



**HAL**  
open science

## **SMN2 exon 7 splicing is inhibited by binding of hnRNP A1 to a common ESS motif that spans the 3' splice site**

Thomas Koed Doktor, Lisbeth Dahl Schroeder, Anne Vested, Johan Palmfeldt, Henriette Skovgaard Andersen, Niels Gregersen, Brage S. Andresen

### ► **To cite this version:**

Thomas Koed Doktor, Lisbeth Dahl Schroeder, Anne Vested, Johan Palmfeldt, Henriette Skovgaard Andersen, et al.. SMN2 exon 7 splicing is inhibited by binding of hnRNP A1 to a common ESS motif that spans the 3' splice site. *Human Mutation*, 2011, 32 (2), pp.220. 10.1002/humu.21419 . hal-00610797

**HAL Id: hal-00610797**

**<https://hal.science/hal-00610797>**

Submitted on 25 Jul 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**SMN2 exon 7 splicing is inhibited by binding of hnRNP A1 to a common ESS motif that spans the 3' splice site**



Journal:	<i>Human Mutation</i>
Manuscript ID:	humu-2010-0396.R2
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	29-Oct-2010
Complete List of Authors:	<p>Doktor, Thomas; University of Southern Denmark, Department of Biochemistry and Molecular Biology; Aarhus University Hospital, Research Unit for Molecular Medicine; Aarhus University, Department of Human Genetics</p> <p>Schroeder, Lisbeth; Aarhus University Hospital, Research Unit for Molecular Medicine; Aarhus University, Department of Human Genetics</p> <p>Vested, Anne; Aarhus University Hospital, Research Unit for Molecular Medicine</p> <p>Palmfeldt, Johan; Aarhus University Hospital, Research Unit for Molecular Medicine</p> <p>Andersen, Henriette; University of Southern Denmark, Department of Biochemistry and Molecular Biology</p> <p>Gregersen, Niels; Aarhus University Hospital, Research Unit for Molecular Medicine</p> <p>Andresen, Brage; University of Southern Denmark, Department of Biochemistry and Molecular Biology</p>
Key Words:	SMN1, SMN2, hnRNP A1, exon skipping, splicing, exon splicing silencer, ESS

SCHOLARONE™  
Manuscripts

1  
2 **SMN2 exon 7 splicing is inhibited by binding of hnRNP A1 to a common ESS motif that**  
3  
4 **spans the 3' splice site**  
5  
6

7  
8 Thomas Koed Doktor<sup>1,2,3</sup>, Lisbeth Dahl Schroeder<sup>2,3</sup>, Anne Vested<sup>2</sup>, Johan Palmfeldt<sup>2</sup>,  
9  
10 Henriette Skovgaard Andersen<sup>1</sup>, Niels Gregersen<sup>2</sup> and Brage Storstein Andresen<sup>1\*</sup>  
11

12  
13  
14  
15 <sup>1</sup>*Department of Biochemistry and Molecular Biology, University of Southern Denmark,*  
16 *Campusvej 55, 5230, Odense M, Denmark;* <sup>2</sup>*Research Unit for Molecular Medicine, Aarhus*  
17 *University Hospital, Skejby, Brendstrupgaardsvej 100, 8200 Aarhus N, Denmark;*  
18  
19 <sup>3</sup>*Department of Human Genetics, Aarhus University, 8000 Aarhus C, Denmark.*  
20  
21

22  
23  
24  
25 *\*Corresponding author:* Brage Storstein Andresen,  
26

27 Department of Biochemistry and Molecular Biology,  
28

29 University of Southern Denmark,  
30

31 Campusvej 55, DK-5230, Odense M, Denmark  
32

33 Phone: 45-65502413  
34

35 Fax: 45-65502467  
36

37 **e-mail: bragea@bmb.sdu.dk**  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**ABSTRACT**

Spinal Muscular Atrophy is caused by homozygous loss of *SMN1* with phenotypic modulation by *SMN2*. *SMN2* expresses only limited amounts of full-length transcript due to skipping of exon 7 caused by disruption of an SF2/ASF binding ESE. Additionally, hnRNP A1 has been reported to inhibit inclusion of *SMN2* exon 7.

We previously reported high similarity between the sequence spanning the 3'ss of *SMN1* and *SMN2* exon 7 and an hnRNP A1 binding ESS, which regulates *MCAD* exon 5 splicing. We show here that this 3'ss motif indeed functions as a crucial hnRNP A1 binding ESS, which inhibits inclusion of *SMN1/2* exon 7 and is antagonized by the *SMN1* ESE, but not by the inactive *SMN2* sequence. Pull-down experiments revealed a specific interaction between hnRNP A1 and the 3'ss AG-dinucleotide, which could be disrupted by mutations shown to improve splicing in reporter minigenes.

Genomic analyses revealed that in the human genome, 3'ss matching the *SMN1/2* ESS motif region are much less abundant than 3'ss with a disrupted ESS motif. This indicates that this ESS may be a general splicing inhibitory motif, which binds hnRNP A1 and inhibits exon inclusion by binding to 3'ss harboring this ESS motif.

**Key words**

*SMN1*; *SMN2*; hnRNP A1; exon skipping; splicing; exonic splicing silencer; ESS

Formatted: Font: Italic

Formatted: Font: Italic

**Deleted: Databases**  
 SMN1 – OMIM: 600354; GENBANK: NG\_008691.1, NM\_000344.3  
 SMN2 – OMIM: 601627; GENBANK: NG\_008728.1, NM\_017411.3  
<http://genes.mit.edu/pictogram.html> (Pictogram)  
<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home> (ESEfinder)  
<http://genes.mit.edu/burgelab/rescue-ese/> (RESCUE-ESE)  
<http://cubweb.biology.columbia.edu/pesx/> (PESX)  
<http://katahdin.cshl.edu/SpliceRack/index.cgi?database=spliceNew> (SpliceRack)  
<http://genome.ucsc.edu/> (UCSC Genome Browser)

Formatted: English (U.S.)

Formatted: English (U.S.)

## INTRODUCTION

Spinal Muscular Atrophy (SMA) type I (OMIM #253300), type II (OMIM #253550) and type III (OMIM #253400) are neurodegenerative disorders characterized by progressive loss of lower motor neurons (Crawford and Pardo, 1996). SMA is caused by decreased levels of survival of motor neuron (SMN) protein usually caused by deletion of the functional *SMN1* gene (Lefebvre, et al., 1995). A secondary gene, *SMN2*, provides a low level of SMN protein, which is sufficient for survival, but not adequate for proper development and continued survival of the lower motor neurons (Lefebvre, et al., 1997; Monani, et al., 1999). Inter-individual variation in the copy numbers of the *SMN2* gene account for most of the phenotypic differences in the clinical subtypes of SMA (Feldkotter, et al., 2002; Lefebvre, et al., 1997; Velasco, et al., 1996; Wirth, et al., 1999).

*SMN2* differs from *SMN1* primarily by a silent c.840C>T transition (corresponding to position +6 in exon 7), which disrupts an exonic splicing enhancer (ESE) within exon 7. The result is predominant skipping of exon 7 (Cartegni, et al., 2002; Cartegni, et al., 2006; Lorson and Androphy, 2000; Monani, et al., 1999) and production of an unstable protein with decreased activity (Burnett, et al., 2009; Butchbach, et al., 2007; Cho and Dreyfuss; Le, et al., 2005). It has also been suggested that the c.840C>T transition instead creates an hnRNP A1 binding exonic splicing silencer (ESS), which causes exon 7 skipping in *SMN2* (Kashima and Manley, 2003; Kashima, et al., 2007a). Additionally, several other hnRNP A1 binding elements in intron 6 and intron 7 of *SMN2* have been reported to affect splicing (Hua, et al., 2010; Hua, et al., 2007; Hua, et al., 2008; Kashima, et al., 2007b), and inactivation of one or several of these elements increases inclusion of *SMN2* exon 7.

We initially noted that there is a strong similarity between the region surrounding the 6<sup>th</sup> nucleotide of *SMN1* and *SMN2* exon 7 and a region located in exon 5 in the *MCAD* gene (*ACADM*). In this region *MCAD* exon 5 harbors a SF2/ASF binding ESE and exon 5 is

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

1  
2 skipped when this ESE is disrupted by mutations, resulting in loss of activity and MCAD  
3  
4 deficiency (OMIM: 607008) (Nielsen, et al., 2007). A synonymous single nucleotide  
5  
6 polymorphism (SNP) (c.351A>C) located upstream of the ESE in *MCAD* exon 5 confers  
7  
8 immunity towards a mutation (c.362C>T) that disrupts the SF2/ASF binding ESE causing  
9  
10 exon skipping and ultimately MCAD deficiency. We have demonstrated that the c.351C  
11  
12 variant in *MCAD* disrupts a CAGGGG ESS element, which functions by binding of hnRNP  
13  
14 A1 (Nielsen, et al., 2007). In the *MCAD* gene, the function of the SF2/ASF binding ESE is to  
15  
16 suppress binding of hnRNP A1 to the ESS. When the ESS function is disrupted, the ESE is  
17  
18 no longer required for exon inclusion. Interestingly, a sequence very similar to the ESS  
19  
20 element in *MCAD* is present in *SMN1* and *SMN2* exon 7 at the exact same position within the  
21  
22 region of high similarity (Fig 1). The potential ESS sequence element in *SMN1* and *SMN2* is  
23  
24 located across the 3'ss of exon 7 and includes nucleotides previously associated with an  
25  
26 extended splicing inhibitory context, which has been suggested to be mediated by formation  
27  
28 of a terminal stem-loop structure that inhibited access to the 3'ss (Singh, et al., 2004).

Field Code Changed

29  
30 Based on our observations from the *MCAD* gene, we hypothesized that instead of an  
31  
32 inhibitory stem loop structure, the sequence spanning the 3' splice site of exon 7 in *SMN1* and  
33  
34 *SMN2* may function as an ESS and that binding to the ESS is counteracted by the functional  
35  
36 ESE at position +6 in *SMN1*, whereas loss of the ESE in *SMN2* shifts the balance, allowing  
37  
38 increased binding of hnRNP A1 to the ESS, thereby blocking the 3'ss and splicing of exon 7.

39  
40 The principal protein recognizing the 3'ss is the U2AF<sup>35</sup> protein, the smaller subunit in the  
41  
42 U2AF complex (Merendino, et al., 1999; Wu, et al., 1999). It resembles an SR protein by  
43  
44 having an RRM motif and an RS-domain but also interacts specifically with the larger  
45  
46 subunit of the U2AF complex, U2AF<sup>65</sup>, which binds to the polypyrimidine tract (Wu and  
47  
48 Maniatis, 1993; Zamore, et al., 1992; Zhang, et al., 1992). The U2AF complex is thought to  
49  
50 be recruited by SR proteins and together with the branch site binding protein SF1 it recruits

Field Code Changed

Field Code Changed

1 the U2 snRNP to the branch site to initiate spliceosome assembly (Brow, 2002; Graveley, et  
2 al., 2001; Kent, et al., 2005). Prevention of binding of the U2AF complex to the 3'ss inhibits  
3  
4  
5 formation of the spliceosome and usage of the 3'ss (Domsic, et al., 2003; Masuda, et al.,  
6  
7  
8 2008; Warf, et al., 2009), and ESS motifs in the immediate vicinity of the 3'ss may  
9  
10 consequently cause skipping of the exon. Although some splice sites are AG-independent due  
11  
12 to a strong polypyrimidine tract and able to undergo splicing in the absence of U2AF<sup>35</sup>, direct  
13  
14 blocking of a 3'ss by binding of splicing inhibitory factors to an overlapping ESS would still  
15  
16 be expected to inhibit splicing. Furthermore, the polypyrimidine tract of the 3'ss of *SMN1/2*  
17  
18 exon 7 has been shown to be functionally suboptimal as substitutions of purines for  
19  
20 pyrimidines within the PPT have been shown to increase inclusion of *SMN2* exon 7 (Lorson  
21  
22 and Androphy, 2000).

23  
24 In the present study we show that the hypothesized ESS element functions as an ESS in a  
25  
26 heterologous context. It binds hnRNP A1 *in vitro* and mutations that disrupt the ESS element  
27  
28 partially restores splicing of *SMN2* minigenes in transfection assays. We show by siRNA  
29  
30 mediated knock-down that hnRNP A1 can regulate splicing of *SMN2* transcripts and that this  
31  
32 effect is decreased in an *SMN2* minigene with a disrupted ESS element. Thus, the present  
33  
34 study identifies a new important splicing regulatory element spanning the 3'ss of *SMN* exon 7  
35  
36 and shows that it is part of the fine balance between positive and negative regulatory  
37  
38 elements, which are decisive for *SMN* exon 7 splicing. Additionally, this is an example of  
39  
40 how hnRNP A1 may function as a general inhibitor of splicing by binding to the 3'ss of  
41  
42 canonical exons. By binding directly to a 3'ss, hnRNP A1 may inhibit exon inclusion either  
43  
44 by directly blocking the 3'ss or by propagating along the exon or along the PPT to prevent  
45  
46 spliceosomal recognition of and recruitment to the exon.

Field Code Changed

Field Code Changed

Deleted: has

Field Code Changed

Deleted: is

## MATERIALS AND METHODS

### Minigene constructs

*SMN minigene.* Human genomic DNA was used as template in a PCR reaction with primers located in *SMN1/2* intron 6 (*SMN1VS6s*: GGCCTTTAAGCTCAGCGAGCCACTGCAAGAAAACCTTAACTGCAGC) and exon 8 (*SMNex8as*: GGCCTTTAAGCTCAGCAAGTACTTACTTGATATTTTGAAGAAATGAGGCCAGTTA) with HiFi Phusion DNA polymerase (Finnzymes, Espoo, Finland).

The PCR product contains 230 bp of the 3' part of intron 6, exon 7, intron 7 and the first 111 bp of exon 8 from *SMN1* and *SMN2*. This resulted in an 881 bp fragment, containing 839 bp of the gene, additional restriction sites and an artificial exon 8 5'ss site, which was cloned into the pCR-Blunt II-Topo vector (Invitrogen Co., Carlsbad, CA). After sequence analysis the generated plasmids were digested with XhoI and BamHI and ligated into an *MCAD* minigene (Nielsen, et al., 2007) with an optimized exon 5 3'ss and a BamHI site in intron 5 (for schematic, see Fig 5A). All plasmids were verified by DNA sequencing on the ABI Prism 3100-Avant (Applied Biosystems, Foster City, CA). Variant *SMN1/2* minigenes were generated by site-directed mutagenesis using the megaprimer method previously described (Sarkar and Sommer, 1990) with the *SMN2<sup>wt</sup>* minigene as template.

Field Code Changed

*pSXN13 minigene.* The *pSXN13* minigene consists of an artificial small (<40bp) exon positioned between  $\beta$ -actin exon 1 and exon 2 as well as a downstream exon 18 from *cTNT*. On both sides of the artificial exon is  $\beta$ -actin intron 1 in order to preserve the natural context of the two neighboring exons (Coulter, et al., 1997; Zatkova, et al., 2004). To generate *pSXN13* constructs we used sense and antisense oligonucleotides with desired sequences, which were mixed 1:1, phosphorylated and ligated to the BamHI and SalI sites in the

Field Code Changed



1  
2 artificial exon within pSXN13 (Zatkova, et al., 2004). The sequences of the generated  
3  
4 plasmids were verified by DNA sequencing.

#### 6 **Transient transfection of Chang cells and splicing analysis**

7  
8 Chang cells were seeded in 6 well plates and left to grow until reaching app. 25% confluence.  
9  
10 Cells were then transfected with 0.4 µg expression plasmid, either *SMN* constructs or  
11  
12 pSXN13 constructs, using FuGENE6 transfection reagent (Roche, Mannheim, Germany). All  
13  
14 transfections were performed at least 3 times. 48 hours after transfection cells were washed in  
15  
16 1xPBS-EDTA, cells lysed by adding 900 µL TRIzol reagent (Invitrogen) and incubating on  
17  
18 ice for 10 min and RNA extracted according to manufactory instructions (Invitrogen).  
19  
20 Purified total RNA was used as template in first strand cDNA synthesis using the Advantage  
21  
22 MMLV RT-PCR kit (BD Biosciences Clontech, Franklin Lakes, NJ) with an oligo (dT)<sub>18</sub>  
23  
24 primer. App. 1/10 of the cDNA synthesis product corresponding to 100 ng RNA was used as  
25  
26 template in each PCR reaction using Tempase DNA polymerase (Ampliqon Aps, Skovlunde,  
27  
28 Denmark). For the *SMN* constructs we used a primer spanning the exon 4 – exon 5 junction  
29  
30 of the *MCAD* gene (*MCex4ex5*: CATTCCAGAGAACTGTGGAGGT) and a primer located  
31  
32 in *SMN1/2* exon 8 (*SMNtestex8as*: GTGGTGTCATTTAGTGCTGCTC). For the pSXN13  
33  
34 constructs we used a primer located in β-actin exon 1 (*pSXN12s2*:  
35  
36 AAGGTGAACGTGGATGAAGTTGGTGGTG) and a primer spanning the β-actin exon 2 –  
37  
38 cTNT exon 18 splice junction (*pSXN13as*: CCCACGTGCAGCCTTTGACCTAGTA). PCR  
39  
40 products were separated and visualized on agarose gels containing ethidium bromide (EtBr)  
41  
42 on an Epi II Darkroom UVP Transilluminator.

#### 44 **siRNA reverse siRNA transfection and transient transfection of HEK-293 cells**

45  
46 App. 200,000 HEK-293 cells were seeded in each well in a 24 well plate (TPP AG) and 100  
47  
48 pmol of the relevant siRNA (hnRNP A1: CAGCUGAGGAAGCUCUUCU; hnRNP H: A  
49  
50 pool of 4 different siRNA oligonucleotides was used: GAUCCACCACGAAAGCUUAU,  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2 CAGAGUAACUGGUGAAGCAUU, GUUCGCAACUCAUGAAGAUUU and  
3  
4 UAACAUUGCCGGUGGACUUUU; Scrambled control: UAGCGACUAAACACAUCAA)  
5  
6 was reverse transfected into the HEK-293 using Dharmafect1 siRNA transfection reagent  
7  
8 (Dharmacon, Chicago, IL). After 24 hours the cells were transfected in triplicates with 0.09  
9  
10 µg *SMN* plasmid using FuGENE6 transfection reagent (Roche). 48 hours after transfection  
11  
12 cells were harvested. Of the triplicate transfections, 2 cell pellets were used for total RNA  
13  
14 isolation by TRIzol (Invitrogen) and cDNA synthesis using Advantage MMLV RT-PCR kit  
15  
16 (BD Biosciences Clontech). Splicing analysis was carried out by PCR amplification and  
17  
18 agarose gel electrophoresis with EtBr staining. Splicing analysis of endogenous *SMN1* and  
19  
20 *SMN2* transcripts were performed by PCR with primers located in exon 6 and exon 8  
21  
22 (*SMN.endo.ex6s*: ATATGTCCAGATTCTCTTGATG and *SMN.endo.ex8as*:  
23  
24 ACCATTTGAAACATTTTAAGAC) followed by restriction digest of the PCR product with  
25  
26 DdeI. The third cell pellet was used for protein analysis by adding 75 µL lysis buffer,  
27  
28 containing 1% Triton-X100 and 3 µL PMSF protease inhibitor, and by incubating at 37°C for  
29  
30 60 min. The lysate was spun and the supernatant extracted. The protein concentration in the  
31  
32 supernatant was estimated using Bradford reagent (Bio-Rad, Hercules, CA, USA) and  
33  
34 knockdown of target proteins was evaluated by western blotting, using affinity purified  
35  
36 hnRNP A1 goat polyclonal antibody (sc-10029), an affinity purified hnRNP H goat  
37  
38 polyclonal antibody (sc-10042) or an actin antibody (sc-1616). All antibodies were from  
39  
40 Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA),

Deleted:

#### 41 42 **Protein pull down by biotin coupled RNA oligonucleotides**

43  
44 We used biotinylated RNA oligonucleotides spanning the area containing the ESS motif in  
45  
46 *MCAD* and *SMN2*: *MCAD*-351A (wt): 5'-AUGGAUGUACAGGGGUUCAG-biotin, *MCAD*-  
47  
48 351C: 5'-AUGGAUGUACCGGGGUUCAG-biotin, *SMN2* 3'ss wt: 5'-  
49  
50 AUUUUCCUUACAGGGUUUUA-biotin, *SMN2* 3'ss -2C: 5'-

8

1  
2 AUUUUCCUUACCGGGUUUA-biotin and *SMN2* 3'ss 2A3C: 5'-  
3  
4 AUUUUCCUUACAGGACUUUA. Pull-down assay and immunodetection of purified  
5  
6 proteins were carried out as previously described (Nielsen, et al., 2007). For  
7  
8 immunodetection we used antibodies from Santa Cruz Biotechnologies against hnRNP A1  
9  
10 (sc-10029) or hnRNP H (sc-10042).

Deleted:

Field Code Changed

### 11 Tandem mass spectrometry (MS/MS) protein analysis

12 Protein eluate from RNA pull down (ca 15  $\mu$ L) was loaded on a 12% SDS-PAGE gel and run  
13  
14 for 60 min at 220V, 4°C. The gel was stained with Coomassie Blue and individual sample  
15  
16 lanes divided into 3 to 5 samples. After destaining the proteins were in-gel trypsin digested,  
17  
18 peptides purified on C18 PepClean spin columns (Pierce Biotechnology Inc., Rockford, IL,  
19  
20 USA) and vacuum-dried. The peptides were then subjected to reverse phase nano liquid  
21  
22 chromatography (nLC) and gradually eluted to the MS, as previously described (Hansen, et  
23  
24 al., 2008). Individual proteins were identified ( $p < 0.001$ ) by Mascot (Matrix Science Inc.,  
25  
26 Boston, MA, USA) and the Mascot score was used as a rough measure of protein abundance  
27  
28 in the sample (Kikuchi, et al., 2004). For a more detailed quantification of protein abundance,  
29  
30 the corresponding peptides were quantified from the nLC chromatogram (Hansen, et al.,  
31  
32 2008). 2 or more peptides per protein were used (in the case of hnRNP A1 4 peptides were  
33  
34 used) and each peptide was identified with better than  $p < 0.01$ . The amount of individual  
35  
36 peptides across the samples were normalized relative to the *SMN2* 3'ss wt sample and taken  
37  
38 as a relative measure of protein abundance. In this way, several independent measurements of  
39  
40 individual protein abundances were obtained through the peptides produced by trypsin  
41  
42 digestion. The whole procedure was repeated twice and data pooled accordingly between  
43  
44 experiments.

Field Code Changed

Field Code Changed

Field Code Changed

### 47 Generation of U2AF<sup>35</sup> Position Weight Matrix

Sequences reported to bind the U2AF complex (Wu, et al., 1999) were aligned and the pictogram program (<http://genes.mit.edu/pictogram.html>) used to calculate information content bit scores for each position as well as base frequencies and a graphical depiction of the motif based on the frequencies. Using the bit scores, positions -3 to +2 relative to the splice site were selected as the U2AF<sup>35</sup> defining motif. We assumed that the high bit scores observed at position -12 to -5 were due to U2AF<sup>65</sup> binding since Wu and co-workers purified the oligonucleotides using the entire U2AF complex. Using the default background frequencies (0.25 for each base), a position weight matrix (PWM) was constructed (Fig 3) using the method previously described (Cartegni, et al., 2003).

Field Code Changed

Field Code Changed

### ***In silico genomic analysis***

For initial analyses of splice sites we used SpliceRack (<http://katahdin.cshl.edu/SpliceRack/index.cgi?database=spliceNew>) (Sheth, et al., 2006).

For further analysis we used publicly available hg19 data tracks from the UCSC browser

(<http://genome.usc.edu>) to obtain the genomic location of constitutive and cassette exons

(Karolchik, et al., 2004; Rhead, et al., 2010). Genomic features were compared using

BEDTools, which was also used to extract sequences (Quinlan and Hall, 2010). We wrote custom Perl scripts to analyze and quantitate motif frequencies.

Field Code Changed

Field Code Changed

### ***Mutation nomenclature***

When nucleotide changes are preceded by a c. (for instance c.351A>C) they refer to the cDNA sequence numbering where +1 corresponds to the A of the ATG translational initiation codon in the reference sequence, according to journal guidelines ([www.hgvs.org/mutnomen](http://www.hgvs.org/mutnomen)).

The initiation codon is codon 1. In all other cases, nucleotide numbering reflects *SMNI/2* exon 7 numbering, where +1 corresponds to the first base of the exon and -1 to the last base of the preceding intron.

Formatted: Font: Italic

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

RefSeq reference sequences are *SMN1* (NG\_008691.1, NM\_000344.3), *SMN2* (NG\_008728.1, NM\_017411.3), *ACADM* (NG\_007045.1, NM\_000016.4) and *HADH2* (NG\_008153.1, NM\_004493.2).

- Formatted: Font: Italic
- Formatted: Font: Italic
- Formatted: Font: Italic
- Formatted: English (U.S.)
- Formatted: Font: Italic
- Formatted: English (U.S.)

For Peer Review

## RESULTS

### The ESS motif spanning the 3'ss of *SMN1* and *SMN2* is functional in a heterologous context and can be abrogated by mutations that decrease the hnRNP A1 binding score.

We hypothesized that a CAGGGT motif that spans the 3'ss of *SMN1/2* exon 7 comprises a hnRNP A1 binding site, like the CAGGGG motif in *MCAD* (Nielsen, et al., 2007, see Fig 1), since both motifs score high in an hnRNP A1 binding matrix and hnRNP A1 has been demonstrated to affect inclusion of *SMN1* and *SMN2* exon 7 (Hua, et al., 2010; Hua, et al., 2008; Kashima and Manley, 2003; Kashima, et al., 2007a; Kashima, et al., 2007b).

The putative ESS motif in *SMN1* and *SMN2* spans the 3'ss, thus only a limited number of positions in the motif can be mutated in the native context without risking that they also disrupt the 3'ss itself. Therefore we first used the pSXN13 splicing reporter minigene, which allows testing of the antagonistic interplay between an ESS and an ESE element (Nielsen, et al., 2007; Zatkova, et al., 2004) to test mutations aimed at disrupting the potential splicing silencer. In the pSXN13-SMN1wt construct, we observed predominant skipping, while the pSXN-SMN2wt construct showed complete skipping of the exon (Figure 2). The inhibitory effect of the ESS motif is abrogated by introducing a -2A>C change, that disrupts the hnRNP A1 consensus sequence. This is analogous to the c.351A>C change that abolishes hnRNP A1 mediated splicing inhibition in *MCAD* (Nielsen, et al., 2007). We generated a position weight matrix (PWM) for U2AF<sup>35</sup> based on earlier binding studies (Wu, et al., 1999) and used this to predict the effect on U2AF<sup>35</sup> binding of various mutations by assigning scores to different motifs. The U2AF<sup>35</sup> PWM (Fig 3) and hnRNP A1 PWM (Cartegni, et al., 2006) were used to select the mutations of +2 and +3 in exon 7 to either AC or CC, which decreased the hnRNP A1 score while keeping the U2AF<sup>35</sup> score essentially unchanged. We tested these mutations in the pSXN13 minigene and both inhibited the function of the ESS with the 2A3C mutation being the most efficient of the two (Fig 2B).

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

1  
2 **The 2A3C mutation decreases binding of hnRNP H and hnRNP A1 to *SMN2* exon 7**  
3  
4 **3'ss.**

5  
6 We used RNA affinity chromatography to investigate changes in hnRNP A1 binding upon  
7  
8 mutation of the ESS motif and to identify additional proteins, which were able to bind to the  
9  
10 3'ss of *SMN2* exon 7. Because the 2A3C mutation decreased the hnRNP A1 score the most  
11  
12 and did not improve the U2AF<sup>35</sup> score relative to the wt 3'ss of *SMN* exon 7, we used this  
13  
14 variant for RNA-Protein affinity purification experiments along with the wild type sequence  
15  
16 and the -2C mutant sequence as a control. Biotin-coupled RNA oligonucleotide sequences are  
17  
18 shown in Fig 4A.

19  
20 MS/MS identification of peptides revealed a large reduction in hnRNP A1 derived peptides in  
21  
22 the -2C sample relative to the wt oligonucleotide (Fig 4B). Western blot analysis also  
23  
24 demonstrated that the -2C mutation drastically decreased hnRNP A1 binding, indicating that  
25  
26 there is a specificity of hnRNP A1 for the CAGGGT motif (Fig 4C). This was corroborated  
27  
28 by the results with the *MCAD* ESS oligonucleotides (Fig 4D), which also showed an equally  
29  
30 dramatic effect of the -2C mutation on hnRNP A1 binding.

31  
32 Although the reduction in hnRNP A1 binding was not as pronounced by the 2A3C mutation,  
33  
34 there was still a significant decrease in the amount of hnRNP A1 bound to the  
35  
36 oligonucleotide both in the MS/MS analysis and in the western blot analysis compared to wt.  
37  
38 Because hnRNP A1 seems to bind preferentially to YAG motifs (Burd and Dreyfuss, 1994),  
39  
40 it is possible that because the seemingly indispensable YAG motif is still present in the 2A3C  
41  
42 mutant, it is still able to bind hnRNP A1 with some affinity. Because both the *MCAD* and  
43  
44 *SMN* ESS are G rich, we suspected that they may potentially bind hnRNP H (Caputi and  
45  
46 Zahler, 2001; Schaub, et al., 2007). We observed strong binding of hnRNP H to the wild type  
47  
48 motif and hnRNP H binding to the variants were decreased and showed an inverse binding-  
49  
50 pattern compared to that observed for hnRNP A1 (Fig 4B and 3C). This is consistent with the  
51  
52

Field Code Changed

Field Code Changed

1  
2 fact that hnRNP H preferentially binds to GGG triplet motifs (Caputi and Zahler, 2001;  
3  
4 Dobrowolski, et al., 2010; Schaub, et al., 2007) and the -2C mutation does not disrupt the  
5  
6 central GGG motif, whereas the 2A3C mutant disrupts the central GGG motif and thus  
7  
8 dramatically decreased hnRNP H binding. MS/MS analysis revealed that not only hnRNP H,  
9  
10 but also other members of the hnRNP H family of hnRNP proteins exhibited reduced binding  
11  
12 to the -2C and 2A3C mutant sequences (data not shown).  
13

#### 14 **Mutations disrupting the ESS increase inclusion of *SMN2* exon 7.**

15  
16 In order to investigate an effect of the ESS disrupting mutations in a native context, we  
17  
18 constructed *SMN* minigenes containing a partial intron 6, a complete exon 7 and intron 7 and  
19  
20 a partial exon 8. In this context the 2A3C mutation resulted in an increase of exon 7 inclusion  
21  
22 from both the *SMN1* and *SMN2* minigenes (Fig 5B) indicating that there is an ESS at the  
23  
24 intron-6/exon-7 junction, and that the ESS function can be disrupted by the 2A3C mutation.  
25  
26 Notably the 2C3C constructs behaved almost identical to the 2A3C constructs (Fig 5B).  
27

28 In order to examine the effect of having an *MCAD* like ESS in *SMN* exon 7, we designed  
29  
30 *SMN1* and *SMN2* constructs with a +3G mutation which converts the ESS to one identical to  
31  
32 the *MCAD* ESS. The CAGGGT *SMN* ESS and the CAGGGG *MCAD* ESS behaved  
33  
34 identically in the *SMN* minigene context (Fig 5B). Additionally, we designed constructs  
35  
36 where the *MCAD* ESS motif was instead generated by insertion of an additional G at position  
37  
38 +1, and these constructs exhibited a similar splicing pattern as the constructs with the *SMN*  
39  
40 CAGGGT ESS or the *MCAD* ESS CAGGGG generated by converting the +3T to a G (Fig  
41  
42 5B). This corroborates the hypothesis of a shared splicing mechanism between *MCAD* exon 5  
43  
44 and *SMN1* and *SMN2* exon 7.  
45

46 Since both the 2A3C and the 2C3C mutations would disrupt formation of the potential  
47  
48 splicing inhibitory terminal stem loop (Singh, et al., 2004) it was possible that the observed  
49  
50 increase in inclusion seen in the 2A3C and 2C3C constructs was due to disruption of this  
51  
52

Field Code Changed

Field Code Changed



1  
2 terminal stem loop, thereby improving access to the 3'ss by the spliceosome. This is,  
3  
4 however, contradicted by the fact that the MCAD CAGGGG ESS motif disrupts formation of  
5  
6 the stem loop, but still causes a degree of exon skipping comparable to the normal CAGGGT  
7  
8 motif. To further test the potential role of an inhibitory stem loop, minigenes containing a  
9  
10 2A3G mutation were constructed. This mutation would also remove the terminal stem loop  
11  
12 by eliminating base-pairing in the stem. If a terminal stem loop is responsible for the skipping  
13  
14 of *SMN2* exon 7 this mutation would increase inclusion to similar levels as that of the 2A3C  
15  
16 mutation. However, exon 7 inclusion was not improved in the 2A3G constructs (Fig 5B)  
17  
18 further supporting that disruption of a potential terminal stem loop structure is not the  
19  
20 underlying mechanism behind the increased inclusion demonstrated by the 2A3C and 2C3C  
21  
22 constructs. The 2A3G represents a curious example, however. Like the 2A3C construct it has  
23  
24 a moderately high U2AF<sup>35</sup> score and a very low hnRNP A1 score, but unlike the 2A3C it fails  
25  
26 to promote inclusion of exon 7. Using Human Splicing Finder (<http://www.umd.be/HSF/>)  
27  
28 (Desmet, et al., 2009), we examined the 2A3G mutation and found that it gave rise to several  
29  
30 new potential ESS elements, which may explain why it failed to restore splicing of the *SMN2*  
31  
32 minigene.

Deleted: This might reflect the limitations of the score matrices.

Formatted: English (U.S.)

Formatted: Font: Italic

### hnRNP A1 knockdown decreases the effect of an ESS disrupting mutation

35  
36 The identification of hnRNP A1 and hnRNP H as potential mediators of the inhibitory  
37  
38 activity of the CAGGGT motif in *SMN1* and *SMN2* indicated that knocking down hnRNP A1  
39  
40 or hnRNP H by siRNA would result in an increase of inclusion of exon 7. We expected this  
41  
42 effect to be less pronounced in the 2A3C constructs, since a reduction in binding affinity is  
43  
44 analogous to a reduction in protein level. To test this hypothesis, we conducted siRNA  
45  
46 experiments in HEK-293 cells in which we knocked down cellular levels of either hnRNP A1  
47  
48 or hnRNP H. We chose to use HEK-293 cells for these experiments as we have obtained  
49  
50 more efficient siRNA mediated knock-down in HEK-293 cells as opposed to Chang cells.

Deleted: wn

1  
2 Surprisingly, knockdown of hnRNP H did not increase inclusion of *SMN2* exon 7, but instead  
3  
4 seemed to result in an increase in skipping (Fig 6A). The effect of the 2A3C mutation was  
5  
6 still evident (Fig 6A).  
7

8 Knockdown of hnRNP A1 resulted in an increase in inclusion of *SMN2* wt exon 7 transcripts  
9  
10 (Fig 6A) and also an increase in inclusion from the 2A3C mutant construct compared to the  
11  
12 scrambled siRNA control, despite the fact that the hnRNP A1 PWM indicates a reduction in  
13  
14 binding of hnRNP A1 to the 2A3C variant motif. It is likely that the positive effect of hnRNP  
15  
16 A1 knockdown is due to the fact that the 2A3C mutant motif still retains some capability for  
17  
18 hnRNP A1 binding (as observed in the pull down experiments) and possibly also that other  
19  
20 hnRNP A1 binding sites (Hua, et al., 2010; Hua, et al., 2007; Hua, et al., 2008; Kashima, et  
21  
22 al., 2007b) still play a role in splicing inhibition. Comparing the splicing pattern of the wt  
23  
24 *SMN2* construct to the 2A3C construct when both are expressed in hnRNP A1 knocked down  
25  
26 cells reveals that the 2A3C mutation still increases inclusion of exon 7. This might be caused  
27  
28 by additional hnRNP proteins which are still able to bind to the 3'ss of *SMN2*, for instance  
29  
30 members of the hnRNP A/B family.

31 All *SMN1* constructs were unaffected by the siRNA mediated knockdown of interacting  
32  
33 proteins.  
34

### 35 **hnRNP A1 but not hnRNP H can regulate splicing of endogenous *SMN2* exon 7.**

36  
37 In order to investigate the effects of hnRNP A1 and hnRNP H knockdown on endogenous  
38  
39 *SMN1* and *SMN2* transcripts, we performed a PCR with primers located in exon 6 and exon 8  
40  
41 of *SMN2* and *SMN1*. To discriminate between *SMN1* and *SMN2* derived transcripts, we  
42  
43 digested the PCR products with DdeI, which cleaves *SMN2* exon 8, but not *SMN1* exon 8,  
44  
45 allowing for discrimination between *SMN1* and *SMN2* according to size. This splicing  
46  
47 analysis revealed that hnRNP A1, but not hnRNP H knock-down induces a shift in the  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Field Code Changed

1  
2 inclusion of endogenous *SMN2* exon 7 (Fig 6B), indicating that although hnRNP H might  
3  
4 bind the ESS *in vitro*, it is not functionally linked *in vivo*.

5  
6 **The 3'ss spanning hnRNP A1 motif is prevalent in human exons and may reduce splice**  
7  
8 **site strength**

9  
10 Because the CAGG portion of the CAGGGT/G ESS motif is present in the majority of  
11  
12 human 3' splice sites we decided to investigate if the subset of splice sites harboring this  
13  
14 CAGGGT/G ESS motif represent weak splice sites because they may be suppressed by  
15  
16 hnRNP A1 binding.

Deleted: ,

Deleted: which

Deleted: thus

17  
18 First we used the SpliceRack web-tool (Sheth, et al., 2006) to investigate the relative  
19  
20 abundance of GT-AG U2 splice sites similar to the 3' splice site of *SMN* exon 7 with the 6 nt  
21  
22 long CAGGGT motif, with the *MCAD* CAGGGG ESS motif, the mutant CAGGAC (2A3C)  
23  
24 motif and the mutant CAGGCC (2C3C) motif. Interestingly, the relative distribution of the 6  
25  
26 nt long motifs did not show clear differences, although CAGGGT and CAGGGG motifs were  
27  
28 less abundant than the other motifs. However, when we looked at the 7 nt long motifs it was  
29  
30 clear that the sequence corresponding to the *SMN* exon 7 is dramatically underrepresented.  
31  
32 The native CAGGGTT sequence was observed in 597, the *MCAD* ESS CAGGGGT in 411,  
33  
34 whereas the positive CAGGCCT and CAGGACT were observed in 1032 and 857, splice sites  
35  
36 respectively.

37  
38 We proceeded to extract the sequence of 29150 cassette exons and 175814 constitutive exons  
39  
40 plus 20 bp of intron sequence, as determined by the Alt Events track in the UCSC genome  
41  
42 browser from the hg19 scaffold (Karolchik, et al., 2004; Rhead, et al., 2010). We defined  
43  
44 constitutive exons as exons in the UCSC knownGene track, which did not overlap with any  
45  
46 of the exons in the Alt Events track. When we searched for the different 3' splice site  
47  
48 sequences in these, we found the *SMN* CAGGGTT motif in 92 cassette exons and in 559  
49  
50 constitutive exons whereas the mutant CAGGACT motif was found in 123 and 769 exons  
51  
52

Field Code Changed

Deleted: looked

1  
2 respectively, an increase compared to the CAGGGTT motif of 33.7% and 37.6%  
3  
4 respectively. The *MCAD* ESS motif CAGGGGT was found in 74 and 391 exons, a decrease  
5  
6 of 19.5% and 30.1%, while the mutant CAGGCCT motif was found in 124 (+34.8%) and 961  
7  
8 (+71.9%) exons respectively, demonstrating that the *MCAD* motif CAGGGGT and the *SMN*  
9  
10 motif CAGGGTT are less frequent in both cassette and constitutive exons than the positive  
11  
12 motifs. Notably, the *MCAD* motif was less frequent in 3'ss of constitutive exons than  
13  
14 alternative exons while the 2C3C was more frequent in 3'ss of constitutive exons than  
15  
16 alternative exons, reflecting the negative and positive roles of the motifs in determining splice  
17  
18 site strength. This indicates that also on a genomic scale these splice sites constitute less  
19  
20 preferable splice sites than the CAGGACT and CAGGCCT type splice site and point towards  
21  
22 a possible general regulatory mechanism.  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## DISCUSSION

All patients with SMA possess the partly functional *SMN2* gene and could therefore benefit from therapies, which correct or improve splicing of *SMN2* exon 7. Therefore it is very important that the underlying molecular defect mechanisms and the involved splicing regulatory elements, that suppress correct splicing of *SMN2* exon 7, are identified and characterized. Moreover, because the affected splicing regulatory elements and the molecular pathological mechanisms involved are general and there is increasing evidence suggesting that the regulatory elements are organized in common architectures between human disease genes this knowledge is also important for a better understanding of other inherited diseases.

In this study we report the identification of an ESS element positioned across the 3'ss, which is important for regulation of splicing of *SMN1* and *SMN2* exon 7. We have previously identified a nearly identical ESS, which binds the splicing inhibitory protein hnRNP A1 and together with a flanking ESE determines splicing of *MCAD* exon 5 (Nielsen, et al., 2007).

Furthermore, a synonymous c.574C>A mutation in *HADH2* exon 5 has been reported to cause X-linked mental retardation through exon skipping (Lenski, et al., 2007). The c.574C>A mutation creates a CAGGGT motif, identical to the ESS in *SMN1/2*. This indicates that the CAGGGT motif can also function as an ESS in a different gene and that creation of this ESS motif is sufficient to disrupt splicing. Based on nearly identical sequences between *MCAD* and *SMN1/2* in a longer region harboring both the *MCAD* ESS and ESE and the region harboring both the critical c.840C>T change that distinguishes *SMN2* from the functional *SMN1* and the flanking splicing inhibitory region (Hua, et al., 2007; Singh, et al., 2004), we speculated whether *SMN* exon 7 splicing could be regulated in a way similar to *MCAD* exon 5, involving the potential ESS that spans the 3'ss of both *SMN* genes.

When we compared the effect of the *SMN* ESS with that of the *MCAD* ESS on the *SMN1* ESE or the corresponding *SMN2* sequence in the heterologous context of the pSXN splicing

Deleted: s

Deleted: s

Field Code Changed

Field Code Changed

Field Code Changed

reporter minigene (Zatkova, et al., 2004) we observed that the splicing inhibitory effect was indistinguishable and that the *SMN* ESS could be inactivated by mutations, which disrupts hnRNP A1 binding in pull down experiments. Moreover, functional testing of mutations, which were designed on the basis of the constructed binding matrices for U2AF35 and hnRNP A1, and testing of mutations substituting the *SMN* ESS with the *MCAD* ESS in *SMN1* and *SMN2* minigenes supports the notion that a hnRNP A1 binding ESS is located across the 3' splice site of *SMN1* and *SMN2*. This, together with experiments, where the effect of siRNA mediated knock down of hnRNP A1 on the splicing of *SMN1* and *SMN2* minigenes with either the native ESS or a mutated ESS with a compromised hnRNP A1 binding motif, also supported the notion that *SMN1/2* exon 7 splicing is inhibited by the presence of an hnRNP A1 binding ESS spanning the 3'ss.

Field Code Changed

Formatted: Font: Italic

Several of the mutations tested in the *SMN1* and *SMN2* minigenes and/or the pSXN splicing reporter indicate that the splicing inhibitory activity located near the 3'ss of exon 7 is not caused by the potential to form an inhibitory stem loop structure as proposed by Singh and coworkers (Singh, et al., 2004). The -2C mutation is located outside the potential stem loop structure, but still improves splicing dramatically in the pSXN splicing reporter minigene, while the 2A3G disrupts the potential to form the stem loop structure, but does not alter the splicing pattern of the *SMN1* and *SMN2* minigenes. Moreover, both the +3T>G and insertion of G at +1 (which both create the CAGGGG ESS) disrupt the potential to form the stem loop structure, but still inhibit splicing at a level indistinguishable from the native *SMN* CAGGGT ESS.

Deleted: the potential

Field Code Changed

Deleted: s

Deleted: s

The 2A3G mutation failed to restore inclusion of exon 7 in the *SMN2* minigene despite displaying a similar decrease in hnRNP A1 score as the 2A3C and 2C3C mutations. Analysis by the Human Splicing Finder tool (Desmet, et al., 2009) indicates that this could be due to the fact that the 2A3G mutation creates 2 new potential ESS motifs (cagGAGTT and

Formatted: Font: Italic

Formatted: English (U.S.)

Formatted: Underline, English

Formatted: English (U.S.)

AGTTTA) although it removes several other ESS elements. Another alternative, but not  
mutually exclusive. explanation could be that in contrast to the 2A3G mutation, the 2C3C and  
2A3C mutations are predicted by the Human Splicing Finder tool to result in formation of  
new ESE motifs, which span or are very closely located to the 3'ss and that these ESEs  
facilitate increased inclusion. Several problems arise with this hypothesis, however. First of  
all, while not impossible, it is hard to imagine how an SR protein might efficiently recruit  
U2AF<sup>35</sup> to a 3'ss, which it is already blocking. Secondly, both the 2C3C and 2A3C mutation  
increase inclusion of exon 7, but the 2C3C mutation only forms one additional ESE motif, an  
SC35 motif, which spans the 3'ss (gGCCTTTA) and both the 2A3C and 2A3G mutations  
form a 9G8 motif spanning the 3'ss (cagGAC and cagGAG). Finally, *SMN2* already has a  
naturally occurring SF2/ASF ESE motif spanning the 3'ss (tacagGG), which is either not  
functional or not sufficient to compensate for the loss of the downstream ESE. This indicates  
that even though a potential ESE motif may occur across the 3'ss it is likely not functional or  
at least not sufficient alone for an increase in inclusion of exon 7 and that the best explanation  
of increased inclusion is the disruption of an ESS motif. The possibility of new or remaining  
functional ESS motifs in the case of the 2A3G mutant thus seems the most likely explanation  
for the splicing pattern of this construct.

Formatted: Underline

Formatted: English (U.S.)

Formatted: Superscript

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: Font: Italic

Deleted:

Our in vitro pull down experiments with RNA oligonucleotides with the wild type and mutant *SMN1/2* exon 7 3'ss sequence further supported that splicing inhibition is mediated by hnRNP A1 binding. The wild type 3'ss sequence bound hnRNP A1 strongly, whereas introduction of the -2C substitution abolished hnRNP A1 binding, consistent with results obtained with a RNA oligonucleotides with the *MCAD* ESS wild type and the -2C mutant sequence. The 2A3C mutation also decreased binding of hnRNP A1, but not as dramatic as the -2C mutation. This shows that the 2A3C mutant sequence retains some affinity for hnRNP A1 and it is consistent with our observation that splicing of the 2A3C mutant *SMN2*

1  
2 minigene is also improved by siRNA mediated knockdown of hnRNP A1, although to a  
3 much lesser extent than the *SMN2* minigene with the wild type high score hnRNP A1 binding  
4 motif. The fact that the wild type motif can be completely as well as partially abrogated by  
5 mutations indicates that hnRNP A1 binds to this region both strongly and specifically as  
6 opposed to binding indiscriminately along the oligonucleotide. We therefore propose that  
7 splicing of *SMN2* exon 7 is inhibited by the ability of the sequence across the 3' splice site to  
8 bind hnRNP A1.  
9

10 The *in vitro* binding studies additionally revealed binding of hnRNP H to the 3'ss region of  
11 *SMN2* exon 7, but to our surprise siRNA mediated knock-down decreased inclusion of *SMN2*  
12 exon 7 in a model minigene. The endogenous transcripts expressed in the same cells were,  
13 however, unaffected by the decreased levels of hnRNP H.  
14

15 These conflicting results might simply be due to minigene specific motifs, which enable  
16 hnRNP H to influence splicing of *SMN2* exon 7 in the minigene specific context and since  
17 these motifs are not present in the endogenous transcript, hnRNP H is unable to regulate  
18 splicing of endogenous *SMN2* exon 7.  
19

20 While we cannot rule out that hnRNP H might, in some circumstances, also regulate splicing  
21 of endogenous *SMN2* exon 7, further studies would be needed in order to clarify this.  
22

23 It has recently been suggested that the c.840C>T transition in *SMN2* exon 7 forms a new  
24 binding motif for the negative splicing regulator Sam68 and that *SMN2* exon 7 splicing  
25 inhibition by Sam68 is mediated by recruitment of hnRNP A1 (Pedrotti, et al., 2010). Such a  
26 model would fit well with our finding that hnRNP A1 can bind directly to the 3'ss, competing  
27 with U2AF<sup>35</sup> for binding. Binding of Sam68 to the motif immediately downstream would  
28 thus improve hnRNP A1 binding to the 3' splice site of *SMN2* thereby inhibiting exon 7  
29 inclusion. This is also compatible with a model where the SF2/ASF binding ESE in *SMN1*,  
30 stimulates recruitment of U2AF to the 3'splice site directly through SR domain interactions  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

Field Code Changed



1  
2 and at the same time also antagonizes hnRNP A1 binding to the 3' splice site in *SMN1*. The  
3  
4 c.8406C>T substitution in *SMN2* would thus have a dual effect by disrupting SF2/ASF  
5  
6 binding and at the same time create a Sam68 binding motif, which would both increase  
7  
8 binding of hnRNP A1 to the 3' splice site of *SMN2* exon 7 and decrease U2AF binding  
9  
10 resulting in splicing inhibition. It has been shown that other hnRNP A1 binding sites in intron  
11  
12 6 and intron 7 also play a fundamental role in mediating exon 7 skipping in *SMN2* (Hua, et  
13  
14 al., 2010; Hua, et al., 2008; Kashima, et al., 2007b). Binding to these sites could further  
15  
16 facilitate hnRNP A1 binding at the 3' splice site through cooperative binding and/or by  
17  
18 propagation of hnRNP A1 binding proteins along the RNA either in a 3'-5' direction or less  
19  
20 efficiently in a 5'-3' direction (Okunola and Krainer, 2009). The proximal hnRNP A1  
21  
22 binding site in intron 7 (Hua, et al., 2008) could for instance initiate propagation of hnRNP  
23  
24 A1 binding in a 3'-5' direction towards the 3'ss in exon 7. This process may be antagonized  
25  
26 by SF2/ASF binding to the ESE in *SMN1* and stimulated by Sam68 binding in *SMN2*. These  
27  
28 proposed mechanisms would imply that U2AF<sup>35</sup> and the U2AF complex would have a higher  
29  
30 affinity for the 3'ss of exon 7 of *SMN1* than that of *SMN2* and this has in fact recently been  
31  
32 reported (Martins de Araujo, et al., 2009).

33  
34 The observation that hnRNP A1 can inhibit splicing of *SMN2* exon 7 by direct binding to the  
35  
36 3'ss sequence is in agreement with earlier studies. The similarity of the motifs binding  
37  
38 hnRNP A1 and the 3'ss consensus sequence has previously been noted (Burd and Dreyfuss,  
39  
40 1994), but the possible role of hnRNP A1 acting as a splicing inhibitor by direct binding to  
41  
42 the 3'ss has not been investigated in detail. Binding of hnRNP A1 to 3'ss was initially  
43  
44 thought to stimulate recognition of the 3'splice site (Buvoli, et al., 1990; Swanson and  
45  
46 Dreyfuss, 1988a; Swanson and Dreyfuss, 1988b). Subsequent studies of hnRNP A1 binding  
47  
48 to the 3'end of an intron indicated, however, that the observed interaction was non-specific or  
49  
50 at least that the function of hnRNP A1 is not dependent on specific interactions with the 3'ss  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

(Abdul-Manan and Williams, 1996). The oligonucleotides used in these studies did not contain any exonic nucleotides, but only the polypyrimidine tract and the AG dinucleotide, and are therefore not in conflict with our observation, since we show that hnRNP A1 binding is dependent on the exonic nucleotides, as illustrated by the 2A3C mutation. It is thus likely that hnRNP A1 can function as a repressor of exon recognition by specific binding to the 3'ss, when the first few nucleotides in the exon completes a binding motif and that this is dependent on contextual conditions, such as the strength of the polypyrimidine tract, presence of nearby antagonizing ESEs and hnRNP A1 binding sites. Increased amounts of hnRNP A1 has previously been reported to affect some alternatively spliced exons dependent on the strength of the polypyrimidine tract (Mayeda, et al., 1993). In agreement with this is the fact that although the polypyrimidine tract appears very conserved, which could even point towards an AG-independent intron, and the strength of *SMN2* exon 7 3'ss as estimated by several different splice site prediction programs is quite high (NNsplice score is 0.99 (Reese, et al., 1997), MaxEntScan score is 10.92 (Yeo and Burge, 2004)), strengthening of the polypyrimidine tract alone induces complete inclusion of exon 7 in an *SMN2* minigene (Lorson and Androphy, 2000). This indicates that while the 3'ss of *SMN2* exon 7 looks like a strong splice site, it is functionally weak and the binding of U2AF<sup>65</sup> to the PPT is suboptimal and dependent on U2AF<sup>35</sup> binding to the AG dinucleotide. This model is supported by the results of a paper describing PUF60 as a regulator of *SMN2* exon 7 splicing (Hastings, et al., 2007). The authors examined *SMN2* exon 7 splicing when U2AF<sup>65</sup>, PUF60 or both U2AF<sup>65</sup> and PUF60 were knocked down by siRNA and reported that *SMN2* exon 7 inclusion increased upon knockdown of these factors. As the authors suggest, this could be explained by less competition from *SMN2* exon 8 while *SMN2* exon 7 3'ss strength is less affected by depletion of U2AF<sup>65</sup> as the binding of U2AF<sup>65</sup> to *SMN2* exon 7 is already compromised (Hastings, et al., 2007). Further, this very clearly demonstrates that splice site scores should

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

be used with caution in determining weak and strong splice sites, as other factors contribute to overall splice site strength. The presence of an hnRNP A1 binding ESS spanning the *SMN1/2* 3'ss would compromise recognition of the 3' splice site and represents a plausible explanation for this. We therefore speculated if the *SMN1/2* 3'ss represent a generally weak 3'ss sequence and decided to investigate globally for the presence of the *SMN1/2* 3'ss motif among 3'splice sites. Comparing only the occurrence of the 6 nt CAGGGT motif in 3'ss to the occurrence of the corresponding 2A3C and 2C3C mutant motifs only revealed minor differences in frequency, but when we included the flanking exonic nucleotide, differences in frequency became apparent, and when we compared the frequencies of these motifs in the 3'ss of constitutive and cassette exons, there were clear indications that 3'ss with either the *SMN1/2* CAGGGTT motif or the *MCAD* CAGGGGT motif were less frequent (i.e. weaker splice sites) than those with the CAGGACT and CAGGCCT motifs. The fact that these differences only became apparent when we included the nucleotide flanking the core 6 nt motif could either indicate that the hnRNPA1 binding motif in fact constitutes 7 nt with a requirement for a T at position 7, or that binding of hnRNP A1 to these 3' ss is strongly dependent on the flanking sequence or alternatively that some other unrelated feature determining 3'ss recognition is influenced by the presence of a T nucleotide at this position (+4 in the exon). It will require further experiments to elucidate the precise mechanism underlying this observation.

In conclusion, we show that an important ESS overlaps the 3'ss of SMN exon 7 and provide evidence to suggest a combination of both the ESE loss model and the ESS gain model for the splicing of exon 7 in *SMN1* and *SMN2*.

We suggest that the ESS motif which is present in the *MCAD* and the *SMN* genes, and which is created by a disease causing mutation in the *HADH2* gene, is an example of a general splicing regulatory motif that binds hnRNP A1 and can inhibit splicing. In particular, we

**Deleted:** strength of the

**Deleted:**

**Formatted:** Font: Italic

**Formatted:** Font: Italic

**Deleted:** that of

**Deleted:** did not

**Deleted:** significant

**Deleted:** one more

**Deleted:** comparing

**Deleted:** the

**Formatted:** Font: Italic

**Formatted:** Font: Italic

**Deleted:** are stronger splice sites than either the *SMN* CAGGGTT motif or the *MCAD* CAGGGGT motif

**Deleted:** clear

**Deleted:** following

**Deleted:** , not only illustrates the limitations of the binding score matrices, but it

**Deleted:** s

**Deleted:** that the flanking immediate context of hnRNP A1 motifs may be decisive for motif strength or

**Deleted:** Finally, w

**Deleted:** found

**Formatted:** Font: Italic

**Deleted:** may be

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

propose that hnRNP A1 may function as a general inhibitor of exon inclusion by direct binding to the 3'ss of exons, where this inhibitory ESS motif constitute part of the 3'ss. This also illustrates how the third and fourth nucleotide of an exon may contribute to modulate splice site strength. Importantly, this means that variations within this region, which have previously been assumed to be neutral, may cause partial or complete skipping of an exon, or confer disease susceptibility depending on cell type or environmental changes in splicing factor concentrations.

Deleted: and

### Acknowledgments

This work was supported by grants from The Danish Medical Research Council (FSS grant no. 271-07-342 to BSA), The Riisfort Foundation (BSA) and The Augustinus Foundation (BSA). Tandem Mass Spectrometry analysis was supported by a grant from the John and Birthe Meyer Foundation to NG. We are grateful to Jane Serup Pedersen and Lone Sundahl for expert technical assistance.

## REFERENCES

- 1  
2  
3  
4 Abdul-Manan N, Williams KR. 1996. hnRNP A1 binds promiscuously to  
5 oligoribonucleotides: utilization of random and homo-oligonucleotides to discriminate  
6 sequence from base-specific binding. Nucleic Acids Res 24(20):4063-70.  
7 Brow DA. 2002. Allosteric cascade of spliceosome activation. Annu Rev Genet 36:333-60. Formatted: Danish  
8 Burd CG, Dreyfuss G. 1994. RNA binding specificity of hnRNP A1: significance of hnRNP  
9 A1 high-affinity binding sites in pre-mRNA splicing. EMBO J 13(5):1197-204.  
10 Burnett BG, Munoz E, Tandon A, Kwon DY, Sumner CJ, Fischbeck KH. 2009. Regulation  
11 of SMN protein stability. Mol Cell Biol 29(5):1107-15.  
12 Butchbach ME, Edwards JD, Burghes AH. 2007. Abnormal motor phenotype in the  
13 SMNDelta7 mouse model of spinal muscular atrophy. Neurobiol Dis 27(2):207-19.  
14 Buvoli M, Cobianchi F, Biamonti G, Riva S. 1990. Recombinant hnRNP protein A1 and its  
15 N-terminal domain show preferential affinity for oligodeoxynucleotides homologous  
16 to intron/exon acceptor sites. Nucleic Acids Res 18(22):6595-600.  
17 Caputi M, Zahler AM. 2001. Determination of the RNA binding specificity of the  
18 heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. J Biol Chem  
19 276(47):43850-9.  
20 Cartegni L, Chew SL, Krainer AR. 2002. Listening to silence and understanding nonsense:  
21 exonic mutations that affect splicing. Nat Rev Genet 3(4):285-98. Formatted: Danish  
22 Cartegni L, Hastings ML, Calarco JA, de Stanchina E, Krainer AR. 2006. Determinants of  
23 exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. Am J Hum  
24 Genet 78(1):63-77.  
25 Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. 2003. ESEfinder: A web resource to  
26 identify exonic splicing enhancers. Nucleic Acids Res 31(13):3568-71.  
27 Cho S, Dreyfuss G. A degren created by SMN2 exon 7 skipping is a principal contributor to  
28 spinal muscular atrophy severity. Genes Dev 24(5):438-42.  
29 Coulter LR, Landree MA, Cooper TA. 1997. Identification of a new class of exonic splicing  
30 enhancers by in vivo selection. Mol Cell Biol 17(4):2143-50.  
31 Crawford TO, Pardo CA. 1996. The neurobiology of childhood spinal muscular atrophy.  
32 Neurobiol Dis 3(2):97-110.  
33 Desmet FO, Hamroun D, Lalonde M, Collod-Beroud G, Claustres M, Beroud C. 2009.  
34 Human Splicing Finder: an online bioinformatics tool to predict splicing signals.  
35 Nucleic Acids Res 37(9):e67.  
36 Dobrowolski SF, Andersen HS, Doktor TK, Andresen BS. 2010. The phenylalanine Formatted: Danish  
37 hydroxylase c.30C>G synonymous variation (p.G10G) creates a common exonic  
38 splicing silencer. Mol Genet Metab.  
39 Domsic JK, Wang Y, Mayeda A, Krainer AR, Stoltzfus CM. 2003. Human  
40 immunodeficiency virus type 1 hnRNP A/B-dependent exonic splicing silencer ESSV  
41 antagonizes binding of U2AF65 to viral polypyrimidine tracts. Mol Cell Biol  
42 23(23):8762-72.  
43 Feldkotter M, Schwarzer V, Wirth R, Wienker TF, Wirth B. 2002. Quantitative analyses of  
44 SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier  
45 testing and prediction of severity of spinal muscular atrophy. Am J Hum Genet  
46 70(2):358-68.  
47 Graveley BR, Hertel KJ, Maniatis T. 2001. The role of U2AF35 and U2AF65 in enhancer-  
48 dependent splicing. RNA 7(6):806-18.  
49 Hansen J, Corydon TJ, Palmfeldt J, Durr A, Fontaine B, Nielsen MN, Christensen JH,  
50 Gregersen N, Bross P. 2008. Decreased expression of the mitochondrial matrix  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- proteases Lon and ClpP in cells from a patient with hereditary spastic paraplegia (SPG13). *Neuroscience* 153(2):474-82.
- Hastings ML, Allemand E, Duelli DM, Myers MP, Krainer AR. 2007. Control of pre-mRNA splicing by the general splicing factors PUF60 and U2AF65. *PLoS ONE* 2(6):e538.
- Hua Y, Sahashi K, Hung G, Rigo F, Passini MA, Bennett CF, Krainer AR. 2010. Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev* 24(15):1634-44.
- Hua Y, Vickers TA, Baker BF, Bennett CF, Krainer AR. 2007. Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon. *PLoS Biol* 5(4):e73.
- Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. 2008. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *Am J Hum Genet* 82(4):834-48.
- Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ. 2004. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 32(Database issue):D493-6.
- Kashima T, Manley JL. 2003. A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nat Genet* 34(4):460-3.
- Kashima T, Rao N, David CJ, Manley JL. 2007a. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. *Hum Mol Genet* 16(24):3149-59.
- Kashima T, Rao N, Manley JL. 2007b. An intronic element contributes to splicing repression in spinal muscular atrophy. *Proc Natl Acad Sci U S A* 104(9):3426-31.
- Kent OA, Ritchie DB, Macmillan AM. 2005. Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway. *Mol Cell Biol* 25(1):233-40.
- Kikuchi M, Hatano N, Yokota S, Shimozawa N, Imanaka T, Taniguchi H. 2004. Proteomic analysis of rat liver peroxisome: presence of peroxisome-specific isozyme of Lon protease. *J Biol Chem* 279(1):421-8.
- Le TT, Pham LT, Butchbach ME, Zhang HL, Monani UR, Coover DD, Gavrulina TO, Xing L, Bassell GJ, Burghes AH. 2005. SMNDelta7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. *Hum Mol Genet* 14(6):845-57.
- Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M and others. 1995. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80(1):155-65.
- Lefebvre S, Burlet P, Liu Q, Bertrand S, Clermont O, Munnich A, Dreyfuss G, Melki J. 1997. Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet* 16(3):265-9.
- Lenski C, Kooy RF, Reyniers E, Loessner D, Wanders RJ, Winnepenninckx B, Hellebrand H, Engert S, Schwartz CE, Meindl A and others. 2007. The reduced expression of the HADH2 protein causes X-linked mental retardation, choreoathetosis, and abnormal behavior. *Am J Hum Genet* 80(2):372-7.
- Lorson CL, Androphy EJ. 2000. An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN. *Hum Mol Genet* 9(2):259-65.
- Martins de Araujo M, Bonnal S, Hastings ML, Krainer AR, Valcarcel J. 2009. Differential 3' splice site recognition of SMN1 and SMN2 transcripts by U2AF and U2 snRNP. *RNA* 15(4):515-23.
- Masuda A, Shen XM, Ito M, Matsuura T, Engel AG, Ohno K. 2008. hnRNP H enhances skipping of a nonfunctional exon P3A in CHRNA1 and a mutation disrupting its binding causes congenital myasthenic syndrome. *Hum Mol Genet* 17(24):4022-35.

Formatted: Danish

- 1  
2  
3 Mayeda A, Helfman DM, Krainer AR. 1993. Modulation of exon skipping and inclusion by  
4 heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF.  
5 Mol Cell Biol 13(5):2993-3001.
- 6 Merendino L, Guth S, Bilbao D, Martinez C, Valcarcel J. 1999. Inhibition of msl-2 splicing  
7 by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. Nature  
8 402(6763):838-41.
- 9 Monani UR, Lorson CL, Parsons DW, Prior TW, Androphy EJ, Burghes AH, McPherson JD.  
10 1999. A single nucleotide difference that alters splicing patterns distinguishes the  
11 SMA gene SMN1 from the copy gene SMN2. Hum Mol Genet 8(7):1177-83.
- 12 Nielsen KB, Sorensen S, Cartegni L, Corydon TJ, Doktor TK, Schroeder LD, Reinert LS,  
13 Elpeleg O, Krainer AR, Gregersen N and others. 2007. Seemingly neutral  
14 polymorphic variants may confer immunity to splicing-inactivating mutations: a  
15 synonymous SNP in exon 5 of MCAD protects from deleterious mutations in a  
16 flanking exonic splicing enhancer. Am J Hum Genet 80(3):416-32.
- 17 Okunola HL, Krainer AR. 2009. Cooperative-binding and splicing-repressive properties of  
18 hnRNP A1. Mol Cell Biol 29(20):5620-31.
- 19 Pedrotti S, Bielli P, Paronetto MP, Ciccocanti F, Fimia GM, Stamm S, Manley JL, Sette C.  
20 2010. The splicing regulator Sam68 binds to a novel exonic splicing silencer and  
21 functions in SMN2 alternative splicing in spinal muscular atrophy. EMBO J  
22 29(7):1235-47.
- 23 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic  
24 features. Bioinformatics 26(6):841-2.
- 25 Reese MG, Eeckman FH, Kulp D, Haussler D. 1997. Improved splice site detection in Genie.  
26 J Comput Biol 4(3):311-23.
- 27 Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M, Smith  
28 KE, Rosenbloom KR, Raney BJ and others. 2010. The UCSC Genome Browser  
29 database: update 2010. Nucleic Acids Res 38(Database issue):D613-9.
- 30 Sarkar G, Sommer SS. 1990. The "megaprimer" method of site-directed mutagenesis.  
31 Biotechniques 8(4):404-7.
- 32 Schaub MC, Lopez SR, Caputi M. 2007. Members of the heterogeneous nuclear  
33 ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by  
34 promoting formation of ATP-dependent spliceosomal complexes. J Biol Chem  
35 282(18):13617-26.
- 36 Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R. 2006.  
37 Comprehensive splice-site analysis using comparative genomics. Nucleic Acids Res  
38 34(14):3955-67.
- 39 Singh NN, Androphy EJ, Singh RN. 2004. An extended inhibitory context causes skipping of  
40 exon 7 of SMN2 in spinal muscular atrophy. Biochem Biophys Res Commun  
41 315(2):381-8.
- 42 Swanson MS, Dreyfuss G. 1988a. Classification and purification of proteins of heterogeneous  
43 nuclear ribonucleoprotein particles by RNA-binding specificities. Mol Cell Biol  
44 8(5):2237-41.
- 45 Swanson MS, Dreyfuss G. 1988b. RNA binding specificity of hnRNP proteins: a subset bind  
46 to the 3' end of introns. EMBO J 7(11):3519-29.
- 47 Velasco E, Valero C, Valero A, Moreno F, Hernandez-Chico C. 1996. Molecular analysis of  
48 the SMN and NAIP genes in Spanish spinal muscular atrophy (SMA) families and  
49 correlation between number of copies of cBCD541 and SMA phenotype. Hum Mol  
50 Genet 5(2):257-63.

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Warf MB, Diegel JV, von Hippel PH, Berglund JA. 2009. The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. Proc Natl Acad Sci U S A 106(23):9203-8.
- Wirth B, Herz M, Wetter A, Moskau S, Hahnen E, Rudnik-Schoneborn S, Wienker T, Zerres K. 1999. Quantitative analysis of survival motor neuron copies: identification of subtle SMN1 mutations in patients with spinal muscular atrophy, genotype-phenotype correlation, and implications for genetic counseling. Am J Hum Genet 64(5):1340-56.
- Wu JY, Maniatis T. 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell 75(6):1061-70.
- Wu S, Romfo CM, Nilsen TW, Green MR. 1999. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature 402(6763):832-5.
- Yeo G, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 11(2-3):377-94.
- Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. Nature 355(6361):609-14.
- Zatkova A, Messiaen L, Vandenbroucke I, Wieser R, Fonatsch C, Krainer AR, Wimmer K. 2004. Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. Hum Mutat 24(6):491-501.
- Zhang M, Zamore PD, Carmo-Fonseca M, Lamond AI, Green MR. 1992. Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. Proc Natl Acad Sci U S A 89(18):8769-73.

Formatted: Danish

**Deleted:** Abdul-Manan N, Williams KR. 1996. hnRNP A1 binds promiscuously to oligoribonucleotides: utilization of random and homo-oligoribonucleotides to discriminate sequence from base-specific binding. Nucleic Acids Res 24(20):4063-70.¶

Brow DA. 2002. Allosteric cascade of spliceosome activation. Annu Rev Genet 36:333-60.¶

Burd CG, Dreyfuss G. 1994. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. EMBO J 13(5):1197-204.¶

Burnett BG, Munoz E, Tandon A, Kwon DY, Sumner CJ, Fischbeck KH. 2009. Regulation of SMN protein stability. Mol Cell Biol 29(5):1107-15.¶

Butchbach ME, Edwards JD, Burghes AH. 2007. Abnormal motor phenotype in the SMNDelta7 mouse model of spinal muscular atrophy. Neurobiol Dis 27(2):207-19.¶

Buvoli M, Cobiainchi F, Biamonti G, Riva S. 1990. Recombinant hnRNP protein A1 and its N-terminal domain show preferential affinity for oligodeoxynucleotides homologous to intron/exon acceptor sites. Nucleic Acids Res 18(22):6595-600.¶

Caputi M, Zahler AM. 2001. Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. J Biol Chem 276(47):43850-9.¶

Cartegni L, Chew SL, Krainer AR. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3(4):285-98.¶

Cartegni L, Hastings ML, Calarco JA, de Stanchina E, Krainer AR. 2006. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. Am J Hum Genet 78(1):63-77.¶

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31(13):3568-71.¶

Cho S, Dreyfuss G. A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. Genes Dev 24(5):438-42.¶

Coulter LR, Landree MA, Cooper TA. 1997. Identification of a new class of exonic splicing enhancers by in vivo selection. Mol Cell Biol 17(4):2143-50.¶

Crawford TO, Pardo CA. 1996. The neurobiology of childhood spinal muscular atrophy. Neurobiol Dis 3(2):97-110.¶

Dobrowolski SF, Andersen HS, Doktor TK, Andresen BS. 2010. The phenylalanine hydroxylase c.30C>G synonymous variation (p.G10G) c... [1]

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)



**FIGURE LEGENDS**

**Figure 1.** Similarities between *SMN1* and *SMN2* exon 7 and *MCAD* exon 5. The region comprising the ESS and ESE motifs in *SMN1* and *SMN2* exon 7 is compared to the corresponding region in *MCAD* exon 5, showing the extensive sequence similarity between the two exons. Both contain an SF2/ASF binding ESE as well as an hTra2- $\beta$ 1 binding ESE further downstream. In the context of *MCAD* exon 5, we have previously described the antagonistic functions of the ESE and the hnRNP A1 binding ESS immediately upstream. This antagonism is depicted in the figure and we hypothesize that a similar antagonism exists in the case of *SMN1* exon 7 and that it is disrupted in the case of *SMN2* exon 7. In addition, position of the putative stem loop structure is indicated (Singh, et al., 2004), as well as the sequence of the *SMN1/2* polypyrimidine tract. Purines in the PPT are underlined.

Field Code Changed

**Figure 2.** Mutational analysis of the putative ESS. **A)** Schematic overview of the pSXN13 minigene and constructs used in this study. Capitals indicate exonic bases. **B)** RT-PCR analysis of splicing of pSXN13 constructs in Chang cells. Inclusion or exclusion of the alternative exon is indicated on the left. hnRNP A1 and U2AF<sup>35</sup> motif scores are indicated below.

**Figure 3.** Binding consensus of U2AF<sup>35</sup>. **A)** Top: A pictogram depiction of the base composition of the U2AF<sup>35</sup> binding consensus. Bases that are more prevalent than the background frequency (25%) are shaded. Bottom: The Position Weight Matrix (PWM) of the U2AF<sup>35</sup> binding consensus. The individual scores of the most frequent bases at any given position are indicated in bold. We have included the 6<sup>th</sup> base for easy comparison with the hnRNP A1 motif in both the pictogram representation and the scoring matrix representation,

Deleted: coloured

1  
2 but since the information content in this position is very low it has not been included when  
3 calculating U2AF<sup>35</sup> scores. **B)** The PWM and scoring matrix of hnRNP A1 previously  
4 published (Cartegni, et al., 2006) is additionally included for comparison.  
5  
6  
7  
8  
9

Field Code Changed

10 **Figure 4.** Protein binding analysis. **A)** The RNA oligonucleotides used to bind proteins from  
11 a HeLa cell Nuclear Extract. X indicates the position of covalently bonded biotin. **B)** Peptide  
12 quantification of selected proteins. Individual peptides were identified by nLC-MS/MS  
13 analysis and quantitated by measuring the area of the corresponding peak in the nLC  
14 chromatogram. Peptides were individually normalized to the wt sample and relative protein  
15 quantities estimated by taking the average of the normalized peptide quantities. Error bars  
16 indicate standard error of mean. **C)** Western blots of protein eluates from the samples  
17 containing *SMN2* sequences. NE indicate control lane with Nuclear Extract. **D)** Western blot  
18 of protein eluates from the samples containing *MCAD* sequences. NE indicate control lane  
19 with Nuclear Extract.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

32 **Figure 5.** Analysis of an *SMN1/2* exon 7 spanning ESS by mutation. **A)** Schematic overview  
33 of the *SMN1/2* minigene and constructs used in this study. Capitals indicate exonic bases. The  
34 minigene is a hybrid construct consisting of both *MCAD* and *SMN* sequence, but retains the  
35 splicing pattern of *SMN1* and *SMN2* wt genes. Endogenous *SMN2* was used as template and  
36 the minigene retains the *SMN2* specific intron variations. **B)** RT-PCR analysis of splicing of  
37 *SMN1/2* minigenes in Chang cells. Inclusion or exclusion of exon 7 are indicated on the left.  
38 hnRNP A1 and U2AF<sup>35</sup> motif scores are indicated below.  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Figure 6.** Analysis of hnRNP A1 and hnRNP H function by siRNA. **A)** RT-PCR analysis of  
splicing of *SMN1/2* minigenes in HEK293 cells. hnRNP A1 and hnRNP H were knocked

1  
2 down by siRNA while a scrambled siRNA duplex was used as control. Band sizes  
3  
4 corresponding to transcripts either including or excluding exon 7 are indicated on the left. **B)**  
5  
6 RT-PCR analysis of endogenous *SMN1* and *SMN2* transcripts in HEK293 cells. RT-PCR was  
7  
8 performed on the same cDNA used for RT-PCR analysis in panel A. *SMN1* and *SMN2* were  
9  
10 discriminated by digesting the PCR products with DdeI. Band sizes corresponding to *SMN1*  
11  
12 and *SMN2* transcripts with or without exon 7 are indicated on the left. Also indicated is the  
13  
14 band corresponding to the 3' end of *SMN2* exon 8 which is a result of the digestion with  
15  
16 DdeI. **C)** Western blot analysis of protein knockdown in HEK293 cells. Antibodies against  
17  
18 Actin, hnRNP A1 and hnRNP H were used for staining. Actin was used as loading control.  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Page 30: [1] Deleted Thomas Koed Doktor 10/28/2010 10:38:00 PM

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- Abdul-Manan N, Williams KR. 1996. hnRNP A1 binds promiscuously to oligoribonucleotides: utilization of random and homo-oligonucleotides to discriminate sequence from base-specific binding. *Nucleic Acids Res* 24(20):4063-70.
- Brow DA. 2002. Allosteric cascade of spliceosome activation. *Annu Rev Genet* 36:333-60.
- Burd CG, Dreyfuss G. 1994. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J* 13(5):1197-204.
- Burnett BG, Munoz E, Tandon A, Kwon DY, Sumner CJ, Fischbeck KH. 2009. Regulation of SMN protein stability. *Mol Cell Biol* 29(5):1107-15.
- Butchbach ME, Edwards JD, Burghes AH. 2007. Abnormal motor phenotype in the SMNDelta7 mouse model of spinal muscular atrophy. *Neurobiol Dis* 27(2):207-19.
- Buvoli M, Cobianchi F, Biamonti G, Riva S. 1990. Recombinant hnRNP protein A1 and its N-terminal domain show preferential affinity for oligodeoxynucleotides homologous to intron/exon acceptor sites. *Nucleic Acids Res* 18(22):6595-600.
- Caputi M, Zahler AM. 2001. Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. *J Biol Chem* 276(47):43850-9.
- Cartegni L, Chew SL, Krainer AR. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 3(4):285-98.
- Cartegni L, Hastings ML, Calarco JA, de Stanchina E, Krainer AR. 2006. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *Am J Hum Genet* 78(1):63-77.
- Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31(13):3568-71.
- Cho S, Dreyfuss G. A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes Dev* 24(5):438-42.
- Coulter LR, Landree MA, Cooper TA. 1997. Identification of a new class of exonic splicing enhancers by in vivo selection. *Mol Cell Biol* 17(4):2143-50.
- Crawford TO, Pardo CA. 1996. The neurobiology of childhood spinal muscular atrophy. *Neurobiol Dis* 3(2):97-110.
- Dobrowolski SF, Andersen HS, Doktor TK, Andresen BS. 2010. The phenylalanine hydroxylase c.30C>G synonymous variation (p.G10G) creates a common exonic splicing silencer. *Mol Genet Metab*.
- Domsic JK, Wang Y, Mayeda A, Krainer AR, Stoltzfus CM. 2003. Human immunodeficiency virus type 1 hnRNP A/B-dependent exonic splicing silencer ESSV antagonizes binding of U2AF65 to viral polypyrimidine tracts. *Mol Cell Biol* 23(23):8762-72.
- Feldkotter M, Schwarzer V, Wirth R, Wienker TF, Wirth B. 2002. Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am J Hum Genet* 70(2):358-68.
- Graveley BR, Hertel KJ, Maniatis T. 2001. The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA* 7(6):806-18.

- 1  
2  
3 Hansen J, Corydon TJ, Palmfeldt J, Durr A, Fontaine B, Nielsen MN, Christensen JH,  
4 Gregersen N, Bross P. 2008. Decreased expression of the mitochondrial matrix  
5 proteases Lon and ClpP in cells from a patient with hereditary spastic paraplegia  
6 (SPG13). *Neuroscience* 153(2):474-82.
- 7  
8 Hastings ML, Allemand E, Duelli DM, Myers MP, Krainer AR. 2007. Control of pre-  
9 mRNA splicing by the general splicing factors PUF60 and U2AF65. *PLoS ONE*  
10 2(6):e538.
- 11 Hua Y, Sahashi K, Hung G, Rigo F, Passini MA, Bennett CF, Krainer AR. 2010.  
12 Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III  
13 SMA mouse model. *Genes Dev* 24(15):1634-44.
- 14 Hua Y, Vickers TA, Baker BF, Bennett CF, Krainer AR. 2007. Enhancement of SMN2  
15 exon 7 inclusion by antisense oligonucleotides targeting the exon. *PLoS Biol*  
16 5(4):e73.
- 17  
18 Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. 2008. Antisense masking of  
19 an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic  
20 mice. *Am J Hum Genet* 82(4):834-48.
- 21  
22 Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ.  
23 2004. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res*  
24 32(Database issue):D493-6.
- 25  
26 Kashima T, Manley JL. 2003. A negative element in SMN2 exon 7 inhibits splicing in  
27 spinal muscular atrophy. *Nat Genet* 34(4):460-3.
- 28 Kashima T, Rao N, David CJ, Manley JL. 2007a. hnRNP A1 functions with specificity in  
29 repression of SMN2 exon 7 splicing. *Hum Mol Genet* 16(24):3149-59.
- 30 Kashima T, Rao N, Manley JL. 2007b. An intronic element contributes to splicing  
31 repression in spinal muscular atrophy. *Proc Natl Acad Sci U S A* 104(9):3426-31.
- 32  
33 Kent OA, Ritchie DB, Macmillan AM. 2005. Characterization of a U2AF-independent  
34 commitment complex (E') in the mammalian spliceosome assembly pathway. *Mol*  
35 *Cell Biol* 25(1):233-40.
- 36  
37 Kikuchi M, Hatano N, Yokota S, Shimozawa N, Imanaka T, Taniguchi H. 2004.  
38 Proteomic analysis of rat liver peroxisome: presence of peroxisome-specific  
39 isozyme of Lon protease. *J Biol Chem* 279(1):421-8.
- 40 Le TT, Pham LT, Butchbach ME, Zhang HL, Monani UR, Coover DD, Gavrilina TO,  
41 Xing L, Bassell GJ, Burghes AH. 2005. SMNDelta7, the major product of the  
42 centromeric survival motor neuron (SMN2) gene, extends survival in mice with  
43 spinal muscular atrophy and associates with full-length SMN. *Hum Mol Genet*  
44 14(6):845-57.
- 45  
46 Lefebvre S, Burglen L, Reboullet S, Clermont O, Burllet P, Viollet L, Benichou B,  
47 Cruaud C, Millasseau P, Zeviani M and others. 1995. Identification and  
48 characterization of a spinal muscular atrophy-determining gene. *Cell* 80(1):155-  
49 65.
- 50  
51 Lefebvre S, Burllet P, Liu Q, Bertrand S, Clermont O, Munnich A, Dreyfuss G, Melki J.  
52 1997. Correlation between severity and SMN protein level in spinal muscular  
53 atrophy. *Nat Genet* 16(3):265-9.
- 54  
55 Lenski C, Kooy RF, Reyniers E, Loessner D, Wanders RJ, Winnepenninckx B,  
56 Hellebrand H, Engert S, Schwartz CE, Meindl A and others. 2007. The reduced  
57  
58  
59  
60

- 1  
2  
3 expression of the HADH2 protein causes X-linked mental retardation,  
4 choreoathetosis, and abnormal behavior. *Am J Hum Genet* 80(2):372-7.  
5  
6 Lorson CL, Androphy EJ. 2000. An exonic enhancer is required for inclusion of an  
7 essential exon in the SMA-determining gene SMN. *Hum Mol Genet* 9(2):259-65.  
8  
9 Martins de Araujo M, Bonnal S, Hastings ML, Krainer AR, Valcarcel J. 2009.  
10 Differential 3' splice site recognition of SMN1 and SMN2 transcripts by U2AF  
11 and U2 snRNP. *RNA* 15(4):515-23.  
12  
13 Masuda A, Shen XM, Ito M, Matsuura T, Engel AG, Ohno K. 2008. hnRNP H enhances  
14 skipping of a nonfunctional exon P3A in CHRNA1 and a mutation disrupting its  
15 binding causes congenital myasthenic syndrome. *Hum Mol Genet* 17(24):4022-  
16 35.  
17  
18 Mayeda A, Helfman DM, Krainer AR. 1993. Modulation of exon skipping and inclusion  
19 by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor  
20 SF2/ASF. *Mol Cell Biol* 13(5):2993-3001.  
21  
22 Merendino L, Guth S, Bilbao D, Martinez C, Valcarcel J. 1999. Inhibition of msl-2  
23 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site  
24 AG. *Nature* 402(6763):838-41.  
25  
26 Monani UR, Lorson CL, Parsons DW, Prior TW, Androphy EJ, Burghes AH, McPherson  
27 JD. 1999. A single nucleotide difference that alters splicing patterns distinguishes  
28 the SMA gene SMN1 from the copy gene SMN2. *Hum Mol Genet* 8(7):1177-83.  
29  
30 Nielsen KB, Sorensen S, Cartegni L, Corydon TJ, Doktor TK, Schroeder LD, Reinert LS,  
31 Elpeleg O, Krainer AR, Gregersen N and others. 2007. Seemingly neutral  
32 polymorphic variants may confer immunity to splicing-inactivating mutations: a  
33 synonymous SNP in exon 5 of MCAD protects from deleterious mutations in a  
34 flanking exonic splicing enhancer. *Am J Hum Genet* 80(3):416-32.  
35  
36 Okunola HL, Krainer AR. 2009. Cooperative-binding and splicing-repressive properties  
37 of hnRNP A1. *Mol Cell Biol* 29(20):5620-31.  
38  
39 Pedrotti S, Bielli P, Paronetto MP, Ciccocanti F, Fimia GM, Stamm S, Manley JL, Sette  
40 C. 2010. The splicing regulator Sam68 binds to a novel exonic splicing silencer  
41 and functions in SMN2 alternative splicing in spinal muscular atrophy. *EMBO J*  
42 29(7):1235-47.  
43  
44 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing  
45 genomic features. *Bioinformatics* 26(6):841-2.  
46  
47 Reese MG, Eeckman FH, Kulp D, Haussler D. 1997. Improved splice site detection in  
48 Genie. *J Comput Biol* 4(3):311-23.  
49  
50 Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M,  
51 Smith KE, Rosenbloom KR, Raney BJ and others. 2010. The UCSC Genome  
52 Browser database: update 2010. *Nucleic Acids Res* 38(Database issue):D613-9.  
53  
54 Sarkar G, Sommer SS. 1990. The "megaprimer" method of site-directed mutagenesis.  
55 *Biotechniques* 8(4):404-7.  
56  
57 Schaub MC, Lopez SR, Caputi M. 2007. Members of the heterogeneous nuclear  
58 ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by  
59 promoting formation of ATP-dependent spliceosomal complexes. *J Biol Chem*  
60 282(18):13617-26.

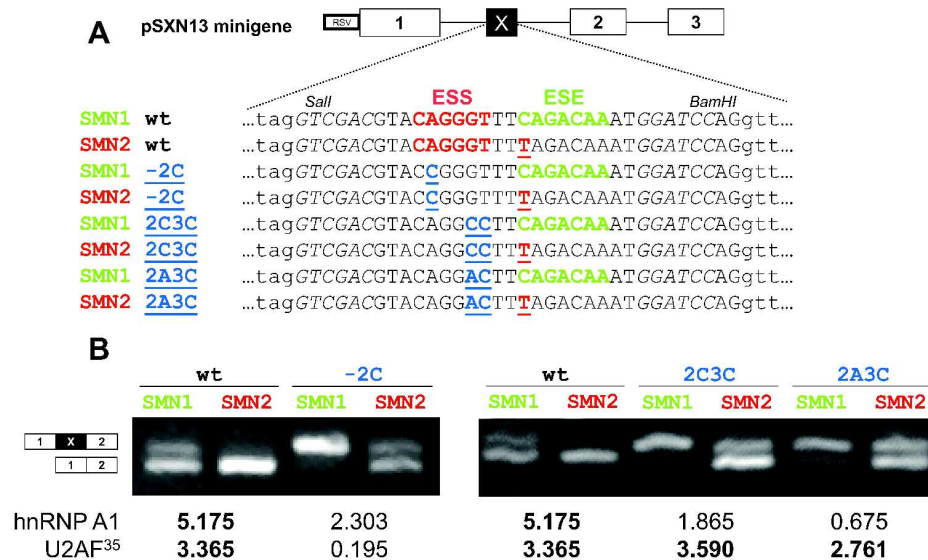
- 1  
2  
3 Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R. 2006.  
4 Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids*  
5 *Res* 34(14):3955-67.  
6  
7 Singh NN, Androphy EJ, Singh RN. 2004. An extended inhibitory context causes  
8 skipping of exon 7 of SMN2 in spinal muscular atrophy. *Biochem Biophys Res*  
9 *Commun* 315(2):381-8.  
10 Swanson MS, Dreyfuss G. 1988a. Classification and purification of proteins of  
11 heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities.  
12 *Mol Cell Biol* 8(5):2237-41.  
13 Swanson MS, Dreyfuss G. 1988b. RNA binding specificity of hnRNP proteins: a subset  
14 bind to the 3' end of introns. *EMBO J* 7(11):3519-29.  
15 Velasco E, Valero C, Valero A, Moreno F, Hernandez-Chico C. 1996. Molecular analysis  
16 of the SMN and NAIP genes in Spanish spinal muscular atrophy (SMA) families  
17 and correlation between number of copies of cBCD541 and SMA phenotype.  
18 *Hum Mol Genet* 5(2):257-63.  
19 Warf MB, Diegel JV, von Hippel PH, Berglund JA. 2009. The protein factors MBNL1  
20 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc Natl Acad*  
21 *Sci U S A* 106(23):9203-8.  
22 Wirth B, Herz M, Wetter A, Moskau S, Hahnen E, Rudnik-Schoneborn S, Wienker T,  
23 Zerres K. 1999. Quantitative analysis of survival motor neuron copies:  
24 identification of subtle SMN1 mutations in patients with spinal muscular atrophy,  
25 genotype-phenotype correlation, and implications for genetic counseling. *Am J*  
26 *Hum Genet* 64(5):1340-56.  
27 Wu JY, Maniatis T. 1993. Specific interactions between proteins implicated in splice site  
28 selection and regulated alternative splicing. *Cell* 75(6):1061-70.  
29 Wu S, Romfo CM, Nilsen TW, Green MR. 1999. Functional recognition of the 3' splice  
30 site AG by the splicing factor U2AF35. *Nature* 402(6763):832-5.  
31 Yeo G, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with  
32 applications to RNA splicing signals. *J Comput Biol* 11(2-3):377-94.  
33 Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the  
34 mammalian splicing factor U2AF. *Nature* 355(6361):609-14.  
35 Zatkova A, Messiaen L, Vandenbroucke I, Wieser R, Fonatsch C, Krainer AR, Wimmer  
36 K. 2004. Disruption of exonic splicing enhancer elements is the principal cause of  
37 exon skipping associated with seven nonsense or missense alleles of NF1. *Hum*  
38 *Mutat* 24(6):491-501.  
39 Zhang M, Zamore PD, Carmo-Fonseca M, Lamond AI, Green MR. 1992. Cloning and  
40 intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary  
41 factor small subunit. *Proc Natl Acad Sci U S A* 89(18):8769-73.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



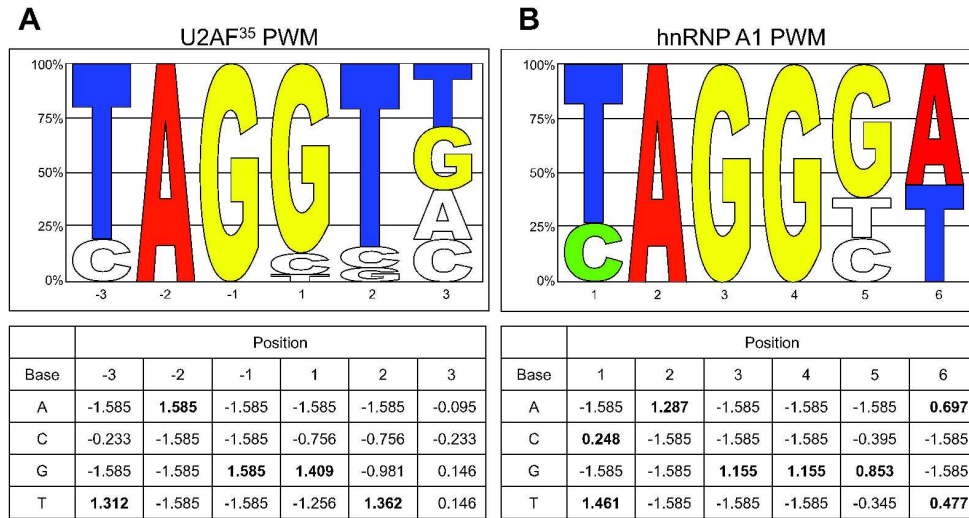
**Figure 1.** Similarities between *SMN1* and *SMN2* exon 7 and *MCAD* exon 5. The region comprising the ESS and ESE motifs in *SMN1* and *SMN2* exon 7 is compared to the corresponding region in *MCAD* exon 5, showing the extensive sequence similarity between the two exons. Both contain an SF2/ASF binding ESE as well as an hTra2-β1 binding ESE further downstream. In the context of *MCAD* exon 5, we have previously described the antagonistic functions of the ESE and the hnRNP A1 binding ESS immediately upstream. This antagonism is depicted in the figure and we hypothesize that a similar antagonism exists in the case of *SMN1* exon 7 and that it is disrupted in the case of *SMN2* exon 7. In addition, position of the putative stem loop structure is indicated (Singh, et al., 2004), as well as the sequence of the *SMN1/2* polypyrimidine tract. Purines in the PPT are underlined.

275x190mm (600 x 600 DPI)



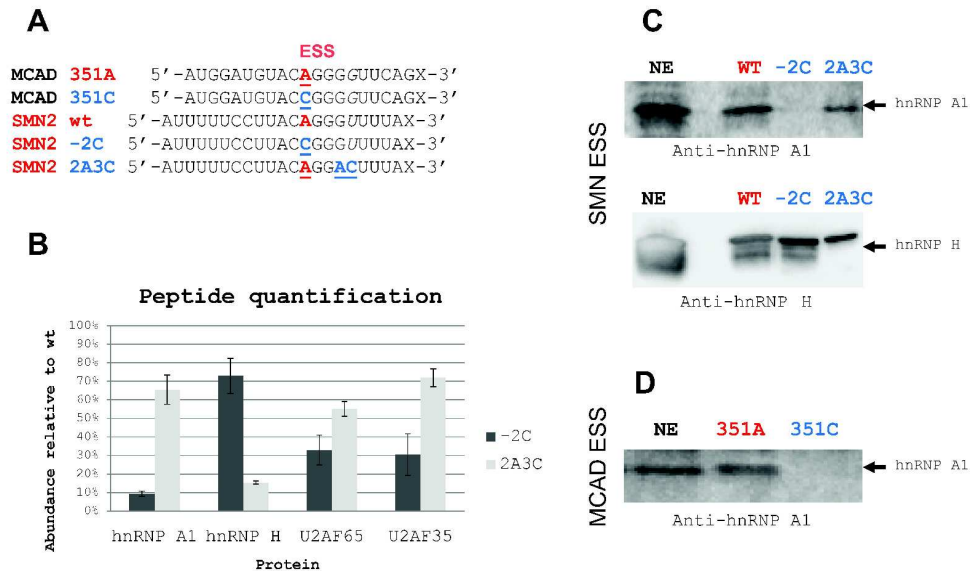


**Figure 2.** Mutational analysis of the putative ESS. **A)** Schematic overview of the pSXN13 minigene and constructs used in this study. Capitals indicate exonic bases. **B)** RT-PCR analysis of splicing of pSXN13 constructs in Chang cells. Inclusion or exclusion of the alternative exon is indicated on the left. hnRNP A1 and U2AF<sup>35</sup> motif scores are indicated below.



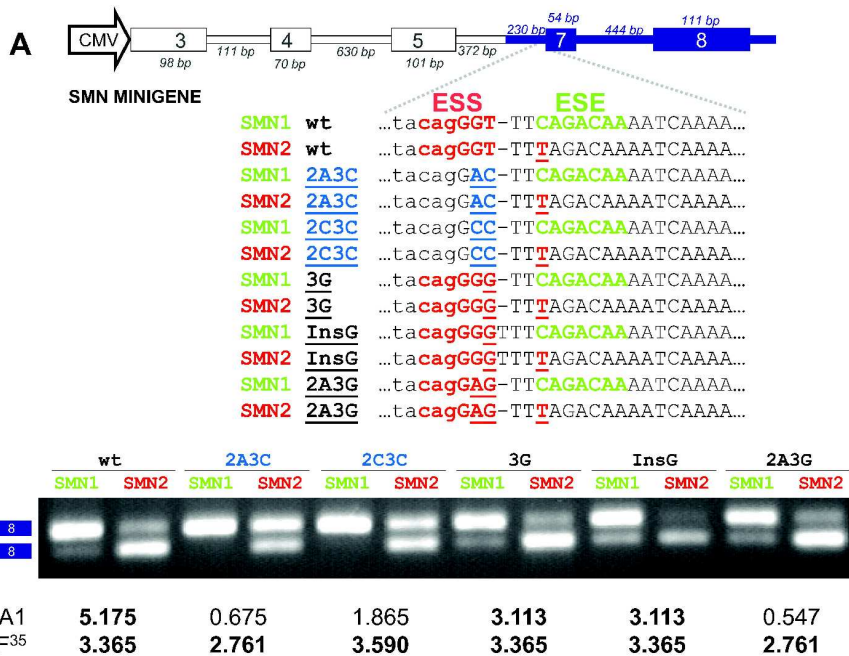
**Figure 3.** Binding consensus of U2AF<sup>35</sup>. **A)** Top: A pictogram depiction of the base composition of the U2AF<sup>35</sup> binding consensus. Bases that are more prevalent than the background frequency (25%) are shaded. Bottom: The Position Weight Matrix (PWM) of the U2AF<sup>35</sup> binding consensus. The individual scores of the most frequent bases at any given position are indicated in bold. We have included the 6<sup>th</sup> base for easy comparison with the hnRNP A1 motif in both the pictogram representation and the scoring matrix representation, but since the information content in this position is very low it has not been included when calculating U2AF<sup>35</sup> scores. **B)** The PWM and scoring matrix of hnRNP A1 previously published (Cartegni, et al., 2006) is additionally included for comparison.

275x190mm (600 x 600 DPI)



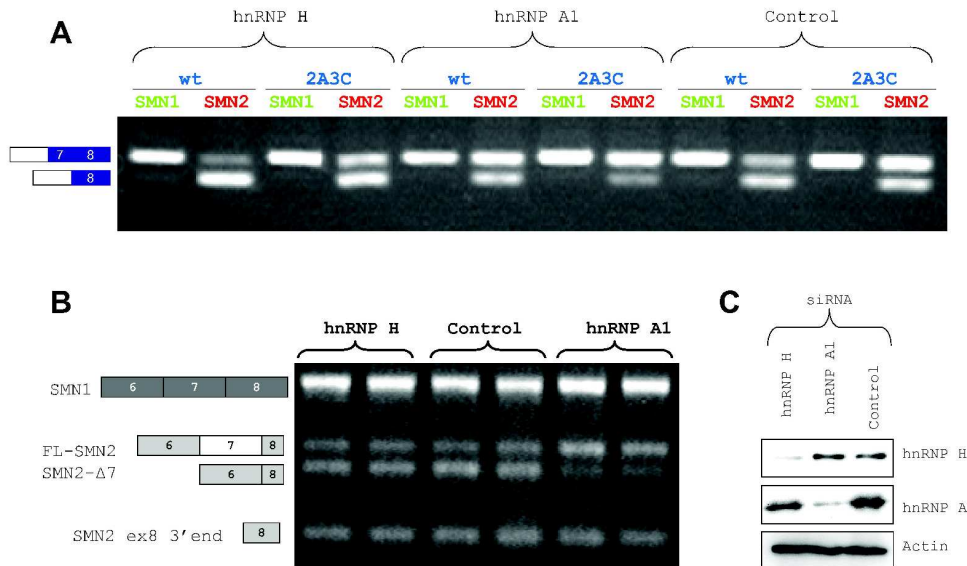
**Figure 4.** Protein binding analysis. **A)** The RNA oligonucleotides used to bind proteins from a HeLa cell Nuclear Extract. X indicates the position of covalently bonded biotin. **B)** Peptide quantification of selected proteins. Individual peptides were identified by nLC-MS/MS analysis and quantitated by measuring the area of the corresponding peak in the nLC chromatogram. Peptides were individually normalized to the wt sample and relative protein quantities estimated by taking the average of the normalized peptide quantities. Error bars indicate standard error of mean. **C)** Western blots of protein eluates from the samples containing *SMN2* sequences. NE indicate control lane with Nuclear Extract. **D)** Western blot of protein eluates from the samples containing *MCAD* sequences. NE indicate control lane with Nuclear Extract.

275x190mm (600 x 600 DPI)



**Figure 5.** Analysis of an *SMN1/2* exon 7 spanning ESS by mutation. **A)** Schematic overview of the *SMN1/2* minigene and constructs used in this study. Capitals indicate exonic bases. The minigene is a hybrid construct consisting of both *MCAD* and *SMN* sequence, but retains the splicing pattern of *SMN1* and *SMN2* wt genes. Endogenous *SMN2* was used as template and the minigene retains the *SMN2* specific intron variations. **B)** RT-PCR analysis of splicing of *SMN1/2* minigenes in Chang cells. Inclusion or exclusion of exon 7 are indicated on the left. hnRNP A1 and U2AF<sup>35</sup> motif scores are indicated below.

275x190mm (600 x 600 DPI)



**Figure 6.** Analysis of hnRNP A1 and hnRNP H function by siRNA. **A)** RT-PCR analysis of splicing of SMN1/2 minigenes in HEK293 cells. hnRNP A1 and hnRNP H were knocked down by siRNA while a scrambled siRNA duplex was used as control. Band sizes corresponding to transcripts either including or excluding exon 7 are indicated on the left. **B)** RT-PCR analysis of endogenous *SMN1* and *SMN2* transcripts in HEK293 cells. RT-PCR was performed on the same cDNA used for RT-PCR analysis in panel A. *SMN1* and *SMN2* were discriminated by digesting the PCR products with DdeI. Band sizes corresponding to *SMN1* and *SMN2* transcripts with or without exon 7 are indicated on the left. Also indicated is the band corresponding to the 3' end of *SMN2* exon 8 which is a result of the digestion with DdeI. **C)** Western blot analysis of protein knockdown in HEK293 cells. Antibodies against Actin, hnRNP A1 and hnRNP H were used for staining. Actin was used as loading control.

275x190mm (600 x 600 DPI)