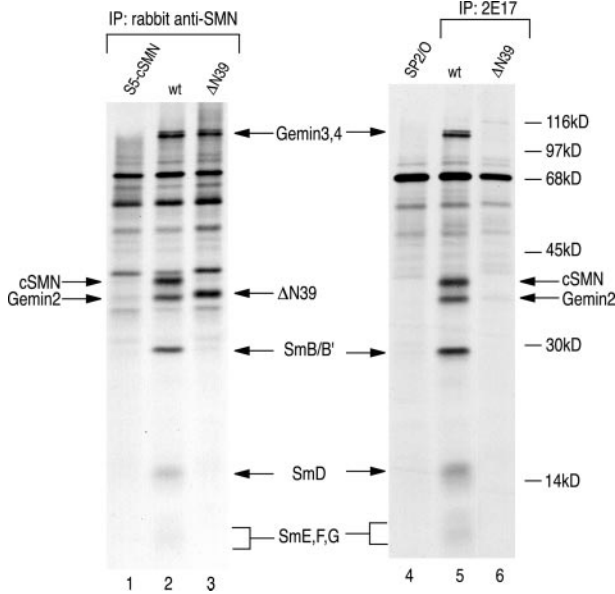


FIG. 7. SMN Δ N39 does not interact with Gemin2 or Sm proteins *in vivo*. Total proteins of DT40 cells (*wt*; lanes 2 and 5), S5 cells depleted of cSMN (*S5-cSMN*; lane 1), and S5 cells expressing only Δ N39 (lanes 3 and 6) were radioactively labeled with [³⁵S]methionine and subjected to immunoprecipitation (*IP*) using rabbit anti-SMN polyclonal (lanes 1–3) or anti-Gemin2 mAb 2E17 (lanes 5 and 6). Immunoprecipitation of DT40 cell extracts by control antibody SP2/O was used as a control for antibody specificity (lane 4). The positions of protein bands of interest and protein molecular mass markers are indicated.

amount of Δ Tudor that was immunoprecipitated was consistently much higher than that of the full-length protein, although Western blotting showed that the overall levels of the two proteins in the cells were similar (Fig. 5). It appears that free Δ Tudor protein is more soluble in the buffer we used to produce the cell extract for the immunoprecipitation experiments (data not shown). The amount of SmB/B' in lane 4 is lower than that seen in the control lane. However, in duplicate experiments, SmB/B' was co-immunoprecipitated with Δ Tudor as efficiently as with the full-length protein (data not shown). Note also that the amount of SmB/B' shown in lane 7 is comparable to that of controls shown in lanes 5 and 6. Since Gemin2 does not interact with Sm proteins, an anti-Gemin2 antibody co-immunoprecipitated Sm proteins through their interactions with Δ Tudor.

Since the epitope of mAb 62E7 lies in the N-terminal 39 amino acids of cSMN, we used a rabbit polyclonal antibody that reacts with the Tudor domain of SMN to perform the immunoprecipitation from Δ N39-rescued cells. Immunoprecipitation from extracts of S5 cells depleted of cSMN was used as a control for antibody specificity (Fig. 7, lane 1). Consistent with the inability of Δ N39 to interact with Gemin2, the antibody-captured Δ N39 (but Gemin2) was not co-immunoprecipitated from cells expressing Δ N39 (Fig. 7, lane 3). Furthermore, mAb 2E17 immunoprecipitated only a small amount of Gemin2 (Fig. 7, lane 6), confirming that Gemin2 does not interact with Δ N39 *in vivo* and therefore is not stable. Although Gemin3 and Gemin4 were still associated with Δ N39, surprisingly, Sm proteins were not co-immunoprecipitated with Δ N39 (Fig. 7, lane 3). As a control, the steady-state level of Sm proteins in Δ N39-rescued cells was identical to that in DT40 cells (data not shown).

Ex7 has been shown to be critical for SMN to oligomerize and to bind to Gemin3 and the Sm proteins *in vitro* (18, 27, 28). Gemin4 is probably recruited to the SMN complex via its interaction with Gemin3 (22). In line with these findings, very small amounts of Sm proteins, Gemin3, or Gemin4 were co-immunoprecipitated with cSMN Δ Ex7 by mAb 62E7 from

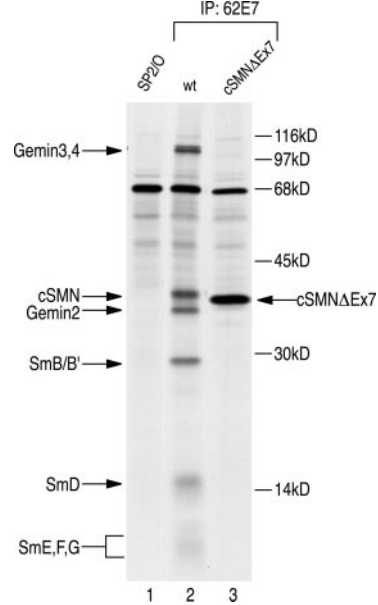


FIG. 8. Significant loss of the protein complex associated with SMN Δ Ex7. Total proteins from DT40 cells (*wt*; lane 2) and S5 cells expressing only cSMN Δ Ex7 (lane 3) were radioactively labeled with [³⁵S]methionine and subjected to immunoprecipitation (*IP*) using anti-SMN mAb 62E7 (lanes 2 and 3). Immunoprecipitation of DT40 cell extracts by control antibody SP2/O was used as a control for antibody specificity (lane 1). The positions of protein bands of interest and protein molecular mass markers are indicated.

cSMN Δ Ex7-rescued cells (Fig. 8, lane 3). In addition, reproducibly less Gemin2 was copurified with cSMN Δ Ex7 compared with immunoprecipitation from DT40 cell extracts. This phenomenon was not due to less input of immunoprecipitation since the amount of cSMN Δ Ex7 was considerably more than that of full-length cSMN (Fig. 8, compare lanes 2 and 3). Consistent with this observation, the Gemin2 level was reduced in cSMN Δ Ex7-rescued cells (data not shown). Taken together, these data suggest that the stoichiometry of the SMN complex that is observed in normal cells may not be essential for maintaining DT40 cell viability.

DISCUSSION

SMN is an essential gene in diverse organisms (13–15, 25, 26, 36). In humans, who uniquely contain two copies of SMN, a reduction of the SMN protein due to deletion or mutation in the telomeric copy causes SMA. Thus, motor neurons may have a particularly high requirement for the fundamental cellular process(es) in which SMN plays an essential role. However, it is also conceivable that SMN performs a unique tissue-specific function in motor neurons. By creating a conditional SMN knockout cell line, we have established a genetic system for characterization of the function of SMN that is essential for cell viability (25). Since depletion of SMN is lethal, restoration of cell viability and proliferation can be used as an unbiased functional assay to evaluate SMN activity.

Five of the six missense mutations that we have analyzed are within or closely adjacent to the conserved Y/G box of SMN. One mutation found in type I SMA patients, Y272C, reduces the ability of SMN to self-oligomerize and to bind to Sm proteins (27, 28). E134K in the SMN Tudor domain, another mutation reportedly associated with type I SMA, has an effect on the interaction between SMN and Sm proteins (32, 33). However, in our *in vivo* assay, both Y272C and E134K restored cell viability and proliferation and thus appear to be indistinguishable from wild-type SMN. In addition, cells expressing mutant proteins showed a growth rate identical to that of wild-type

huSMN-rescued cells when all proteins were expressed at a comparable level. We have previously shown that the growth rate of DT40 cells is dependent on the level of SMN in these cells (25). A partial loss of function, such as Δ Ex7, was readily distinguished. Therefore, we conclude that Y272C and E134K do not significantly affect the general function of SMN. It has been reported that the SMN level in a patient carrying the Y272C mutation was reduced to the same extent as that in patients homozygous for a large-scale deletion of *SMN1* (10). Thus, Y272C may render the protein relatively unstable in SMA patients, which may be a consequence of the reduced ability of the mutant protein to oligomerize (27). Our results further suggest that the functionality of Y272C is largely unaffected when the mutant protein reaches a level comparable to that of the wild-type protein. Thus, stabilizing the mutant protein may be considered a therapeutic strategy for SMA. On the other hand, since type III SMA can also be triggered by a marginal reduction of the SMN level, the mutations associated solely with type III SMA may have subtle effects on the function of SMN that were not detected in our assay.

In SMA patients with homozygous *SMN1* deletion, the SMN protein is expressed exclusively from the *SMN2* gene. Alternative splicing of *SMN2* pre-mRNA generates full-length and mostly Δ Ex7 mRNAs. Mice homozygous in the Ex7 deletion die at an early embryonic stage (40). However, mosaic mice expressing Δ Ex7 in neurons were viable, but displayed phenotypes similar to those of SMA (40). Although defects in motor neurons in the spinal cord were detected, no abnormality was reported in the brain and cerebellum, where the majority of cells also expressed only Δ Ex7. Mosaic mice expressing Δ Ex7 in skeletal muscle were also viable, but showed muscular atrophy late after birth (41). The defect was closely correlated with a severe reduction of the level of SMN in muscle cells, indicating that Δ Ex7 is an unstable protein. In our experiments, both huSMN Δ Ex7 and cSMN Δ Ex7 restored cell viability and behaved in a manner consistent with a partial loss of function. These findings indicate that Δ Ex7 retains a partial activity of SMN and suggest that the instability of Δ Ex7 is likely a major factor in the pathogenesis of SMA. Thus, a potential therapeutic strategy for SMA may be to stabilize the Δ Ex7 protein and/or to enhance the activity of Δ Ex7. To this end, Δ Ex7-rescued S5 cells may be employed for high throughput screening for compounds that increase cell growth. Further deletion of the Y/G box completely abolished the function of SMN, which, along with the high evolutionary conservation of the Y/G box, suggests that the Y/G box is a critical functional domain. Intriguingly, compared with the Y/G box, the sequence of Ex7 is poorly conserved, yet it must be functionally important, at least for murine SMN, cSMN, and huSMN. Since Ex7 is adjacent to the critical Y/G box, it is conceivable that the peptide encoded by Ex7 helps maintain the structural conformation of the adjacent Y/G box near the C terminus of SMN. Determination of the structure of SMN will shed light on this question.

The Tudor domain was first defined as a repeat element in the *Drosophila* Tudor protein. Thereafter, this domain has been found in several proteins, most of which are capable of binding to RNA or single-stranded DNA (29). SMN has also been shown to have the capacity to bind RNA (42). The function of the Tudor domain is unknown, although it has been proposed to make direct contact with nucleic acids or to mediate protein-protein interactions. The NMR structure of the Tudor domain of SMN, which bears some resemblance to the Sm domain of Sm proteins, is consistent with the latter hypothesis (32). The Tudor domain of SMN alone has been shown to bind to Sm proteins *in vitro*, and the E134K mutation, located within the

Tudor domain, was shown to disrupt the interaction with Sm proteins (32, 33). Similarly, injection of an antibody generated against the SMN Tudor domain into *Xenopus* oocytes strongly inhibits snRNP assembly, possibly by preventing SMN from binding to Sm proteins (33). Despite these implications, we found that the Tudor domain of SMN is functionally dispensable *in vivo*. In addition, the self-oligomerization of SMN and the association of SMN and Sm proteins were largely unaffected by deletion of this domain (Fig. 6 and data not shown). This appears to be consistent with our observation that the E134K mutation has no detectable effect on the function of SMN. We also note that the sequence of the Tudor domain in SMN is not evolutionarily conserved and is not present at all in *S. pombe* Smn.

The specific amino acid sequence in SMN that mediates Gemin2 binding has not been determined. A peptide corresponding to amino acids 13–44 of huSMN could compete for Gemin2 binding (16), whereas a recent report suggested that the sequence encoded by exon 2b (amino acids 52–91 of huSMN) mediates the binding to Gemin2 (43). Here we showed that deletion of amino acids 1–39 of cSMN, corresponding to amino acids 1–53 of huSMN, completely disrupted binding to Gemin2 *in vivo*. In contrast, the level of Gemin2 was not changed in cells expressing Δ K (data not shown), a mutant lacking almost the entire exon 2b, suggesting that Δ K still binds to Gemin2 *in vivo*. Therefore, our results indicate that the Gemin2-binding domain lies within amino acids 13–44 of huSMN, an evolutionarily highly conserved domain of SMN.

Surprisingly, deletion of the Gemin2-binding domain of SMN diminished the association of SMN and Sm proteins (Fig. 7). Although the C terminus of SMN has been shown to be necessary for the SMN-Sm association, it does not appear to be sufficient. Thus, it is not clear which part of SMN directly contacts Sm proteins. Our data suggest that the N terminus of SMN contributes to its interaction with Sm proteins. Similarly, Gemin2 immunoprecipitated with SMN Δ Ex7 was greatly reduced (Fig. 8). Perhaps diminishing oligomerization of SMN by Ex7 deletion also affects the interaction between SMN and Gemin2.

Expression of Δ N39, which does not interact with Gemin2, substituted for full-length cSMN and restored cell viability and proliferation. In cells expressing Δ N39, the Gemin2 level was markedly reduced. Taken together, these results indicate that Gemin2 is not required for Δ N39 to perform the function of SMN that is needed for DT40 cell viability. However, we do not know at this point whether Gemin2 is required for the function of full-length protein, *i.e.* whether the first 39 amino acids render cSMN dependent on Gemin2. Nevertheless, this unexpected finding provides new insights into the function of SMN *in vivo*. The SMN complex (containing Gemin2) has been shown to play a critical role in snRNP metabolism (24). However, a combination of disruption of Gemin2 binding and a large decrease in the Gemin2 level did not affect the function of SMN that is required for DT40 cell viability. In addition, Δ N39 was not stably associated with Sm proteins, some of the known substrates of the SMN complex. These findings suggest that SMN plays an essential role in a fundamental pathway very likely other than in snRNP metabolism in DT40 cells. This idea is further supported by our finding that Sm proteins, Gemin3, and Gemin4 are not associated with cSMN Δ Ex7, which also rescued cells depleted of full-length SMN. Gemin3, a putative RNA helicase, is also likely to play a role in the function of SMN in snRNP metabolism (18). The possibility exists that cSMN Δ Ex7 is a partial loss-of-function mutant, which accounts for loss of the association of Sm proteins, Gemin3, or Gemin4. We showed that the partial loss of function was compensated

by expression of a higher level of the mutant. Under this condition, we still could not detect formation of a complex with cSMN Δ Ex7. We therefore conclude that SMN may perform a function unrelated to snRNPs, but essential for DT40 cell viability. This function is also conserved between chicken and human, as huSMN restored cell viability in the absence of cSMN.

Since DT40 is a pre-B cell line, it is possible, however, that the role of SMN in snRNP metabolism, pre-mRNA splicing, or gene transcription is essential for other cell types or tissues. Perhaps DT40 cells have a redundant mechanism that can substitute for these functions of the SMN complex. In line with this idea, both the cellular localization of Sm proteins and the splicing patterns of several pre-mRNAs remained unchanged following depletion of SMN (data not shown). Furthermore, an huSMN mutant lacking the N-terminal 27 amino acids (huSMN Δ N27) that has been shown to be a dominant-negative mutant in HeLa cells (34, 35) functioned like the wild-type protein in DT40 cells since it rescued cells depleted of cSMN (data not shown). Thus, SMN appears to be a pleiotropic factor. Although one function of SMN is critical for some cell types, another activity may be critical for others. However, it is unclear which of these functions, if any, contributes to the development of SMA. Nevertheless, characterization of the full range of SMN activities will further the understanding of the molecular pathogenesis of SMA.

Acknowledgments—We thank Linda Abel for producing the anti-SMN monoclonal antibody 62E7. We also thank Drs. Livio Pellizzoni and Amelie Gubitz for critical reading of and for comments on this manuscript and our laboratory members for helpful discussions.

REFERENCES

- Roberts, D. F., Chavez, J., and Court, S. D. (1970) *Arch. Dis. Child* **45**, 33–38
- Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Violette, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., Le Paslier, D., Frezal, J., Cohen, D., Weissenbach, J., Munnich, A., and Melki, J. (1995) *Cell* **80**, 155–165
- Bussaglia, E., Clermont, O., Tizzano, E., Lefebvre, S., Burglen, L., Cruaud, C., Urtizberea, J. A., Colomer, J., Munnich, A., Baiget, M., and Melki, J. (1995) *Nat. Genet.* **11**, 335–337
- Hahnen, E., Forkert, R., Marke, C., Rudnik-Schoneborn, S., Schonling, J., Zerres, K., and Wirth, B. (1995) *Hum. Mol. Genet.* **4**, 1927–1933
- 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
- Rodrigues, N. R., Owen, N., Talbot, K., Ignatius, J., Dubowitz, V., and Davies, K. E. (1995) *Hum. Mol. Genet.* **4**, 631–634
- Cobben, J. M., van der Steege, G., Grootscholten, P., de Visser, M., Scheffer, H., and Buys, C. H. (1995) *Am. J. Hum. Genet.* **57**, 805–808
- Velasco, E., Valero, C., Valero, A., Moreno, F., and Hernandez-Chico, C. (1996) *Hum. Mol. Genet.* **5**, 257–263
- Hahnen, E., Schonling, J., Rudnik-Schoneborn, S., Zerres, K., and Wirth, B. (1996) *Am. J. Hum. Genet.* **59**, 1057–1065
- Lefebvre, S., Burlet, P., Liu, Q., Bertrand, S., Clermont, O., Munnich, A., Dreyfuss, G., and Melki, J. (1997) *Nat. Genet.* **16**, 265–269
- Covert, D. D., Le, T. T., McAndrew, P. E., Strasswimmer, J., Crawford, T. O., Mendell, J. R., Coulson, S. E., Androphy, E. J., Prior, T. W., and Burghes, A. H. (1997) *Hum. Mol. Genet.* **6**, 1205–1214
- Wirth, B. (2000) *Hum. Mutat.* **15**, 228–237
- 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
- Paushkin, S., Charroux, B., Abel, L., Perkinson, R. A., Pellizzoni, L., and Dreyfuss, G. (2000) *J. Biol. Chem.* **275**, 23841–23846
- Owen, N., Doe, C. L., Mellor, J., and Davies, K. E. (2000) *Hum. Mol. Genet.* **9**, 675–684
- Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997) *Cell* **90**, 1013–1021
- Giesemann, T., Rathke-Hartlieb, S., Rothkegel, M., Bartsch, J. W., Buchmeier, S., Jockusch, B. M., and Jockusch, H. (1999) *J. Biol. Chem.* **274**, 37908–37914
- Charroux, B., Pellizzoni, L., Perkinson, R. A., Shevchenko, A., Mann, M., and Dreyfuss, G. (1999) *J. Cell Biol.* **147**, 1181–1194
- Strasswimmer, J., Lorson, C. L., Breiding, D. E., Chen, J. J., Le, T., Burghes, A. H., and Androphy, E. J. (1999) *Hum. Mol. Genet.* **8**, 1219–1226
- Williams, B. Y., Hamilton, S. L., and Sarkar, H. K. (2000) *FEBS Lett.* **470**, 207–210
- Friessen, W. J., and Dreyfuss, G. (2000) *J. Biol. Chem.* **275**, 26370–26375
- Charroux, B., Pellizzoni, L., Perkinson, R. A., Yong, J., Shevchenko, A., Mann, M., and Dreyfuss, G. (2000) *J. Cell Biol.* **148**, 1177–1186
- Campbell, L., Hunter, K. M., Mohaghegh, P., Tinsley, J. M., Brasch, M. A., and Davies, K. E. (2000) *Hum. Mol. Genet.* **9**, 1093–1100
- Fischer, U., Liu, Q., and Dreyfuss, G. (1997) *Cell* **90**, 1023–1029
- Wang, J., and Dreyfuss, G. (2001) *J. Biol. Chem.* **276**, 9599–9605
- Hannus, S., Buhler, D., Romano, M., Seraphin, B., and Fischer, U. (2000) *Hum. Mol. Genet.* **9**, 663–674
- 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
- Pellizzoni, L., Charroux, B., and Dreyfuss, G. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11167–11172
- Ponting, C. P. (1997) *Trends Biochem. Sci.* **22**, 51–52
- Mushagian, A. R., Bassett, D. E., Boguski, M. S., Bork, P., and Koonin, E. V. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5831–5836
- Talbot, K., Miguel-Aliaga, I., Mohaghegh, P., Ponting, C. P., and Davies, K. E. (1998) *Hum. Mol. Genet.* **7**, 2149–2156
- Selenko, P., Sprangers, R., Stier, G., Buhler, D., Fischer, U., and Sattler, M. (2001) *Nat. Struct. Biol.* **8**, 27–31
- Buhler, D., Raker, V., Luhrmann, R., and Fischer, U. (1999) *Hum. Mol. Genet.* **8**, 2351–2357
- Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998) *Cell* **95**, 615–624
- Pellizzoni, L., Charroux, B., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001) *J. Cell Biol.* **152**, 75–85
- 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
- Hsieh-Li, H. M., Chang, J. G., Jong, Y. J., Wu, M. H., Wang, N. M., Tsai, C. H., and Li, H. (2000) *Nat. Genet.* **24**, 66–70
- Monani, U. R., Sendtner, M., Covert, D. D., Parsons, D. W., Andreassi, C., Le, T. T., Jablonka, S., Schrank, B., Rossol, W., Prior, T. W., Morris, G. E., and Burghes, A. H. (2000) *Hum. Mol. Genet.* **9**, 333–339
- Burns, J. C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J. K. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8033–8037
- Frugier, T., Tiziano, F. D., Cifuentes-Diaz, C., Miniou, P., Roblot, N., Dierich, A., Le Meur, M., and Melki, J. (2000) *Hum. Mol. Genet.* **9**, 849–858
- Cifuentes-Diaz, C., Frugier, T., Tiziano, F. D., Lacene, E., Roblot, N., Joshi, V., Moreau, M. H., and Melki, J. (2001) *J. Cell Biol.* **152**, 1107–1114
- Lorson, C. L., and Androphy, E. J. (1998) *Hum. Mol. Genet.* **7**, 1269–1275
- Young, P. J., Man, N. T., Lorson, C. L., Le, T. T., Androphy, E. J., Burghes, A. H., and Morris, G. E. (2000) *Hum. Mol. Genet.* **9**, 2869–2877