Targeting of SMN to Cajal bodies is mediated by self-association

Robert Morse, Debra J. Shaw, Adrian G. Todd and Philip J. Young*

Clinical Neurobiology, Institute of Biomedical and Clinical Science, Peninsula College of Medicine and Dentistry, Exeter EX1 2LU, UK

Received May 17, 2007; Revised and Accepted July 14, 2007

The childhood autosomal recessive disorder spinal muscular atrophy (SMA) is caused by mutations in the survival motor neuron (SMN) gene. SMN localizes diffusely in the cytoplasm and in distinct nuclear structures called Cajal bodies. Cajal bodies are believed to be the storage and processing sites of several ribonucleoproteins. Here, using a novel panel of SMN exon deletion constructs, we report a systematic analysis of internal targeting domains in the SMN protein. We demonstrate that the peptides encoded by exons 2b, 3 and 6 perform an integral role in the cellular targeting of SMN. In addition, we identify a nine amino acid motif within the highly conserved sequences of the exon 2b encoded domain that mediates Cajal body targeting and self-association. Deletion of this domain dramatically affects SMN activity and results in a dominant-negative clone. These results identify critical domains within the SMN protein and have an impact on our understanding of the SMN protein with regards to SMA as well as cellular biology.

INTRODUCTION

Spinal muscular atrophy (SMA) is a common autosomal recessive disorder that is caused by mutations of the survival motor neuron (SMN) gene (1). The SMN protein is a ubiquitously expressed, multi-functional protein that has potential roles in RNA splicing (2–6), transcription (7), cytoplasmic transport (8–10), apoptosis (11–13) and ribosomal processing (14,15).

Several functional domains have been identified within the SMN protein, including (i) an exon 2-encoded nucleic acid-binding domain (16); (ii) an exon 2b-encoded self-association site (17); (iii) an exon 3-encoded Tudor domain that interacts directly with RG-motif proteins (18–21); (iv) an exon 5-encoded poly-proline-rich domain that interacts directly with Profilin (22,23) and (v) an exon 6-encoded self-association site (24). The self-association domains are important for SMN function, as most of the identified binding partners will not interact with monomeric SMN (11,19,24–26).

The multifunctional nature of SMN is highlighted by the number of identified SMN-binding partners. These include the Gemin proteins (Gemin2–8), seven Sm core proteins, Fibrillarin, EWS, GAR1, Profilin, p53, hnRNP Q, R and U, RNA helicase A, collin and ZPR1 (27). A subset of these, collin, Sm core proteins, Fibrillarin, EWS and GAR1, are RG-motif proteins and have been shown to interact directly with the exon 3-encoded Tudor domain of SMN (18–21). In addition, symmetrical dimethylation of the arginines in the RG motif increases the affinity to SMN for several of these binding partners (20,28,29).

SMN localizes diffusely throughout the cytoplasm, and in nuclear bodies called Cajal bodies, which are believed to be ribonucleoprotein (RNP) processing and storage sites. The nature of nuclear bodies is dynamic; a point highlighted by the relationship between SMN and Cajal bodies. During fetal development and in some cultured cell lines, SMN localizes to nuclear bodies closely associated with, but distinct from Cajal bodies (30,31). These bodies are called Gems, or Gemini of Cajal bodies, due to their close relationship with Cajal bodies. However, later in development and in adult tissues, Gems become indistinguishable from Cajal bodies (30,31). This is due to an increase in binding affinity between SMN and a Cajal body marker protein, collin (32,33), with the increase attributed to arginine symmetrical dimethylation that occurs later in development (32,33).

To date, the mechanisms that mediate targeting of SMN to Cajal bodies remain unclear. To address this, we have utilized an exon deletion panel that we have used to identify subcellular targeting domains in the SMN protein. We demonstrate that localization of SMN in Cajal bodies is controlled by self-association, with removal of exon 2b, 3 or 6 independently...
sufficient to disrupt the cellular distribution of SMN. Additionally, removal of exon 1, 2a, 4, 5, or 7 does not prevent SMN from accumulating in Cajal bodies, suggesting that these domains do not contain localization signals. Furthermore, removal of a nine amino acid motif within exon 2b, ‘PAKKNKSQL’, produces a dominant-negative phenotype, resulting in extensive cell death. We demonstrate that this sequence not only plays an integral role in Cajal body targeting, but also encodes the self-association domain. Possible mechanisms and implications for SMA will be discussed.

RESULTS

Removal of exons 2b, 3 and 6 prevents SMN accumulation in Cajal bodies

In the nucleus, SMN localizes to distinct nuclear organelles called Cajal bodies, which function as RNP processing and storage sites (34–36). However, the mechanisms that mediate localization of SMN to these bodies remain unclear. To identify functional targeting domains in the SMN protein, a panel of exon deletion constructs was produced: SMN-full-length (SMN-FL), SMNΔ1, SMNΔ2a, SMNΔ2b, SMNΔ3, SMNΔ4, SMNΔ5, SMNΔ6 and SMNΔ7 (Fig. 1A). This panel was cloned into the pEGFP mammalian expression vector, allowing transient expression of the GFP-tagged truncated and full-length SMN protein in cultured cell lines. In initial studies, expression in HeLa cells identified several domains that do not appear to be involved in Cajal body targeting, since GFP-tagged SMN lacking exon 1, 2a, 4, 5, or 7 was efficiently targeted to Cajal bodies (Fig. 1A).

In contrast, removal of exon 2b, 3 or 6 prevented accumulation of expressed protein in Cajal bodies (Fig. 1A). The SMNΔ2b protein appears to exert a dominant-negative effect on endogenous protein, triggering cell death in all transfected cells (Fig. 1A and B). Cells were counterstained with...
4',6'-diamidino-2-phenylindole (DAPI), to stain nuclear DNA, and cell death was identified through nuclear shrinkage and fragmentation (Fig. 1A). The SMNΔ3 protein accumulates predominantly in the nucleolus, suggesting that exon 3 may contain a nucleolar export sequence (Fig. 1A, yellow arrows). The SMNΔ6 protein accumulates predominantly in the cytoplasm, suggesting that removal of exon 6 prevents either nuclear import or nuclear retention (Fig. 1A). In each experiment, an Sm core antibody was used to confirm that nuclear bodies were true Cajal bodies (Fig. 1A). These observations suggest that exons 2b, 3 and 6 contain Cajal body targeting domains.

SMN contains two distinct self-association domains, encoded by exons 2b and 6. Potentially, each deletion construct that contains these domains could associate with endogenous full-length SMN. To determine whether this was occurring, immunofluorescence experiments were performed on HeLa cells transfected with the deletion panel (Fig. 1B). All constructs that were targeted to Cajal bodies (SMNΔ1, Δ2a, Δ4, Δ5 and Δ7; Fig. 1A) also co-localized with endogenous SMN (Fig. 1B). In addition, the SMNΔ2b construct also co-localized in dying cells with endogenous SMN (Fig. 1B), suggesting that SMNΔ2b is sequestering endogenous SMN and exerting a dominant-negative effect. Interestingly, the SMNΔ6 construct that is mainly retained in the cytoplasm did not co-localize with endogenous SMN (Fig. 1B). In cells expressing SMNΔ6, no altered distribution of endogenous SMN was observed, with SMN localizing in distinct nuclear Cajal bodies (Fig. 1B). Like SMNΔ6, the SMNΔ3 protein failed to co-localize with or prevent the Cajal body targeting of endogenous SMN (Fig. 1B). This observation is interesting as the SMNΔ3 construct contains both of the identified self-association domains present in full-length SMN.

Although removal of exons 1, 2a, 4, 5 and 7 did not appear to prevent targeting to Cajal bodies, as these deletion constructs each contain both the exons 2b and 6 encoded self-association domains, their targeting to Cajal bodies could be through an association with endogenous SMN. To address this, these constructs were transiently expressed in SMA Type I (3813; Fig. 1C) and SMA Type II (7319; data not shown) patient fibroblasts. As in HeLa cells, each of these constructs was targeted efficiently to Cajal bodies, suggesting...
that exons 1, 2a, 4, 5 and 7 do not contain Cajal body targeting sequences (Fig. 1C).

Amino acids 74–82 rescue cells and target SMN to Cajal bodies

To determine a minimal domain encoded by exon 2b that enables correct cellular targeting, a panel of constructs containing successive deletions in SMN2b was produced: 2bΔ1, 2bΔ2, 2bΔ3, 2bΔ4 and 2bΔ5 (Fig. 2A). All constructs containing amino acids 74–82 (2bΔ4; Fig. 2A) localized in Cajal bodies. However, removal of amino acids 74–82 (2bΔ5; Fig. 2A) prevented Cajal body targeting, caused a decrease in the number of Sm core protein nuclear bodies (data not shown) and triggered cell death in all transfected cells (Fig. 2A), suggesting that the nine amino acids, PAKKNKSQK (Fig. 2B), are needed for functional SMN and are involved in cellular targeting. In keeping with this, a GFP–SMN construct in which the complete exon 2b was replaced solely with the PAKKNKSQK (SMN Δ2b + PAKK) when expressed in HeLa cells did not trigger cell death and localized to correctly Cajal bodies (Fig. 2C). These bodies also contain the Sm core proteins and endogenous SMN (data not shown), confirming they are bona fide Cajal bodies.

Removal of exon 6 prevents import of SMN into the nucleus, suggesting that self-association may play an important role in the import process (Fig. 1A). SMN contains two independent self-association domains; one encoded by exon 6 and another encoded by exon 2b (17,24). Potentially, therefore, the PAKKNKSQK sequence could contain the exon 2b encoded self-association domain. To establish whether this is the case, recombinant binding studies were performed using in vitro transcribed and translated proteins (Fig. 2D). In keeping with previous studies (17), SMN Δ2b failed to associate with SMN exon 2 (Fig. 2B; pull-down). However, in contrast, SMN Δ2b + PAKK efficiently associated with the exon 2 peptide (Fig. 2D; pull-down), suggesting the PAKKNKSQK contains the 2b encoded self-association domain. In control experiments, neither the SMN Δ2b nor the SMN Δ2b + PAKK proteins interacted non-specifically with the GST-resin (Fig. 2D; mock).

An intact Tudor domain is needed for Cajal body targeting

Removal of exon 3 causes accumulation of SMN diffusely throughout the cytoplasm, the nucleus and also in the nucleolus (Fig. 1A and B). Exon 3 encodes a Tudor domain that interacts with the RG-rich motifs found in several SMN-binding partners, including Sm core proteins, EWS, p80 and fibrillarin. To determine whether an intact Tudor domain is needed for Cajal body targeting, a panel of constructs containing successive deletions in SMN3a was produced: 3aΔ1, 3aΔ2, 3aΔ3, 3aΔ4 and 3aΔ5 (Fig. 2A). All constructs containing amino acids 74–82 (3aΔ4; Fig. 2A) localized to Cajal bodies. However, removal of amino acids 74–82 (3aΔ5; Fig. 2A) prevented Cajal body targeting, caused a decrease in the number of Sm core protein nuclear bodies (data not shown) and triggered cell death in all transfected cells (Fig. 2A), suggesting that the nine amino acids, PAKKNKSQK (Fig. 2B), are needed for functional SMN and are involved in cellular targeting. In keeping with this, a GFP–SMN construct in which the complete exon 3 was replaced solely with the PAKKNKSQK (SMN Δ3 + PAKK) when expressed in HeLa cells did not trigger cell death and localized to correctly Cajal bodies (Fig. 2C). These bodies also contain the Sm core proteins and endogenous SMN (data not shown), confirming they are bona fide Cajal bodies.
domain is required for subcellular targeting, or whether a minimal exon 3-domain can rescue Cajal body targeting, a panel of constructs containing successive deletions of exon 3 was produced: 3\_D\_1, 3\_D\_2, 3\_D\_3, 3\_D\_4, 3\_D\_5 and 3\_D\_6 (Fig. 3). In all of the exon 3 deletion constructs, SMN failed to properly localize within Cajal bodies (Fig. 3), suggesting that an intact Tudor domain is essential for SMN targeting to Cajal bodies. Exon 3 deletion constructs were readily detectable; however, the expressed proteins were all observed aggregating in the nucleolus (Fig. 3). Sm core proteins and endogenous SMN, however, still accumulate in normal bodies (data not shown).

An intact exon 6-encoded domain is needed for the correct nuclear targeting of SMN

The potent self-association domain encoded by SMN exon 6 is central to SMN function. Inhibition of SMN self-association triggers protein instability and formation of non-functional SMN monomers. In keeping with this, several pathogenic missense mutations occur in exon 6. Removal of exon 6 causes the cytoplasmic accumulation of SMN (Fig. 1A and B). To determine whether an intact exon 6 domain is needed, or whether a minimal exon 6 domain can rescue nuclear targeting, a panel of constructs containing successive deletions of exon 6 was produced: 6\_D\_1, 6\_D\_2, 6\_D\_3, 6\_D\_4, and 6\_D\_5 (Fig. 4). As with the exon 3 deletion panel (Fig. 3), removal of any part of exon 6 prevented the correct nuclear targeting, with the expressed protein localizing predominantly in the cytoplasm (Fig. 4). Removal of exon 6 did not exert a dominant-negative effect on endogenous SMN, suggesting that removal of the main self-association domain prevents association with the endogenous full-length SMN protein.

Single pathogenic missense mutations in exons 3 and 6 occur in various SMA patients. These include the E134K mutations in exon 3, and the S262I and Y272C mutations in exon 6. To determine whether these mutations prevent Cajal body targeting, a panel of constructs containing each of these mutations was produced: 6\_D\_1, 6\_D\_2, 6\_D\_3, 6\_D\_4, and 6\_D\_5 (Fig. 4). Each construct is represented schematically with the deleted nucleotides indicated (left-hand panel). All nuclei were counterstained with DAPI (blue). Bars represent 30 \( \mu \text{m} \).

DISCUSSION

Using a novel panel of SMN exon deletion cDNA constructs we have performed a systematic analysis of the SMN protein and demonstrated that the removal of the domains encoded by exons 2b, 3 and 6 prevents the efficient targeting of SMN to nuclear Cajal bodies (Fig. 1A–C). Various functions have previously been assigned to each of these domains: exons 2b and 6 encode self-association domains (17), whereas exon 3 encodes a Tudor domain that interacts directly with RG domains identified in various SMN binding partners, including Sm core protein,
EWS, coilin and fibrillarin (5,6,14,32). This suggests that self-association and RG-motif binding could play an important role in efficient Cajal body targeting of SMN. These observations are in keeping with a previous report (37), which states that subnuclear targeting of SMN is mediated by the exons 3 and 6 encoded domains. In this previous study, multiple domains were deleted from expressed proteins and the combinatory effects were monitored (37). Here we extend this study, using a novel exon deletion construct panel, through which we have individually deleted each exon-encoded domain. This has allowed us to individually monitor the effect removal of each domain has on SMN cellular distribution.

Our most interesting finding is the dominant-negative effect expression of SMNΔ2b exerts on endogenous SMN (Figs 1A and B and 2B). Using our exon 2b panel, we have successfully identified a nine amino acid sequence that is independently sufficient to rescue cells from the dominant-negative effect of SMNΔ2b. The incorporation of these amino acids (PAKKNKSQK; amino acids 74–82) into the SMNΔ2b construct prevents cell death and enables the correct targeting of SMN to Cajal bodies (Fig. 2C). In addition, we have further characterized this sequence and demonstrated that it contains the exon 2b-encoded self-association domain, with the SMNΔ2b + PAKK protein associating with SMN exon 2, whereas SMNΔ2b does not (Fig. 2D). It is important to note that the first five amino acids of this sequence (‘PAKKN’) are missing in the 2bΔ3 construct, which is correctly targeted to Cajal bodies and does not trigger a dominant-negative effect in endogenous SMN (Fig. 2A). This suggests that the true minimal domain is the last four amino acids of our identified sequence (‘KSQK’). However, it has previously been reported that the lysines in the ‘PAKKNKSQK’ sequence are involved in nucleolar targeting of SMN (discussed subsequently) (37). Therefore, we decided to study the ‘PAKKNKSQK’ in its entirety.

Figure 6. A working model for the dominant-negative effect exerted by SMNΔ2b on endogenous SMN. In our model, wild-type dimers are held together by the exon 6-encoded self-association domain, with dimers associating through the 2b-encoded self-association domain to form multimers. Our model suggests that SMNΔ2b could form homomeric dimers through the exon 6-encoded self-association domain that are incapable of forming multimers. The dominant-negative effect could be due to formation of heterodimers consisting of SMNΔ2b/endogenous SMN. These heterodimers would have a reduced ability to form functional multimeric complexes due to the absence of the exon 2b self-association domain in SMNΔ2b.

Of the two identified SMN self-association domains, the exon 6–6 interaction is stronger than the 2b–2b interactions (17). Therefore, we have previously suggested a working model through which minimal functional SMN dimers are formed through the 6–6 association, with these dimers associating through the 2b–2b site to form the previously reported multimeric complexes (17) (Fig. 6). This fits in with our working hypothesis (Fig. 6) in which SMNΔ2b triggers cell lethality by sequestering endogenous SMN through the exon 6-encoded self-association domain, but prevents multimeric complex formation through the absence of the 2b-encoded domain (Fig. 6). Under this working hypothesis, re-insertion of the ‘PAKKNKSQK’ sequence rescues cell viability because this sequence contains the self-association domain and therefore enables the heterodimers containing endogenous and SMNΔ2b + PAKK to form the larger multimeric complexes that have previously been identified (38) (Fig. 6). It is interesting to note that a recent study has demonstrated that formation of these large macromolecular SMN complexes is dependent on the ability of SMN to interact directly with Gemin2 (39) and that SMN contains two Gemin2-binding sites, a C-terminal one encoded by exons 6–7 and a second
in the N-terminus. Although the N-terminal binding domain was originally identified as an exon 2a-encoded-domain (2), we have previously reported that the domain is actually encoded by exon 2b (17). If our mapping data are correct, removal of 2b would not only prevent the N-terminal SMN self-association, but also prevent the interaction with Gemin2, which would essentially prevent the formation of these larger functional complexes (39). We are currently performing sucrose gradient studies to confirm whether SMN\(\Delta 2b\) is capable of forming the previously described larger multimeric complex and whether SMN\(\Delta 2b\) is not.

The ‘PAKKKNKSQK’ sequence has previously been identified as a K-rich nucleolar localization sequence (NoLS) (37). Similar NoLSs are found in MDM2, ARF and coilin (37). Here, in reverse experiments, we demonstrate here that removal of the domain encoded by the adjacent exon 3 triggers a nucleolar accumulation of the SMN protein (Fig. 3). This may suggest that not only does the ‘PAKKKNKSQK’ sequence mediate self-association, it also enables efficient targeting to both the nucleolus and Cajal bodies. There is a clear association between Cajal bodies and the nucleolus (31,40). SMN has been identified as a nucleolar protein (31,40), and various nucleolar proteins, including fibrillarin (40), GAR1 (15) and nucleolin (14), have been reported in Cajal bodies. In addition, Cajal bodies have been identified associated with the nucleolar periphery (30), and under certain conditions Cajal bodies have been shown to enter the nucleolus (41).

Therefore, it is not surprising that Cajal body proteins contain nucleolar-targeting sequences (37). This close association is thought to aid the nuclear processing of the various RNP, including the splicing U snRNPs (42–44). SMN is involved in both the cytoplasmic assembly of the U snRNPs and their subsequent nuclear import (18,38,45–47). In the nucleus, the U snRNPs are initially taken to the nucleolus, then to Cajal bodies and finally to the interchromatin granule clusters (42–44). The identification of nuclear import sequences in both SMN and coilin, and also Cajal body targeting domains in SMN, could suggest that SMN, as well as coilin, is involved in the nuclear distribution and trafficking of the U snRNPs. This is in keeping with a previous report that suggests SMN is involved in the recycling of Sm core proteins and U snRNPs (5). It is interesting that the removal of the exon 3-encoded Tudor domain mediates the nucleolar accumulation of SMN (Fig. 3). This suggests that the Tudor domain functions as a nucleolar export sequence. We were unable to identify a minimal domain that functioned as the export sequence, with removal of any section from the N- or C-terminus of exon 3 causing the nucleolar localization. This suggests that rather than containing an actual export sequence, exon 3 may interact with an associated factor that mediates this export. Interestingly, with regard to a potential role in the U snRNP maturation pathway, the Tudor domain interacts directly with the RG motif found on Sm D1, D3 and B core proteins (18,45), and the association with the U snRNP could function as an export signal. With reference to this, it would be interesting to see the effect that the recently reported W92S, exon 3 SMN mutation that is found close to the 2b/3 exon boundary has on the nuclear distribution of SMN (48).

Our data also suggest that the removal of exons 1, 2a, 4, 5 and 7 alone does not prevent the accumulation of expressed GFP-tagged protein in Cajal bodies in HeLa cells (Fig. 1A and B) and human skin fibroblasts derived from SMA Type I (Fig. 1C) and II (data not shown) patients. Each of these constructs displays a similar distribution to full-length GFP-SMN in patient fibroblasts (Fig. 1C), suggesting removal of these domains does not alter the in vivo characteristics of the expressed protein. However, although 3813 fibroblasts display a reduced level of endogenous SMN protein (49), we cannot rule out the possibility that endogenous levels are still high enough to ensure the formation of endogenous SMN/truncated SMN heterodimers and that through the presence of endogenous SMN, these heterodimers are targeted to the correct cellular organelles. To address this, we are validating a working model through which we can knock-out endogenous SMN expression using RNAi technology in cell lines stably expressing our deletion constructs. These experiments will conclusively demonstrate the importance of the individual exons and will be reported in due course.

The importance of the exon 6-encoded self-association domain is highlighted by our SMN\(\Delta 6\) construct, which failed to associate with endogenous SMN and accumulated mainly in the cytoplasm (Fig. 1A and B). As a similar cytoplasmic accumulation was identified for all the exon 6 deletion constructs (Fig. 4), the presence of a complete and intact exon 6 domain appears essential for the correct subcellular localization of SMN (Fig. 4). This suggests that SMN transport into the nucleus is directly linked to its ability to self-associate. It is possible that the exon 6-encoded domain is involved in the nuclear import of SMN and/or the nuclear retention of SMN. Although all our exon 6 deletion constructs localize predominantly to the cytoplasm, they all display a weaker diffuse nuclear staining (Figs 1A and B and 4), suggesting that while the SMN\(\Delta 6\) protein can enter the nucleus it does not accumulate there. This again may suggest that instead of mediating nuclear import, exon 6 may play a functional role in nuclear retention. In addition, it has previously been shown that an SMN protein lacking the domains encoded by exons 6 and 7 (SMN\(\Delta 6–7\)) is imported into the nucleus but is not targeted to Cajal bodies (37). This report adds additional weight to the hypothesis that the exon 6-encoded domain is involved in nuclear retention rather than nuclear import. Also, the difference between the SMN\(\Delta 6–7\) construct, which is predominantly nuclear (37), and our SMN\(\Delta 6\) protein, which is predominantly in the cytoplasm, suggests two things: (i) exon 7 may encode a nuclear export sequence and (ii) although exon 6 is not essential for nuclear import, it is involved in Cajal body targeting. Removal of exon 7 has previously been shown to cause the nuclear accumulation of SMN in neuronal cells (10). This suggests that although the exon 7 domain is not essential for the nuclear import of SMN, it plays a pivotal role in nuclear export. However, in the experiments presented here, removal of exon 7 does not appear to alter the cellular distribution of SMN (Fig. 1A). It is important to note that our experiments have been performed in HeLa cells, whereas the experiments by Zhang et al. (10) were performed in cultured neuronal cells, suggesting that the nuclear accumulation of SMN\(\Delta 7\) could be a neuronal specific feature. This, therefore, could highlight a fundamental difference
between neuronal and non-neuronal cells. Recent experiments in SMNΔ7 mice have demonstrated that SMNΔ7 is capable of prolonging mouse survival (50), suggesting that SMNΔ7 is partially functional. However, the failure of SMNΔ7 to compensate fully for reductions in full-length SMN protein may be explained by the nuclear accumulation of SMNΔ7 in neuronal cells. It is possible that SMNΔ7 is efficiently targeted to the cytoplasm in most cell types, whereas it is retained in the nucleus in motor neurons. If this is the case, it would suggest that SMA arises through the functional loss of cytoplasmic SMN.

We have previously demonstrated that an exon 3 missense mutation (E134K) and two exon 6 missense mutations (S262I and Y272C) that each trigger SMA are capable of prolonging mouse survival (50), suggesting that SMN protein potentially capable of compensating for SMN1 point mutations.

**MATERIALS AND METHODS**

**cDNA constructs**

SMN exon deletion constructs were cloned into the pEGFP (BD Bioscience) mammalian expression vector in a stepwise manner using primers listed in Table 1 to produce the constructs listed in Table 2. All products were amplified by PCR using a previously reported full-length SMN cDNA cloned into pET 32 (Novagen) as a template (11,30). SMNΔ1 was amplified using Exon 2a forward (Xhol)/Exon 7 reverse (BamHI) primers and cloned into the respective vector sites. SMNΔ7 was amplified using Exon 1 forward (Xhol)/Exon 6 reverse (HindIII) primers and cloned into the respective vector sites. The additional deletion constructs were produced in stepwise manner. Briefly, the 5' fragments (exon 1, exons 1–2a, exons 1–2b, exons 1–3 and exons 1–4) were cloned into the 5' Xhol/3' HindIII vector sites and the 3' fragments (exons 2b–7, exons 3–7, exons 4–7, exons 5–7 and exons 6–7, respectively) were cloned downstream into the 5' HindIII/3' BamHI vector sites to produce SMNΔ2a, SMNΔ2b, SMNΔ3, SMNΔ4 and SMNΔ5. SMNΔ6 was produced by cloning exons 1–5 into the 5' Xhol/3' HindIII vector sites and then commercially produced coding and non-coding oligonucleotides corresponding to exon 7 were then annealed, digested and cloned downstream (Table 2).

**Cell culture and immunohistochemistry**

Human cervical carcinoma (HeLa) cells, Type I SMA patient fibroblasts (3813) and Type II SMA patient fibroblasts (7319) were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) fetal calf serum (FCS) and 1% (w/v) each of penicillin and streptomycin in 5% (v/v) CO2 at 37°C. Subconfluent cells were grown on coverslips and transfected with 1 μg cDNA using GeneJuice (Novagen) according to manufacturer's recommendations. Briefly, for a 2 μg cDNA transfected, 100 μl of DMEM lacking FCS was incubated with 6 μl of GeneJuice at room temperature for 5 min. Two micrograms of supercoiled cDNA were then added, mixed by pipetting and incubated for a further 15 min at room temperature. The mixture was then added dropwise to the cells in DMEM lacking FCS. Cells were incubated for 4 h in 5% (v/v) CO2 at 37°C, then washed and re-fed with DMEM containing 10% FCS and incubated for a further 24 h. Transfected cells were washed three times with phosphate-buffered saline (PBS) and then fixed with 50% acetone/50% methanol. Immunohistochemistry staining was carried out as previously described. Sm core proteins were identified using monoclonal antibody (Y12;
Table 2. Primer combinations used for each pEGFP cDNA construct

<table>
<thead>
<tr>
<th>Construct</th>
<th>Primers (number in Table 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMN</td>
<td>1, 33 (exons 1–7)</td>
</tr>
<tr>
<td>SMN1</td>
<td>3, 33 (exons 2a–7)</td>
</tr>
<tr>
<td>SMN2a</td>
<td>1, 2 (exon 1); 5, 33 (exons 2b–7)</td>
</tr>
<tr>
<td>SMN2b</td>
<td>1, 4 (exons 1–2a); 7, 33 (exons 3–7)</td>
</tr>
<tr>
<td>SMN3</td>
<td>1, 6 (exons 1–2b); 9, 33 (exons 4–7)</td>
</tr>
<tr>
<td>SMN4</td>
<td>1, 8 (exons 1–3); 11, 33 (exons 5–7)</td>
</tr>
<tr>
<td>SMN5</td>
<td>1, 10 (exons 1–4); 13, 33 (exons 6–7)</td>
</tr>
<tr>
<td>SMN6</td>
<td>1, 12 (exons 1–5); 15, 16 (snap anneal; exon 7)</td>
</tr>
<tr>
<td>SMN7</td>
<td>1, 14 (exons 1–6)</td>
</tr>
<tr>
<td>2hΔ1</td>
<td>1, 4 (exons 1–2a); 17, 33 (exons 2Δ1–7)</td>
</tr>
<tr>
<td>3Δ6</td>
<td>1, 27 (exons 1–3Δ6); 9, 33 (exons 4–9)</td>
</tr>
<tr>
<td>6Δ1</td>
<td>1, 12 (exons 1–5); 28, 33 (exons 6Δ1–7)</td>
</tr>
<tr>
<td>6Δ2</td>
<td>1, 12 (exons 1–5); 29, 33 (exons 6Δ2–7)</td>
</tr>
<tr>
<td>6Δ3</td>
<td>1, 12 (exons 1–5); 30, 33 (exons 6Δ3–7)</td>
</tr>
<tr>
<td>6Δ4</td>
<td>1, 31 (exons 1–6Δ4); 15, 16 (snap anneal; exon 7)</td>
</tr>
<tr>
<td>6Δ5</td>
<td>1, 32 (exons 1–6Δ5); 15, 16 (snap anneal; exon 7)</td>
</tr>
</tbody>
</table>

Santa Cruz) and visualized using secondary antibodies conjugated to TRITC. Cell nuclei were counterstained with DAPI.

**TNT quick coupled system and recombinant binding study**

Recombinant GST-Exon 2 protein which has previously been reported (17), and recombinant His-SMNΔ2b and His-SMNΔ2b + PAKK were expressed using the non-radioactive TNT quick coupled system, following manufacturer’s protocol (Promega). Briefly, supercoiled plasmid constructs encoding exon 2, SMN2a and SMNΔ2b + PAKK were incubated with TNT mastermix containing biotinylated lysine and methionine for 90 min at 30°C. Resulting recombinant protein was analysed via western blot using a streptavidin (SA)-HRP conjugate. For recombinant binding studies, 0.5 µg of recombinant His-SMN2a and His-SMNΔ2b + PAKK were incubated with 0.5 µg GST-Exon 2 for 1 h at 4°C. Twenty-five microlitres of washed GST Bind-resin was then added to each reaction for 1 h at 4°C. GST–protein complexes were then washed three times in PBS containing 0.5% Triton X-100. Negative control pull-downs were used, in which His-SMNΔ2b and SMNΔ2b + PAKK were incubated with GST Bind-resin in the absence of GST-Exon 2, to eliminate non-specific cross-reaction between the His-proteins and the GST Bind-resin. All pull-down experiments were analysed by western blot using an SA-HRP.

**ACKNOWLEDGEMENTS**

P.J.Y. and R.M. were supported by a fellowship from the Vandervell Foundation and the Northcott Devon Medical Foundation. D.S. was supported by an Andrews’ Buddies/Fight SMA, IBCS joint studentship. A.G.T. was supported by IBCS. P.J.Y. thanks Dr Belina Hall (Imperial College, London) for the 7319 SMA patient fibroblasts, Professor Glenn E. Morris for the 3813 SMA patient fibroblasts and Dr Christian L. Lorson (University of Missouri) for comments and support.

Conflict of Interest statement. None declared.


