

7-16-2019

How RNA structure dictates the usage of a critical exon of spinal muscular atrophy gene

Natalia N. Singh

Iowa State University, natalias@iastate.edu

Ravindra N. Singh

Iowa State University, singhr@iastate.edu

Follow this and additional works at: https://lib.dr.iastate.edu/bms_pubs

 Part of the [Genetic Structures Commons](#), [Medical Genetics Commons](#), and the [Molecular Genetics Commons](#)

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/bms_pubs/72. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Biomedical Sciences at Iowa State University Digital Repository. It has been accepted for inclusion in Biomedical Sciences Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

How RNA structure dictates the usage of a critical exon of spinal muscular atrophy gene

Abstract

Role of RNA structure in pre-mRNA splicing has been implicated for several critical exons associated with genetic disorders. However, much of the structural studies linked to pre-mRNA splicing regulation are limited to terminal stem-loop structures (hairpins) sequestering splice sites. In few instances, role of long-distance interactions is implicated as the major determinant of splicing regulation. With the recent surge of reports of circular RNA (circRNAs) generated by backsplicing, role of Alu-associated RNA structures formed by long-range interactions are taking central stage. Humans contain two nearly identical copies of *Survival Motor Neuron* (SMN) genes, *SMN1* and *SMN2*. Deletion or mutation of *SMN1* coupled with the inability of *SMN2* to compensate for the loss of *SMN1* due to exon 7 skipping causes spinal muscular atrophy (SMA), one of the leading genetic diseases of children. In this review, we describe how structural elements formed by both local and long-distance interactions are being exploited to modulate *SMN2* exon 7 splicing as a potential therapy for SMA. We also discuss how Alu-associated secondary structure modulates generation of a vast repertoire of *SMN* circRNAs. This article is part of a Special Issue entitled: RNA structure and splicing regulation edited by Francisco Baralle, Ravindra Singh and Stefan Stamm.

Keywords

RNA structure, Splicing, Small molecule, ISS-N1, U1 snRNP, SMN, SMA

Disciplines

Genetic Structures | Medical Genetics | Molecular Genetics

Comments

This article is published as Singh, Natalia N., and Ravindra N. Singh. "How RNA structure dictates the usage of a critical exon of spinal muscular atrophy gene." *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* (2019). DOI: [10.1016/j.bbagr.2019.07.004](https://doi.org/10.1016/j.bbagr.2019.07.004). Posted with permission.

Creative Commons License

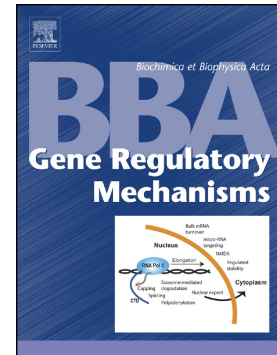


This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Accepted Manuscript

How RNA structure dictates the usage of a critical exon of spinal muscular atrophy gene

Natalia N. Singh, Ravindra N. Singh



PII: S1874-9399(18)30510-8

DOI: <https://doi.org/10.1016/j.bbagr.2019.07.004>

Reference: BBAGRM 94403

To appear in: *BBA - Gene Regulatory Mechanisms*

Received date: 17 May 2019

Accepted date: 9 July 2019

Please cite this article as: N.N. Singh and R.N. Singh, How RNA structure dictates the usage of a critical exon of spinal muscular atrophy gene, *BBA - Gene Regulatory Mechanisms*, <https://doi.org/10.1016/j.bbagr.2019.07.004>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

How RNA structure dictates the usage of a critical exon of spinal muscular atrophy gene

Natalia N. Singh and Ravindra N. Singh*

Department of Biomedical Science, Iowa State University, Ames, IA, 50011

Running Title: RNA structure of SMA gene

*Corresponding author.

Department of Biomedical Science, Iowa State University, 2035 Veterinary Medicine, Ames, IA, 50011, Tel.: 515-294-8505, Fax: 515-294-2315, Email: singhr@iastate.edu

Keywords: RNA structure, splicing, small molecule, ISS-N1, U1 snRNP, SMN, SMA,

ABSTRACT

Role of RNA structure in pre-mRNA splicing has been implicated for several critical exons associated with genetic disorders. However, much of the structural studies linked to pre-mRNA splicing regulation are limited to terminal stem-loop structures (hairpins) sequestering splice sites. In few instances, role of long-distance interactions is implicated as the major determinant of splicing regulation. With the recent surge of reports of circular RNA (circRNAs) generated by backsplicing, role of Alu-associated RNA structures formed by long-range interactions are taking central stage. Humans contain two nearly identical copies of *Survival Motor Neuron (SMN)* genes, *SMN1* and *SMN2*. Deletion or mutation of *SMN1* coupled with inability of *SMN2* to compensate for the loss of *SMN1* due to exon 7 causes spinal muscular atrophy (SMA), one of the leading genetic diseases of children. In this review, we describe how structural elements formed by both local and long-distance interactions are being exploited to modulate *SMN2* exon 7 splicing as a potential therapy for SMA. We also discuss how Alu-associated secondary structure modulate generation of a vast repertoire of *SMN* circRNAs.

1. Introduction

Alternative pre-mRNA splicing enhances the coding potential of eukaryotic genome by generating multiple transcripts from a single gene [1-4]. All human genes with two or more exons have potential to be alternatively spliced. Complexity of alternative splicing multiplies with the increasing number of exons within a gene. Pre-mRNA splicing can occur in both directions, forward and backward [5]. While forward splicing generates linear transcripts, including messenger RNAs (mRNAs) and long non-coding RNAs (lncRNAs), backsplicing produces circular RNAs (circRNAs) [5-7]. Based on the experimental validations in case of *Survival Motor Neuron (SMN)* gene, a 30-kb long human gene has capability to generate hundreds of transcripts if forward and back splicing are appropriately factored [8-11]. Not all transcripts generated by alternative splicing are stable due to RNA surveillance mechanisms, which preferentially degrade transcripts with premature termination codons (PTCs) and/or other suboptimal features [12,13]. Hence, the abundance of transcripts generated by a gene cannot be accurately captured. Arrangement of cis-elements on a pre-mRNA coupled with the relative expression of transacting factors are the drivers of alternative splicing [14,15]. Depending upon the context, similar cis-element may have positive or negative effect on alternative splicing of an exon [16,17].

Context of an exon is defined by multiple factors, including size of the exon, strength of the splice sites, positioning of cis-elements within the exon and the flanking intronic sequences [14-21]. RNA structure adds an important dimension to the context of an exon as it modulates accessibility of the cis-elements and shortens the distance between cis-elements located far apart [22-28]. In case of backsplicing that generates circRNAs, RNA structure plays a pivotal role in bringing the downstream 5' splice site (5'ss) in close proximity to the upstream 3'ss [29]. RNA structure-mediated alternative splicing events are sensitive to the surrounding environment that affects structure, including temperature, ionic strength and the interacting molecules [25,30-32]. Consequently, RNA structure may serve as a trigger for differential splicing of exons under changed environment such as thermal, osmotic and oxidative stress. In living cells splicing is coupled to transcription and vice versa [33]. Chromatin structure can influence the context by modulating the rate of transcription and consequently determining the outcome of alternative splicing [34,35]. Protein factors and small molecules that modulate the rate of transcription

through modification of chromatin structure can have effect on the context of an exon as well [33]. There is also evidence to suggest that the nascent pre-mRNA structure influences transcription and consequently modulates co-transcriptional splicing regulation [36].

Humans have two nearly identical copies of *Survival Motor Neuron (SMN) gene*: *SMN1* and *SMN2* [37]. Both *SMN* genes code for SMN, an essential protein linked to diverse functions, including snRNP assembly, pre-mRNA splicing, transcription, translation, stress granule formation, macromolecular trafficking and maintenance of the cytoskeletal dynamics [38-40]. A critical C-to-T mutation at the 6th position (C6U substitution in RNA) of exon 7 distinguishes *SMN2* from *SMN1* [41,42]. Although C6U substitution alone is sufficient to trigger *SMN2* exon 7 skipping [41], more than fifty cis-elements and transacting factors have been implicated in regulation of *SMN* exon 7 splicing (Fig.1A). While a complete description of all regulatory elements of *SMN* is beyond the scope of this report, reviews published on this topic over the recent years capture how our understanding of *SMN* exon 7 splicing regulation is continuing to evolve [43-46]. The overriding evidence support that the effect of a disease-causing exonic mutation (such as C6U substitution within *SMN2* exon 7) could be overcome by compensatory mutations within both exonic and intronic sequences. In this review, we focus on a structure as a modulator of *SMN* exon 7 splicing.

Due to predominant skipping of *SMN2* exon 7, *SMN2* contributes very little towards the generation of the overall pool of SMN in human cells (Fig. 1B). Consequently, reduction in SMN levels due to the loss of *SMN1* causes spinal muscular atrophy (SMA), a leading genetic disease of children and infants [47]. While degeneration of motor neurons is the first manifestation of the disease, peripheral tissues are also affected in SMA [48,49]. In addition, based on the severe impairment of male reproductive organ development and male infertility observed in mild SMA mouse models, the disease appears to be gender-specific [50-52]. Incidentally, testis produces high levels of SMN due to an adult-specific splicing switch from exclusion to mostly inclusion of *SMN2* exon 7 [50]. The mechanism of this splicing switch remains unknown. In general, *SMN2* is dispensable for survival of humans. However, testis-specific switch of *SMN2* exon 7 splicing to fulfill the demand for an elevated levels of SMN during testicular development offers solid rationale to why humans have retained copies of

SMN2 during evolution. Considering *SMN2* is almost universally present in SMA patients, correction of *SMN2* exon 7 splicing provides one of the best avenues for SMA therapy.

Nusinersen (SpinrazaTM), the first drug approved by FDA in 2016 for the treatment of SMA, acts through correction of *SMN2* exon 7 splicing [53,54]. Nusinersen is an antisense oligonucleotide (ASO) that sequesters Intronic Splicing Silencer N1 (ISS-N1) located within intron 7 (Fig. 1B) [55,56]. Positioned within a complex structural context, ISS-N1 harbors multiple cis-elements, including a GC-rich sequence and two hnRNP motifs and a cryptic 5'ss, Cr1 (Fig. 1A) [57-59]. Two putative TIA1 binding sites are located immediately downstream of ISS-N1 (Fig. 1A) [60]. Another cryptic 5'ss, Cr2, located 28 nucleotides (nts) downstream of the Cr1 falls within the TIA1 binding site (Fig. 1A) [59]. Similar to ISS-N1 targeting ASOs, engineered U1 snRNPs (eU1s) annealing to ISS-N1 or sequences in its vicinity also promote *SMN2* exon 7 inclusion [59,61-63]. It is proposed that the stimulatory effect of ISS-N1 targeting ASO is at least in part due to the structural rearrangement at the 5'ss of exon 7 [56]. Similar mechanism could hold true for the stimulatory effects of eU1s that target ISS-N1 or nearby sequences. In case when the authentic 5'ss of exon 7 is mutated, eU1s annealing to ISS-N1 or sequences in the vicinity preferentially activate Cr1 [59,63]. Here again, eU1-induced structural arrangement could play an important role in bringing Cr1 and the 3'ss of intron 7 in close proximity. A number of small molecules that modulate *SMN2* exon 7 splicing with high specificity are at different stages of the pre-clinical and/or clinical development [64-67]. At least two of these small molecules are shown to interact with RNA structures within *SMN* pre-mRNA [66,67].

2. Role of local structures formed by exon 7

As indicated by the results of in vitro structure probing, most of the 54 nts of human *SMN* exon 7 is engaged in the formation of RNA secondary structures comprised of two terminal stem loops (TSLs) and an internal stem (IS) (Fig. 2) [68,69]. TSL1 and TSL2 sequester the 3'ss and 5'ss, respectively (Fig. 2) [23]. The IS1 is formed by the central region of exon 7 and intron 6 sequence (Fig. 2) [23]. Here we discuss the role of all probed structures in modulation of *SMN* exon 7 splicing. Where relevant, we also discuss potential role of the predicted secondary structures involved in *SMN* exon 7 splicing regulation.

2.1 Role of TSL1

The TSL1 is formed by the first sixteen nucleotides of exon 7 (Fig. 2A) [68,69]. Six of these sixteen nucleotides are located in the loop of TSL1; whereas, remaining ten residues form the stem. C6U substitution that distinguishes *SMN2* from *SMN1* is located at the loop-closing position and is predicted to increase the stem size of TSL1 from 5-bp to 6-bp duplex (Fig. 2) [68,69]. However, the results of enzymatic structure probing showed unpaired nature of 6U, suggesting that the stem size of TSL1 is not changed in *SMN2* [23]. In this scenario, the UAGACA motif created by C6U substitution falls within the loop of TSL1. It has been proposed that UAGACA motif serves as a strong binding site for hnRNP A1, a negative regulator of *SMN2* exon 7 splicing [70]. Hence, UAGACA location within the loop makes this motif highly accessible for interactions with hnRNP A1. In case of *SMN1*, the loop of TSL1 is represented by the CAGACA motif that overlaps with CAGACAA motif, which is proposed to interact with ASF/SF2, a positive regulator of *SMN1* exon 7 splicing [71]. Aside from the loop harboring protein-binding motifs, mutations within the stem region of TSL1 promote *SMN2* exon 7 inclusion, suggesting that the entire TSL1 structure is inhibitory [69]. Based on the results of extensive point mutations, we concluded that TSL1 is located within an extended inhibitory context (Exinct) spanning from the 3rd to 15th positions of exon 7 [69]. An antisense microwalk independently validated the presence of Exinct [72]. In particular, ASOs disrupting TSL1 promoted *SMN2* exon 7 inclusion [72]. In vivo selection serves as one of the unbiased methods to interrogate the position-specific significance of nucleotide residues in alternative splicing [73]. Incidentally, *SMN1* exon 7 was the first entire exon subjected to in vivo selection [74]. The results of in vivo selection of *SMN1* exon 7 supported the presence of Exinct [74]. Of note, unlike *SMN1* exon 7, *SMN2* exon 7 has a strengthened Exinct due to C6U substitution. More recently, a machine-learning-based computational analysis validated the findings of in vivo selection and suggested that the Exinct-like motifs created by genetic mutations are associated with genetic diseases [75].

Employing in vitro Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE)-based RNA structure probing method, an alternative TSL structure was proposed at the location previously designated as TSL1 (Fig. 2A) [23, 66]. Disparity between these two structures at the beginning of exon 7 could be due to two different in vitro methods used for

structure probing [23,66]. Also, differences in structures could be due to different lengths of sequences analyzed for in vitro structure probing. Interestingly, only a part of the alternative structure at the beginning of exon 7 was found to be preserved when sequences were analyzed by in-cell SHAPE structure probing method [66]. Future studies will determine if RNA structures adopt different confirmations based on the presence of the exon 7-interacting factors in the cell.

2.2 Role of TSL2

The TSL2 is the first experimentally validated structure of *SMN* pre-mRNA (Fig. 2) [23]. TSL2 is formed by a 19-nt long sequence encompassing the last seventeen and the first two residues of exon 7 and intron 7, respectively. Based on the results of in vivo selection of the entire exon 7, TSL2 encompasses an inhibitory region called the “3'-Cluster” [73,74]. Antisense microwalk as well as the machine-learning-based computational analysis validate the inhibitory nature of the 3'-Cluster [72,75]. Supporting the negative effect of TSL2, point mutations that abrogate TSL2 stimulate *SMN2* exon 7 inclusion (Fig. 2C) [23]. Conversely, compensatory mutations that restore TSL2 promote *SMN2* exon 7 skipping (Fig. 2C) [23]. Located within TSL2, wild-type A residue at the last position (54A) of exon 7 has been determined to be highly inhibitory [74]. Consistently, an A-to-G substitution (A54G) at the last position of exon 7 fully restores *SMN2* exon 7 inclusion. The stimulatory effect of A54G substitution happens to be so pronounced that absence of most known positive cis-elements within exon 7 is tolerated [74]. A54G substitution is predicted to destabilize TSL2 as it changes the canonical U-A base pair into wobble U-G base pair. In addition, 54G increases the base pairing between endogenous U1 snRNA and the 5'ss of exon 7. An eU1 with extended base-pairing with 5'ss of exon 7 fully restored *SMN2* exon 7 inclusion [23]. These findings support that the poor recruitment of U1 snRNP is one of the limiting factors for *SMN2* exon 7 splicing. Recently, a small compound PK4C9 that promotes *SMN2* exon 7 inclusion, has been shown to interact with TSL2 [67]. Based on this interaction, authors propose that PK4C9 favors recruitment of U1 snRNP through remodeling of the 5'ss of exon 7.

2.3 Role of IS1

Based on the enzymatic structure probing, IS1 is a 10-bp long duplex in the middle of *SMN* exon 7 (Fig. 2A) [23]. While the purine-rich 3'-strand of IS1 is provided by exon 7, the pyrimidine-rich 5'-strand of IS1 comes from intron 6. IL1, a 4-nt (ACAG) internal loop, bridges the 5'-strand of IS1 with the 5'-strand of TSL1 (Fig. 2A). IL2, a 6-nt purine-rich (AAAAG) internal loop, bridges the 3'-strand of TSL1 with the 3'-strand of IS1 (Fig. 2A). IL3, another 5-nt internal loop, bridges the 3'-strand of IS1 with the 5'-strand of TSL2 (Fig. 2A). Based on the results of in vivo selection, sequences involved in the formation of IL2, IS1 and IL3 appear to be stimulatory [74]. Incidentally, IL2 overlaps with the binding site of Tra2- β 1, which recruits other factors to exon 7 (Fig. 1A). List of factors interacting with exon 7 is continuing to grow, although no effort has yet been made to associate these interactions with specific structures. It is likely that the single stranded regions, including IL2 and IL3 provide primary binding sites for positive regulators. Supporting this argument, mutations within either IL2 or IL3 have been found to promote *SMN1* exon 7 skipping [74].

Unlike enzymatic structure probing, results of a recent in vitro chemical structure probing by SHAPE revealed a shorter IS1 (Fig. 2B) [23,66]. Further, findings of in vitro SHAPE showed different structural context upstream and downstream of IS1. However, results of in cell SHAPE only partially validated the structural context revealed by in vitro SHAPE. Interestingly, a small compound, SMN-C3, that stimulates *SMN2* exon 7 inclusion has been proposed to interact with IS1. However, this interaction is proposed to be modulated by protein factor(s) [66].

3. Role of structural context of intron 7

Secondary structure of *SMN2* intron 7 deduced from the chemical structure probing was reported in 2013 (Fig. 3) [27]. The key findings of the probed secondary structure are the presence of internal stems formed by unique long-distance interactions (LDIs). Here we describe functional significance of some of these LDIs and as well as other structures formed by intron 7.

3.1 Role of ISTL1

Internal stem formed by the long-distance interaction 1 (ISTL1) is a unique intra-intronic structure located within intron 7 [27]. The two strands of the 8-bp ISTL1 are separated from each other by 279 nts. Together with TSL2, ISTL1 sequesters the entire 5'ss of exon 7. In addition, the

5'-strand of ISTL1 encompasses half of the GC-rich motif that has been confirmed to be an effective target for a short (8-mer) ASO [57,76]. The last base of the 5'-strand of ISTL1 is a C residue located at the 10th intronic position (¹⁰C), which also happens to be the first residue of the 15-nt long ISS-N1 [27]. The significance of ¹⁰C first emerged when two 14-nt long ASOs, F14 and L14, targeting ISS-N1 produced opposite effects on *SMN2* exon 7 splicing [18]. While F14 sequestered the first 14 residues (including ¹⁰C) of ISS-N1 and promoted exon 7 inclusion, L14 sequestered the last 14 residues (excluding ¹⁰C) of ISS-N1 and promoted exon 7 skipping [18]. Results were validated using ASOs with two different chemistries (locked nucleic acid and 2'-O-methyl modifications), ruling out any chemistry-specific effect [18]. Subsequent experiments confirmed that F14 destabilizes ISTL1 due to sequestration of ¹⁰C, whereas, L14 stabilizes ISTL1 due to formation of a duplex with the remainder of the ISS-N1 sequence. Supporting this argument, deletion of deep intronic sequences encompassing the 3'-strand of ISTL1 eliminated the negative effect associated with L14 [27]. Also, increasing the length of ISTL1 duplex substantially enhanced the skipping of *SMN2* exon 7. These results provided one of the rarest examples of how a single nucleotide shift in the annealing position of an ASO could flip the effect on splicing. Findings also suggested that the effect of an ISS-N1-annealing ASO may not be due to the displacement of the negative factor(s) (hnRNP A1/A2) only, it could also be due to the change in structural context upstream and/or downstream of the ASO annealing position. While ISTL1 is the first validated regulatory structure formed by an LDI in the context of a human genetic disease, a similar structure in intron 3 of *Proteolipid Protein 1 (PLP1)* gene has been implicated in Pelizaeus-Merzbacher disease (PMD) [77].

3.2 Role of other ISTLs

In addition to ISTL1, additional structures formed by LDIs within intron 7 include ISTL2, ISTL3 and ISTL4 (Fig. 3) [27]. Two strands of ISTL2, ISTL3 and ISTL4 are separated from each other by 226, 85 and 55 nts, respectively. Similar to ISTL1, ISTL2 and ISTL3 exert negative effect on *SMN2* splicing, although the effect is not as pronounced as in case of ISTL1. The negative effect of ISTL2 could at least in part be attributed to the sequestration of TIA1 binding site located immediately downstream of the ISS-N1. Interestingly, a continuous intronic sequence, termed ISS-N2, provides the 3'-strands of ISTL1, ISTL2 and ISTL3. Deletion or an ASO-mediated sequestration of ISS-N2 confers stimulatory effect on *SMN2* exon 7 inclusion [27]. Unlike

ISTL1, ISTL2 and ISTL3, the significance of ISTL4 in exon 7 splicing regulation has not yet been established.

3.3 Other structures within intron 7

Results of structure probing supported the formation of several TSLs, ILs and bulges (Fig. 3) [27]. Significance of some of these structures could be inferred from the location of the critical cis-elements within these structures. For instance, one of the putative hnRNP A1 binding sites within ISS-N1 is located within an IL. This loop also harbors four of the eight residues of the GC-rich motif that serves as the target for the shortest ASO. Presence of this loop makes ISS-N1 and GC-rich sequence highly accessible for targeting by ASO. Other putative hnRNP A1 motif associated with ISS-N1 is sequestered within TSL3, which also locks one of the TIA1 binding sites located downstream of the ISS-N1. It is likely that the stimulatory effect of the ISS-N1-targeting ASO is linked at least in part to freeing of the TIA1 motif. Sequences downstream of ISTL4 form two TSLs, TSL6 and TSL7. Deletion of intron 7 sequences encompassing TSL6 and TSL7 moderately stimulated *SMN2* exon 7 inclusion [27]. Future studies will determine if these and other structures within intron 7 play critical role in regulation of exon 7 splicing.

4. Role of RNA structure in backsplicing

Primate-specific Alu elements are the most abundant short interspersed nuclear elements (SINEs) totaling more than one million copies and accounting for ~11% of the human genome [78,79]. Surprisingly, a whopping 39% sequence of human *SMN* genes are represented by Alu elements as inverted repeats in the intronic regions [80]. Presence of inverted Alu repeats results in formation of double stranded structures that are conducive for the generation of circRNAs through backsplicing [81,82]. Indeed, a recent report confirmed the existence of a large repertoire of circRNAs generated from *SMN* genes (Fig. 4) [11]. Based on the nature of exons, *SMN* circRNAs are broadly categorized into four types [11]. While type 1 circRNAs harbor one or more of the early exons (exons 1 through 4), type 2 circRNAs harbor middle exons in addition to early exons (exons 1 through 7). Some of the type 1 circRNAs contain novel exons derived from sequences within intron 1. Type 3 circRNAs harbor the 3'-terminal exons with or without upstream exons. Some of the type 3 circRNAs also contain one or more of the four novel exons derived from intergenic region located downstream of the annotated *SMN* gene. Type 4

circRNAs are generated by a combination of backsplicing and trans-splicing, such that one or more exons in these circRNAs originate from other genes. Overall, more than fifty *SMN* circRNAs are universally produced in all cell types examined, with type 1 circRNAs being the most predominant [11]. All *SMN* circRNAs are generated by backsplicing from canonical splice sites. Both *SMN1* and *SMN2* generate circRNAs with near equal frequencies. Most common type 1 circRNAs produced by *SMN* are C2A-2B-3-4, C2B-3-4 and C3-4 (Fig. 4). These three circRNAs share the same 5'ss of exon 4 suggesting a random pairing between the 5'ss of exon 4 with the 3'ss of any of the upstream exons 2A, 2B and 3. Predicted secondary structures support such pairing through RNA:RNA duplex formation between sequences at the 5'ss of exon 4 and the 3'ss of the upstream exons (Fig. 4A) [11]. C5-6 is the predominant circRNA among type 2 *SMN* circRNAs (Fig. 4) [11]. Formation of C5-6 is likely to be facilitated by inverted Alu repeats that are predicted form secondary structures bringing the 5'ss of exon 6 and the 3'ss of the exon 5 in close proximity (Fig. 4A). Among type 3 circRNAs of *SMN*, C6-7-8A and C6-7-8A-9tr1-10 are the most common. Here again, generation of type 3 circRNAs is likely facilitated by Alu-associated secondary structures that bring the 3'ss of exon 6 in close proximity with 5'ss of any of the downstream exons (Fig. 4B). Type 4 circRNAs of *SMN* are generated by incorporation exonic sequences from the neighboring Alu-rich genes, including *ERBIN* and *SERFI* [11].

Generation of *SMN* circRNAs appears to be a posttranscriptional event, as suppression of transcription elongation had no appreciable effect on the relative abundance of *SMN* circRNAs. At the same time, depletion of DHX9, an RNA helicase that specifically unwinds Alu-associated secondary structures, enhanced expression of several *SMN* circRNAs, including C3-4, one of the most predominant circRNAs of *SMN* [11]. Interestingly, depletion of DHX9 also promoted skipping of exons 3 and 4 in various combinations, suggesting that the C3-4 is generated from intronic lariats harboring these exons. Considering the 5'ss of exon 4 is used to generate several circRNAs, including C2A-2B-3-4, C2B-3-4 and C3-4, an enhanced expression of C3-4 under the DHX9-depleted conditions consequently led to decreased expression of C2A-2B-3-4 and C2B-3-4. These findings also suggested that the base pairing between inverted Alu elements located upstream and downstream of exons 3 and 4, respectively, is under the control of DHX9. Many of the types 2 and 3 circRNAs showed elevated expressions under the conditions of depleted DHX9, suggesting that the base pairing between Alu elements located toward the 3'-terminal

region of *SMN* pre-mRNA are also under the control of DHX9 [11]. Future studies will determine the role of specific secondary structures that are primary drivers for the generation of *SMN* circRNAs.

6. Conclusions

Generation of mature RNAs, including mRNAs, lncRNAs and circRNAs requires an accurate removal of introns during pre-mRNA splicing. Information to remove an intron is distributed among cis-elements located within both intronic and exonic sequences. Studies conducted for *SMN* pre-mRNA splicing regulation are instructive in suggesting that a large part of essential information for intron removal is locked into structural elements. The inhibitory structural context is one of the primary causes of *SMN2* exon 7 skipping linked to SMA. The currently approved drug, Spinraza, works at least in part through abrogation of the inhibitory structural context at the 5'ss of exon 7. Many of the small compounds that have recently shown promise for SMA therapy appear to also work through modulation of RNA structure at the 5'ss of exon 7. A vast repertoire of the newly reported *SMN* circRNAs are generated through structural elements abundantly present within Alu-rich *SMN*. While functions of *SMN* circRNAs are not yet clear, generation of these circRNAs come at the expense of linear transcripts required to produce SMN, an essential protein involved in multiple aspects of RNA metabolism. Future studies will determine how transacting factors and structural elements collaborate to maintain a fine balance between linear transcripts and circRNAs generated by *SMN*.

Acknowledgements

Authors acknowledge Dr. Joonbae Seo for the critical reading of the manuscript and for valuable suggestions. While authors have attempted to include most contributions on role of RNA structure in splicing regulation of SMA gene, they regret for not being able to include several related references due to the lack of space.

Funding: This work was supported by grants from the National Institutes of Health (R01 NS055925 and R21 NS101312).

Disclosures and competing interests: ISS-N1 target (US patent # 7,838,657) mentioned in this review was discovered in the Singh lab at UMASS Medical School (Worcester, MA, USA). Inventors, including RNS, NNS and UMASS Medical School, are currently benefiting from licensing of ISS-N1 target (US patent # 7,838,657) to IONIS Pharmaceuticals/Biogen, which is marketing Spinraza™ (Nusinersen), the FDA-approved drug, based on ISS-N1 target.

References

- [1] T.W. Nilsen, B.R. Graveley BR, Expansion of the eukaryotic proteome by alternative splicing, *Nature* 463 (2010) 457-463.
- [2] O. Kelemen, P. Convertini, Z. Zhang, Y. Wen, M. Szhén, M. Falaleeva, S. Stamm, Function of alternative splicing, *Gene* 514 (2013) 1-30.
- [3] B. Raj, B.J. Blencowe, Alternative splicing in the mammalian nervous system: recent insights into mechanisms and functional roles, *Neuron* 87 (2015) 14-27.
- [4] F.E. Baralle, J. Giudice J, Alternative splicing as a regulator of development and tissue identity, *Nat. Rev. Mol. Cell. Biol.* 18 (2017) 437-451.
- [5] T.J. Chuang, Y.J. Chen, C.Y. Chen, T.L. Mai, Y.D. Wang, C.S. Yeh, M.Y. Yang, Y.T. Hsiao, T.H. Chang, T.C. Kuo, H.H. Cho, C.N. Shen, H.C. Kuo, M.Y. Lu, Y.H. Chen, S.C. Hsieh, T.W. Chiang, Integrative transcriptome sequencing reveals extensive alternative trans-splicing and cis-backsplicing in human cells. *Nucleic Acids Res.* 46 (2018) 3671-3691.
- [6] I.W. Deveson, S.A. Hardwick, T.R. Mercer, J.S. Mattick, The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome, *Trends Genet.* 33 (2017) 464-478.
- [7] L.S. Kristensen, T.L.H. Okholm, M.T. Venø, J. Kjems, Circular RNAs are abundantly expressed and upregulated during human epidermal stem cell differentiation, *RNA Biol.* 15 (2018) 280-291.
- [8] N.N. Singh, J. Seo, S.J. Rahn, R.N. Singh, A multi-exon-skipping detection assay reveals surprising diversity of splice isoforms of spinal muscular atrophy genes. *PLoS One.* 7 (2012) e49595. doi: 10.1371/journal.pone.0049595.

- [9] J. Seo, N.N. Singh, E.W. Ottesen, S. Sivanesan, M. Shishimorova, R.N. Singh, Oxidative Stress Triggers Body-Wide Skipping of Multiple Exons of the Spinal Muscular Atrophy Gene. *PLoS One*. 11 (2016) e0154390.
- [10] J. Seo, N.N. Singh, E.W. Ottesen, B.M. Lee, R.N. Singh, A novel human-specific splice isoform alters the critical C-terminus of Survival Motor Neuron protein. *Sci. Rep.* 6 (2016) 0778. doi: 10.1038/srep30778.
- [11] E.W. Ottesen, D. Luo, J. Seo, N.N. Singh, R.N. Singh, Human Survival Motor Neuron genes generate a vast repertoire of circular RNAs, *Nucleic Acids Res.* 47 (2019) 2884-2905
- [12] P.J. da Costa, J. Menezes, L. Romão, The role of alternative splicing coupled to nonsense-mediated mRNA decay in human disease, *Int. J. Biochem. Cell. Biol.* 91(2017) 168-175.
- [13] S. Zheng, Alternative splicing and nonsense-mediated mRNA decay enforce neural specific gene expression, *Int. J. Dev. Neurosci.* 55 (2016) 102-108.
- [14] Y. Lee, D.C. Rio, Mechanisms and Regulation of Alternative Pre-mRNA Splicing, *Annu. Rev. Biochem.* 84 (2015) 291-323.
- [15] H. Dvinge, Regulation of alternative mRNA splicing: old players and new perspectives, *FEBS Lett.* 592 (2018) 2987-3006.
- [16] R. Martinez-Contreras, J.F. Fisette, F.U. Nasim, R. Madden, M. Cordeau, B. Chabot , Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing, *PLoS Biol.* 4 (2006) e21.
- [17] Z. Wang, M. Kayikci, M. Briese, K. Zarnack, N.M. Luscombe, G. Rot, B. Zupan, T. Curk, J. Ule, iCLIP predicts the dual splicing effects of TIA-RNA interactions, *PLoS Biol.* 8 (2010) e1000530.
- [18] N.N. Singh, K. Hollinger, D. Bhattacharya, R.N. Singh, An antisense microwalk reveals critical role of an intronic position linked to a unique long-distance interaction in pre mRNA splicing, *RNA* 16 (2010) 1167-1181.
- [19] S. Erkelenz, W.F. Mueller, M.S. Evans, A. Busch, K. Schöneweis, K.J. Hertel, H. Schaal, Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms, *RNA* 19 (2013) 96–102.
- [20] X.D. Fu, M Jr. Ares, Context-dependent control of alternative splicing by RNA-binding proteins, *Nat. Rev. Genet.* 15 (2014) 689-701.

- [21] M.S. Wong, J.B. Kinney, A.R. Krainer, Quantitative Activity Profile and Context Dependence of All Human 5' Splice Sites, *Mol. Cell* 71 (2018) 1012-1026.
- [22] M. Hiller, Z. Zhang, R. Backofen, S. Stamm, Pre-mRNA secondary structures influence exon recognition, *PLoS Genet.* 3 (2007) e204.
- [23] N.N. Singh, R.N. Singh, E.J. Androphy, Modulating role of a RNA structure in skipping of a critical exon in the spinal muscular atrophy genes, *Nucleic Acids Res.* 35 (2007) 371-389.
- [24] E. Buratti, A. Dhir, M.A. Lewandowska, F.E. Baralle, RNA structure is a key regulatory element in pathological ATM and CFTR pseudoexon inclusion events, *Nucleic Acids Res.* 35 (2007) 4369-4383.
- [25] M.B. Warf, J.A. Berglund, Role of RNA structure in regulating pre-mRNA splicing, *Trends Biochem. Sci.* 35 (2010) 169-178.
- [26] D.D. Pervouchine, E.E. Khrameeva, M.Y. Pichugina, O.V. Nikolaienko, M.S. Gelfand, P.M. Rubtsov, A.A. Mironov, Evidence for widespread association of mammalian splicing and conserved long-range RNA structures, *RNA* 18 (2012) 1-15.
- [27] N.N. Singh, M.N. Lawler, E.W. Ottesen, D. Upreti, J.R. Kaczynski, R.N. Singh, An intronic structure enabled by a long-distance interaction serves as a novel target for splicing correction in spinal muscular atrophy, *Nucleic Acids Res.* 41 (2013) 8144-8165.
- [28] N.N. Singh, B.M. Lee, R.N. Singh, Splicing regulation in spinal muscular atrophy by a RNA structure formed by long distance interactions, *Annals of NY Academy of Sciences* 1341 (2015) 176-187.
- [29] J.E. Wilusz, A 360° view of circular RNAs: From biogenesis to functions, *Wiley Interdiscip. Rev. RNA* 9 (2018), e1478.
- [30] Y. Wan, M. Kertesz, R.C. Spitale, E. Segal, H.Y. Chang, Understanding the transcriptome through RNA structure, *Nat. Rev. Genet.* 12 (2011) 641-655.
- [31] J. Peschek, D. Acosta-Alvear, A.S. Mendez, P. Walter, A conformational RNA zipper promotes intron ejection during non-conventional XBP1 mRNA splicing. *EMBO Rep.* 16 (2015) 1688-1698.
- [32] M. Preußner, G. Goldammer, A. Neumann, T. Haltenhof, P. Rautenstrauch, M. Müller-McNicoll, F. Heyd, Body Temperature Cycles Control Rhythmic Alternative Splicing in Mammals, *Mol. Cell.* 67 (2017) 433-446.e4.

- [33] T. Saldi, M.A. Cortazar, R.M. Sheridan, D.L. Bentley, Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *J. Mol. Biol.*, 428 (2016), 2623-2635.
- [34] G. Dujardin, C. Lafaille, E. Petrillo, V. Buggiano, L.I. Gómez Acuña, A. Fiszbein, M.A. Godoy Herz, N. Nieto Moreno, M.J. Muñoz, M. Alló, I.E. Schor, A.R. Kornblihtt, Transcriptional elongation and alternative splicing, *Biochim. Biophys. Acta.* 1829 (2013) 134-40.
- [35] N. Nieto Moreno, L.E. Giono, A.E. Cambindo Botto, M.J. Muñoz, A.R. Kornblihtt, Chromatin, DNA structure and alternative splicing, *FEBS Lett.* 589 (2015) 3370-3378.
- [36] J.W. Zhang, R. Landick, A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure, *Trends Biochem. Sci.*, 41 (2016), 293-310.
- [37] S. Lefebvre, L. Bürglen, S. Reboullet, O. Clermont, P. Burlet, L. Viollet, B. Benichou, C. Cruaud, P. Millasseau, M. Zeviani, Identification and characterization of a spinal muscular atrophy-determining gene, *Cell* 80 (1995) 155-165.
- [38] R.N. Singh, M.D. Howell, E.W. Ottesen, N.N. Singh, Diverse role of survival motor neuron protein, *Biochim Biophys Acta Gene Regul Mech.* 1860 (2017) 299-315.
- [39] E.W. Ottesen, N.N. Singh, D. Luo D, R.N. Singh, High-affinity RNA targets of the Survival Motor Neuron protein reveal diverse preferences for sequence and structural motifs, *Nucleic Acids Res.* 46 (2018) 10983-11001.
- [40] P.L. Price, D. Morderer, W. Rossoll, RNP Assembly Defects in Spinal Muscular Atrophy. *Adv. Neurobiol.* 20 (2018) 143-171.
- [41] C.L. Lorson, E. Hahnen, E.J. Androphy, B. Wirth, A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy, *Proc. Natl. Acad. Sci. U.S.A.* 11 (1999) 6307-6311.
- [42] U.R. Monani, C.L. Lorson, D.W. Parsons, T.W. Prior, E.J. Androphy, A.H. Burghes, J.D. McPherson, A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2, *Hum Mol Genet.* 8 (1999) 1177-1183.
- [43] R.N. Singh, Evolving concepts on human SMN pre-mRNA splicing, *RNA Biol.* 4 (2007) 7-10.
- [44] N.N. Singh, R.N. Singh, Alternative splicing in spinal muscular atrophy underscores the role of an intron definition model, *RNA Biol.* 8 (2011) 600-606.

- [45] N.N. Singh, M.D. Howell, R.N. Singh, Transcription and splicing regulation of spinal muscular atrophy genes, in: C.J. Sumner, S. Paushkin, C.-P. Ko (Eds.), *Spinal Muscular Atrophy: Disease Mechanisms and Therapy*, Academic Press, London, (2017) 75-97.
- [46] R.N. Singh, N.N. Singh, Mechanism of Splicing Regulation of Spinal Muscular Atrophy Genes, *Adv Neurobiol.* 20 (2018) 31-61.
- [47] S. Ahmad, K. Bhatia, A. Kannan, L. Gangwani, Molecular Mechanisms of Neurodegeneration in Spinal Muscular Atrophy. *J. Exp. Neurosci.* 10 (2016) 39-49.
- [48] L.A. Nash, J.K. Burns, J.W. Chardon, R. Kothary, R.J. Parks, Spinal Muscular Atrophy: More than a Disease of Motor Neurons? *Curr. Mol. Med.* 16 (2016) 779-792.
- [49] S.L. Lipnick, D.M. Agniel, R. Aggarwal, N.R. Makhortova, S.G. Finlayson, A. Brocato, N. Palmer, B.T. Darras, I. Kohane, L.L. Rubin, Systemic nature of spinal muscular atrophy revealed by studying insurance claims, *PLoS One.* 14 (2019) e0213680.
- [50] E.W. Ottesen, M.D. Howell, N.N. Singh, J. Seo, E.M. Whitley, R.N. Singh, Severe impairment of male reproductive organ development in a low SMN expressing mouse model of spinal muscular atrophy, *Sci. Rep.* 6 (2016) 20193.
- [51] M.D. Howell, E.W. Ottesen, N.N. Singh, R.L. Anderson, J. Seo, S. Sivanesan, E.M. Whitley, R.N. Singh, TIA1 is a gender-specific disease modifier of a mild mouse model of spinal muscular atrophy, *Sci. Rep.* 7 (2017) 7183.
- [52] M.D. Howell, E.W. Ottesen, N.N. Singh, R.L. Anderson, R.N. Singh, Gender-Specific Amelioration of SMA Phenotype upon Disruption of a Deep Intronic Structure by an Oligonucleotide, *Mol Ther.* 25 (2017) 1328-1341.
- [53] E.W. Ottesen, ISS-N1 makes the First FDA-approved Drug for Spinal Muscular Atrophy, *Transl Neurosci.* 8 (2017) 1-6.
- [54] N.N. Singh, M.D. Howell, E.J. Androphy, R.N. Singh, How the discovery of ISS-N1 led to the first medical therapy for spinal muscular atrophy, *Gene Ther.* 24 (2017) 520-526.
- [55] N.K. Singh, N.N. Singh, E.J. Androphy, R.N. Singh, Splicing of a critical exon of human Survival Motor Neuron is regulated by a unique silencer element located in the last intron, *Mol. Cell. Biol.* 26 (2006) 1333-1346.
- [56] N.N. Singh, B.M. Lee, C.J. DiDonato, R.N. Singh, Mechanistic principles of antisense targets for the treatment of spinal muscular atrophy, *Future Med. Chem.* 7 (2015) 1793-1808.

- [57] N.N. Singh, M. Shishimorova, L.C. Cao, L. Gangwani, R.N. Singh, A short antisense oligonucleotide masking a unique intronic motif prevents skipping of a critical exon in spinal muscular atrophy, *RNA Biol.* 6 (2009) 341-350.
- [58] Y. Hua, T.A. Vickers, H.L. Okunola, C.F. Bennett, A.R. Krainer, Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice, *Am. J. Hum. Genet.* 82 (2008) 834-848.
- [59] N.N. Singh, J.B. Del Rio-Malewski, D. Luo, E.W. Ottesen, M.D. Howell, R.N. Singh, Activation of a cryptic 5' splice site reverses the impact of pathogenic splice site mutations in the spinal muscular atrophy gene, *Nucleic Acids Res.* 45 (2017) 12214-12240.
- [60] N.N. Singh, J. Seo, E.W. Ottesen, M. Shishimorova, D. Bhattacharya, R.N. Singh, TIA1 prevents skipping of a critical exon associated with spinal muscular atrophy, *Mol. Cell. Biol.* 1 (2011) 935-954.
- [61] A. Dal Mas, M.E. Rogalska, E. Bussani, F. Pagani, Improvement of SMN2 pre-mRNA processing mediated by exon-specific-U1-small nuclear RNA, *Am J Hum Genet.* 96 (2015) 93-103.
- [62] M.E. Rogalska, M. Tajnik, D. Licastro, E. Bussani, L. Camparini, C. Mattioli, F. Pagani, Therapeutic activity of modified U1 core spliceosomal particles, *Nat. Commun.* 7 (2016) 11168.
- [63] R.N. Singh, N.N. Singh, A novel role of U1 snRNP: Splice site selection from a distance. *Biochim. Biophys. Acta.* 1862 (2019) 634-642.
- [64] J. Palacino, S.E. Swalley, C. Song, A.K. Cheung, L. Shu, X. Zhang, M. Van Hoosear, Y. Shin, D.N. Chin, C.G. Keller, M. Beibel, N.A. Renaud, T.M. Smith, M. Salcius, X. Shi, M. Hild, R. Servais, M. Jain, L. Deng, C. Bullock, M. McLellan, S. Schuierer, L. Murphy, M.J. Blommers, C. Blaustein, F. Berenshteyn, A. Lacoste, J.R. Thomas, G. Roma, G.A. Michaud, B.S. Tseng, J.A. Porter, V.E. Myer, J.A. Tallarico, L.G. Hamann, D. Curtis, M.C. Fishman, W.F. Dietrich, N.A. Dales, R. Sivasankaran, SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice, *Nat. Chem. Biol.* 11 (2015) 511-517.
- [65] M. Sivaramakrishnan, K.D. McCarthy, S. Campagne, S. Huber, S. Meier, A. Augustin, T. Heckel, H. Meistermann, M.N. Hug, P. Birrer, A. Moursy, S. Khawaja, R. Schmucki, N. Berntenis, N. Giroud, S. Golling, M. Tzouros, B. Banfai, G. Duran-Pacheco, J. Lamerz, Y.

- Hsiu Liu, T. Luebbers, H. Ratni, M. Ebeling, A. Cléry, S. Paushkin, A.R. Krainer, F.H. Allain, F. Metzger, Binding to SMN2 pre-mRNA-protein complex elicits specificity for small molecule splicing modifiers, *Nat. Commun.* 8 (2017) 1476.
- [66] J. Wang, P.G. Schultz, K.A. Johnson, Mechanistic studies of a small-molecule modulator of SMN2 splicing, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) E4604-E4612.
- [67] A. Garcia-Lopez, F. Tessaro, H.R.A. Jonker, A. Wacker, C. Richter, A. Comte, N. Berntenis, R. Schmucki, K. Hatje, O. Petermann, G. Chiriano, R. Perozzo, D. Sciarra, P. Konieczny, I. Faustino, G. Fournet, M. Orozco, R. Artero, F. Metzger, M. Ebeling, P. Goekjian, B. Joseph, H. Schwalbe, L. Scapozza, Targeting RNA structure in SMN2 reverses spinal muscular atrophy molecular phenotypes, *Nat Commun.* 9 (2018), 2032.
- [68] N.N. Singh, E.J. Androphy, R.N. Singh, The regulation and regulatory activities of alternative splicing of the SMN gene. *Critical Reviews in Eukaryotic Gene Expression* 14 (2004), 271-285.
- [69] N.N. Singh, E.J. Androphy, R.N. Singh, An extended inhibitory context causes skipping of exon 7 of *SMN2* in spinal muscular atrophy. *Biochem. Biophys. Res. Commun.* 315 (2004), 381-388.
- [70] T. Kashima, J.L. Manley, A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy, *Nat. Genet.* 34 (2003) 460-463.
- [71] L. Cartegni, A.R. Krainer, Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1, *Nat. Genet.* 30 (2002) 377-384.
- [72] Y. Hua, T.A. Vickers, B.F. Baker, C.F. Bennett, A.R. Krainer, Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon, *PLoS Biol.* 5 (2007) e73.
- [73] R.N. Singh, Unfolding the mystery of alternative splicing through a unique method of in vivo selection, *Front. Biosci.* 12 (2007) 3263-3272.
- [74] N.N. Singh, E.J. Androphy, R.N. Singh, In vivo selection reveals features of combinatorial control that defines a critical exon in the spinal muscular atrophy genes, *RNA* 10 (2004) 1291-1305.
- [75] H.Y. Xiong, B. Alipanahi, L.J. Lee, H. Bretschneider, D. Merico, R.K. Yuen, Y. Hua, S. Gueroussov, H.S. Najafabadi, T.R. Hughes, Q. Morris, Y. Barash, A.R. Krainer, N. Jovic,

- S.W. Scherer, B.J. Blencowe, B.J. Frey, RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease, *Science* 347 (2015) 1254806.
- [76] J.M. Keil, J. Seo, M.D. Howell, W.H. Hsu, R.N. Singh, C.J. DiDonato, A short antisense oligonucleotide ameliorates symptoms of severe mouse models of spinal muscular atrophy, *Mol. Ther. Nucleic Acids* 3 (2014) e174.
- [77] J.R. Taube, K. Sperle, L. Banser, P. Seeman, B.C. Cavan, J.Y. Garbern, G.M. Hobson, PMD patient mutations reveal a long-distance intronic interaction that regulates PLP1/DM20 alternative splicing, *Hum. Mol. Genet.* 23 (2014) 5464-5478.
- [78] P. Deininger, Alu elements: know the SINEs, *Genome Biol.* 12 (2011) 236.
- [79] C. Daniel, G. Silberberg, M. Behm, M. Ohman, Alu elements shape the primate transcriptome by cis-regulation of RNA editing, *Genome Biol.* 15 (2014) R28.
- [80] E.W. Ottesen, J. Seo, N.N. Singh, R.N. Singh, A Multilayered Control of the Human *Survival Motor Neuron* Gene Expression by Alu Elements, *Front. Microbiol.* 8 (2017) 2252.
- [81] J.E. Wilusz, Repetitive elements regulate circular RNA biogenesis, *Mob. Genet. Elem.* 5 (2015) 1-7.
- [82] X.O. Zhang, H.B. Wang, Y. Zhang, X.H. Lu, L.L. Chen, L. Yang, Complementary sequence-mediated exon circularization, *Cell*, 159 (2014), 134-147.

Figure Legends:

Figure 1. Regulation of *SMN* exon 7 splicing. (A) Diagrammatic representation of intronic and exonic cis-elements as well as trans-acting factors that modulate *SMN* exon 7 splicing. Upper-case letters signify exonic sequences, small-case letters, intronic sequences. Exons and introns are also shown as colored boxes and lines, respectively. Numbering of nucleotides, neutral and positive, starts from the first exonic and intronic position, respectively. The 5' and 3' splice sites (5'ss and 3'ss) are indicated by arrows. Cr1 and Cr2 represent cryptic 5' splice sites recently reported [59]. Negative and positive regulators of exon 7 splicing are indicated by (-) and (+), respectively. For details see [28,46]. (B) Effect of Spinraza™ on *SMN2* exon 7 splicing. Exons are represented by colored boxes, introns by lines. First 27 nts of intron 7 with highlighted ISS-N1 element are given. Annealing positions of Spinraza™ to ISS-N1 target are shown. *SMN2* pre-mRNA splicing results in predominant skipping of exon 7 in mRNA (*SMN2Δ7*) and production of truncated SMN protein (*SMNΔ7*). Sequestration of ISS-N1 by Spinraza™ promote *SMN2* exon 7 inclusion (*SMN2FL*) and as a consequence increases levels of full-length SMN protein (SMN).

Figure 2. Local structures formed by *SMN* exon 7 and adjacent upstream/downstream intronic sequences. Interactions of small exon 7-splicing modifying compounds with a local structure. (A) Secondary structure of *SMN2* exon 7 and its surrounding intronic sequences derived from enzymatic structure probing [23]. Existence of TSL2 and its effect on exon 7 splicing was also confirmed by mutational analysis. Intron 6 and intron 7 sequences are shown in small-case green and blue letters, respectively. Exon 7 sequence is shown in upper-case black letters. Numbering of nucleotides, neutral, positive and negative, starts from the first position in exon 7, first position of intron 7 and the last position in intron 6, respectively. The splice sites of exon 7 are indicated by the red arrows. IS1 structure is boxed. Nucleotides in magenta indicate site-specific mutations, single and compensatory, used to confirm the existence of TSL2 and its effect on exon 7 splicing [23]. Abbreviation: IL, internal loop; IS, internal stem; TSL, terminal stem loop. (B) Interactions of a small exon 7-splicing modifying compound, SMN-C3, with a local structure. Secondary structure of *SMN2* exon 7 and its surrounding intronic sequences was derived from in-cell SHAPE-mutation profiling [66]. Markings and abbreviations are the same as

in (A). Chemical structure of the small compound, SMN-C3, that increases inclusion of exon 7 is shown. Binding site of SMN-C3 (in the context of a double strand) is highlighted with red box. Residues, whose sensitivity to a structure probing agent, NAI, is increased and decreased upon SMN-C3 binding are highlighted in a pink and a blue circle, respectively [66]. (C) Interactions of a small exon 7-splicing modifying compound, PK4C9, with TSL2 structure. Upper-case letters signify exonic sequence, small-case letters in blue, intron 7 sequence. Nucleotides in red indicate site-specific mutations, single and double, used to confirm the effect of destabilization or stabilization TSL2 on exon 7 splicing. The 5'ss of exon 7 is indicated by the red arrows. Chemical structure of PK4C9 is shown. In the lower panel PK4C9 is represented as a red star. Upon PK4C9 binding to TSL2, several residues at the base of the stem highlighted with pink circles form hydrogen bonds, hydrophobic interactions or stack with the compound [67]. These interactions destabilize TSL2 and increase the accessibility of the 5'ss. In addition, binding of PK4C9 to TSL2 stabilizes the triloop of this RNA structure.

Figure 3. Secondary structure of *SMN* intron 7. (A) SHAPE-derived structure of intron 7 [27]. The structure is based on combined probing results obtained with multiple extension primers. An exon 7/intron 7 junction as well as the 5'ss are indicated. Exon 8 is represented by a green box. Numbering of nucleotides, neutral and negative, starts from the first position of intron 7 and the last position of exon 7, respectively. Binding sites of hnRNP A1/A2 and TIA1 are highlighted in pink and green, respectively. Residues with normalized 1M7 reactivity > 0.5 are marked with large circles, while residues with normalized 1M7 reactivity between 0.3 and 0.5 are indicated with small circles. Broken lines indicate potential base-pairing. Positions corresponding to Reverse Transcriptase (RTase) falloffs are marked with "F". Residues in red constitute a region whose structure was not determined due to RTase fall off. Abbreviations: ISTL, internal stem formed by a long-distance interaction; LS1, long-distance interaction site 1; TSL, terminal stem-loop; ISS-N2, intronic splicing silencer 2. (B) Site-specific mutations in ISTL1 and adjacent regions that effect this structure and cause changes in a pattern in exon 7 splicing with/without antisense oligonucleotides [27]. Arrows and letters in red indicate positions and identities of mutations, single and compensatory, that break (when single) and restore (when combined) ISTL1 structure. Arrows and letters in green indicate positions and identities of mutations deep within intron 7, which increase the size of the stem and cause increase in exon 7 skipping.

Figure 4. Repertoire of circRNAs generated from the *SMN* genes. (A) Diagrammatic representation of circRNAs identified using primers that anneal to exons 2a, 2b, 3, 4 and 5. A genomic overview of the *SMN* gene layout is given. Exons are indicated by colored shapes, introns by lines. Exon sizes are given using numbers in black located below exons, intron sizes are shown using numbers in grey located above introns. A newly identified 5'ss within exon 8 is indicated with black arrow. Colored arrows show canonical backsplicing events. Arrow thickness reflects the estimated prevalence of each event. Schematic of the likely role of the Alu-associated secondary structures in the generation of the most abundant types 1 and 2 circRNAs are given (diagram not to the scale). (B) Diagrammatic representation of circRNAs identified using primers that anneal to exons 6, 6b, 7 and 8. A genomic overview of the 3' portion of the *SMN* gene layout is given. Labeling and color coding are the same as in (A). Schematic of the likely role of the Alu-associated secondary structures in the generation of the most abundant type 3 circRNAs are given (diagram not to the scale).