Vectors for therapeutic antisense sequences delivery and the modification of messenger RNA processing

Julie Tordo

Thesis for the Degree of Doctor of Philosophy

UNIVERSITY COLLEGE LONDON

2012

Supervisor: Prof. Olivier Danos

Second supervisor: Prof. Mary Collins

UCL Cancer Institute
Department of Gene Therapy
Paul O'Gorman Building
72 Huntley Street - London WC1E 6BT

Acknowledgments

The end of the road is now fast approaching and this journey has been a tremendous experience full of lessons learned and challenges. Writing this thesis has finalised this experience, which would never have happened without the help and support of so many.

I would like to deeply thank Prof. Olivier Danos for all these years passed under his supervision. From the first day I met him he made me fall in love with Gene Therapy and I thoroughly enjoyed working for him ever since. His constant encouragements and help as well as his trust enabled me to gain confidence and to complete this project. I thank him as well for his endless patience with my sometimes naive questions and mistakes; he was always able to re-direct my interest and motivation in the hardest times, even across continents. I won't thank him enough for the opportunity he gave me four years ago and I sincerely hope we will be able to work together again in the future.

Thank you very much to Prof. Mary Collins and Dr. Yasu Takeuchi for their help and comments on my work, as well as for welcoming us to join their lab meetings and celebrations for the past year. I learned a lot and had wonderful times during those moments.

I also would like to thoroughly thank Agathe Eckenfelder for teaching me everything during my master and for helping me throughout my PhD project. She has always been here to help, always with a smile and a laugh, and she really encouraged me to go on in a research project and to never give up even during the hardest times. Thanks also to all my previous colleagues at Necker hospital in Paris who made my first experience in a research lab so enjoyable. Thanks as well to Aurelie Goyenvalle for answering many of my questions and, of course, for her help on this work.

I also really would like to thank all my colleagues from 206 that made my time at the UCL Cancer Institute so enjoyable and unforgettable. In particular, Veronica and Martin, you brightened my days there and I thank you so much for the continuous entertainment, laughs and support you gave me, as well as all the helpful talks about

my work and problem faced. My PhD would have been so much different and difficult without you, thanks for having been so great friends! Thanks also to all the people from the third floor that welcomed me as one of theirs and made me feel at home when I first arrived to open the lab. Sarah, Chrys, Mora, Ghassem, Jenny, Marco and so many more: this would have been so difficult without your help and I will never forget the fantastic times I had with all of you.

Thanks also to all my dear friends scattered everywhere, Paris, Strasbourg and London, who helped me to switch off and have fantastic times with them when it was much needed. Thanks for their patience as well, they were always here when I needed them.

Last, but truly not least, thanks SO MUCH to my family for their endless support and encouragements during these hard times. In particular my parents and my sister, I don't even know how to thank them enough! They always believed in me (even when I was not), always brightened my days with words of support, long talks, amazing moments and parcels full of delicacies. Your support is so important to me, you were always here to encourage me and I really, really would not be here without you! Thanks to the rest of the family too, for the good times and of course for their support, especially "Uncle Christmas" who always had a fitting joke for the situation! Finally, thanks a lot Samir for bearing with me during these years. These have been long hard times but you have been incredibly patient and supportive, always pushing me forward when I was doubting. You made this journey so much easier by just being here for me.

I, Julie Tordo, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Synthetic antisense oligoribonucleotides can be used to modulate gene splicing by masking key motifs on the pre-mRNA required for spliceosome assembly. Yet, intracellular expression of oligoribonucleotides generates only a transitory effect whereas stable delivery of antisense sequences can be achieved by linking them to chimeric small RNAs delivered and expressed by viral vectors.

In the murine model of Duchenne Muscular Dystrophy a chimeric U7 snRNA (U7Dtex23) induces skipping of the mutated exon 23 and restores the Dystrophin mRNA reading frame. The main limitation of this approach remains the large amount of snRNA vector needed to be produced and administered to patients. To optimize this system we used self-complementary AAV vectors (scAAV) to express the U7snRNA shuttles. ScAAV vectors were tested in mouse myoblast cultures and we observed an increase in U7Dtex23 expression and in Dystrophin exon 23 skipping compared to single-stranded AAV, highlighting the potential for this strategy to reduce the vector dose. Alternatively, we have used a muscle and heart-specific enhancer (MHCK) to drive the expression of U7Dtex23 cassettes delivered with AAV vectors and our results showed that MHCK improves chimeric U7snRNA expression and increases dystrophin exon 23 skipping *in vitro* and *in vivo*. However, additional U7snRNA species were produced following gene transfer, pointing at a possible limitation of the cellular processing machinery capability with saturating levels of U7 shuttles.

We have also explored the possibility of using small nucleolar (sno) RNAs as novel molecular platforms for antisense delivery. We replaced the original antisense of MBII-52 snoRNA with the Dtex23 sequence and observed low levels of exon 23 skipping in AAV-transduced myotubes. While our observation validates the approach, the efficiency of skipping is still considerably lower than with the U7snRNA cassette.

As a last approach, we engineered the human C/D box U24snoRNA to specifically target the 2'-O-ribose methylation of an adenosine branch point in a luciferase reporter pre-mRNA in order to induce the skipping of the downstream exon. We were not able to observe any modulation of splicing using this strategy.

TABLE OF CONTENTS

ABBREVIATIONS

2'OMePS	2'O-methyl Phosphorothioate
A	Adenosine
AAV	Adeno-associated virus
- ssAAV	Single stranded AAV
- scAAV	Self-complementary AAV
AON	Antisense Oligoribonucleotide
AS	Alternative splicing
BMD	Becker Muscular Dystrophy
BP	Branch point
Bp	Base pair
C	Cytosin
cDNA	Complementary DNA
CMV	cytomegalo virus
DMD	Duchenne Muscular Dystrophy
CTD	Carboxy terminal domain
DNA	Desoxiribonucleic acid
DSE	Distal sequence element
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
G	Guanosine
GFP	Green Fluorescent Protein
GRMD	Golden-Retriever muscular Dystrophy
hMW	High molecular weight
hnRNP	Heterogeneous ribonucleoprotein particle
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
ITR	Inverted terminal repeat
kb	Kilobase
kDa	Kilodalton
LNA	Locked-nucleix acid
LTR	Long terminal repeat
NMD	Non-sense mediated decay
nt	Nucleotide
PCR	Polymerase Chain Reaction
PMO	phosphoroamidate morpholino oligomer
PSE	Proximal sequence element
PTB	Polypyrimidine-tract binding protein
PTM	Pre-trans-splicing molecule
R	Purine
RNA	Ribonucleic acid
- lncRNA	Long non-coding RNA
- mRNA	Messenger RNA
- miRNA	Micro-RNA
- ncRNA	Non-coding RNA
- rRNA	Ribosomal RNA
- snRNA	Small nuclear RNA
- snoRNA	Small nucleolar RNA
- tRNA	Tranfer RNA
RNAi	RNA interference
RNAP	RNA polymerase
RRM	RNA recognition motif
RT-PCR	Reverse transcription polymerase chain reaction
RQ-PCR	Real-time quantitative polymerase chain reaction
RSV	Roux sarcomma virus
SA	Splice acceptor
SD	Splice donnor
SMA	Spinal Muscular Atrophy
SMART	Spliceosome-mediated RNA trans-splicing
SMN	Survival motor neuron
smOPT	Optimised Sm site
snRNP	Small nuclear ribonucleoparticle
SR	Serine-arginine
ss	Splice site
T	Thymidine
TBP	TATA binding protein
TMG	Tri-méthyle guanosine
U	Uridine
U2AF	U2 auxillary factor
UsnRNP	Uridin-rich small ribonucleoprotein
vg	Viral genome
Y	Pirimidine

LIST OF FIGURES

The missing figures have been removed for copyright reasons.

INTRODUCTION

I. The centrality of RNA in cellular processes and gene regulation.

In the central dogma of molecular biology proposed by Francis Crick in 1958, RNA was depicted as an intermediate in the transfer of genetic information, suspected to occur unidirectionally from DNA to RNA to proteins [1]. The characterisation of messenger RNA in 1961 by Jacob and Monod confirmed this attribute by demonstrating that "structural genes" could form a transcript of themselves, which in turn would synthesise the protein in the cytoplasm [2]. At this time, DNA and proteins still appeared to be the dominant molecules in cell function and gene regulation. However the consideration of RNA as a central molecule started to emerge in 1968: Crick [3], Orgel [4] and Woese [5] hypothesised that RNA could have preceded DNA in being the first genetic molecule in evolution and suggested that, beyond being a simple message or transfer molecule, RNA might also act as an enzyme. As RNA dual role as both genetic material and catalytic molecule was later established [6], the notion of a "RNA world" was introduced by Gilbert in 1986 [7]. Although the idea that RNA could have been the primordial molecule is nowadays disputed, this theory importantly emphasized the centrality of RNA in cellular processes [8].

Transcription is at the core of gene regulation as the messenger RNA (mRNA) carries the genetic information from the DNA to the cell. However, the notion that a simple gene encodes a single transcript to produce a function is misleading, as it cannot account for the diversity of cellular processes. The definition of a gene must therefore be more intricate. An additional layer of regulation of the genetic information became evident with the observation of the split structure of genes: RNA splicing is critical for gene expression and more than 90% of human genes are alternatively spliced [9]. Tissue-specific alternative splicing provides diversity between cell types by modifying the information flowing from DNA to the cells depending on the cellular state or environment factors. Thus, a gene can henceforth be described as a "transcription unit": it is not defined by its primary sequence but by the array of transcribed sequences that will be joined through splicing to encode the information needed in the cell [10]. RNA as the final product of the transcription unit is therefore fundamental in gene regulation.

Furthermore, multifunctional sequences are common in eukaryotes as pervasive transcription can take place at numerous cryptic sites on one gene. A single locus can consequently produce different transcripts from overlapping reading frames to code for

proteins with completely distinct functions [11-12]. This overlapping aspect of the genetic information provides a high level of transcript diversity. Moreover, a single RNA can be bifunctional and act as both protein coding RNA and regulatory RNA [13-14]. The regulatory RNAs do not code for proteins but have a considerable importance in gene regulation. Indeed, as much as 98% of the human transcriptome could represent non-coding RNAs (ncRNA) [15]. In the RNA interference (RNAi) pathway, small double-stranded ncRNAs can act as trans-acting regulatory factors to mediate transcriptional regulation, translational silencing and mRNA degradation, or direct assembly of heterochromatin leading to silencing at the level of transcription. The newly discovered long non-coding RNAs (lncRNA) constitute another category of regulatory RNA ranging in size from 200 nt to 100 kb. Although most of them still don't have an identified function, the few characterised human lncRNA have demonstrated roles in important biological processes such as epigenetics, alternative splicing or as precursor of small RNAs [16]. Long ncRNAs can also upregulate or downregulate gene expression, as is the case of the X-inactive-specific transcript (XIST) lncRNA which drives the inactivation of the X-chromosome [17]. The non-coding RNAs therefore prove to be very important components of genetic regulation.

Adding to this vast network of regulatory RNAs, post-transcriptional cleavage can affect different kind of mature transcripts and generate small RNA fragments with their own distinct functions. The mRNAs, tRNAs or snRNAs can all be processed into shorter fragments that can act as transcriptional activators, mRNA silencers or translational inhibitors amongst other roles [18]. It is not yet clear how many of these processed species are functionally important, but nonetheless RNA fragmentation further expands the diversity of transcripts available in the cells to regulate biological functions.

This overall level of transcriptome diversity and intricacy, associated with the vast array of regulatory functions that transcripts can perform and their ability to be multifunctional, indisputably establish RNA as a fundamental and central molecule in gene regulation. A gene can no longer be considered as a single entity but as a part of a highly complicated network of regulatory molecules and interactions, in which RNAs play a critical role.

RNA and diseases

Considering the central position of RNA in gene regulation, it is not surprising that mutations affecting production, processing or function of either protein-coding or non-coding RNAs can be deleterious and trigger disease. Numerous RNA-associated defects have been linked to human disorders in literature [19]. As pre-mRNA splicing is a highly complex mechanism involving numerous partners, it is probably the most affected process in RNA-related diseases. Mutations in core components of the splicing machinery occur occasionally, for example a form of retinitis pigmentosa is caused by a mutation in the splicing factor PRPF31 [20]. However, pre-mRNA alterations affecting the alternative splicing process have been more frequently described and linked to human diseases [21]. Point mutations can abolish important splice sites or create new splice signals on the pre-mRNA. These missplicing events, which will be discussed further in this thesis, usually result in the assembly of incorrectly processed mRNAs and the production of defective proteins, ultimately leading to a disease. In other cases, a mutation can trigger a change in the ratio of protein isoforms produced by alternative splicing of a gene. The proportion of each isoform is spatially and temporally regulated in cells and a change in this ratio can be deleterious and lead to diseases such as Tauopathies [22]. Numerous alterations of alternative splicing are also associated with cancers, although it is not clear whether these changes are the cause or the consequence of cancer development [23].

Besides pre-mRNA splicing defects, there are many well-established examples of diseases arising from alterations of non-coding RNAs [24]. The micro-RNAs (miRNAs), components of the RNAi pathway, are estimated to regulate approximately 50% of all mRNAs in vertebrates. Their disruption can therefore be highly deleterious, and epigenetic or genetic defects affecting their production, processing or interaction with target mRNAs have been associated with various disorders. The miRNAs post-transcriptionally down-regulate gene expression by binding to complementary sequences in the 3' UTRs of their target mRNAs. It has been shown that mutations in 3' UTRs of mRNAs can lead to dysregulation of gene expression by creating aberrant miRNA target sites, such as in the myostatin mRNA in sheep [25], or by disrupting important miRNA target sites, such as in the human SLITRK1 mRNA where a mutation in 3'UTR is associated with Tourette syndrome [26]. Furthermore, a general dysregulation of miRNAs expression and function is observed in cancer [27], and specific miRNAs

defects have been observed in neurological and cardiovascular disorders and have even been related to deafness. Defects in other non-coding RNAs can also contribute to the onset of diseases: the long non-coding RNAs display aberrant regulation in a variety of diseases including cancer [16], and defects in specific snoRNAs expression have been linked to imprinted disorders such as Prader–Willi syndrome [28].

RNA and therapy

To counter this type of RNA-based disorders, therapeutic intervention at the transcript level is appealing as disrupted features of RNAs provide potentially more accessible targets than the defective proteins they encode. Different strategies using oligonucleotides that target RNAs through specific base pairing are suitable to modulate RNA expression or function and prevent the onset of disorders [29]. The small interfering RNAs (siRNA), part of the RNAi pathway, and the antisense oligonucleotides trigger an enzymatic degradation of the target RNA leading to gene silencing. This technique is useful to inhibit disease-associated miRNA expression or to reduce the level of abnormally upregulated mRNAs. Another RNA-based technology consists of using steric-blocking oligonucleotides to prevent the access of the cellular machinery to specific sites on pre-mRNAs and mRNAs. This strategy presents the advantage of allowing the correction of a defective RNA without inducing its degradation. It can be used to inhibit mRNA translation or to interfere with RNA folding to reduce the expression of a gene, but also enables to modulate pre-mRNA splicing to correct a mutated transcript and restore the production of a protein.

It can therefore be valuable to use RNA as a therapeutic molecule for intervention at the transcript level. However, these strategies still present some challenges: the cellular uptake of oligonucleotides is generally poor and systemic delivery is relatively inefficient. Chemical modifications can be used to stabilise the molecules and improve cell targeting and penetration, but can impair the performance of the oligonucleotide in some cases. Administration of exogenous oligonucleotides can also have a toxic effect and trigger strong immunostimulatory responses. The potential of RNA-based therapeutics to treat otherwise virtually unreachable defects is therefore very encouraging but still faces the need for optimisations.

II. Splicing and associated diseases

As mentioned earlier, gene expression is largely regulated by different molecular machines that direct messenger RNA maturation, transport and translation. One important step is the RNA splicing process where the intervening sequences, or introns, are precisely removed from the mRNA precursor to keep only the coding sequences, or exons, in the mature messenger. Different types of splicing exist depending on the species and on the class of introns to be spliced. Four major classes of introns can be distinguished [30-32]: the autocatalytic group I and group II introns, the tRNA and/or archaeal introns, and the spliceosomal introns. The latter constitute the most represented class of introns and are observed in the majority of mammalian precursor mRNAs. We will focus on the splicing mechanism of this type of introns hereafter.

Pre-mRNA splicing is a highly regulated mechanism and, together with 5' capping and 3' polyadenylation, constitutes an important post-transcriptional RNA processing step. The relevance of this mechanism is considerable in humans as more than 90% of the pre-mRNA is removed as introns after processing, and only 10% of the precursor sequences are joined in the mature mRNA by splicing [21]. The complex splicing mechanism can be altered by many factors: the strength of splice signals, the presence of enhancer or silencer sequences and the concentration of important splicing factors directly affect the selection of sequences to be removed. Other factors such as the pre-mRNA secondary structure, the transcription rate, the coupling between splicing and other steps of gene expression and external stimuli can also have an influence on the outcome of splicing. Depending on the cellular context, different mRNAs can therefore be produced from the same precursor through alternative splicing and at least 90% of human genes have alternatively spliced variants [9]. To decide on a pattern of alternative splicing the spliceosome machinery must therefore take into account all the variables and a defect in any of the factors or signals involved in pre-mRNA splicing regulation is therefore highly deleterious and result in splicing diseases [33].

II. 1. Pre-mRNA splicing: principal actors and mechanism

II. 1. 1. Invariant splice sites

The most challenging issue during the splicing process is the selection of splice sites. Most introns are primarily defined by the GU and AG dinucleotides at both ends which, combined with other precise signals spread on the mRNA precursor, recruit the spliceosome and define the regions to splice. The major splice sites are invariant and are composed of the 5' or donor splice site (SD), the branch point (BP), the polypyrimidine (poly-Y) tract, and the 3' or acceptor splice site (SA).

In vertebrate, a consensus for these splice sites is difficult to predict as their sequences are highly degenerate. The general consensus sequence for the 5' splice site (5'ss) is AG|**GU**RAGU, where R is a purine. For the 3' splice site (3'ss) the vertebrate consensus is (Y)nNC**AG**|GN, where Y is a pyrimidine and N is any nucleotide. The characters in bold represent the invariant GU and AG dinucleotides at both ends, and | represents the exon-intron junctions.

The polypyrimidine tract is an intronic sequence encompassing the region between the branch point sequence and the 3' splice site, and represented by (Y)n in the 3' splice site consensus. The Y-string occur very close to the 3'ss end and can range from 2 to 32 nucleotides in length, with an average stretch of 9 pyrimidines [35]. Finally, a vertebrate branch site consensus can be defined by CUR**A**Y, with R being a purine and Y a pyrimidine. The bold adenosine constitutes the branch point itself. Most branch points are located 15–30 bp upstream of the 3'ss end, the average distance being 26 nt [35]. These major splice sites constitute strong signals that are detected by the splicing machinery and are very important to delimitate the regions to be spliced. It is therefore not surprising that variation in their composition or location have an influence on the outcome of splicing: the proximity of splice sites to each other or to the branch point [36-37] as well as the length and composition of the polypyrimidine tracts [38] can affect the splicing process.

II. 1. 2. spliceosome assembly and mechanism

The splicing mechanism takes place in the nucleus and is catalyzed by a macromolecular complex, the spliceosome, consisting of more than 200 proteins and 5 uridine-rich small nuclear ribonucleoprotein particles (snRNPs) assembled in an intricate network of RNA–RNA, RNA–protein and protein–protein interactions. Most mammalian introns are U2-type introns, which are processed by the major spliceosome. In this spliceosome, each snRNP is composed of a small nuclear RNA (snRNAs U1, U2, U4, U5 or U6), a complex of Sm or Sm-like proteins and other specific proteins [39]. Another class of introns, the U12-type introns, presents AT/AC dinucleotides at the splice sites instead of the canonical GT/AG splice site consensus. This rare class of introns is spliced by the minor spliceosome composed of U11, U12, U4atac, U6atac, and U5 snRNPs [40-41]. This type of introns can only be observed at a frequency of 0,1% in few eukaryotes and will not be detailed further here.

In vertebrate pre-mRNAs the exons are usually short (50-300 nt) and the introns are in general much larger, reaching sometimes more than one hundred thousand nucleotides. It has been demonstrated that in transcripts with such large introns the splicing process follows the "Exon Definition model", where the first recognition unit of splicing is an exon [42] (**Figure 2**). In this model, splicing factors bind to the 3' end of an intron during early assembly and initiate a search for a downstream 5' splice site across the exon. If this downstream exon is less than 300 nucleotides long, the factors binding the 3' and 5' splice sites interact and communicate across the exon and define it as the initial unit of spliceosome assembly.

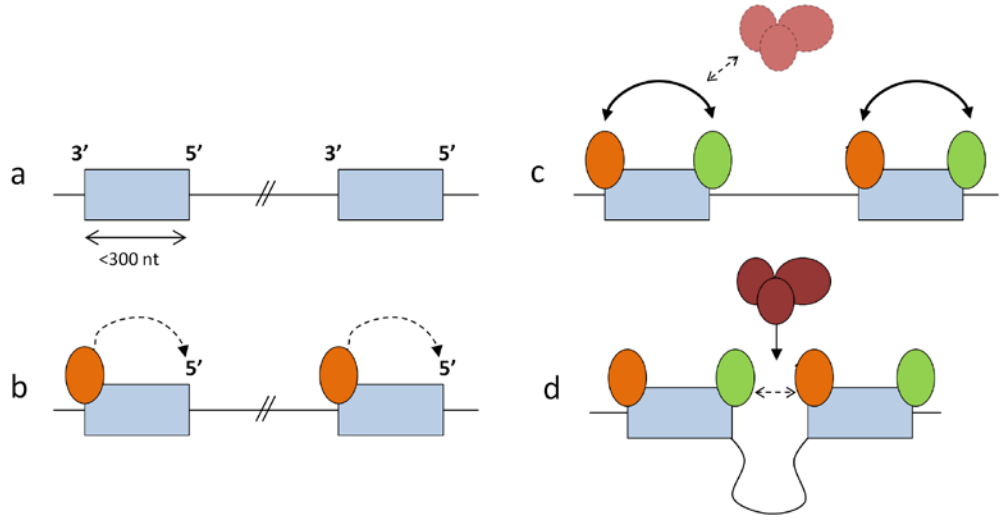

Figure 1: exon definition model

In vertebrate pre-mRNAs containing exons shorter than 300 nt and large introns (a), each exon is recognized as an independent unit prior to pairing of two exons across an intron, following the exon definition model: (b) the first splicing factor binds to the 3' end of the intron and searches for a 5' ss across the downstream exon. (c) Factors on the 3' and 5' splice sites within 300 nt of each other then communicate across the exon to define it as the first unit of splicing. (d) Intron definition then takes place, where two exons across the long intron can communicate and initiate spliceosome assembly.

The spliceosome is built sequentially on the transcript by assembly of complexes E, A, B and C on specific splice signals on the mRNA precursor [43] The process begins with the simultaneous binding of U1snRNP to the 5' splice site, of splicing factor 1 (SF1) to the branch site, and the recognition of the 3’ splice site by the U2 snRNP auxiliary factor (U2AF). The association of U2AF with the 3'ss occurs through the binding of its subunits U2AF65 and U2AF35 to the pyrimidine tract and the 3' AG dinucleotide, respectively. This protein arrangement, called the early (E) complex, triggers the recruitment of U2 snRNP to the branch point to form the pre-spliceosome complex A.

Subsequently, a pre-assembled tri-snRNP containing U4, U5 and U6 associates with the developing spliceosome and triggers several rearrangements of the spliceosome: U4 is dissociated from the complex, U6 replaces U1 at the 5' splice site, and the newly formed U6–U2 interaction brings the 5'-splice site and the branch point close together (**Figure 3**). These structural and compositional rearrangements enable the formation of a catalytically activated spliceosome, which is necessary to support the two trans-esterification steps of the splicing reaction.

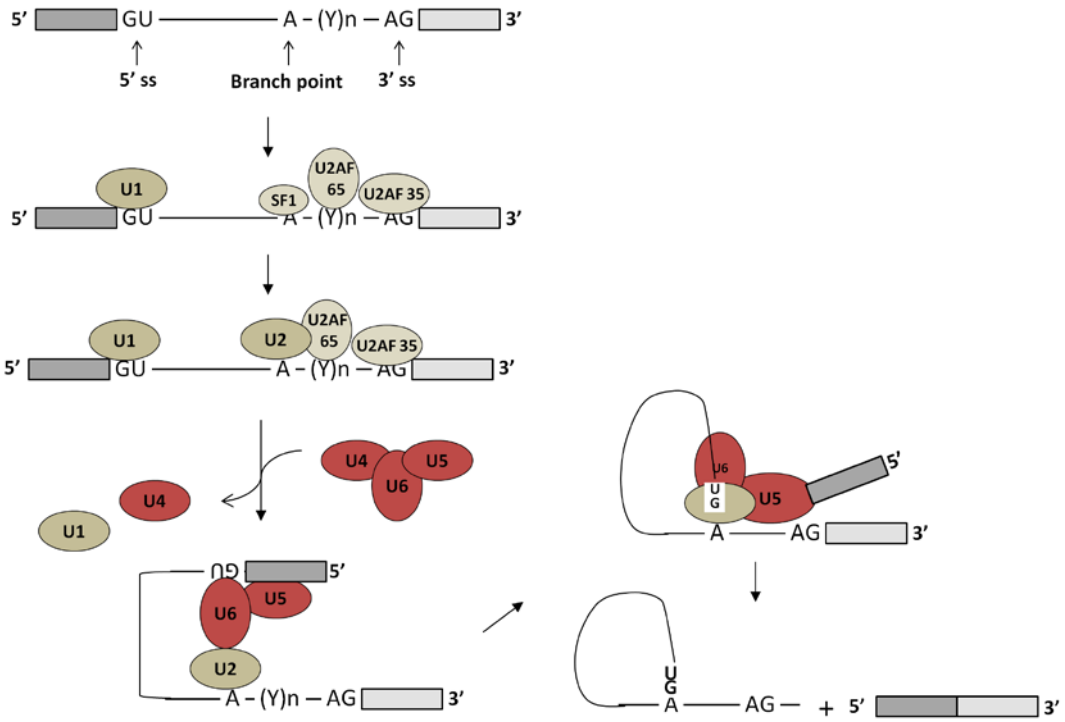

Figure 2: spliceosome assembly and mechanism of action

Representation of the sequential assembly of splicing factors on the pre-mRNA and of the transesterification reaction during the splicing process. U1 and U2 snRNPs first assemble on the 5' and 3'ss respectively and trigger the recruitment of the U4-U5-U6 tri-snRNP, which induces rearrangements of the spliceosome. The 5'ss and the adenosine branch point thereby brought close together can then initiate the transesterification reactions, leading to the release of the spliced mature mRNA and the debranched intron.

Introns are removed through a two-step splicing pathway set off by the catalytic core formed by U2, U5 and U6 snRNPs [44]. Upon base-pairing between U2 snRNA and the branch point sequence a complex conformation change occurs: the adenosine nucleotide rotates around the sugar-phosphate backbone and bulges out from the branch site-U2 snRNA duplex. This conformational change allows the branch site adenosine to present its 2'-hydroxyl (2'OH) directly into the catalytic site for the first step of splicing [45]. The activated complex B initiates the first step where the 2'-hydroxyl of the branch point adenosine attacks the phosphate of the 5'-splice site. This splicing intermediate is then submitted to a second spliceosome rearrangement, giving rise to the complex C which catalyses the second trans-esterification step: the free 3'-hydroxyl of the cleaved 5' exon attacks the 5' phosphate of the second, downstream exon. This ultimately leads to the excision of the intron and the ligation of the exons. The spliceosome is disassembled with the release of the final products of the splicing reaction: the spliced mRNA and the lariat intron.

The importance of the conserved small RNAs in the splicing mechanism is crucial and there is growing evidence that some of them, in addition to binding the splice sites, also play a role in catalysis of the splicing reactions [47]. Two important catalytic domains where identified in U6 snRNA, where mutations trigger an arrest of the splicing reaction. The catalytic potential of spliceosomal snRNAs was further demonstrated as a protein-free U6–U2 complex was able to catalyse a covalent linkage with a short transcript containing the branch site consensus in vitro, without any help of proteins [48-49]. Although the exact nature of the spliceosome catalytic center is not yet clearly elucidated, the implication of snRNAs is becoming more and more apparent.

II. 1. 3. Secondary splice signals

The major splice signals are highly degenerate and contribute only in part to the specificity of splicing: they are required for the recognition of exons by the spliceosome but contain only about half of the information needed to accurately define exon/intron boundaries [50]. Despite this concern, eukaryotic pre-mRNA splicing occurs with very high fidelity suggesting the widespread participation of additional sequences for accurate splicing. Numerous cis-regulatory elements that play important roles in the selection of splice sites have been identified [51]. These secondary splice signals are short and conserved sequences widely distributed within the pre-mRNA, in exons or in introns, and serve as either splicing enhancers or silencers. They are conventionally classified as exonic splicing enhancers (ESEs) or silencers (ESSs) and intronic splicing enhancers (ISEs) or silencers (ISSs). The secondary splice sites function by recruiting trans-acting regulators that activate or suppress splice site recognition or spliceosome assembly. Two major classes of splicing regulators are involved: the serine-arginine-rich (SR) proteins usually bind enhancer sequences and promote the recruitment of snRNPs to the spliceosome to favour splicing, while heterogeneous nuclear ribonucleoproteins (hnRNP) bind to silencer elements and inhibit the recognition of alternative exons. These proteins, through their RNA-binding domains, bind with low specificity to accessible pre-mRNA sequences that are preferentially in a single-stranded conformation [52], and use their protein interaction domains to bind to other proteins and overcome this weak specificity. The splicing regulatory elements lack precise consensus sequences and are therefore difficult to identify by simple analysis of their primary sequences. Different bioinformatic or experimental approaches have been used to identify potential ESEs and ESSs in exons [53-56], and a limited number of ISE and ISS elements have been identified in introns [57-58].

The potential for exon sequences to have an effect on splice site selection was first demonstrated 25 years ago [59], and it is nowadays widely recognised that many if not all exons contain splicing-regulatory sequences [51, 60]. The exonic splicing enhancers are the most characterised of these elements and most of them have a high purine content, although some non purine-rich sequences, such as the A/C-rich ESEs, can also act as enhancers [53, 61]. Most ESEs function by recruiting members of the serine-arginine-rich (SR) protein family, which bind the exonic enhancers through their N-terminal RNA recognition motifs (RRM domain) and facilitate spliceosome assembly by

mediating protein-protein interactions through their C-terminal RS domains [62-63]. SR proteins that are bound to ESEs can also promote exon identification by antagonizing the action of silencer elements bound to nearby ESSs [64] (**Figure 5**).

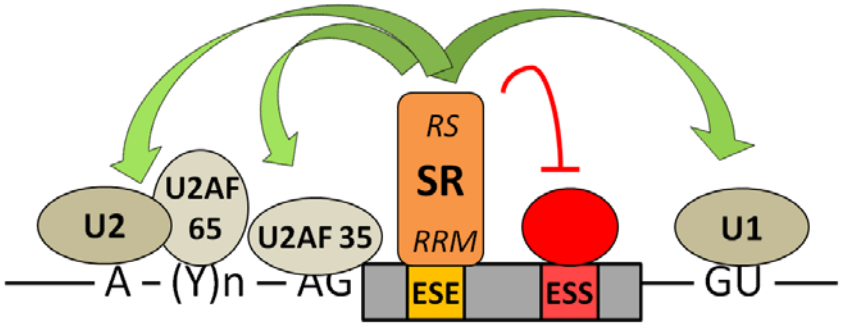

Figure 3: mechanisms of splicing activation

The exonic splicing enhancer (ESE) sequence can recruit activators of splicing such as the SR proteins that bind to the RNA through their RNA recognition motifs (RRM) and can promote spliceosome assembly as well as block the action of adjacent silencers that are bound to nearby exonic splicing silencer sequences (ESS).

Exonic splicing silencers (ESSs) are less well described than ESEs, although several of them have been characterised [64-65]. Their mechanisms of action involve the recruitment of splicing repressors of the heterogeneous nuclear ribonucleoprotein (hnRNP) class [66-67]. HnRNP proteins bind to silencer elements via their RNA-binding domains (RRMs) and inhibit the splicing reaction by different mechanisms [34]: they can be in direct competition with splicing stimulatory factors by overlapping enhancer-binding sites and preventing essential interactions between ESEs and snRNPs (**Figure 6a**). HnRNP proteins can also bind to either side of an exon and inhibit splicing by "looping out" the portion of pre-mRNA to be silenced, making it inaccessible for splicing [68] (**Figure 6b**). A hnRNP, by binding to a strong inhibitory element, can also trigger the localised recruitment of more hnRNPs, which polymerise into complexes to coat the pre-mRNA and prevent spliceosome assembly [69] (**Figure 6c**).

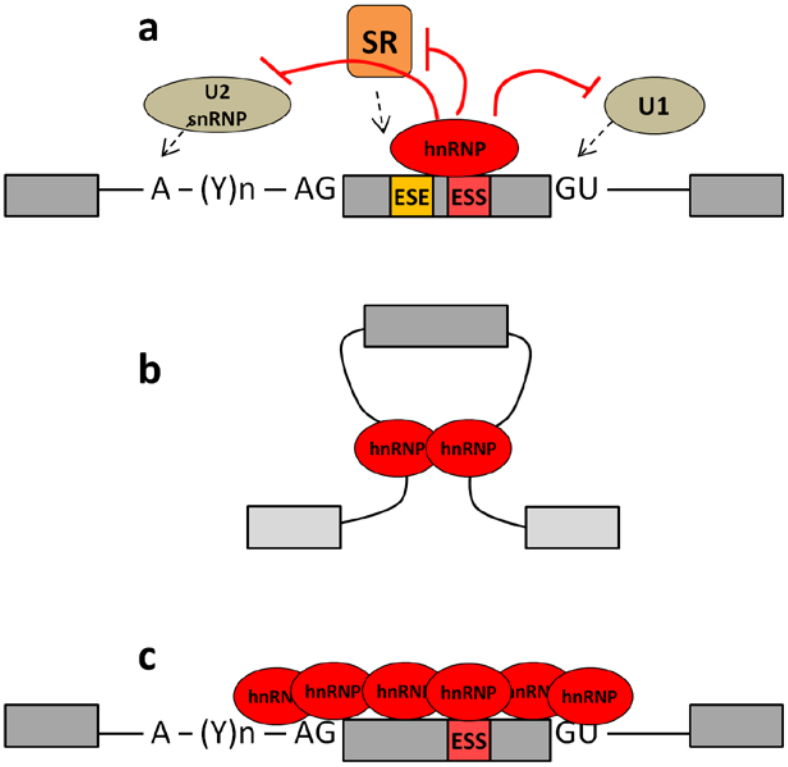

Figure 4: mechanisms of splicing repression by hnRNP proteins

a) hnRNPs can overlap ESE sequences when binding to an ESS and prevent the binding of splicing activating factors
b) hnRNPs can loop out the portion of pre-mRNA to be silenced, making it inaccessible for splicing
c) hnRNPs can polymerise into complexes to coat the pre-mRNA and prevent spliceosome assembly.

The intronic regulatory elements are less well characterised. Their highly degenerate sequences make them difficult to detect and it is likely that many more intronic elements remain to be identified. A limited number of ISE and ISS elements have however been documented. A well characterized ISE is the G triplet (GGG) that can promote recognition of 5'ss or 3'ss [70-71], and it has been shown that intronic CA repeats can also constitute positive or negative regulatory elements of splicing [72].

The overall contribution of secondary sequences in splicing regulation has therefore been widely demonstrated but their mechanism of action is very complex and still not fully comprehended. Indeed, the various effects of these regulatory elements cannot be described too simplistically as it has been shown that their activities could depend on their relative locations in pre-mRNAs: for example the G triplets mentioned before,

known to enhance splicing from intronic locations, can also act as splicing silencers when located in exons [73]. Moreover, the effect of the regulatory sequences seems to be manifested only in the presence of relatively weak splicing signals as it has been shown that an exonic element can be easily overcome by strong flanking splicing elements [74]. Enhancer and silencer sequences can be adjacent or even overlapping on the pre-mRNA and, consequently, a competition between trans-acting factors for binding these regulatory sequences and for interacting with the spliceosome occurs [75]. A definite splicing event is thus the outcome of the cooperative or antagonistic interactions between the various regulatory elements, combined with the intrinsic strength of the flanking splice sites.

II. 1. 4. Coupling between splicing regulation and other steps of gene expression

The splicing process is part of a large network of gene expression mechanisms which are often tightly connected, both physically and functionally [77]. A well-described example is the coupling between splicing events and RNA transcription. The first evidence for cotranscriptional splicing in vertebrate has been demonstrated in the human dystrophin gene in 1995 [78], and nowadays many introns are thought to be spliced cotranscriptionally [79-81]. Splicing factors can have extensive physical interactions with the core transcription machinery, as has been shown for SR proteins that associate with the C-terminal domain of RNA polymerase II [82-84], and the transcription elongation rate can directly influence the splicing pattern of a pre-mRNA. Conversely, splicing factors that bind to the nascent messenger RNA can stimulate transcription by promoting polymerase elongation [85]. These two processes are therefore tightly coupled and can be reciprocally influenced [86].

There are also possible connections between splicing and polyadenylation processes [87]. For the definition of the 3'-terminal exon at the end of transcription, splicing factors on the upstream 3'splice site communicate across the terminal exon with the polyadenylation machinery on the downstream poly(A) signal. The result is a mutual enhancement of both processes, where the polyadenylation apparatus assists in the last exon definition and enables last intron splicing, whilst the splicing machinery also helps the polyadenylation activity [88-89].

II. 2. Alternative splicing

In higher eukaryotes there is an important inconsistency between gene number and complexity of the organism, implying that the number of proteins must be much higher than the number of genes. Indeed, in humans up to ten times more functional proteins than expected could be produced from the 30 000 genes in the human genome [76]. The mechanism most widely used to expand protein diversity is alternative splicing. Alternative pre-mRNA splicing (AS) is a central mode of genetic regulation in higher eukaryotes which enables to increase the number of proteins that can be synthesized from each gene by modulating the inclusion or exclusion of a portion of coding sequence in the mRNA. A single pre-mRNA can have multiple splicing patterns that generate numerous mature mRNAs, ultimately giving rise to functionally distinct protein isoforms that differ in their peptide sequence, enzymatic activity, ligand binding, subcellular localization, stability, post-translational modification, and protein folding [90]. This source of protein diversity is important in many cellular and developmental processes and is often regulated in a tissue-specific manner. A recent estimate showed that more than 90% of human transcripts are alternatively spliced [9]. However, the functional significance of most of these numerous AS events is still unknown and only a small proportion of them have characterised biological roles [91].

There are different patterns of alternative splicing described in nature [92] (**Figure 8**). A single alternative exon situated between two constitutive exons can be either skipped or included in the mRNA. Alternatively, multiple cassette exons can reside between two constitutive exons and be included in a mutually exclusive manner so that only one exon in the group is included at a time. Exons can also have multiple 5' or 3' splice sites that are alternatively used to produce an extended or shortened exon. Additionally, the use of alternative promoters or alternative polyadenylation sites can alter the selection of the 5' or 3' most exon and modify the length and composition of the transcript. Introns, which are usually removed from the mature messenger, can be retained in the mRNA and become translated. All these different patterns of alternative splicing are not exclusive and several can be combined in a single transcript to produce a vast number of distinct mRNAs.

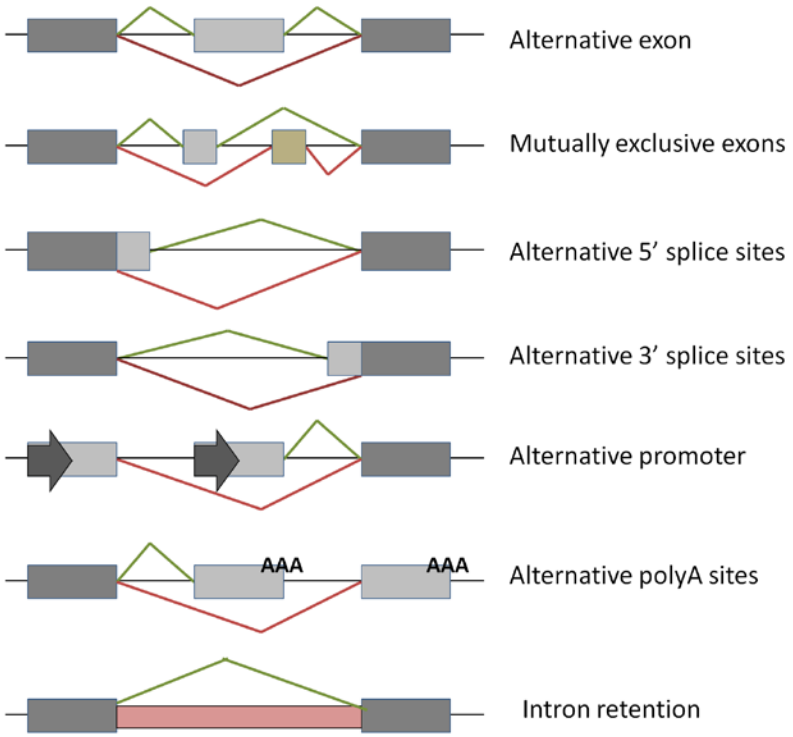

Figure 5: patterns of alternative splicing

Representation of the different models of alternative splicing that can give rise to various mature mRNAs producing functionally distinct protein isoforms from the same precursor transcript. The green lines represent normal splicing patterns and the red lines depict alternative patterns.

- Mechanism and regulation of AS

Alternative splicing is predominantly regulated by the combinatorial recruitment of trans-acting splicing factors to specific RNA motifs on the pre-mRNA, the regulatory enhancers or silencers sequences mentioned earlier [93]. The auxiliary splice sites are highly degenerate but are important in defining constitutive and alternative exons. The decision to include or skip an exon in an mRNA primarily depends on the strength of the flanking splice signals. If the major splice sites flanking an exon are weak, the relative enrichment of this exon in ESE/Ss or ISE/Ss enables to define its frequency of inclusion. The exonic elements have been reported to mostly play roles in constitutive splicing but many have also been identified in tissue specific or developmentally regulated exons, which usually have weak splice sites and require the ESE for inclusion [53]. On the other hand, intronic regulatory elements seem to play a critical role in alternative splicing regulation as the intronic regions surrounding alternative exons are

more conserved in mammals than those surrounding constitutive exons. The activity of these regulatory elements depends on the availability of the associated splicing regulatory proteins, which can be expressed at different levels in different tissues. During alternative splicing, a particular splice signal can therefore be selected or ignored depending on the cellular state and the combination of splicing factors, and this phenomenon is probably responsible for most of the tissue-specific splicing.

Although it is clear that the splice signals contribute largely to the regulation of alternative splicing, this mechanism is a multifactorial process and other cellular events are implicated [94]. As pointed out earlier, most introns are thought to be spliced cotranscriptionally and the splicing machinery is physically linked to the transcriptional apparatus. A significant role for the transcription machinery in the regulation of alternative splicing is therefore emerging. Various studies demonstrated that alternative splicing is sensitive to RNA polymerase II elongation rate: pausing or slowing down the polymerase increases the time an exon can recruit the splicing machinery and favours exon inclusion [81]. One can therefore hypothesised that some splicing regulatory sequences could act as transcription pause sites to enable the inclusion of a weak exon [51]. Possible connections have also been suggested between alternative splicing and polyadenylation. Patterns of alternative splicing and alternative cleavage and polyadenylation are highly correlated across tissues, suggesting that some common tissue-specific RNA-binding factors could coordinate the regulation of both processes [9].

The chromatin can also act as a regulator of alternative splicing. In regions dense with nucleosomes, such as compacted chromatin, the elongation rate of the transcription is slower than in open chromatin regions. Interestingly, alternative exons that are not always included are less enriched in nucleosomes than constitutive exons [95], which suggests a link between chromatin structure and alternative splicing regulation. It can be speculated that a high density of nucleosomes in an exon promotes the pausing of the RNA polymerase II and enables the recruitment of the spliceosome. Histone modifications and DNA methylation may also contribute to splicing regulation. Different histone modification signatures have been shown to correlate with particular splicing patterns in several human genes, and an alteration of these signatures can induce a splice site switch [96-97]. For example for splicing events dependent on the splicing factor PTB, it has been shown that histone marks can promote the recruitment

of trans-factors to the pre-mRNA via formation of a chromatin-splicing adaptor complex. The chromatin-binding adaptor protein MRG15 acts as a new regulator of alternative splicing by recognizing the histone modification H3-K36me3 and recruiting the splicing repressor PTB to the pre-mRNA, inducing forced exclusion [98].

In addition, the secondary structure of the pre-mRNA can also influence splice site selection by masking or revealing particular splice signals. The formation of stem-loop structures between flanking sequences of an alternatively spliced exon enables to regulate the usage of the looped-out exon and lead to its skipping [99-100].

Alternative splicing events can also be regulated by cell signalling and external stimuli, such as growth factors, cytokines, hormones, and stress [101]. The activation of cellular signal transduction pathways can trigger the phosphorylation of trans-acting splicing factors, which induces a change in their activity, expression and interaction with the pre-mRNA. The modulation of these kinase cascades can alter splice site selection and therefore influence alternative exon usage [102-103].

A role for non-coding RNAs as novel regulators of alternative splicing is also emerging. It has been shown that microRNAs can regulate the expression of important splicing factors during development and differentiation. In early myoblasts, the splicing factors PTB and nPTB repress the splicing of some muscle-specific exons. When the differentiation occurs, the muscle-restricted microRNA miR-133 is upregulated and silences the nPTB repressor, thus allowing the inclusion of PTB-repressed exons [104]. Another example of ncRNAs implicated in alternative splicing is the brain-specific small nucleolar RNA HBII-52 that can bind by sequence complementarity to an exonic splice silencer on the serotonin receptor 5-HT2C pre-mRNA. This direct sequence-specific association seems to prevent the binding of a splicing repressor on the transcript and favours the inclusion of the alternatively spliced exon [105-106].

It is therefore evident that alternative pre-mRNA splicing is a very complex and tightly regulated mechanism that controls important developmental and tissue-specific programs and that can be modified by numerous factors. The downside of this overall complexity, which is necessary for proteome diversity in complex organisms, is that it also increases the potential for errors. Various mis-splicing events can arise from defects in regulatory factors or splice site selection, which can have dramatic effects and lead to numerous genetic diseases.

II. 3. Diseases caused by splicing defect

Considering the wide array of factors and sequences playing crucial roles in regulation of pre-mRNA splicing, it is important to consider that variations occurring virtually anywhere within a gene could potentially alter or impair the splicing process [107]. Simple point mutations in the pre-mRNA can inactivate splice signals or create new cryptic ones, and even a silent mutation can result in splicing dysfunction and have important pathological effects [33, 108]. More than 15% of point mutations inducing a genetic disease affect pre-mRNA splicing [109] and the mRNAs then produced are often out of frame and contain premature stop codons. These incorrectly spliced transcripts are subjected to two different outcomes: their translation can still take place and result in the production of aberrant or truncated proteins, some of which would not be functional but others displaying dominant functions that can negatively affect the cells. Another fate for mis-spliced mRNAs is their direct degradation in nonsense-mediated decay (NMD) [110]. This mRNA surveillance system in eukaryotes is able to detect premature stop codons in newly spliced mRNAs and to induce their rapid degradation [111].

Different types of pre-mRNA mutations affect the splicing process and can cause disease and they can typically be classified in two groups: mutations that affect the canonical splice signals, or mutations that affect other important regulatory elements, either in *cis* or in *trans*.

II. 3. 1. Mutations in the consensus splice signals

It has been estimated that 60% of all the splicing-associated mutations occur in the invariant splice site sequences [112]. Mutations that destroy natural donor splice site (5' ss), acceptor splice site (3' ss) or branch point abolish exon recognition and usually result in complete exon skipping [109]. The resulting mRNAs often contain premature stop codons, which prevent the synthesis of the corresponding protein and therefore trigger a disease (**Figure 10a**). This is the case for a substitution in the 5' splice site of exon 20 of I-kappa-B kinase complex associated protein (IKBKAP) gene that induces exon 20 skipping and results in familial dysautonomia [113].

However, in some cases the inactivation of a consensus splice signal leads to other mis-splicing events such as the activation of cryptic splice sites, the retention of an intron or the inclusion of a pseudoexon. The occurrence of these alternative events depends on the local context of the mutation as similar mutations in other contexts can cause exon skipping.

The activation of cryptic splice sites is after exon skipping the most frequent outcome that follows the inactivation of strong natural splice site [114]. The use of cryptic sites, which are generally weak signals, often leads to the exclusion of an exon or the inclusion of an intron, in part or in totality, ultimately inducing a frameshift in the protein coding sequence (**Figure 10b**). A well known example occurs in the β-globin gene after mutation of consensus donor or acceptor splice sites and is associated with forms of β-thalassemia [115-116]. Alternatively, a mutation of a donor splice site can induce aberrant intron retention in the transcript, which has been associated with numerous mis-splicing diseases such as Leigh syndrome [117] (**Figure 10c**). Consensus splice site mutations can also result in the inclusion of pseudoexons, which are intronic sequences that are flanked by consensus splicing signals but are not recognized by the splicing machinery and are normally excluded from the mRNA [118]. The splice sites surrounding a pseudoexon are generally weak but various mutations can either increase the strength of these sites or instead weaken the invariant splice sites of nearby exons, increasing the probability of pseudoexon inclusion in the processed mRNA (**Figure 10d**). Pseudoexon inclusions often generate premature termination codons and result in human diseases. This is the case in some forms of Becker muscular dystrophy where a mutation strengthening a potential 3' splice site of an intronic "false exon" in the Duchenne muscular dystrophy (DMD) gene triggers the aberrant inclusion of this pseudoexon in the mature transcript, which becomes out-of-frame [119].

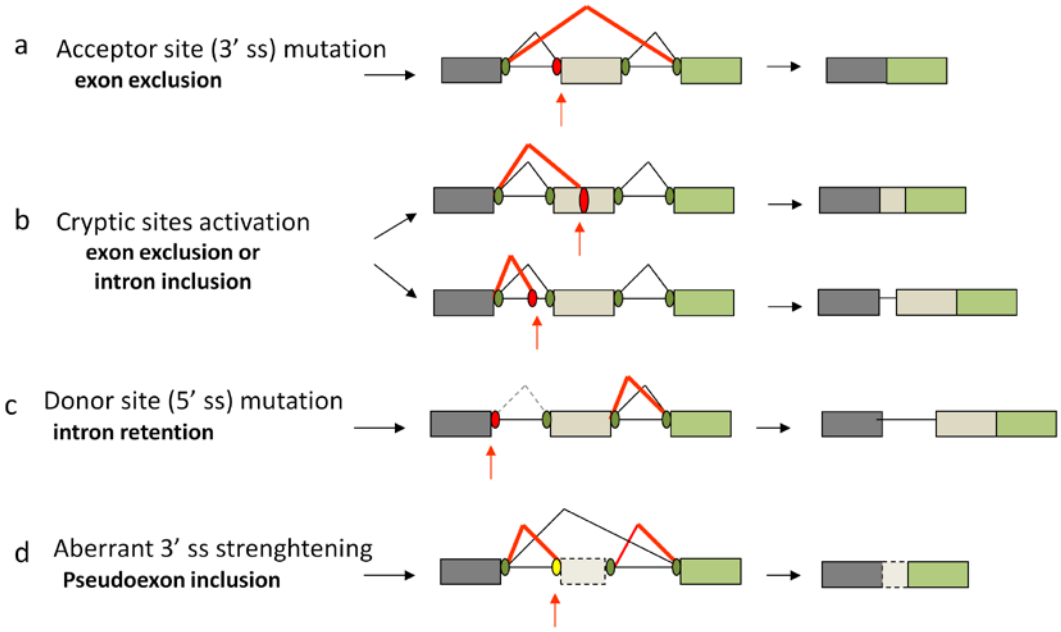

Figure 6: mutations affecting the consensus splice signals and consequences on the mature transcript

(a) A mutation of the acceptor splice site at the 3' end of an intron can lead to the exclusion of the downstream exon in the mature mRNA;
(b) The creation of intronic or exonic cryptic splice sites in the premature transcript can induce the inclusion of an intron or the exclusion of an exon in the mRNA, in part or in totality;
(c) The inactivation of a donor splice signal leads to the inclusion of intronic sequences in the transcript;
(d) A mutation in an intronic sequence with pre-existing weak splice signals can alter one of these sites into consensus signals, leading to the recognition of the intronic sequence as pseudoexon and to its inclusion into the mRNA.
These various events often lead to the creation of premature stop codons in the mature mRNA, which prevent the synthesis of the corresponding protein.

II. 3. 2. Mutations in the regulatory elements of introns and exons

The second type of pre-mRNA mutations that affect the splicing process targets the variable intronic or exonic splicing motifs. These mutations can lead to aberrantly spliced transcripts by weakening or strengthening splicing recognition motifs and are associated with a large number of diseases [34].

A well-described example of mutations in the regulatory elements of exons is a point mutation found in the survival of motor neuron (SMN) gene that does not change the amino acid code but is associated with spinal muscular atrophy (SMA). In humans, there

are two copies of the SMN gene, SMN1 and SMN2, and inactivation of the SMN1 gene leads to SMA as the SMN2 isoform is unable to compensate the loss of SMN1. This is due to the fact that the SMN2 gene contains a silent C to T substitution in exon 7 that disrupts an exonic enhancer (ESE) normally recognized by the splicing activator SF2/ASF [120] and enables the activation of a silencer element (ESS) recognized by the splicing suppressor hnRNP A1 [121-122]. This mutation significantly alters the splicing pattern of the SMN2 pre-mRNA and predominantly induces the skipping of exon 7. An average of 80% of SMN2 mRNAs lack exon 7, leading to the production of a truncated protein, and the severity of SMA in patients depends on the level of residual functional SMN protein.

Some mutations can also alter the balance of alternatively spliced transcripts. For example, the microtubule-associated protein tau (MAPT) gene contains four enhancer and three silencer regions in exon 10, which is alternatively spliced in a highly regulated manner both spatially and temporally. This alternative splicing event give rise to two mRNA isoforms, containing or not the exon 10, and the ratio of the corresponding proteins is precisely controlled in the brain. Mutations in MAPT exon 10 can disrupt either a silencer or an enhancer region and therefore increase or decrease the frequency of exon 10 inclusion in the mRNA. Such mutations alter the ratio of protein isoforms generated by alternative splicing and induce various neuropathologies, called tauopathies, such as frontotemporal dementia with parkinsonism associated with chromosome 17 (FTDP-17) [22, 123].

Some diseases can also stem from a change in length of short repeat elements that can change the balance of alternative splicing regulation. These short sequence repeats, called triplet repeats, are naturally present in exons and enable to increase their recognition by certain splicing factors. The number of these repeats varies between individuals but can cause disease when they expand beyond a normal threshold. One form of Myotonic dystrophy, DM1, is caused by the extension of a CUG repeat in the DMPK gene, inducing a gain of function of the transcript. The CUG element normally binds to the Muscleblind-like1 (Mbnl1) binding protein but an extension of the repeats induces an abnormal sequestration of this protein and reduces its cellular concentration (**Figure 12**). The sequestrated protein is therefore unavailable to modulate the Mbnl1-dependent alternative splicing of a set of genes, which ultimately leads to aberrant splicing and disease [124].

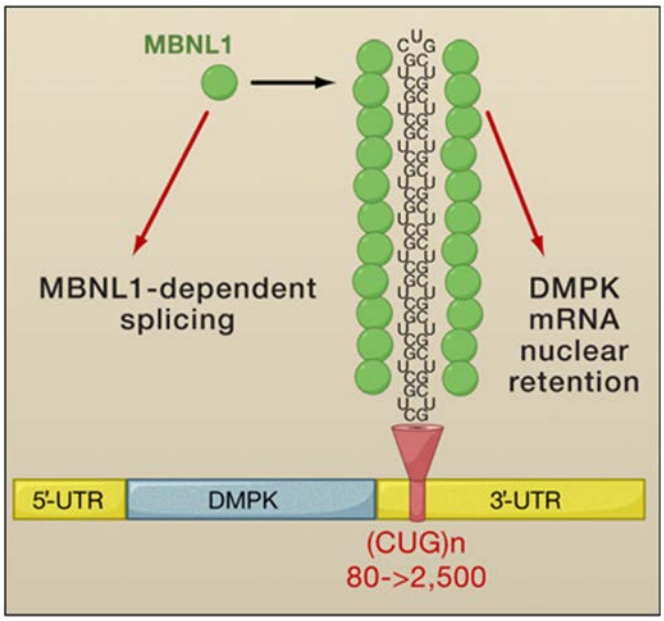

Figure 7: microsatellite expansions in DMPK gene cause MBNL1 protein loss-of-function

In myotonic dystrophy DM1, abnormal expansions of CUG repeats in DMPK transcripts induce the sequestration of MBNL1 proteins that become unavailable for splicing regulation of various transcripts.

II. 3. 3. Mutations in trans-acting splicing factors

Only few mutations in core elements of the splicing machinery have been described, even though the spliceosome is composed of a large number of proteins, which could imply that such mutations are often lethal. However, some diseases have been associated with mutations in genes encoding pre-mRNA splicing factors. For example, a severe autosomal dominant form of retinitis pigmentosa is caused by mutation of the splicing factor PRP8 [125], and Frontotemporal lobar dementias develop from the loss of TDP43, a member of the heterogeneous nuclear ribonucleoproteins (hnRNP) family of proteins [126]. Numerous studies have also demonstrated specific defects in the expression of splicing factors that are associated with cancer [23]. For example, an overexpression of the SR protein SF2/ASF in several cancers has been shown to promote cell transformation by inducing mis-splicing of different genes: SF2/ ASF controls alternative splicing of the oncogene Ron by binding to regulatory elements within exon 12, and the overexpression of SF2/ASF stimulates the skipping of

upstream exon 11 and increases the level of a tumor-associated isoform of Ron, which is sufficient to transform cells in culture [127].

Another example is the modulation of alternative splicing of the Pyruvate Kinase (PK) pre-mRNA in tumor cells through the activation of hnRNP splicing repressors. The PK gene catalyses the production of Pyruvate in the glycolysis pathway. PK produces two mRNA isoforms, PKM1 and PKM2, by alternative splicing of two mutually exclusive exons, exons 9 and 10 respectively. In differentiated non-dividing cells, exon 9 is predominantly included to produce PKM1 mRNA, which channels Pyruvate towards the Krebs cycle. Conversely, in actively replicating cells and tumor cells exon 10 is included and high levels of PKM2 species are produced, leading to the degradation of Pyruvate into Lactate, which is advantageous for rapidly replicating cells [128]. The relative amounts of PKM1 and PKM2 expressed in cells therefore determine the metabolic status of the cells, which in turn influence their state of differentiation, proliferation and oncogenic transformation. The PK pre-mRNA alternative splicing process is regulated by splicing factors interacting with splicing enhancers or inhibitors located in exon 9 and 10 [129-130]: in tumor cells the oncogenic transcription factor c-Myc upregulates the transcription of splicing repressors hnRNP A1, A2 and PTB, which then associate with exonic splicing silencers signals located in exon 9, resulting in the repression of exon 9 in favour of exon 10 inclusion and in the PKM2-induced proliferation and transformation of cancer cell lines [129-130].

Numerous other splicing factors display an aberrant expression in cancer, it is however not clear whether these changes are the cause or consequence of cancer formation and tumor progression [131]. Additionally, mutations that affect splicing of tumor suppressor genes have been reported as cancers predisposition factors. This is the case of the BRCA1 gene where a point mutation disrupting an ESE sequence induces the skipping of exon 18, thereby predisposing to breast and ovarian cancer [132].

Finally, a snoRNA (HBII-52) has been described as a novel *trans*-acting regulator of alternative splicing and shown to influence the splicing of the serotonin receptor 2C mRNA by binding to a silencing element in exon Vb [105] (**Figure 13**). Patients with the Prader–Willi syndrome have lost paternal expression of some maternally imprinted genes, including the C/D-box snoRNA HBII-52, which triggers the formation of an isoform of the serotonine receptor with diminished activity and causes the disease [28].

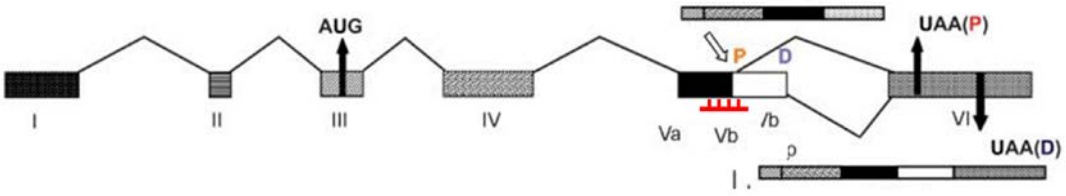

Figure 8: regulation of alternative splicing by the snoRNA HBII-52

HBII-52 snoRNA partially blocks a silencer element located in alternative exon Vb of human 5-HT2CR receptor and promote its inclusion to produce a functional protein. The snoRNA is represented in red in exon Vb.

P and D indicate the proximal and distal splice sites in exon Vb of 5-HT2CR pre-mRNA, and UAA(P) and UAA(D) represent the stop codons resulting from their usage.

Numerous defects in a wide array of factors are therefore likely to affect splicing and result in various genetic diseases. As this type of diseases present a real challenge for conventional therapies, it is becoming more and more important to find strategies that can overcome the deleterious effects of these mutations and restore or correct the processing of pre-mRNAs.

III. Strategies to modulate splicing

Knowledge of the molecular basis of splicing defects has offered the possibility of intervening on mRNA and various strategies have emerged over the years aimed at modulating pre-mRNA splicing in order to correct splicing-associated defects. These include the use of chemical substances that can act as modulators of alternative splicing and influence splice site selection [133]. The mechanisms of action of these substances are varied: kinase or phosphatase inhibitors can be used to modulate the state of phosphorylation of SR protein domains and alter SR proteins interactions, thereby modifying their action on alternative splice sites [103, 134-135]. Some other compounds act as histone deacetylases (HDAC) inhibitors and enable to maintain the chromatin in an open, transcriptionally active structure, which can allow increasing the transcription of genes such as splicing regulatory factors [136-137]. Alternatively, some substances are able to induce the readthrough of premature stop codon mutations present in disease-associated pre-mRNAs through their association with the ribosome. The readthrough of stop codons can be beneficial within an in-frame mRNA in order to maintain the production of a partly functional protein and this strategy has shown some successes for various splicing-associated diseases such as Duchenne muscular dystrophy [138-139]. However, the major downside of most of these chemical substances is their lack of specificity: their broad mechanism of action implies that genes or transcripts unrelated to the disease might also be aberrantly altered, which can have dramatic consequences.

On the other hand, other splicing modulation strategies present a high degree of specificity for the targeted pre-mRNA and seem therefore favoured and more promising for the treatment of splicing-associated diseases. Here, we detail the antisense oligoribonucleotide technology that is used for the steric blocking of splice signals in order to modulate the splicing pattern of mutated pre-mRNAs and recover mature transcripts able to produce functional proteins. The development of a targeted methylation strategy used to inactivate splice signals through the use of modified small nucleolar RNAs is also considered.

III. 1. Antisense RNA to modulate gene expression at the post-transcriptional level

The correction of aberrant pre-mRNA splicing can be achieved with the use of antisense oligoribonucleotides (AON) [140]. AONs are small synthetic antisense RNA molecules that bind to a specific target sequence by Watson-Crick complementarity. These antisense sequences can be used to modulate gene expression at the post-transcriptional level by masking important regulatory sequences in the pre-mRNA; in particular they have been extensively used to redirect and correct the splicing of a mutated pre-mRNA in order to restore the function of a defective gene [141]. AONs designed to mask key splice motifs can alter the splicing pattern of a pre-mRNA by preventing spliceosome assembly on the precursor and forcing it to recognize alternative signals. This strategy, which was first demonstrated to correct the aberrant splicing of β-globin pre-mRNA associated with β-thalassemia [142], presents the remarkable advantage of allowing therapeutic intervention on native transcripts within their normal regulatory environments. This RNA-based approach has therefore become widely used as a tool in molecular genetic studies or in therapeutic approaches to counter the deleterious effects of splicing-associated mutations [143].

III. 1. 1. Antisense RNA: design and chemistry

- Antisense oligoribonucleotides design

The design of AONs for therapeutic pre-mRNA processing modulation can prove very challenging [144]. Many factors have to be taken into account when conceiving antisense sequences against new pre-mRNA targets to ensure a strong and stable interaction between the AON and the target and effective splicing modulation [145].

First, the antisense needs to be able to reach its target for efficient base-pairing, implying that the two molecules must co-localize in the cell and that the secondary structure of the target transcript must be favourable. The accessibility of the target is a major determinant of AONs efficacy and an accurate prediction of the secondary structure of the pre-mRNA is therefore really important. Several software programs give an estimate of all possible structures of a particular RNA sequence, one of the most widely used being the mfold program [146]. AONs are generally unable to target a mRNA region

with very stable local structure, such as a stem-loop, and usually show the highest activity against targets with the least amount of secondary structure [147]. It has also been suggested that AONs that target partially closed (but still accessible) regions could be more potent than antisense sequences against unstructured, highly accessible regions. This observation implies that, in addition to the steric blocking of regulatory factors, the binding of an AONs to its target could also function by disrupting important secondary structures [148].

The second important factor to take into consideration is the thermodynamic energy of the AON-target complex, which depends on AON length, sequence composition and the free energy of local structures. To efficiently bind a target sequence, the binding energy between the AON and the target must be thermodynamically stable and higher than the free energy of the predicted secondary structures of both the target and the oligonucleotide [145]. Moreover, the energy of the AON-target mRNA must be high enough to effectively compete with the natural binding of splicing factors on the pre-mRNA. These parameters can be evaluated *in silico* with the help of various softwares, such as OligoWalk that predicts the thermodynamic affinity of complementary oligonucleotides to a structured mRNA target [149].

It is however difficult to assess and validate all these parameters when designing a new antisense molecule and it is not guaranteed that a molecule that finally meets these requirements will actually be effective to induce the desired splicing switch. To increase the chances of designing a potent AON, various computational algorithms can prove very useful. However in practice the efficiency of these molecules needs to be addressed by screening many different AONs targeting adjacent sites on a pre-mRNA in order to determine the best antisense sequence.

To add to the complexity of antisense sequences design, one needs to consider the high variability in expression of splice variants between individuals. It has indeed been shown that up to 30% of splicing events display individual-specific differences [9], and it may eventually be important to take into account the effects of genetic variation on the splicing code in order to provide efficient personalized therapy.

- Antisense oligoribonucleotides chemistry

Unmodified AONs penetrate through the cell membrane with difficulty and are rapidly targeted and degraded by nucleases before being able to act on splicing. Various chemical modifications of AONs were therefore developed, enabling to enhance nuclease resistance, increase the intracellular stability, improve target affinity and enhance cell penetration and nuclear accumulation [145].

The first generation of AONs contained a phosphorothioate (PS) modification on the backbone, which improved the resistance to nuclease activity but also decreased the affinity to the target mRNA. A second generation of antisense chemistry was therefore designed, which involved the addition of residues at the 2' position of the ribose in the oligonucleotide. 2'-O-Methyl (2'-O-Me) and 2'-O-methoxyethyl (2'-MOE) modifications provided a higher nuclease resistance to the AONs and enhanced their binding affinity for the target. These carbohydrate modifications are probably the most widely used AON chemistries for splicing modulation [150]. Finally, a later generation of AONs was fashioned to further improve the characteristics of the molecules. These developments involved various chemical modifications of the nucleotide backbone and further enhanced the affinity for target sequences and improved biostability and resistance to nucleases [151]. Amongst these new-generation AONs, the three most studied and used are phosphoroamidate morpholino oligomers (PMO) [152], peptide nucleic acids (PNA) [153], and locked nucleic acids (LNA) [154].

The steric blocking of mutations or splice signals on the premature transcript can have different applications (**Figure 15**): AONs can be used to inactivate cryptic splice sites or to block the recognition of pseudoexons, thereby forcing the alternative splicing of the pre-mRNA to recover the normal splicing pattern; they can also mask natural splice sites to induce the skipping of a mutated exon and recover an in-frame mRNA depleted of an exon, or to induce the re-inclusion of aberrantly excluded exons. Finally, a novel technology based on the use of bifunctional antisense molecules enables to trigger a trans-splicing mechanism, where a sequence exchange between a mutated target pre-mRNA and a therapeutic tailed antisense molecule occurs.

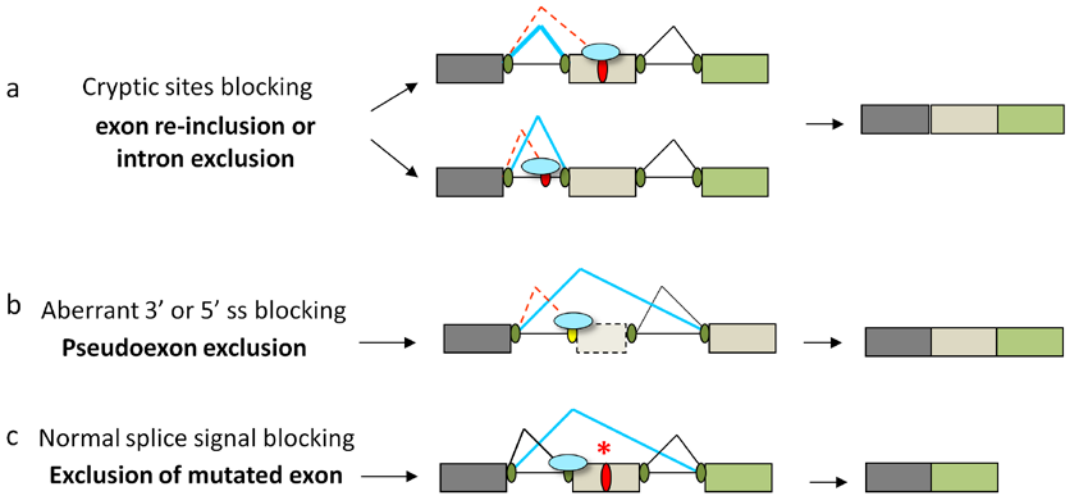

Figure 9: antisense RNA strategies to modulate splicing

Various splice signals, aberrant or normal, can be masked by AONs in order to modulate the splicing of disease-associated pre-mRNAs and restore the synthesis of functional proteins. (a) The masking of cryptic splice sites enables to recover the correct splicing pattern of the pre-mRNA; (b) blocking aberrant intronic splice signals enables to inhibit the inclusion of a pseudoexon; (c) masking a natural splice signal can promote the skipping of the downstream mutated exon.

III. 1. 2. Cryptic splice sites inactivation

Successful AON-mediated restoration of correct splicing has been demonstrated for various diseases triggered by cryptic sites activation (**Figure 15a**). A mutation generating an aberrant 5' ss in exon 55 of the Ataxia telangiectasia (ATM) gene causes exon 55 truncation and generates an out-of-frame transcript. The effect of this mutation was countered by blocking the aberrant splice site with PMO oligonucleotides, thereby restoring some correctly spliced transcripts [155]. This approach has also been successfully used to correct intronic cryptic site activation in the β-globin gene causing thalassemia [156] or in the CFTR gene in cystic fibrosis [157].

A similar strategy can be used to counter the aberrant inclusion of a pseudoexon in an mRNA (**Figure 15b**). In the same way as for the inactivation of cryptic splice sites, AONs can sterically block the abnormal splice site responsible for the pseudoexon inclusion and retarget the splicing machinery to the natural sites. This approach has

been successfully used to block the recognition of a cryptic donor site generated by a point mutation in intron 9 of the NPC1 gene that results in the inclusion of a pseudoexon in the mRNA and is associated with Niemann-Pick disease [158]. Similarly, an intronic mutation in CEP290 pre-mRNA that creates an aberrant splice donor site and results in the inclusion of a cryptic exon has been associated with Leber congenital amaurosis (LCA), a severe form of retinal dystrophy. AONs directed against the cryptic splice site were successfully used to prevent the inclusion of this aberrant pseudoexon and restore the normal splicing pattern of CEP290 pre-mRNA in immortalized lymphoblastoid cells or fibroblasts of LCA patients [159-160].

III. 1. 3. Exon re-inclusion

Antisense oligonucleotides can also block intronic or exonic splicing silencers (ESSs or ISSs) and mediate the re-inclusion of aberrantly repressed exons. A classical example of this approach is the reintroduction of exon 7 in SMN2 mRNA to restore SMN expression, which is developed as a treatment for patients with spinal muscular atrophy (SMA). It has been demonstrated that blocking either of two identified ESS elements in SMN2 exon 7 with 2'-MOE antisense sequences efficiently induced the inclusion of exon 7 in vitro [161]. Alternatively, a 2'-MOE oligonucleotide targeting an intronic sequence (ISS) within intron 7 of SMN2 greatly enhanced the inclusion of exon 7 in the mRNA, both *in vitro* and *in vivo* [162-163].

Another novel approach to induce the re-inclusion of an exon consists of using bifunctional RNAs that are composed of an antisense sequence moiety, responsible for binding the complementary target sequence, and a tail effector domain containing a splicing regulatory sequence (ESE) for the recruitment of known trans-acting splicing factors to mediate activation or silencing of the targeted exon [164] (**Figure 17**). This strategy, called TOES (targeted oligonucleotide enhancer of splicing) when involving the recruitment of splicing enhancer factors such as SR proteins, has been tested for the inclusion of SMN2 exon 7 for the treatment of SMA. Tailed 2'-O-methyl antisense oligonucleotides carrying an antisense element binding exon 7 and an ESE tail successfully triggered the recruitment of SR proteins and promoted exon 7 inclusion in SMN2 mRNA, leading to partial recovery of SMN protein in SMA patient fibroblasts [164].

Another approach that uses bifunctional reagents to correct aberrant splicing by inducing exon inclusion is called ESSENSE (exon-specific splicing enhancement by small chimeric effectors). This involves novel hybrid molecules that couple an antisense domain for specific sequence targeting with a peptide effector moiety composed of arginine-serine (RS) repeats that mimic the effect of authentic SR proteins and therefore promote the activation of the targeted exon. These chimeric molecules were proven effective at reproducing the functions of SR proteins and restoring correct splicing by promoting exon inclusion in BRCA1 or SMN2 pre-mRNAs [165]

III. 1. 4. Exon skipping

Another application of antisense-mediated splicing modulation involves the blocking of natural splice sites to induce the skipping of a mutated exon and restore the disrupted reading frame of a transcript [166]. By masking key splicing motifs, the spliceosome is prevented from recognizing the exon as such and the machinery therefore "skips" to the next available set of splice sites on the pre-mRNA, inducing the exclusion of the exon along with its surrounding introns (**Figure 15c**). The requirements for such a strategy limit the range of candidate diseases: first the removal of the exon must still allow the recovery an in-frame mRNA; secondly the skipped exon must not correspond to a crucial region of the protein so that its depletion gives rise to an internally deleted but still functional protein. A highly studied example of diseases that can be targeted by this mechanism is Duchenne Muscular Dystrophy (DMD), where various missense mutations and deletions in the dystrophin gene often result in the generation of premature stop codon and absence of functional protein. The dystrophin gene is a good candidate for AON-mediated exon skipping as its central region is composed of repetitive exons. AONs have been shown to efficiently induce skipping of different mutated exons within the repeated region of the gene, leading to the production of a truncated protein that still significantly improves the phenotype [167]. This application will be further detailed hereafter.

AON-mediated exon skipping has also been used to disrupt the reading frame of an mRNA and silence the expression of disease-associated genes. The blocking of different splice sites or ESE sequences within apolipoprotein B (apoB) exon 27 was shown to efficiently induce exon 27 skipping and produce a mRNA that is translated into a truncated isoform, thereby triggering a downregulation of apoB expression and providing a potential way to reduce cholesterol levels for patients with hypercholesterolemia [168].

As previously mentioned for exon re-inclusion approaches, bifunctional RNAs can also be used for exon skipping applications by introducing an ESS sequence into the tail effector domain of the chimeric RNA, in which case the method is called TOSS (targeted oligonucleotide silencer of splicing). The silencing effector domain thus attracts splice silencers such as hnRNP proteins near the exon to be skipped and provoke its exclusion [169].

III. 1. 5. RNA trans-splicing

Several mis-splicing defects can also be corrected via an alternative splicing modulation approach that involves trans-splicing. This mechanism, which naturally exists in various organisms, consists of the mutual splicing of two independent pre-mRNAs to form a fused final mRNA [170-171]. This biological process has been derived for therapeutic purposes into spliceosome-mediated RNA trans-splicing (SmaRT), a technique that enables the splicing machinery to realise a sequence exchange between a mutated target pre-mRNA and a repair RNA molecule with functional exons, called pre-trans-splicing molecule (PTM) [172]. This molecule contains a binding domain complementary to an intronic sequence situated upstream of the targeted defect and a 3' tail composed of a branch point, a polypyrimidine tract, the AG dinucleotide acceptor splice site and the corrected cDNA 3' end terminated by a poly A site (**Figure 19**).

This technique can be applied to correct many mis-splicing diseases, an example being the correction of SMN2 splicing in Spinal Muscular Atrophy. In this case the PTM molecule targets intron 6 of SMN2 pre-mRNA and contains a replacement SMN1 exon 7 to be incorporated in the final trans-processed mRNA. To enhance the trans-splicing process between the two molecules and repress cis-splicing of the target, an AONs targeting the 3' ss of SMN2 exon 8 was co-administered alongside the PTM to block the recognition of this exon in *cis* [173]. The trans-splicing of the two transcripts occurred, even if with moderate efficiency, as a mouse model carrying human SMN2 transgenes that was administered with this PTM-AON cocktail showed increased levels of SMN protein [174]. SmaRT was also designed to correct the mutated CFTR pre-mRNA in cells from cystic fibrosis (CF) patients [175].

This approach brings the possibility to correct the splicing of a transcript not suitable for single exon modulation applications, such as exon skipping or re-inclusion. Moreover, a single pre-trans-splicing molecule is able to correct multiple mutations located downstream of the hybridization site in a same pre-mRNA. This advantage is valuable for mis-splicing diseases treatment as it abolishes the need to design a different therapeutic molecule for each different mutation of a given mRNA. For many splicing diseases an AON is usually designed to be mutation-specific and has to be considered each time as a new drug that require full pre-clinical and clinical studies , which would

prove to be very expensive for a drug suitable for only a small number of patients. By being able to correct a variety of mutations in a target mRNA with a single therapeutic molecule, SmaRT could therefore provide treatment for a larger population of patients with one drug. Although very promising, the efficiency of this strategy is however still low and could benefit from many improvements.

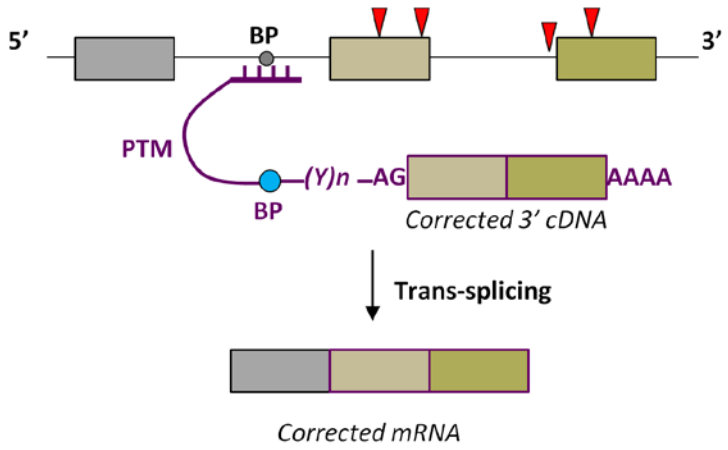

Figure 10: spliceosome-mediated RNA trans-splicing (SmaRT)

The PTM (pre-trans-splicing molecule) repair construct (purple) contains a binding domain complementary to an intronic sequence situated upstream of the defect in the target pre-mRNA, and a 3' tail composed of all the necessary splice signals: branch point (BP), polypyrimidine tract ((y)n) and the 3' ss (AG), preceding the corrected cDNA 3' end.

III. 2. Small nucleolar RNAs as new tools for the modification of pre-mRNA splicing

Another strategy for the modulation of pre-mRNA splicing emerged recently and involves the targeted modification of branch point sites to inhibit their recognition by the splicing machinery. For this purpose, small nucleolar RNAs (snoRNAs) that naturally direct the 2'-O-methylation of various cellular RNAs have been engineered in order to re-direct them towards a pre-mRNA and specifically induce the methylation of the adenosine branch point, thereby triggering the skipping of the downstream exon.

III. 2. 1. The C/D box small nucleolar RNAs

The small nucleolar RNAs (snoRNAs) are 60 to 150 nt long non–protein-coding RNAs that function mainly in guiding the post-transcriptional modifications of ribosomal RNAs (rRNAs) and small nuclear RNAs (snRNAs) [176-177]. The various snoRNAs can be classified into two major groups based on their structural features and functions: the H/ACA box snoRNAs mediate the pseudo-uridylation of cellular RNAs and the C/D box snoRNAs induce the site-specific 2'-O-methylation of ribose on their target RNAs by base pairing through the complementary guide sequence of 10-20nt they contain [178]. Currently, more than 350 human snoRNAs are referenced in the snoRNA-LBME-db database [179], which lists approximately 250 C/D box snoRNAs and 80 H/ACA snoRNAs that have been experimentally confirmed. Amongst these snoRNAs a great number have no identified target RNA and are referred to as "orphan" guides. Moreover, the snoRNA world is a fast growing entity and many novel snoRNA-like sequences have been computationally predicted within human introns but have not been confirmed experimentally yet [180-182], suggesting that thousands of snoRNAs could actually be encoded in the human genome and are yet to be discovered. Many snoRNAs are expressed in a tissue-specific manner whilst others can be submitted to genomic imprinting, which depicts the complexity associated with these non-coding RNAs and the difficulty faced when trying to identify new members of these groups. We will focus here on the C/D box snoRNA group.

The C/D box snoRNAs are characterized by two short conserved sequence motifs: the consensus C Box (RUGAUGA, where R is a purine) and D Box (CUGA) that are positioned close to the 5' and 3' end, respectively. These two motifs are generally brought together through the complementary association of four to five nucleotides from each termini to form a 5'-3' terminal stem structure. Additionally, C/D box snoRNAs often contain in their central portion imperfect copies of box C and box D, called box C' and box D' [183]. Finally, the C/D box snoRNAs possess a 10 to 22 nt long antisense element located upstream of the D box at the 3' extremity, sometimes complemented by a second antisense sequence upstream of the D' box.

These antisense sequences present a perfect complementarity to sequences on their target RNAs to which they associate by direct Watson-Crick base pairing, thereby positioning the nucleotide to be methylated exactly 5 nt upstream of the conserved D (or D') box. The protein fibrillarin that carries the 2'-O-methyltransferase activity specifically recognizes the D box and induces the methylation on the associated RNA at this selected nucleotide position. The antisense-snoRNAs therefore act as guides for the site-specific methylation of intracellular RNAs, and the antisense sequence together with the D/D' box provide all the information necessary for the selection of the targeted nucleotide [184-185].

The majority of vertebrate C/D box snoRNAs are encoded within introns of pre-mRNAs and are co-transcribed with their host genes in the nucleoplasm [187]. A host gene can contain one or multiple different snoRNAs within its introns. Many host genes are involved in ribosome biogenesis or function, although others have been identified that are devoid of protein coding potential and appear to be transcribed by RNAPII primarily to express the snoRNA [188-189]. The intron-encoded snoRNAs are excised from the spliced and debranched host introns by exonuclease digestion in the nucleoplasm [190] (**Figure 21**). It has been shown that all the necessary signals for correct processing of a snoRNA reside within the snoRNA sequence itself as a snoRNA artificially placed into a non-natural intron context is accurately processed [191].

After transcription of the snoRNA-host gene, a snoRNA-specific core of proteins composed of the four evolutionarily conserved proteins fibrillarin (Nop1), Nop56, Nop58 and Snu13 (15,5K), assembles on the newly synthesised transcript by recognising the conserved C and D boxes and the neighbouring 5'-3' stem structure within the pre-snoRNA sequences, which constitute the signal elements essential for

correct processing of the mature snoRNA [192-193]. The assembly of this core of proteins is triggered by the splicing of the host pre-mRNA: the proteins are actively recruited on the snoRNA by a factor that also associates with splicing factors on the pre-mRNA (**Figure 21**). Intronic snoRNA processing is therefore physically and functionally linked to the splicing of the host pre-mRNA intron and these two event occur synergistically [194-195]. The position of the snoRNA in the host intron relative to consensus splice sites is critical for both pre-mRNA splicing and efficient snoRNA synthesis [196-199]. Once the host intron is debranched, exonuclease trimming occurs to produce the mature ends of the snoRNA. The core of proteins bound to the pre-snoRNA therefore define the correct ends of the mature snoRNA and provide a protection against over-digestion by exonucleases as well as stability for the snoRNA [200].

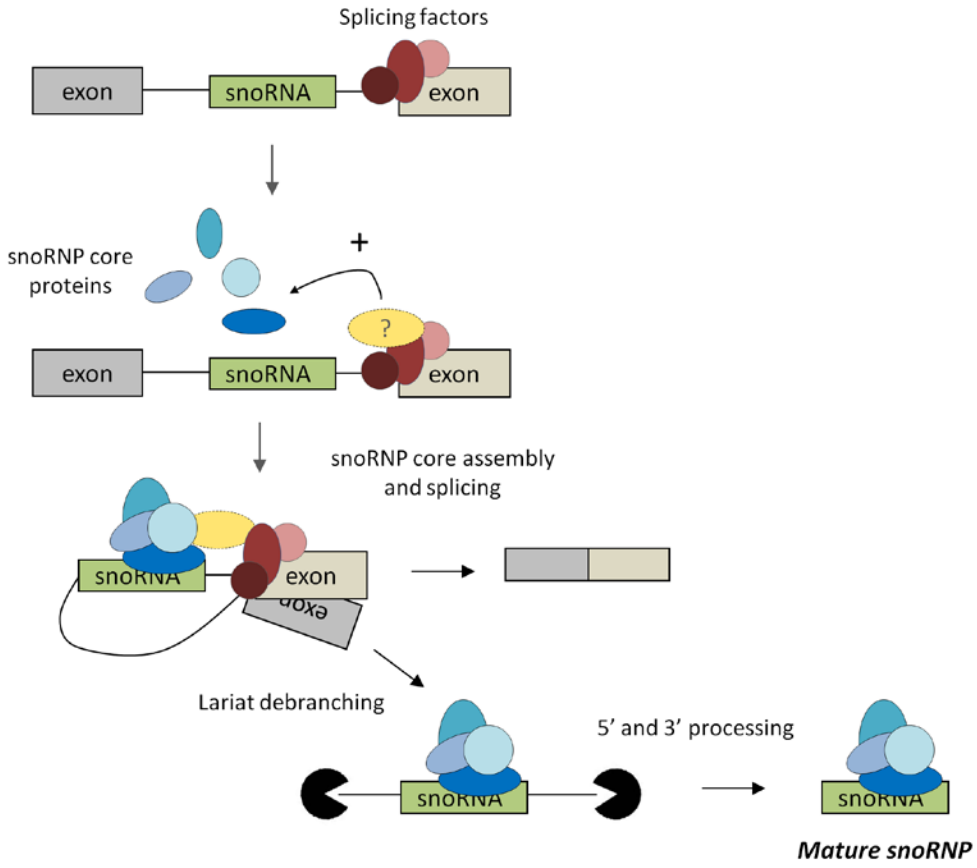

Figure 11: biogenesis of intronic C/D box snoRNAs

During the splicing of the host pre-mRNA, an assembly factor (orange) associates with splicing factors (red) bound to the transcript and recruits 4 core proteins (blue) on the snoRNA. After splicing, the debranched intron containing the snoRNA is released and digested by exonucleases to produce the mature snoRNP.

The correctly processed mature snoRNA associated with the core proteins is then released in the nucleoplasm in the form of a small nucleolar ribonucleoprotein particle (snoRNP). Most C/D box snoRNPs then transit into Cajal bodies before accumulating in the nucleolus, their final destination and site of action. The conserved C/D box motif and its associated core proteins are necessary and required to direct nucleolar localization of the mature guide snoRNPs [201-202].

III. 2. 2. Targeted branch point methylation

During the splicing process, the 2'-OH of the branch point (BP) adenosine is essential as it initiates the first trans-esterification step of the reaction by attacking the phosphate of the 5' splice site. The inactivation of the branch point by eliminating its 2'OH group is therefore likely to prevent the reaction and impair the splicing process. It has indeed been demonstrated that a 2'-O-methylated BP nucleotide is inactivated and unable to pursue the splicing reaction [45].

This new concept led several groups to assess the possibility of engineering a C/D box snoRNA to specifically target the 2'-O-ribose methylation of a chosen pre-mRNA branch point in order to induce the skipping of the downstream exon. The notion that C/D box snoRNAs can interact with pre-mRNAs and modulate their processing is surprising considering that snoRNAs are thought to be predominantly localised in the nucleolus whilst mRNA splicing occurs in the nucleoplasm. It suggests that the sub-nuclear sites to which the guide snoRNAs accumulate do not necessarily correspond to their sites of action. Indeed, several studies have now demonstrated that numerous C/D box snoRNA are devoid of complementarity to rRNAs or snRNAs but instead contain antisense sequences matching messenger RNAs [203], suggesting that the interaction between nucleoplasmic pre-mRNAs and snoRNAs can naturally occur. Accordingly, it has been recently shown that the snoRNA MBII-52 could localise in the nucleoplasm, at least transiently, and this transit between nucleoplasm and nucleolus enables the snoRNA to access the pool of pre-mRNAs in order to function as a regulator of splicing [204].

Altering the original guide sequence of the snoRNA has been shown to be sufficient to re-direct it to the chosen target RNA and to change the specificity of methylation [184]. Zhao et al. used this property to design synthetic C/D snoRNAs that specifically target the 2'-O-methylation of branch point adenosines in various *Saccharomyces cerevisiae* pre-mRNAs in order to block their splicing (**Figure 22**). They demonstrated that the targeted pre-mRNAs became specifically 2'-O-methylated in presence of the modified snoRNAs, resulting in efficient splicing inhibition [186].

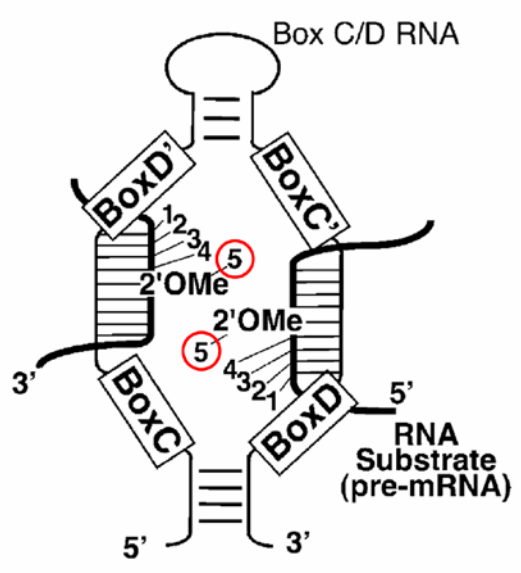

Figure 12: site-specific methylation of pre-mRNA using C/D box snoRNA

The snoRNA antisense sequences are modified to hybridize a target pre-mRNA, perfectly positioning the nucleotide to be methylated 5 nt upstream of the conserved D (and/or D') box (red circles).

Following this proof-of-concept in yeast cells, this strategy was further investigated for its potential to alter the splicing of human genes. An artificial human U24 C/D box snoRNA was engineered to guide the 2'-O-methylation of the branch point of the human heat-shock cognate protein (HSC8) pre-mRNA and its expression in human adenocarcinoma cells induced the exclusion of the exon situated directly downstream of the methylated branch point, albeit with a low efficiency of 6 to 10% [205]. This approach was also validated in another vertebrate system by injection into *Xenopus* oocytes of an artificial C/D box snoRNA re-directed to the adenovirus pre-mRNA branch point. The snoRNA efficiently induced the methylation of the target BP, which subsequently resulted in splicing impairment of the region downstream of the modification [206].

Although the proof-of-concept of this snoRNA-mediated pre-mRNA splicing modulation strategy has been established in vertebrate systems, the efficiency of this approach is low and still needs to be improved. To date, AON-mediated modulation of splicing therefore constitutes a more efficient and more promising strategy applicable to a wide array of diseases. Significant progress and improvements of the technique have been made over the past years and proof-of-concept of AONs efficacy has been demonstrated in various diseases. However the use of antisense strategies for clinical applications still faces several hurdles: it is currently challenging to express long-term therapeutic amounts of antisense sequences in the target tissue as intracellular delivery of synthetic oligoribonucleotides induces only a transitory effect and can be rapidly degraded in the cells, implying the requirement for repeated re-administration. Moreover, the systemic delivery of AONs cannot guarantee a targeted tissue-specific expression of the antisense and can result in potential undesirable off-target effects. Finally, the *in vivo* safety of antisense molecules needs to be rigorously investigated and controlled as some moderate toxic effects dependent on the AON backbone chemistry have been reported [145, 207]. Nevertheless, the potential of AON-mediated therapy for mis-splicing defects is becoming more and more evident and is highlighted by the onset of clinical trials for various diseases, such as the recent Duchenne muscular dystrophy trials that will be detailed later.

IV. Vector technologies for expression of exogenous small RNA sequences

The intracellular delivery of synthetic oligoribonucleotides induces only a local and transitory effect and repeated administration of high doses of oligonucleotides would be needed for the treatment of chronic genetic diseases such as DMD. To circumvent this problem, the introduction of antisense sequences in strong expression cassettes that are delivered by gene transfer vectors could ensure their stable, targeted and efficient expression. One possibility is to use modified small nuclear RNAs (snRNA) as shuttles to obtain a sustained intracellular expression of antisense sequences. SnRNAs are appealing as antisense vectors since they enable to express the antisense sequence in relatively high concentrations from their strong promoters, they induce their colocalisation in the same subcellular compartment as the splicing machinery and they provide protection of the antisense sequences against intracellular nucleases. The validity of this strategy has been established using U1, U2 or U7 snRNAs delivered by viral vectors [208-210].

IV. 1. The small nuclear RNAs: characteristics and biogenesis

The small nuclear RNAs are abundant non-coding transcripts that function in the nucleoplasm and can be divided into two groups. The Sm-class snRNAs, composed of U1, U2, U4, U4atac, U5, U7, U11 and U12 snRNAs, are transcribed by RNA polymerase II (RNAPII) and contain a Sm binding site for the recruitment of seven Sm proteins, a 5'-trimethylguanosine cap and a 3' stem-loop (**Figure 23**). The second group, the Lsm-class snRNAs, is composed of U6 and U6atac snRNAs. They are produced by RNA polymerase III (RNAPIII) and are characterised by a 3' stretch of uridines that constitutes the binding site for Sm-like proteins (termed Lsm proteins), a monomethylphosphate cap and a 3' stem-loop. These Lsm-class snRNAs will not be discussed further here.

All the uridine-rich snRNAs assemble with proteins into small nuclear ribonucleoparticles (snRNPs) and form the core component of the spliceosome, except the U7snRNA which is non-spliceosomal and functions in histone pre-mRNA 3' processing.

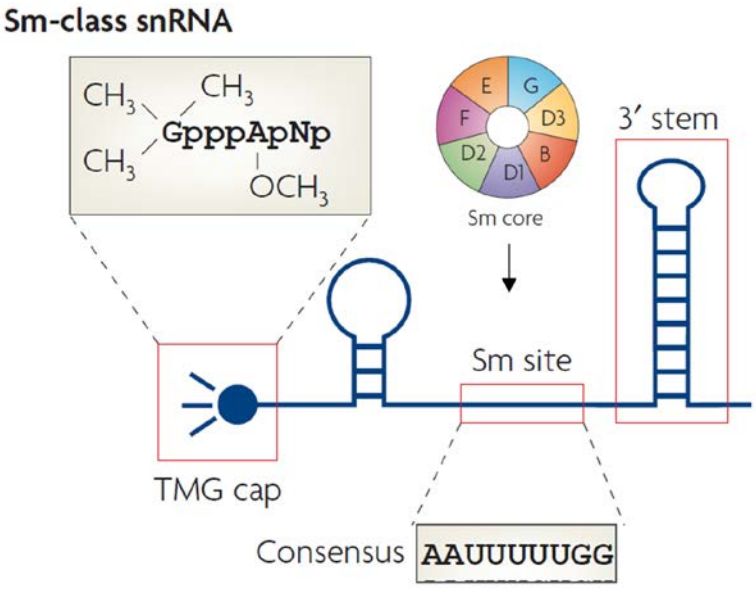

Figure 13: features of Sm-class snRNA transcripts

SnRNA transcripts contain a Sm binding site for the recruitment of the Sm protein ring (Sm core), a 5′-trimethylguanosine cap and a 3′ stem–loop. The general consensus sequence of the Sm binding site and the composition of the general Sm core of proteins are represented.

The vertebrate Sm-class snRNA genes are ubiquitously and highly expressed and are present in multiple copies within genomes, usually in clusters [212]. The structure of snRNAs genes is simple: they contain a snRNA-specific core promoter of around 250 bp that is very active but lacks a TATA box, no intron and no polyadenylation signal. The TATA-less promoters contain two elements: a distal sequence element (DSE), which behaves like an enhancer and is required for high levels of transcription, and a snRNA specific proximal sequence element (PSE) that drives basal levels of transcription [213]. The 3' end of snRNAs contains the 3'-box located 9 to 19 nt downstream of the snRNA coding region, which is required for correct 3' end processing (**Figure 24**).

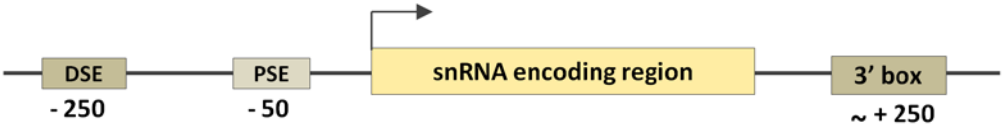

Figure 14: structure of Sm-class snRNA genes transcribed by RNAPII

The promoters of snRNA genes are composed of the proximal (PSE) and distal (DSE) sequence elements situated approximately 250 nt and 50 nt upstream of the transcription start site (arrow). The 3' box important for snRNA transcript processing is located between 9 and 19 nucleotides downstream of the snRNA coding region.

IV. 1. 1. Transcription of Sm snRNA genes and coupling with 3'end processing

The transcription of Sm-class snRNA genes starts with the binding of general transcription factors on the DSE sequence, such as Oct-1 and Staf that are implicated in the transcription of both mRNA and snRNA genes [213]. This interaction triggers the recruitment of a complex of transcription factors specific to snRNA genes on the PSE, which define the +1 transcription start site [214]. This multisubunit complex called SNAPc (snRNA activator protein complex), also known as PTF or PBP, associates with the PSE and mediates the recruitment of RNAPII to the promoter [215-216]. The transcription of snRNA genes by RNAPII, similarly to mRNA genes, involves the general transcription factors TBP, TFIIB, TFIIA, TFIIE, TFIIF and possibly TFIIH [217-218]. The downstream AU-rich cis-acting element called the 3' box does not act as a termination sequence as the pre-snRNAs produced after transcription are 3' extended beyond this element [219], but is instead required for the subsequent 3' processing of the newly formed snRNA and marks the upstream cleavage site [220].

The 3' end processing of snRNAs transcripts is tightly coupled with their transcription and the snRNA promoters play a critical role in this process [221]. Indeed, it has been demonstrated that the replacement of snRNA promoters with non-snRNA promoters abolishes the recognition of the 3' box and causes constitutive misprocessing of the snRNA [222-223]. The correct processing of pre-snRNAs therefore depends on a factor that requires the presence of the 3' box and of a snRNA-specific promoter, highlighting the link between transcription and RNA processing. The factor responsible for snRNA processing is a multiprotein complex called Integrator, which contains two subunits

(Int9 and Int11) homologous to components of the cleavage and polyadenylation specificity factor (CPSF) complex that is essential for the cleavage of mRNA 3' ends [224-225]. The Integrator has been shown to associate with the C-terminal domain (CTD) of the largest subunit of RNAPII [224, 226], which is specifically required to be phosphorylated on serine 2 and serine 7 for efficient recruitment of the Integrator complex and for proper 3' box-dependent 3' end formation [227-228]. Finally, it has also been shown that sequences in the terminal 3' stem-loop present within the coding region of U7 snRNA are required for Drosophila snRNA processing, although the requirement for this stem loop signal has not been demonstrated in vertebrates [223].

The currently accepted model of the coupling between snRNA transcription and processing is that the Integrator complex associates with the phosphorylated CTD of RNAPII at the promoter, travels with it along the gene during transcription until it recognizes signals in the stem loop of the coding sequence as well as the 3' box at the end of the transcript. The co-recognition of these two elements directs the Integrator to the proper cleavage site, approximately 10nt upstream of the 3' box, which then mediates the co-transcriptional 3' processing of the nascent pre-snRNA (**Figure 25**).

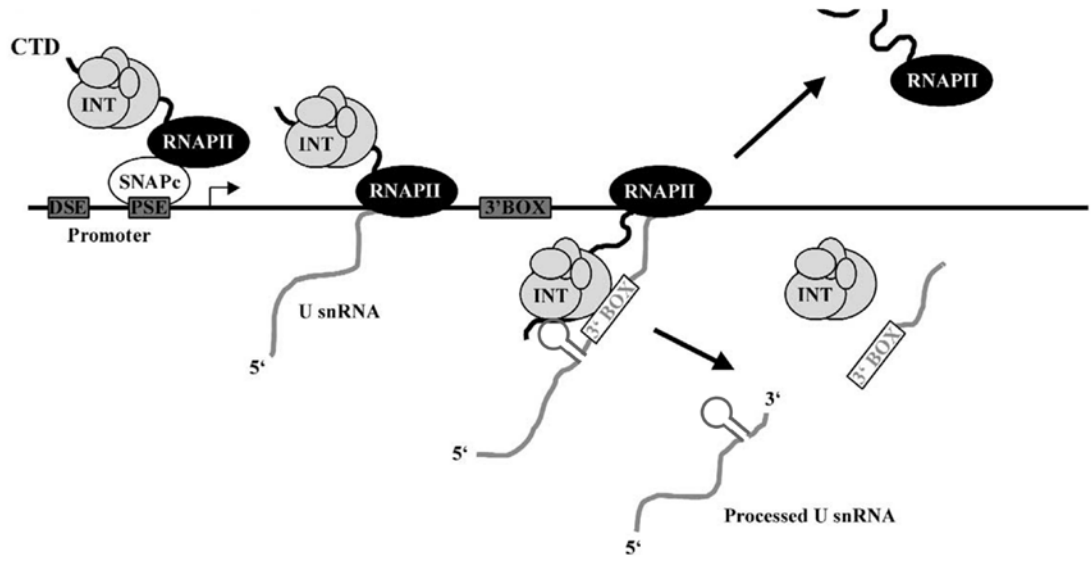

Figure 15: Coupling between snRNA transcription and 3' box-dependent processing through the Integrator Complex.

The Integrator complex is attached to the C-terminal domain of RNAPII, which is fixed to the promoter of snRNA genes. The Integrator complex then travels along the gene with the polymerase during transcription until it recognises the 3' box and the stem loop in the newly synthesised transcript. These signals direct the Integrator to cleave the nascent transcript approximately 10 nt upstream of the 3' box, thereby releasing the pre-snRNA from its template.

Once the Sm-class snRNAs are cleaved off the template, a 7-methylguanosine cap is added on their 5' end and they are exported from the nucleus to the cytoplasm for further maturation. The transport to cytoplasm occurs through the activity of a snRNA-specific export complex that contains the export adaptor protein PHAX, the export receptor chromosome region maintenance-1 (CRM1), the cap-binding complex (CBC) and the Ran GTPase [229-230].

IV. 1. 2. Sm core assembly

The assembly of snRNAs into stable Sm-core particles (snRNP**s)** takes place in the cytoplasm and is directed by a large assemblyosome called the survival of motor neuron (SMN) complex, which is composed of the SMN protein, seven distinct Gemin proteins (Gemin 2-8), and several other protein factors [231]. The Sm core proteins (B/B0, D1, D2, D3, E, F and G) associate into a stable seven-membered ring around a conserved binding sequence called the Sm site (consensus RAUU$^{U}/_{G}$UUGR) to form the core of the snRNP particle [232]. This assembly is facilitated by the SMN complex which binds to the snRNA precursors and to the seven Sm proteins and brings them together. As an exception, the U7 snRNA contains a non-consensus Sm site that associates with a Sm ring of a different composition, which will be detailed later.

The assembly of the Sm core is required for the completion of the snRNPs maturation steps. First, the SMN complex directs the hypermethylation of the snRNA 5' cap to form a 2,2,7-trimethylguanosine cap [233]. The 3' extended pre-snRNA is then submitted to a nucleolytic trimming, which is also Sm-core dependent, to generate the mature length snRNA [234-235]. The Sm core assembly as well as the 3' end trimming events are a prerequisite for the nuclear import of the mature snRNA [236]. Once in the nucleoplasm, the newly made snRNPs transiently accumulate in Cajal bodies where additional modifications steps take place, such as 2'-O-methylation and pseudouridylation of the snRNAs, which are essential for the snRNP to be fully functional [237-238]. Finally, the Sm particles assemble into di-snRNP or tri-snRNP prior to accumulation in nuclear subdomains.

IV. 2. Modification of U snRNAs for antisense sequence expression

The small nuclear RNAs can be engineered and used as tools to express antisense sequences in the cells in order to modulate mis-splicing defects. As mentioned before, the snRNAs are expressed at high level from their strong promoters, are localized in the nucleus and interact with their RNA targets by direct base pairing. It is possible to replace the natural antisense sequence of a snRNA with an antisense sequence chosen to specifically target an mRNA of interest and modulate its splicing, with potentially more efficiency than intracellular delivery of AONs alone. Indeed, the use of snRNA shuttles to express antisense sequences presents several advantages: the sequences are then protected from degradation and produced continuously and at high levels within the cell. Moreover, the modified snRNAs are still able to associate with Sm core proteins into snRNPs, which provides intracellular stability and induces the nuclear localization of the antisense sequence.

Sm Class snRNAs are therefore ideal candidates as antisense vehicles since the only required modification for targeting a desired mRNA is the replacement of their natural RNA-binding sequence. It has been shown that U2 snRNA can constitute an antisense carrier for splicing modulation purposes [210], but so far U1 snRNA- and U7 snRNA-based vectors have been the most commonly used for this purpose.

The 164 bp long human U1 snRNA is transcribed by a strong promoter and is present in multiple copies in the genome, allowing the production of approximately 10^6 U1 transcripts per cell. This high level of expression has been a strong argument for the use of U1snRNA for antisense expression, and chimeric snRNAs targeting various splice signals and delivered by viral vectors have demonstrated some efficiency for the correction of aberrant splicing in multiple diseases. For example in the context of Duchenne Muscular Dystrophy, modified U1 snRNAs redirected towards splice sites of the mutated mouse DMD pre-mRNA and systemically delivered by adeno-associated virus (AAV) vectors successfully induced a body-wide restoration of Dystrophin in the mouse model of the disease [239]. In a more recent study, U1 snRNA also successfully induced exon skipping in the human DMD pre-mRNA in primary patient fibroblasts [240].

Alternatively, modified U1 snRNAs have also been used to correct donor splice site mutations in pre-mRNA by compensatory changes. It has been demonstrated that U1 snRNAs could be engineered into suppressors of donor splice site mutations by introducing compensatory substitutions in their 5' extremities to perfectly match mutated splice donors. Such modified U1 snRNAs are then able to promote normal splicing at the mutated 5' ss and have shown some efficiency for the correction of aberrant splicing in diseases such as β-thalassemia [209], severe factor VII deficiency [241-242] or some cases of Bardet-Biedl syndrome [243].A recent study also reported the development of modified U1 snRNA designed to hybridize into intronic sequences downstream of the donor splice sites of various pre-mRNAs to correct splicing impairments [244]. This study showed that some of these U1 variants, termed exon-specific U1 snRNAs (ExSpeU1), were able to rescue the splicing defects caused by a mutated 5' ss, but also to correct the effects of mutations located at the 3' splice sites or in exonic regulatory elements of mutated pre-mRNAs.

IV. 3. U7snRNA for antisense sequence delivery

U7 snRNA is a non-spliceosomal small nuclear RNA that is not involved in pre-mRNA splicing, but mediates the 3′ end processing of replication-dependent histone pre-mRNAs by base pairing through its 5’ 18 nucleotides antisense sequence [245-246] (**Figure 27**).

U7 snRNP is present at only 10^3 to 10^4 copies in mammalian cells, much less than the spliceosomal snRNPs, but the U7-specific promoter has been shown to be of comparable strength to that of the spliceosomal snRNAs [210, 247-248]. The conserved U7snRNA transcript is 57–71 nt long depending on the species, and the structure of the gene shares structural features with the spliceosomal snRNAs (**Figure 27**).

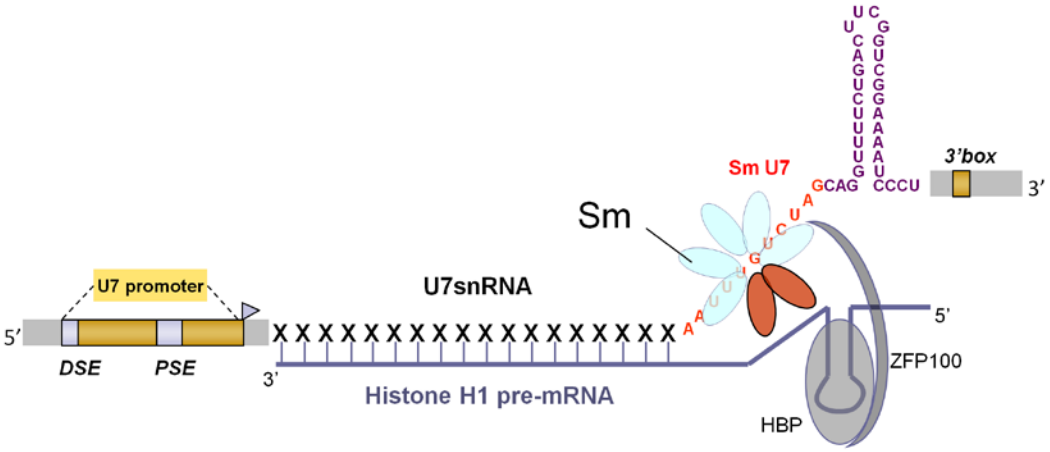

Figure 16: U7 snRNA and hybridization to replication-dependent histone pre-mRNA

The 239-bp U7-specific promoter contains the DSE and PSE sequences. The mouse U7 transcript is composed of a 18-nt antisense element to histone pre-mRNAs at the 5' end, a non-consensus Sm binding site of 11 nt specific to U7, and a 31-nt long 3'-terminal hairpin structure required for stability. The 3' extension present in the premature U7 transcript before nucleolytic trimming contains the 15-nt long 3' box located 13 nt downstream of the mature U7 snRNA 3' end.

The 100-kDa zinc finger protein (ZFP100) is a component of U7 snRNP that interacts with the hairpin binding protein (HBP) associated on the stem-loop of the histone H1 pre-mRNA to stimulate histone mRNA processing [249].

As for the spliceosomal snRNAs, the promoter, stem-loop structure and 3' box element are important for the correct processing of the U7 precursor [221, 223, 225]. The maturation pathway is also similar, including a cytoplasmic phase where it assembles into a snRNP. The main difference resides in the composition of the Sm core of the U7 snRNP [234]. The non-consensus Sm binding site of the U7 snRNA (AAUUUGUCUAG) is responsible for the recruitment of a U7-specific Sm core of proteins during the assembly into snRNP. This particular Sm ring is composed of 5 classical Sm proteins (B/B0, D3, E, F, and G) and of two Sm-like proteins, LSM10 and LSM11 that replace D1 and D2 respectively [250-252].

The assembly of the unique U7 Sm protein core follows several precise requirements. First, four nucleotides located within the 3' half of the Sm site are crucial for the assembly of an active U7 snRNP and need to be conserved. Furthermore, the position and size of the U7 terminal stem-loop structure also affect RNP assembly. Finally, it has been shown that the substitution of a particular nucleotide within the U7 Sm site with a pseudouridine, which confers structural rigidity on the sugar-phosphate backbone of the snRNA, impairs assembly of the U7 Sm core. This highlighted a requirement for backbone flexibility at the Sm site for correct assembly of the Sm ring [253].

The modification of the non-canonical Sm binding site of U7 snRNA (AAUUUGUCUAG) into the consensus sequence derived from spliceosomal snRNPs (AAUUUUUGGAG) greatly increases the intracellular U7 RNA expression levels and significantly improves the nuclear accumulation of U7 snRNA [254]. Conversely, the modification of U1 genes by replacement of their Sm sequence with the U7 Sm binding site reduced RNA levels to those of the wild type U7. Altogether, these observations demonstrate that the sequence of the Sm binding site strongly influences the level of snRNAs in the cell [254]. Moreover, the modification of U7 Sm site into the spliceosomal consensus led to the formation of U7snRNPs that are non-functional in histone RNA processing [234, 254]. This Sm modification triggers the recruitment of the conventional seven-membered Sm core on the U7 snRNA instead of the U7-specific Lsm10 andLsm11 proteins, which are functionally involved in histone pre-mRNA 3' processing [250-251], thereby abolishing the natural U7snRNP function.

Similarly to the spliceosomal snRNAs, U7snRNA can be modified into a mediator of splicing by replacing its original 5' anti-histone sequence with an antisense sequence complementary to a pre-mRNA of interest [208, 255] and converting the wild-type Sm-binding site of U7 snRNA into the canonical Sm sequence (SmOPT). Thereby, the U7snRNA shuttle assembles with Sm proteins into a snRNP and associates with the spliceosome, facilitating the colocalization of the antisense sequences with the target pre-mRNA [256] (**Figure 29**). Moreover, as the U7-SmOPT transcripts are no longer functional in histone pre-mRNA processing, they become fully available to exert their new function of splicing modulation [234, 254].

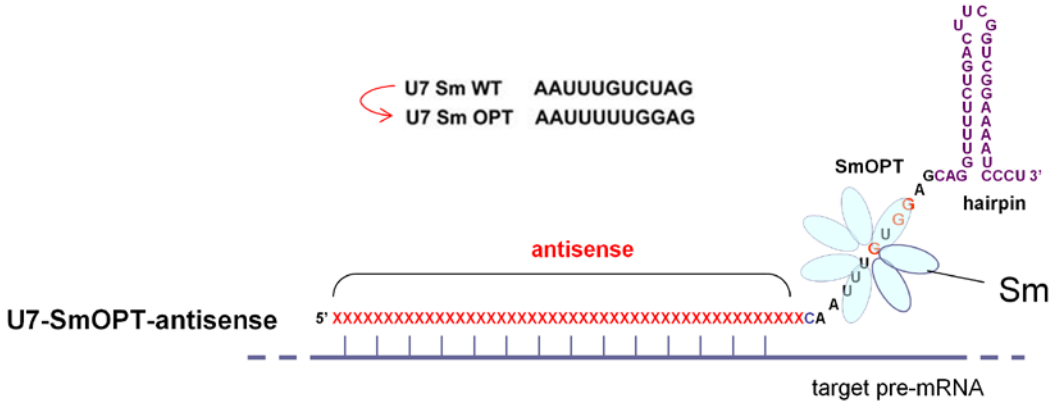

Figure 17: modification U7 snRNA for antisense sequences delivery

U7 snRNA is engineered into an antisense carrier by replacing its 18-nt antisense sequence and by modifying the U7-specific Sm binding site (Sm WT) into the canonical Sm sequence of spliceosomal snRNAs (SmOPT)

This U7 snRNA-mediated expression of antisense sequences has proven efficient to alter splicing in a variety of mis-splicing diseases. The first U7-induced splicing modulation applications demonstrated the successful correction of aberrant β-globin pre-mRNAs in cultured cells model of β-thalassemia [208, 257]. Since then, U7 snRNAs were also used to induce the correction of SMN2 pre-mRNA splicing in the context of spinal muscular atrophy [258-262], to successfully inhibit HIV-1 multiplication through disruption of tat, rev or cyclophilin A expression by exon skipping [263-264], and to restore the splicing pattern of pre-mRNAs for a range of other disorders (**Table 1**). One of the most promising applications of this strategy is the restoration of dystrophin expression by forced exon skipping of the mutated pre-mRNA in Duchenne Muscular Dystrophy. The potential of modified U7 snRNAs to successfully induce the skipping of the mutated

dystrophin pre-mRNA was first demonstrated in muscular cells of the *mdx* mouse model and resulted in de novo protein synthesis [256]. This study also provided evidence that inserting in the U7 cassette two antisense sequences targeting two different sites on the pre-mRNA (double-target antisense) allowed a more efficient exon skipping than when a single sequence was targeted, as previously demonstrated in the β-globin pre-mRNA [257].

Disorder and genetic target	Genetic therapy using modified U7 snRNA	Delivery of U7 snRNA	Ref
β-thalassemia β-globin gene	- U7 snRNAs containing antisense sequences against cryptic splice sites + Double-target U7 snRNAs	- Transfection in Hela cells expressing the mutated β-globin gene - lentiviral vector delivery in erythroid progenitor cells of patients	[208, 257] [265]
Duchenne muscular dystrophy dystrophin gene	**mouse gene:** stop mutation in exon 23 Double-target U7 snRNA for mouse dystrophin exon 23 skipping U7(BP22/5'ss23)	- Transfection of cultured mdx cells - local administration of AAV vectors in mdx muscles - intravenous injection of self-complementary AAV in dKO mice	[256] [266] [267]
	human gene: deletion of exons 49 and 50 - Bifunctional U7 snRNAs carrying an antisense sequence and a splicing silencer sequence for exon 51 skipping - modified U7 snRNAs for skipping of exons 45 to 55 + combination of three U7snRNAs into one vector for multiple-exon skipping of exons 45-47	- lentiviral vectors delivery in human DMD myoblasts + intramuscular injection of AAV in hDMD mice - AAV vectors delivery in DMD myoblasts + intramuscular injection in hDMD	[268] [269]
	canine gene: deletion of exon 7 Double-target U7 snRNAs directed against ESE for double-exon 6-8 skipping	transendocardial delivery of AAV vector in GRMD dog	[270-271]
Spinal muscular atrophy SMN2 gene	- U7 snRNA complementary to the 3' splice site of exon 8 for exon 7 inclusion - bifunctional U7 snRNAs carrying a binding sequence for exon 7 and a splicing enhancer sequence (ESE) for exon 7 inclusion	Transfection in HeLa cells expressing endogenous SMN1 and SMN2. - lentiviral vectors delivery in human fibroblasts from SMA patients - introduction by germline transgenesis into SMA mouse model homozygous for human SMN2	[258] [260] [259, 272]
HIV/AIDS cyclophilin A, tat and rev genes	U7 snRNA induced exon skipping to impair cyPA, tat or rev pre-mRNAs splicing and reduce HIV-1 multiplication	lentiviral vectors delivery in CEM-SS or CEM T-lymphocytes	[263-264]
Dysferlinopathies Dysferlin gene	- double target U7 snRNA against ESE sequences to induce exon 32 skipping	Lentiviral vectors delivery in fibroblast-derived myoblasts of patient	[273]
cancers PTCH1 and BRCA1 genes	U7 snRNAs containing antisense sequences against cryptic splice sites	Transfection in Hela cells containing minigene constructs	[274]
Myotonic dystrophy (DM1) DMPK gene	Human U7-snRNA containing a poly-CAG antisense for degradation of target mRNAs with abnormal CUG expansions	Lentiviral vectors delivery in skeletal muscle cells from DM1 patients	[275]

Table 1: list of applications using the U7 snRNA-mediated expression of antisense sequences to modulate pre-mRNA processing

IV. 4. Viral vectors for gene therapy

The efficient intracellular delivery of antisense-snRNA cassettes requires the use of viral vectors, which are able to enter the target cells and to transfer their genetic information in the nucleus of infected cells. For safety purposes, these viral vectors are engineered to create replication-defective viral particles by replacing the genes required for viral replication with an expression cassette containing the antisense gene. Different viruses have been modified into antisense delivery vehicles, including retrovirus [210], adenovirus [175], lentivirus [265], and adeno-associated virus (AAV) [276].

Retroviral vectors are unable to infect nondividing cells and their ability to integrate their genome into the host DNA has been associated with serious adverse consequences, such as the development of T-cell leukemia in children during a clinical trial for gene therapy of X-linked SCID [277-279]. The use of adenoviral vectors in clinical trials has also raised various concerns such as low efficiency and potential adverse effects [280]. Nowadays, the most widely used viral vectors are vectors derived from adeno-associated viruses and from lentiviruses.

IV. 4. 1. AAV vectors

Adeno-associated viruses (AAV) belong to the Parvovirus family and the Dependovirus gender. They are dependent on the co-infection of a helper virus, such as an adenovirus, to be able to replicate and form viral capsids in the host cell. AAV therefore remain quiescent in the absence of a helper virus and so far no human pathology has been reported associated to AAV infection.

AAV viruses contain a small single-stranded DNA genome of approximately 4.7 kb that is packaged into an icosahedral, non-enveloped capsid. The genome contains two genes, called rep and cap, and two inverted terminal repeat regions (ITR) at the 5' and 3' extremities of the genome. The rep gene encodes the four non-structural Rep proteins involved in viral genome replication, integration and packaging of AAV genomes into preformed empty particles [281-283]. The cap gene encodes three capsid proteins (VP1 to 3), which are identical in their amino acid sequences except for the N-terminal parts [284-285], as well as the assembly-activating protein (AAP), which directs the newly synthesized capsid proteins to the nucleolus where AAV capsid morphogenesis occurs

and is also implicated in the capsid assembly reaction [286]. Depending on the serotype of the AAV virus the amino acid composition of the capsid varies, which results in the use of different cell surface receptors for cell entry and provides distinct tissue tropisms for each AAV serotype. At the present time, 12 different serotypes and many variants have been described. Finally, the ITRs at each end of the genome are folded into a T-shape through base-pairing between palindromic regions. These secondary structures function as origin of replication and are essential for the conversion of the single-stranded viral DNA into double-stranded DNA prior to expression. The ITRs are also important for genome integration of the wild-type AAV into the host DNA [287-288].

AAV viruses can be modified into gene transfer vectors by eliminating all the viral genes (rep and cap) and replacing them by a transgene, such as a U7-antisense cassette, while leaving the ITRs at each end that are required for replication and packaging into viral particles. The major inconvenient of using AAV as vectors is the small packaging capacity of their genome, which limits the transgene size to a maximum of 4,7 kb.

During infection by AAV, the presence or absence of a helper virus determine the fate of the viral genome: if the helper virus is present, induction of gene expression and replication takes place immediately. However in absence of helper virus, the AAV enters a latent life cycle and its genome integrates into human chromosome 19 at a specific locus known as AAVS1 [290-291]. The expression of the AAV Rep gene is necessary for this site-specific integration to occur. Because recombinant AAV (rAAV) vectors lack a Rep coding sequences, they are usually not able to integrate their genome in the host DNA and persist in an extrachromosomal form, or episome, as either linear or circular monomers that can be converted into large concatemers (>12 kb) [292-294].

Another important point to consider for clinical use of recombinant AAV vectors is the potential presence of neutralizing antibodies in humans administered with the vector. It has been shown that anti-AAV antibodies can be naturally found in the sera of approximately 80% of the human population due to previous exposure from natural AAV infections [295], and these specific antibodies can neutralize and impair AAV mediated gene transfer. Moreover, repeated vector administration can induce the *de novo* production of neutralizing antibodies even if none were initially detected in the serum, which would inhibit the efficacy of the vector in subsequent applications to the same patient. One way to circumvent this obstacle is to vary the AAV serotype used in order to achieve immune escape from neutralizing antibodies in patients with anti-AAV

immunity [296]. Another possibility is to prevent the recognition and action of neutralizing antibodies by mutating the relevant capsid epitopes. AAV vectors are considered as good tools for transduction of postmitotic cells and Phase I/II clinical trials for the treatment of various diseases are currently ongoing [297].

IV. 4. 2. Lentiviral vectors

Lentiviruses, which belong to the family of Retroviridae, are RNA viruses that are packaged into a capsid enveloped in a lipid membrane. The lentivirus contains a single-stranded RNA genome that contains its own regulatory elements and is transcribed by the transcription apparatus of the host cell. The lentiviral genome features two long terminal repeats (LTR) at the 5' and 3' ends flanking the protein-coding regions, and a packaging signal located directly downstream of the 5' LTR (ψ). The typical retroviral genes, gag, pol, and env, respectively code for the capsid proteins, the reverse transcriptase and integrase, and the envelope proteins. Additionally, the lentivirus encodes two essential regulatory genes, tat and rev, necessary for the expression of the provirus, and some accessory genes (vpr, vpu, vif, and nef) that are essential for virulence and infectivity (**Figure 31**).

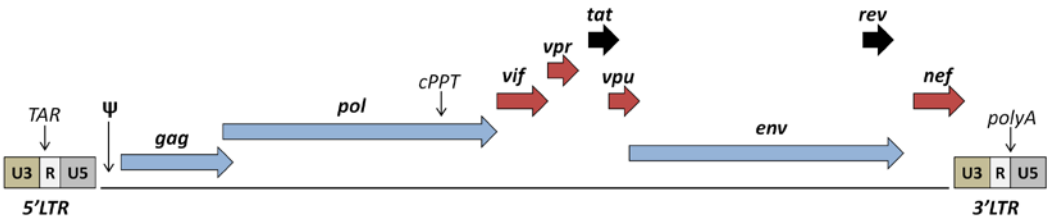

Figure 18: organisation of the HIV-1 (lentiviral) genome

The proviral DNA is represented with all the coding sequences. The 5' and 3' LTR contain the U3-R-U5 elements in the proviral DNA.

The LTRs flanking the coding sequences consist of three regions: the 3' unique element (U3), the repeat element (R), and the 5' unique element (U5), and are important for reverse-transcription, integration and expression of the viral genome. Following cell entry the RNA viral genome is reverse-transcribed into double-stranded DNA, a

process during which the U3-R-U5 Long Terminal Repeats are formed. This double-stranded DNA genome then enters the nucleus and integrates into the host DNA as a provirus (**Figure 31**), which will then be transmitted to daughter cells upon cell division.

Lentiviruses such as HIV-1 virus can be engineered into gene delivery vectors with great potential for human gene therapy [298]. HIV-based lentivectors have been designed and submitted to numerous changes and optimisations over time to minimize the risk associated with the pathogenicity of the parental virus and to produce safe and efficient gene transfer vectors. The coding regions of the gag, pol and env genes, as well as the auxiliary genes (vif, vpr, vpu, env, and nef) and the regulatory genes tat and rev, are removed from the vector sequence leaving only a backbone of the 5' and 3' LTRs and the packaging signal (ψ). The transgene expression cassette can then be inserted between the LTRs to enable its packaging into the new viral particles. The packaging functions (gag, pol, and env) and the rev gene necessary for efficient expression of gag and pol are then provided in *trans* in the cell, split into different transcriptional units to reduce the possibility of replication-competent virus production through recombination. Furthermore, the 5' LTR is modified by replacing the U3 region, which contains the promoter and enhancer elements, with a chosen promoter such as RSV or CMV to abolish the transcriptional dependence on tat [299]. Finally, another modification present in self-inactivating (SIN) vectors involves a deletion in the U3 region of the 3' LTR, including the TATA box [300]. This deletion is transmitted to the 5' LTR of the proviral DNA during reverse transcription, abolishing the transcriptional activity of the LTR and the possibility to generate vector genomic RNA in the target cell (**Figure 32**). This deletion also prevents transcriptional interference between the 5' LTR and either host chromosomal genes or internal promoters driving transgene expression.

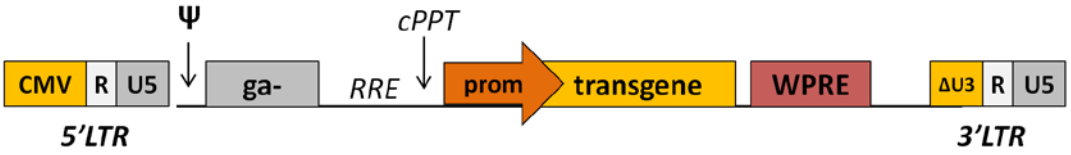

Figure 19: self-inactivating (SIN) lentiviral vector

SIN vectors contain modifications in the LTR regions: the U3 region in the 5' LTR is replaced by a chosen promoter and a deletion in the U3 region of the 3' LTR (ΔU3) abolishes the transcriptional activity of the LTR.

V. Exon skipping for Duchenne muscular dystrophy

V. 1. The Dystrophin gene and the disease

The human dystrophin gene (DMD) is the largest gene in the human genome, spanning 2.4 Mb of DNA on the human X chromosome and containing 79 exons [301]. This gene encodes a 14 kb, low abundance transcript that is predominantly expressed in skeletal and cardiac muscle, and at lower levels in smooth muscle and brain tissue [302]. The transcription of the DMD gene, which lasts approximately 16h [78], is activated during the process of myoblast differentiation when myoblasts begin to fuse into multinucleated myotubes [303]. The large dystrophin protein (427 kDa), located beneath the muscle sarcolemna, is a structural protein that plays a major role in the integrity of muscular fibres as it links the internal actin cytoskeleton to the extracellular matrix via the DAPC, which enables the resistance of the muscular cell against contraction induced stress [304].It contains four domains: an actin-binding domain at the N terminus, a central flexible rod domain composed of 24 spectrin-like repeats [305], a cysteine rich domain, and a C-terminal domain that binds to the dystrophin-associated protein complex (DAPC) at the sarcolemma (**Figure 33**).

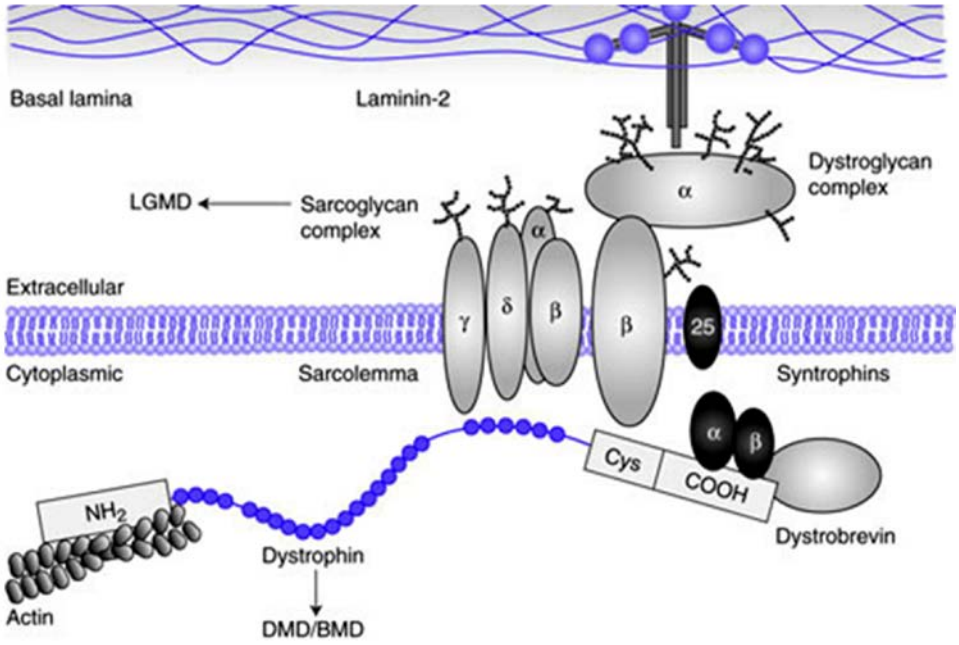

Figure 20: the dystrophin-associated protein complex links the internal cytoskeleton to the extracellular matrix in muscle cells

Duchenne Muscular Dystrophy (DMD) is an X-linked recessive disease caused by loss-of-function mutations in the dystrophin gene. It is the most severe and most common form of muscular dystrophies affecting 1 in 3500 males. As the dystrophin is missing from the membrane of muscular cells in DMD patients, the DAPC is destabilized and the integrity of the muscle fibre is lost during muscle contraction, which leads to progressive degeneration and ultimately necrosis of the myofibers [306-307]. The cardiac muscle, the vascular smooth muscle and the central nervous system are also affected. The onset of the disease usually occurs around 3 to 5 years of age, and progresses rapidly: the muscle weakness and degeneration drive the DMD patients to use a wheelchair in their early teens, and usually culminates in their death around the age of 25 due to cardiac or respiratory failure [304].

The majority of DMD mutations abolish the dystrophin mRNA reading frame, preventing the synthesis of functional dystrophin protein. More than 60% of loss-of-function DMD mutations are large deletions of one or more exons, usually regrouped in a hotspot region between exons 45 and 55. Other mutations are non-sense mutations, insertions, or exonic and intronic point mutations causing aberrant splicing [308-309]. A milder form of the disease, Becker Muscular Dystrophy (BMD), is caused by in-frame mutations in the dystrophin gene that still allow the production of a partially functional dystrophin protein, even if in reduced amounts [310]. These BMD mutations mostly affect the central domain of the dystrophin constituted of 24 central repeats that provide a highly modular structure to the protein. Therefore, truncated proteins where one or several of these central domains are missing can still be functional as they maintain their essential N and C-terminal domains.

Animal model of the disease are available for study of the disease pathogenesis and for pre-clinical applications. A well-established mouse model of DMD, the *mdx* mouse, presents an out-of-frame mutation in dystrophin exon 23 that prematurely stops the protein synthesis [311]. However, the *mdx* mouse only has a mild phenotype compared to the human situation and another mouse model more representative of human DMD is available. This model called double-knockout (dKO) is null for both dystrophin and another muscle protein, utrophin, and therefore presents a severe phenotype, including severe cardiomyopathy [312-313]. Another mouse model, hDMD, carries an integrated copy of the full-length human dystrophin gene and therefore allows investigations on specific human exons on a small experimental model. However, as the hDMD mice

were engineered from a wild-type mouse background and thus still express the functional endogenous mouse dystrophin gene, this model can only be used for analysis at the RNA level [314]. Besides the murine models, a large animal model has been developed: the golden retriever muscular dystrophy (GRMD) dog, which is also dystrophin negative due to a point mutation, is an interesting model for preclinical studies as it enables to estimate more closely the doses necessary to obtain a particular improvement on muscle function before translation to the human situation [315].

There is currently no effective therapy for DMD and the current options consist mainly in symptomatic treatments to ameliorate the patient's condition. Novel experimental strategies are currently being investigated, for which these animal models prove very useful and informative.

V. 2. Gene therapy strategies for Duchenne muscular dystrophy

Various gene therapy strategies for the treatment of DMD are showing promising results. Gene replacement therapy aiming at reintroducing a functional version of the dystrophin gene using AAV or lentiviral vectors has made significant progress. The limited cloning capacity of the recombinant AAV vectors implies that the dystrophin gene cannot be delivered in its full-length form and truncated but functional versions of the gene must be used. These mini- and micro-dystrophin constructs that lack large portions of the central domain of the gene have been designed following the observation that some BMD patients with large genomic deletions of dystrophin present a very mild phenotype [316] (**Figure 35**). The feasibility of delivering small versions of dystrophin to treat DMD patients has been demonstrated in various animal models of DMD. Efficient mini-dystrophin gene transfer has been obtained in *mdx* mice through recombinant AAV injection locally into the muscle [317] or systemically [318] and has been shown to ameliorate the pathological phenotypes. Improvements were shown by systemic injection in the dKO mice, which led to widespread *de novo* dystrophin expression in skeletal and cardiac muscles [319]. However, although AAV vectors do not elicit a cellular immune response in mice, an immune response was developed following AAV-mediated gene transfer in the GRMD dog model [320-321], as well as in monkeys [322] and in humans [323-324]. Sustained dystrophin expression

was however achieved by inducing short-term immunosupression in the dystrophic dogs or in non- human primates at the time of treatment [322, 325-326].

These approaches therefore hold great promises but are still confronted with many barriers. First, the use of internally deleted dystrophin constructs can have limited biochemical efficacy. Moreover, the long-term transduction of all the different muscles of a DMD patient by systemic injection is still a challenge. Finally, there are some concerns that the reintroduction of a previously missing dystrophin gene product in DMD patients could be immunogenic and trigger a dystrophin-specific reaction, hampering the therapeutic benefits of the treatment [328].

V. 2. 1. exon skipping correction of Duchenne muscular dystrophy

Many current efforts for the treatment of DMD are focusing on correcting the endogenous dystrophic pre-mRNA by antisense-mediated exon skipping instead of replacing the gene. The viability of therapeutic exon skipping for DMD was evidenced by the observation that "BMD-like" shortened forms of dystrophin protein could still be functional. Moreover, the detection in some dystrophic muscles of revertant fibres that resume dystrophin protein expression by naturally skipping an exon has laid foundation for AON-mediated gene correction in DMD [329-331]. An out-of-frame point mutation or deletion located in the central domain of dystrophin could therefore be countered by skipping the corresponding or surrounding exon(s) in order to recover an in-frame mRNA producing a shorter but functional dystrophin protein, thus converting the severe DMD phenotype into a much milder BMD phenotype.

Correction of dystrophin mRNA through splicing modulation presents several advantages over classical gene replacement strategies. First, by correcting the endogenous gene rather than providing a new truncated copy this strategy highly diminishes the issue of a potential immunogenic reaction against the dystrophin product. Indeed, the mini-dystrophin constructs used for gene replacement strategies are artificial versions that contain fragments of the normal gene joined together at abnormal places. The proteins produced from these constructs can therefore present

abnormal epitopes and be recognised as new in the treated organism, which is likely to trigger an immune response. On the other hand, correcting the endogenous pre-mRNA by exon skipping generally aims to skip one or few exons of the normal gene and the restored Dystrophin proteins are then very similar to the ones produced in the rare revertant fibers found in many DMD patients and should not activate the immune system. No immune response against the Dystrophin produced by exon skipping has so far been detected, including in patients currently enrolled in clinical trials, and although this possibility still needs to be carefully monitored, the likeliness of an immune reaction to exon-skipped Dystrophin is low. Furthermore, if the mRNA is efficiently corrected through splicing modulation the new truncated product should accumulate for a relatively long period of time as the dystrophin protein has been shown to have a long half-life [333]. Finally, the small snRNA-based antisense cassettes can be easily packaged into AAV vectors, abolishing the difficulties associated with packaging the large dystrophin gene in gene replacement strategies.

Several pre-clinical studies have demonstrated the therapeutic potential of the exon skipping strategy for DMD [167] and various observations regarding potential improvements of the technique emerged. Efficient restoration of dystrophin expression was obtained using various antisense chemistries in the *mdx* mouse model [334-336], in the canine model [337-338] and in DMD patient muscle cells in vitro [339]. It has been suggested that exonic splicing sequences may constitute better AON targets than introns, and AONs blocking different ESE sites were shown to be more effective in inducing the specific skipping of various human dystrophin exons compared to their intronic counterparts [340-341]. Furthermore, two AONs targeting different splice sites, with similar or different chemistries, can also be combined to induce the exclusion of exons. This double targeting strategy has been tested in human DMD pre-mRNA and a significant improvement in exon skipping efficiency was observed when two splicing regulatory sequences within one exon were simultaneously blocked [342-343].

The possibility of realising multiple exon skipping on the dystrophin pre-mRNA using a cocktail of AONs was also demonstrated, which would enable to increase the proportion of patients that could be treated with one same drug, or cocktail, and would circumvent the actual need for a personalised medicine for virtually each patient [344]. This prospect is interesting since, as mentioned before, a great number of DMD mutations are regrouped in a hotspot region between exons 45 and 55. In theory, the

simultaneous skipping of exons 45 to 55 could treat over 60% of patients and would create a deletion associated with a mild phenotype [345-346]. The validity of this approach has been demonstrated using cocktails of synthetic AONs in the dog model of DMD to induce the skipping of four exons simultaneously [337]. However skipping up to 11 exons simultaneously in human dystrophin is still challenging and has not be successful as yet [342, 347].

These promising results provided the background for two recent phase I/IIa clinical trials of antisense-mediated exon skipping in DMD patients. Both selected antisense oligonucleotides that target exon 51 of dystrophin pre-mRNA, the skipping of which would constitute a suitable therapy for 13% of all DMD patients [348].

The first study, carried out by the Dutch consortium from the Leiden University medical School in association with the com pany Prosensa, used a 2-'O-MePS AON chemistry (PRO051). Four DMD patients treated intramuscularly showed specific exon 51 skipping and *de novo* dystrophin expression, associated with the production of dystrophin protein at up to 12% of normal levels four weeks after a single injection of 0.8 mg PRO051 in the tibialis anterior muscle [349]. This encouraging trial led to repeated systemic administration studies where four groups of DMD patients received escalating doses of PRO051 weekly for 5 weeks. Specific exon 51 skipping was observed in patients treated with 2 mg or more per kilogram and novel dystrophin expression was detected in muscle fibers of most of the patients to up to 15.6% of the healthy muscle [350].

The second clinical trial was led by the MDEX consortium in collaboration with AVI Biopharma who selected PMO antisense oligonucleotide targeting exon 51 (AVI-4658). Seven patients were injected into the extensor digitorum brevis muscles at escalating doses and the patients treated with the highest dose of 0.9 mg showed an increased dystrophin expression to up to 32% of normal levels [351]. This consortium then completed a systemic dose-escalation study where 19 patients received weekly intravenous administrations of AVI-4658. Exon 51 skipping was detected in all patients and novel protein expression was observed for patients treated with 2 mg per kg and above, reaching up to 18% of the protein levels in the normal muscle [352].

These clinical trials constitute the first demonstration that successful restoration of dystrophin protein can be achieved by antisense-mediated therapy in patients with DMD. However the systemic delivery of AONs faces the issues of poor cellular uptake and relatively rapid degradation in the cells, requiring regular re-administration. More importantly, the restoration of dystrophin expression in non-skeletal muscle, especially in cardiac muscle, is achieved only with low efficiency following systemic delivery of AONs. Other approaches using snRNAs to vectorize AONs are therefore being investigated to optimize this antisense-mediated strategy.

V. 2. 2. snRNA-mediated exon skipping

It has been reported that modified U7 or U1 snRNAs packaged into AAV vectors and containing a double-target antisense for specific exon 23 skipping in the *mdx* model could efficiently correct the dystrophin transcript and restore high and stable levels of functional proteins after a single injection in skeletal muscle [266, 353]. The same strategy has been shown to induce sustained production of dystrophin in all the muscles, including in the heart, following a single systemic administration of the vector in the tail vein of *mdx mice* [239, 354] or of dKO mice [267]. The efficacy of chimeric snRNAs packaged in AAV vectors was also efficient to rescue dystrophin expression in the brain of *mdx* mice [355], or in the heart of GRMD dogs [270-271] by local injections. The proof-of-concept of this snRNA-mediated exon skipping strategy was also demonstrated on human sequences since, despite strong sequence homology between the human and mouse dystrophin genes, an antisense sequence targeting the same pre-mRNA sequence in both species can result in different exon skipping efficiencies [356]. Chimeric snRNAs carrying antisense sequences against the 5' and 3' splice sites of exon 51 efficiently induced high levels of exon 51 skipping and rescued dystrophin synthesis in myoblasts from patient with deletion of exons 48 to 50 [210].

A downside to the antisense-mediated modulation of pre-mRNA splicing is the need to design and validate a new antisense sequence for each different exon to be skipped, which require the assessment of several different sequences every time and can be fastidious and relatively expensive. An alternative to the conventional modified snRNA approach has been proposed to counter this issue, based on the principle of bifunctional

oligonucleotides using the TOSS technology. It has been recently shown that tailed U7 snRNA, carrying a binding sequence for the target dystrophin exon 51 and a splicing silencer sequence moiety that is recognized by hnRNP proteins, can be packed into lentiviral vectors and induce efficient exon 51 skipping associated with novel dystrophin expression in myoblasts from DMD patients, and in vivo in the hDMD mouse using AAV vectors [268]. Whilst the binding sequence itself probably plays a role in this result, as an antisense oligonucleotide composed of the same sequence targeting the same exonic site has previously been reported to induce exon skipping [357], it was also apparent that the addition of the splicing silencer sequence to this antisense moiety greatly increased the exon skipping efficiency. This observation suggests that the splicing silencer tail is sufficient to induce good levels of exon skipping on its own, even when the binding sequence of the bifunctional oligonucleotide does not mask an important splice signal. This strategy eliminates the need for screening numerous sequences when selecting an efficient steric-blocking oligonucleotide since, as long as the RNA binding moiety hybridize with the chosen exon, the silencer binding moiety will induce the skipping of this exon regardless of the binding site. The design of exon skipping AONs is therefore greatly facilitated by this approach, which is easily adaptable to many different exons.

Another difficulty faced by this AAV-based delivery of antisense sequences is the issue of personalized medicine for DMD patients: as many different mutations in the dystrophin gene can cause DMD, a different AON treatment must be designed for virtually every mutation. The adaptation of AAV-mediated delivery to multiexon skipping would in theory imply the production and co-administration of numerous viral vectors, one for each snRNA-antisense to be used in the cocktail, which is not realistic for clinical purposes. However a recent study showed that multiexon-skipping can be achieved using multiple U7 snRNA cassettes combined into one single AAV vector. Three U7snRNA targeting exons 45, 46, and 47 of human dystrophin were introduced sequentially in a AAV vector and hDMD mice injected intramuscularly showed efficient skipping of these three exons, with an efficiency almost as high as that of single exon skipping [269]. These results are very promising for DMD therapy as they give a proof-of-concept for AAV-mediated multiexon skipping in vivo, which would enable to associate the high-efficiency of AAV vectors for muscle targeting with the possibility of optimizing a single drug to treat a high number of patients.

The difficulty of targeting all muscles in the body still represents one hurdle that prevents AON-directed exon skipping to reach its full potential for the treatment of Duchenne muscular dystrophy. In the case of AAV-mediated delivery of antisense sequences, it is currently estimated that doses of more than 10^{13} vector genome per kg are required to transduce all the muscles in a mouse model [358], which by extrapolation would represent very elevated doses for a DMD boy , likely to trigger an immune response. It is therefore of crucial importance to lower the dose of vector necessary to obtain sustained therapeutic effects. Identifying the appropriate AAV serotype, such as using AAV6, AAV8 or AAV9 that demonstrate natural tropisms for muscle and are the most efficient for crossing the blood vessel barrier for systemic gene transfer [318, 358-359], can help increasing the transduction efficiency of the target tissue and diminish the required doses for treatment. The genetic engineering of AAV vectors, including modification of their capsid proteins, can provide immune evasion for the vector and therefore increase transduction efficiency [359-360]. The work described in this thesis focuses on the development of improved vector systems that would enable to maximise the amount of antisense sequences delivered to a target tissue and would therefore allow a reduction of vector dose.

VI. Objectives of the study

The objective of this project is to find new and improved ways of modulating pre-mRNA processing and splicing. As mentioned earlier, several studies have now demonstrated the efficacy of chimeric snRNAs to induce therapeutic exon skipping in animal models of Duchenne Muscular Dystrophy (DMD) or in cells from Duchenne patients. The main limitation of this approach for clinical translation remains the large amount of snRNA expressing vector which needs to be produced and administered to patients in order to restore therapeutic levels of dystrophin protein. These high doses of snRNA-carrying AAV vectors are potentially toxic for the patients and are currently challenging to produce in vector manufacturing facilities. Improving the therapeutic index of these vectors is therefore of critical importance to allow reducing the vector dose for clinical application and to avoid the need for re-administration. This can be achieved partly by using the appropriate AAV serotype for muscle and heart delivery, but also by optimizing the expression level of chimeric snRNA shuttles.

Here we try to maximize the amount of active antisense sequences delivered to a target cell or tissue and obtain robust and long-term levels of expression using more efficient and tissue-specific expression cassettes. For this purpose we will explore different approaches:

- we will first try to optimize the AAV mediated delivery of antisense sequences for dystrophin exon skipping by using self-complementary AAV vectors (scAAV), which have been demonstrated to present a faster and enhanced transduction efficiency compared to their single-stranded counterparts. U7 snRNA cassettes carrying antisense sequences for mouse dystrophin exon 23 skipping will be packaged in scAAV vectors and the efficiency of this system to increase both the level of antisense sequences readily available in the transduced cells and the resulting skipping of the target mRNA will be tested in mouse skeletal muscle cells cultures.

- Secondly, we will try to increase the level of expression of chimeric U7 snRNA cassettes in a tissue-specific manner through the use of enhancer elements. A muscle and heart-specific enhancer, MHCK, will be assembled and inserted directly upstream of the natural U7 promoter in the snRNA cassette designed for mouse dystrophin exon 23 skipping. The ability of this strong transcriptional enhancer to drive high-level and

tissue-specific expression of modified U7 cassettes will be tested *in vitro* in mouse muscular cells and *in vivo* in the *mdx* mouse model.

- As a third goal in this study, we aimed to identify improved ways of modulating pre-mRNA processing by creating new small RNA cassettes designed for splicing alteration. For this purpose, we explored the possibility of using small nucleolar RNAs (snoRNAs) as novel molecular platforms for the intracellular delivery of antisense sequences. Some C/D box snoRNAs naturally associate with pre-mRNAs and regulate their alternative splicing through complementary base pairing. By replacing native antisense elements of such snoRNAs with antisense sequences for mouse dystrophin exon 23 skipping, we will test the efficiency of this new snoRNA-based antisense delivery system to trigger exon 23 skipping when delivered by AAV vectors in mouse cultured myotubes.

- Finally, the natural property of C/D box snoRNAs to direct the 2'-O-ribose methylation of cellular RNAs on precise nucleotides will be investigated as another strategy to modulate splicing. A human snoRNA will be engineered to specifically guide the 2'-O-ribose methylation of a chosen adenosine branch point in a target pre-mRNA in order to inactivate branch point recognition by the splicing machinery and induce the skipping of the downstream exon. This approach will first be designed to modulate the splicing of an exon skipping reporter construct based on the splicing-dependent activation of a luciferase gene, ultimately aiming to apply this approach for dystrophin exon skipping.

RESULTS

I. Optimising the levels of modified U7 snRNA expression to improve Dystrophin exon skipping [1]

The delivery of antisense sequences inserted into small nuclear RNA cassettes by viral vectors requires further optimisation before being suitable for clinical translation, as considerable amounts of vector would still need to be administered to DMD patients in order to restore therapeutic levels of Dystrophin protein in all muscular tissues. In an attempt to maximise the therapeutic index of such vectors, we first tried to use self-complementary AAV (scAAV) vectors to deliver antisense-expressing U7 snRNA cassettes for dystrophin exon skipping in muscular cells. ScAAV vectors have previously demonstrated enhanced transduction efficiencies over single-stranded AAV vectors, and this increased efficiency could enable to obtain elevated levels of transgene expression at minimal doses of vector, which is a critical property for clinical applications. The small size of U7 snRNA cassettes enables their packaging as inverted repeat genome in scAAV vectors and we therefore expressed the U7Dtex23 cassette for mouse dystrophin exon 23 skipping in such vectors and tested their potential to increase the level of antisense sequences readily available in transduced mouse cells for splicing modulation and to improve Dystrophin restoration in muscular tissues.

Another optimisation strategy explored in this chapter consists in using muscle-specific enhancer elements to drive the expression of the U7Dtex23 cassette, aiming to increase the amount of transcripts produced from a single expression unit. The identification and use of such enhancer elements capable of inducing high-level of transgene expression in both cardiac and skeletal muscle could enable to maximise the amount of product expressed in muscle tissues and to improve the skipping of mouse dystrophin exon 23 whilst diminishing the vector dose needed to be administered for therapeutic correction of dystrophin mRNA splicing.

[1] These results have been published in part in **361. Eckenfelder, A, Tordo, J, Babbs, A, Davies, KE, Goyenvalle, A and Danos, O, The Cellular Processing Capacity Limits the Amounts of Chimeric U7 snRNA Available for Antisense Delivery. *Mol Ther Nucleic Acids*, 2012. 1: p. e31.** This paper is shown in Annex 1.

I. 1. Assessing the best time point to target dystrophin pre-mRNA for exon skipping

In order to test and compare the efficiency and tissue-specificity of the various optimised small RNA-antisense cassettes delivered by AAV vectors, we first needed to identify the optimal time to transduce the C2C12 mouse muscle cell line used in this study, especially regarding the differentiation state of the cells at the time of transduction, in order to target dystrophin pre-mRNA for exon skipping.

It has not yet been clearly established at what stage of the myogenesis process the dystrophin gene (dmd) starts to be transcribed and spliced. Some studies suggest that dmd is not expressed in mononucleated myoblasts and that dystrophin transcription is only initiated as cultured myoblasts begin to differentiate and fusion into multinucleated myotubes [303, 362-363]. Conversely, other studies show that in mouse myogenic cell lines such as C2C12, the dystrophin gene is already activated and expressed at the myoblastic stage before being strongly up-regulated in the course of in vitro myogenesis, increasing to up to 40 times during differentiation into myotubes [364]. Finally it has been estimated that approximately 16 h are required for the polymerase to transcribe the entire dystrophin gene in myogenic cultures and that the transcripts already start to be spliced at the 5'end before transcription is complete [78]. It is therefore really important to determine more precisely when the best time point for C2C12 cells transduction is in order to achieve efficient dystrophin exon skipping, taking into account the time needed for dystrophin to be transcribed and the moment it becomes spliced.

To determine more precisely the kinetics of dmd gene transcription and to assess when the transcript levels reach a steady state in cultured myotubes in the course of in vitro myogenesis, we assessed the dystrophin mRNA levels in the C2C12 muscle cell line by RQ-PCR (Real-time Quantitative Polymerase Chain Reaction) at different time points during myogenesis and differentiation. C2C12 myoblasts were amplified and differentiated for up to 14 days and total mRNA was extracted at different time points, from 1 day of proliferation (prolif+1d) to 14 days of induced differentiation (diff+14d).

A Taqman assay was then performed on reverse-transcribed mRNAs using primers and probes specific for the mouse dystrophin mRNA (see Material and Methods), and the relative amounts of transcript detected were normalised at each time point with the endogenous 18s rRNA expression levels (**Figure 39**). The 18s ribosomal RNA was

considered as a useful endogenous control as the rRNA transcripts are produced by a distinct polymerase [365] and are constitutively expressed. Furthermore, it has been shown that rRNA expression levels are less likely to vary between different tissues, stages of development or in response to variations of treatment conditions than mRNAs, such as β-actin or GAPDH that are also traditionally used as internal controls [366-367].

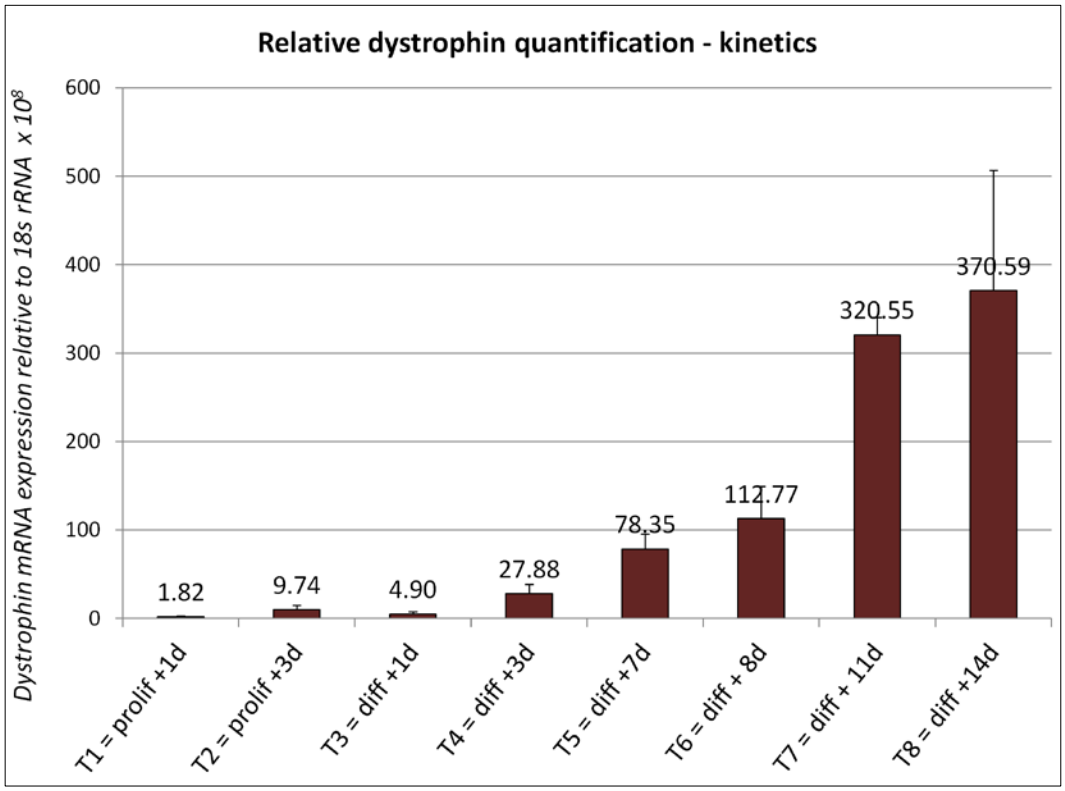

Figure 21: kinetics of dmd gene transcription in the course of in vitro myogenesis in C2C12 muscle cell line.

The relative dystrophin transcript levels were normalised to the endogenous 18s ribosomal RNA and multiplied by a factor of 10^8 for readability. Each time point represents an average of 3 independent experiments. Error bars are shown as mean ± SEM

The results presented here suggest that the dystrophin transcript is close to undetectable at the myoblastic stage whilst the cells are still proliferating, including after one day of induced differentiation. We note that the level of dystrophin mRNA seems to be higher after 3 days of proliferation than after one day in differentiation conditions, which probably represents an artefact of the experiment as the myoblasts reached a confluent

state in culture, inducing them to spontaneously start the differentiation process. We observe that the dystrophin transcript begins to be up-regulated after the muscular cells experienced 3 days of induced differentiation and started to fuse into myotubes. At later time points, the dystrophin levels steadily increase in time until 14 days after differentiation, when the transcript production seems to finally slow-down and plateau in some of the cultures.

As we want to target the dystrophin pre-mRNA whilst it is still being transcribed, considering that the transcription and splicing of this precursor occur simultaneously, we then hypothesised that transducing the myotubes around day 7 of induced differentiation, when the dystrophin transcript levels didn't reach a steady state yet, should enable us to target and alter the dystrophin pre-mRNA splicing. We therefore adapted the transduction protocol of C2C12 muscular cells based on these dystrophin kinetics results and decided to transduce the cells at 7 days post-differentiation for our subsequent dystrophin exon skipping experiments.

I. 2. Optimal levels of antisense sequences using self-complementary AAV vectors

I. 2. 1. Principle and objective

Adeno-associated viruses need to convert their single-stranded DNA (ssDNA) genomes into double-stranded DNA (dsDNA) prior to expression, which can be achieved either by synthesis of the complementary DNA strand [368] or by base pairing of the plus and minus strands from two co-infecting viruses into the same cell when the AAV vectors are administered at high doses [369]. However, this ssDNA to dsDNA conversion is a slow and rate-limiting process during AAV transduction of a target cell, delaying gene expression and hindering efficiency of the viruses in many cell types. AAV vectors containing a genome of half the size of the wild-type AAV genome are occasionally able to package their DNA as dimer molecules that can directly fold into double-stranded DNA [370]. These self-complementary AAV vectors (scAAV) circumvent the requirement for DNA synthesis or base-pairing between multiple vector genomes and therefore present an enhanced transduction efficiency compared to their single-stranded counterparts by inducing a faster and higher level of transgene expression.

It is therefore possible to produce scAAV vectors simply by decreasing the genome size to approximately 2.3 kb or less so that the newly created dimeric inverted repeat DNA does not exceed the maximal packaging capacity of 4.7 kb of wild-type AAV. The vectors produced through this strategy are however not only composed of dsDNA but are usually a combination of ssAAV and dsAAV.

In order to efficiently and predominantly produce self-complementary particles, two different groups have engineered the AAV genome in a way that forced the creation of dimeric genomes. A first approach to generate self-complementary rAAV vectors was to delete the terminal resolution site (trs) of one ITR , leaving the second ITR intact [372]. The second technique involved the deletion of the D-sequence, which constitute the packaging signal, along with removal of the trs on one ITR [373] (**Figure 41**). In both cases, the dimer genomes fail to be resolved into monomers and lead to the production of a high proportion of scAAV vector.

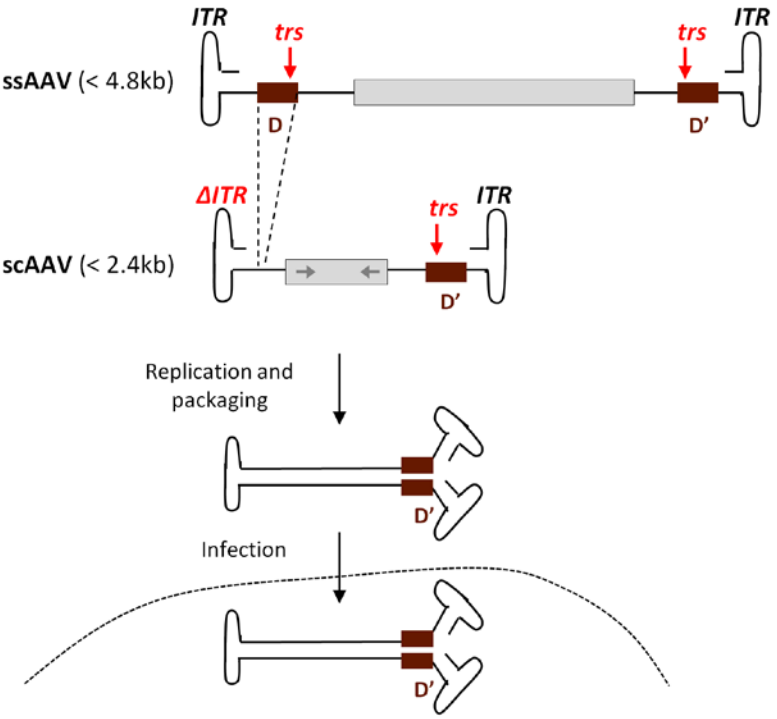

Figure 22: engineering of scAAV vectors

Self-complementary vectors are generated through deletion of the terminal resolution site (trs) of one ITR, with or without deletion of the packaging signal (D) on the same ITR, and leaving the second ITR intact for replication and packaging functions.

The effect of such mutations is that, when the replication that started at the wild-type ITR reaches the mutated ITR, the Rep protein fails to recognize the trs and therefore does not cut the genome at this site to form single-stranded DNA. The replication instead continues past the mutated ITR without terminal resolution and back across the genome to create the dimer DNA until it reaches the initiating wild type ITR. The final product of replication is therefore a dimeric inverted repeat genome that contains the mutated ITR in the middle and two wild-type ITRs at each end [374]. This self-complementary molecule can initiate normal rounds of replication from its wild-type ITR and generate several hairpin-like dimeric genomes that can fold as dsDNA and can be efficiently packaged in the viral particles through their remaining wild-type ITR.

The counterpart to the increased efficiency of scAAV vectors is that the entire vector genome, including the ITRs, has to be smaller than 2.4 kb to meet the requirements for efficient AAV packaging. This very limited cloning capacity prevents the use of scAAV vectors for many gene replacement approaches. Yet, a number of applications based on the use of small transgenes, such as antisense cassettes or small interfering siRNAs can benefit from the enhanced transduction efficiency of these vectors. Self-complementary vectors have demonstrated increased and faster transduction in a variety of tissues [374]: they resulted in more robust and stable transgene expression and more effective DNA circularization in mouse liver, brain and muscle [372-373]. The delivery of small therapeutic agents, such as ribozymes and siRNAs, has also shown improved efficacy using scAAV [375-376]. The systemic administration of scAAV vectors carrying a mini-human factor IX (hFIX) transgene in mice induced a 20-fold increase in hFIX expression in liver compared to conventional ssAAV vectors [377]. The same construct also efficiently transduced the liver of nonhuman primates after systemic injection and allowed the restoration of therapeutic levels of hFIX at a much lower dose than usually needed with ssAAV [377-378]. Finally, a self-complementary AAV vector was recently used to systemically deliver U7snRNA cassettes designed for mouse dystrophin exon 23 skipping and successfully induced the restoration of high levels of dystrophin protein in the severe dKO mouse model [267].

Considering the small size of antisense-expressing U7 snRNA cassettes, it is possible to package them as an inverted repeat genome in scAAV vectors and we therefore decided to use this type of vector in our study, aiming to increase the level of antisense sequence readily available in the transduced cells for splicing modulation.

I. 2. 2. Self-complementary AAV vector to express the U7Dtex23 cassette and improve transduction of myotubes

The read-out in this study is the efficiency of dystrophin exon 23 skipping in mouse skeletal muscle cells as a result of modified U7 snRNA expression. The snRNA cassette used, U7Dtex23, is designed to induce exon 23 skipping in mouse DMD pre-mRNA by specifically recognising two different target sequences: the donor splice site of intron 23 and the branch point of intron 22 [266]. The U7Dtex23 cassette, containing its own promoter and terminator sequences, was amplified by polymerase chain reaction (PCR) and introduced between the inverted terminal repeat sequences of a self-complementary AAV vector (**Figure 42**, and see material and methods). The construct, termed AAV-**sc**-U7Dtex23 was verified and validated by restriction enzymes digestion and sequencing and the efficiency of this system to express high levels of antisense sequences and to increase the correction of the dystrophin transcript compared to ssAAV was tested in mouse skeletal muscle cells cultures.

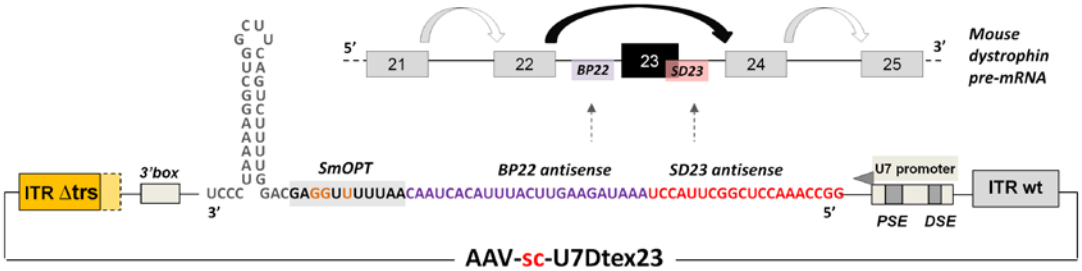

Figure 23: AAV-sc-U7Dtex23 vector for mouse dystrophin exon 23 skipping

U7Dtex23 snRNA contains the optimised SmOPT sequence and a double target antisense sequence directed against the branch point of intron 22 (BP22) and the splice donor of exon 23 (SD23) on the mouse dystrophin pre-mRNA [266]. The U7Dtex23 snRNA cassette, represented here from 3' to 5' and containing the U7-specific promoter and 3'box at each end, is inserted between the ITR of a self-complementary AAV vector, where the 3' ITR is deleted of the terminal resolution site (trs) sequence.

We first compared different AAV serotypes, AAV2/8 and AAV2/5, for their efficiency to transduce C2C12 muscular cells in culture and observed that the AAV2/5 viral serotype was more efficient at expressing a GFP transgene in differentiated myotubes when administered at 7 days after differentiation and at a MOI of 10^4 vg/cell (data not shown). We therefore selected this serotype for all the *in vitro* AAV transduction experiments in this study. We then produced AAV2/5-**sc**-U7Dtex23 and conventional AAV2/5-U7Dtex23 viral vectors by triple transfection of 293T cells and purified the harvested viral particles by affinity chromatography on an AVB sepharose column. The viral titer of these vectors assayed by quantitative PCR enabled to estimate a multiplicity of infection (MOI), corresponding to the number of viral genomes administered per cell, with which to transduce the cells. Cultured C2C12 cells were then transduced at 7 days post-differentiation, according to the results of the kinetics experiment, with either of the AAV2/5 viral preparations at MOI 10^5 or 10^6 particles per cell.

We also included a control self-complementary AAV vector expressing a non-relevant sequence to assess if the mere transduction of C2C12 cells with such vectors would impair the dystrophin splicing. We chose the wild-type (wt) U7 snRNA carrying the natural U7 antisense, and inserted it in the scAAV vector to produce AAV2/5 particles. C2C12 cells were transduced with this vector in parallel with the other transduction experiments. Total RNAs were extracted from transduced cells at 7 days post-transduction and we compared the efficiency of scAAV and conventional AAV vectors to express the U7Dtex23 cassette and to induce dystrophin exon skipping.

- Northern blot analysis of U7Dtex23 expression in transduced cells

We first assessed the RNA extracts from the transduced cells for U7Dtex23 expression by Northern blot in order to match the amount of snRNA actually expressed in the cells with the exon skipping efficiency observed. Total RNAs were loaded on denaturating gels and the resulting blots were hybridized with a probe specific for the Dtex23 antisense sequence expressed in the U7 cassette. An additional probe specific for the endogenous U2 snRNA was used in order to normalise the signals obtained (**Figure 43**).

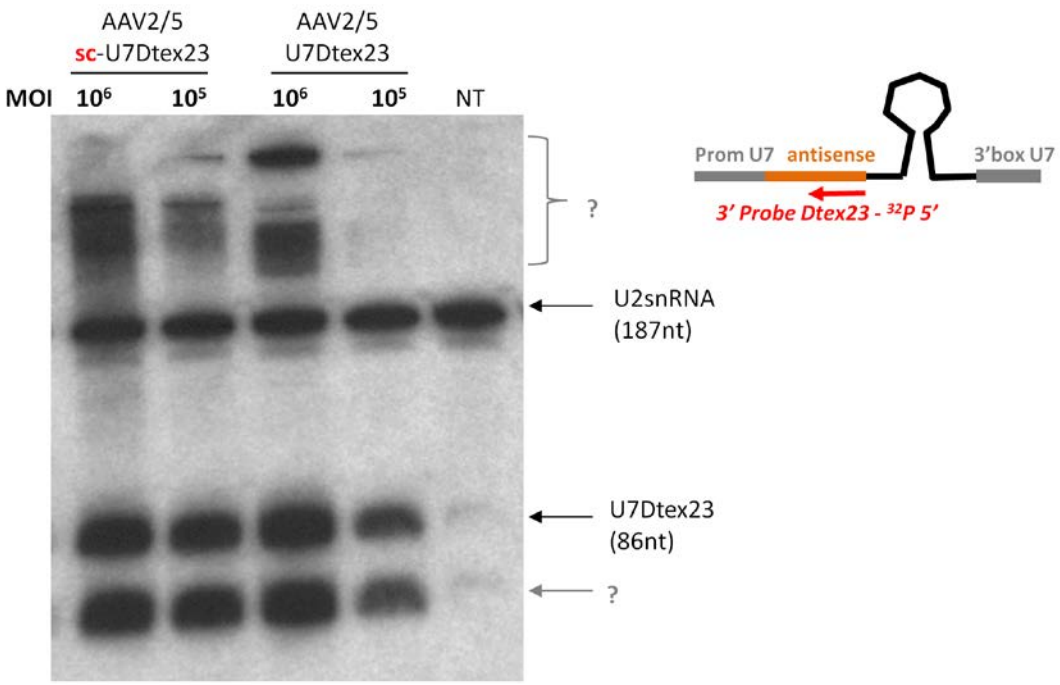

Figure 24: northern blot analysis of U7Dtex23 snRNA expressed in transduced mouse C2C12 cells

The U7Dtex23 chimera is 86 nucleotides long and the wild-type U2 snRNA is detected at 187 nt. A schematic representation of the U7Dtex23 cassette is shown on the right hand side, indicating the approximate position of the ^{32}P end-labeled probe recognizing the Dtex23 antisense sequence in the U7 snRNA. A second band of approximately 65 nt also reacts with the Dtex23 probe (bottom question mark), as well as some high molecular weight species (top question mark). NT represents non-transduced cells. A 15% TBE-Urea gel was used.

The U7Dtex23 transcript was readily detected at 86 nt in all the transduced samples. As expected, the transcript was absent from the sample corresponding to the non-transduced cells (NT). The endogenous U2 snRNA signal indicated that approximately the same amounts of RNAs were loaded for each condition. A dose-response could be observed of the amount of U7Dtex23 transcript expressed in the cells depending on the dose of viral vector used as we could see a slightly stronger U7Dtex23 signal for the cells transduced at MOI 10^6 compared to that of the cells transduced at MOI 10^5. This difference of signal representative of the transcript expression was however not dramatic and suggested that saturating levels of U7 cassette could have been expressed in the cells following treatment at such high doses, thereby hampering the distinction of signal strength on the Northern blot.

It was not clear from this analysis whether the use of the self-complementary vector induced a stronger expression of the transgene than the normal AAV vector at comparable doses. At the highest dose of vectors administered (MOI 10^6), the U7Dtex23 signal obtained at 86 nt from each vectors seemed to be saturated and therefore prevented a comparison of the vectors efficiency. However, at lower dose the self-complementary vector seemed to increase the amount of U7Dtex23 expressed compared to the one observed with the conventional AAV vector (compare lanes 2 and 4). We could also detect a second band on the Northern blot**,** around 20 nucleotides shorter than the expected U7Dtex23 transcript, that also reacted with the U7Dtex23-specific probe. Furthermore, high molecular weight (HMW) species that seemingly increased with the dose were detected. These phenomena have to our knowledge not been described or explained before and will therefore be further investigated in this study (see results sections I. 3. and I. 5)

- Analysis of exon 23 skipping in mouse *Dmd* mRNA by nested PCR

In order to analyze the exon 23 skipping obtained on mouse dystrophin mRNA after transduction of AAV2/5-**sc**-U7Dtex23 or AAV2/5-U7Dtex23 viral vectors, the RNA extracts from these cells were reverse-transcribed and used in a nested PCR reaction, where two successive rounds of PCRs enabled to specifically amplify and detect the dystrophin transcript from exons 20 to 26, thereby differentiating mRNAs still containing exon 23 with mRNAs efficiently depleted of it (**Figure 44**).

The full-length dystrophin transcript containing exon 23 and amplified from exon 20 to exon 26 produces a 901 bp fragment, as illustrated on the right hand side of **figure 44a**. The non-transduced cells (NT), as expected, showed only a 901bp band corresponding to the complete dystrophin transcript from exon 20 to 26. Furthermore, the negative control AAV2/5-sc-U7wt displayed only a 901 bp product.

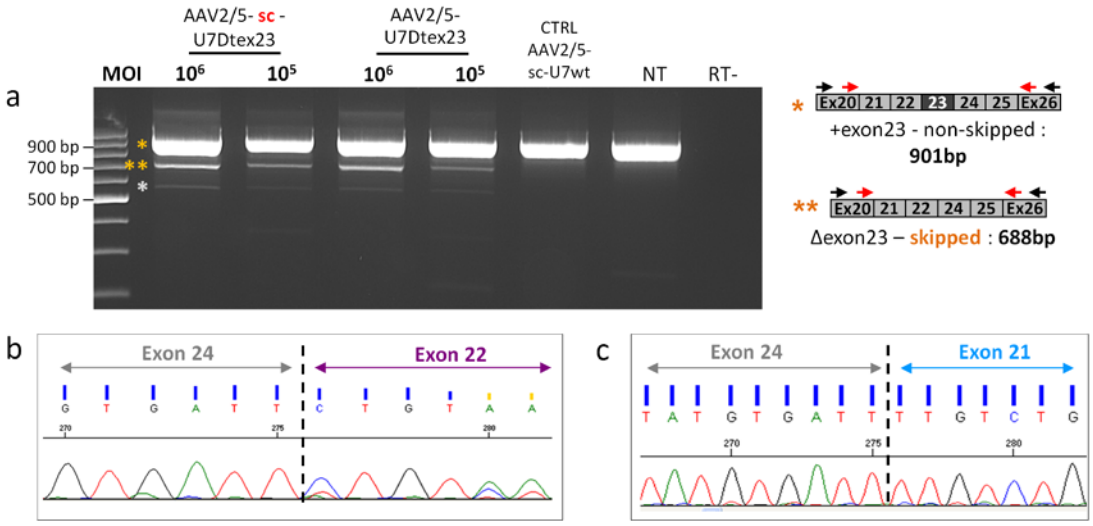

Figure 25: exon 23 skipping analysis in transduced mouse muscular cells

(a) Reverse transcription and nested PCR detection of dystrophin exon 23 skipping after transduction of C2C12 cells with AAV2/5-sc-U7DTex23 or AAV2/5-U7DTex23 at various MOI. A schematic representation of the PCR primers and products is shown on the right hand side. AAV2/5-sc-U7wt is included as a negative control of skipping. NT represents non-transduced cells and RT- is a control without reverse transcriptase. The native dystrophin mRNA is detected at 901 bp (yellow star) and the skipped mRNA yields a 688 bp product (double yellow star). A third product is detected on the gel (white star). (b and c) Sequencing of the 688 bp band (double yellow star) and of the third band (white star) extracted from gel (a), respectively. The exons are inversed on the chromatograms as the sequences were produced from a reverse primer.

In the samples transduced with AAV2/5 viral vectors specific for dystrophin mRNA we observed a band of 688 bp that corresponds to the dystrophin transcript depleted of exon 23 (double yellow star, **Figure 44a**). Gel extraction of the 688 bp band and sequencing confirmed that exon 23 was successfully skipped in these RNA species, creating a correct junction between exon22 and 24 (**Figure 44b**).

A smaller band was also observed on the gel in the samples where exon 23 skipping was induced by U7 snRNA expression (white star, **Figure 44a**). The sequencing of this band indicated that these RNA species are 542 bp long and correspond to a double exon skipping event where dystrophin mRNA lacks both exons 22 and 23, creating a clean junction between exons 21 and 24 (**Figure 44c**). Similar additional amplification products of ~542 bp have been reported previously [256] and their appearance seem to correlate with the advanced age of the culture. These mRNA species are out of frame and cannot yield a functional dystrophin protein.

We observed a good dose-response of the amount of skipped dystrophin transcript depending on the amount of viral vector administered on the cells. This was consistent with the amount of U7Dtex23 transcript expressed in the cells (**Figure 43**).

This experiment suggests that the use of self-complementary AAV vectors enabled to increase the amount of antisense sequences expressed in the cells and the associated level of dystrophin exon skipping compared to the use of a normal AAV vector. To determine more stringently the potential beneficial effects of the self-complementary AAV vector we sought to obtain more quantitative data.

- Analysis of exon 23 skipping in mouse Dmd mRNA by quantitative real-time PCR

The proportion of skipped dystrophin transcripts obtained in each condition was quantified using a Real-Time Quantitative PCR (RQ-PCR) assay in order to compare the efficiency of each vector.

A Taqman assay for dystrophin exon skipping was designed, where a set of primers and probe specific for exon 22-24 junction enables to specifically estimate the amount of transcripts depleted of exon 23. Another set of primers/probe specific for exon 22-23 junction detects full-length dystrophin transcripts containing exon 23 (see material and methods and **Figure 45**, right hand pannel). RNA extracts obtained from C2C12 cells transduced with AAV2/5-sc-U7DTex23 or AAV2/5-U7DTex23 at MOI 10^5 or 10^6 were reverse-transcribed and analysed by Taqman RQ-PCR.

In this assay the emission of a fluorescent signal is directly proportional to the amount of product formed at each PCR cycle and the Ct (cycle threshold) is defined as the number of cycles required for the fluorescent signal to cross the threshold level. A comparative CT quantitative analysis of exon 22-24 junctions (skipped dystrophin) and exon 22-23 junctions (non-skipped dystrophin) was performed on each sample, which involved the subtraction of the Cts for the 22-23 assay from the CTs for the 22-24 assay to obtain a delta Ct value for each condition. The delta Ct can then be used to calculate the relative amounts of skipped dystrophin RNA species relative to full length species (**Figure 45**).

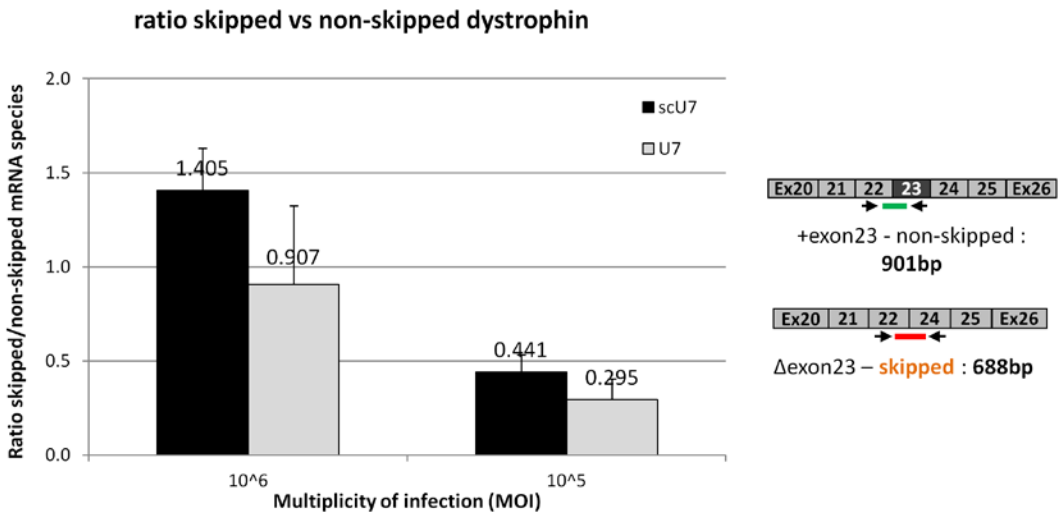

Figure 26: analysis of mouse exon 23 skipping by RQ-PCR

The graph represents the ratio between the exon 22–24 junction (skipped mRNA species) and the exon 22–23 junction (full-length mRNA species). Quantification was performed on mRNA extracts from C2C12 cells transduced with AAV2/5-sc-U7DTex23 (scU7) or AAV2/5-U7DTex23 (U7) at the indicated MOI and harvested at 7 days post-transduction. A schematic representation of the RQ-PCR primers and probe used is shown on the right hand side. Each condition represents an average of 5 independent experiments. No significant difference was observed between the various conditions, according to Student's t-test (P value > 0.05). Error bars are shown as mean ± SEM.

We could observe that the delivery of the U7Dtex23 cassette by a self-complementary AAV vector seemed to produce a slight increase in the proportion of skipped transcripts relative to full-length dystrophin species compared to that obtained from the use of conventional AAV vectors. Of note, we could observe that the ratios between skipped and non-skipped dystrophin species obtained by this Taqman assay seem much higher than those evaluated from the nested PCR realised previously on the same samples (**Figure 44**). Although we could not explain this observation, the moderate increase in skipped dystrophin mRNAs obtained with the scAAV vector was also detected by the nested PCR analysis. This beneficial effect of the self-complementary vector was however very moderate according to the RQ-PCR results and the difference in exon skipping efficiency was not statistically different. Moreover, the important error bars showed on each histogram, especially at the highest dose of virus, indicate that this experiment presents a relatively low reproducibility and that the efficiency of exon skipping can vary from one sample to another. We hypothesised that in order to assess the efficiency of the self-complementary AAV vectors to induce an accelerated

transduction compared to conventional single stranded AAV vectors it would be interesting to compare the exon skipping efficiency obtained from each vector at an earlier time-point after transduction.

- Analysis of scAAV kinetics of transgene expression and resulting exon 23 skipping by RQ-PCR

We therefore transduced more differentiated C2C12 mouse myotubes with AAV5-sc-U7Dtex23 and AAV2/5-U7Dtex23 viral vectors at a dose of 10^5 particles per cell. The cells were harvested at 4 days or at 8 days after transduction and we compared the level of exon skipping obtained in each condition by Taqman RQ-PCR on the various RNA extracts (**Figure 46**). This experiment aimed at assessing if, at 4 days after transduction, the scAAV vector enables to induce higher levels of dystrophin exon skipping than normal AAV vectors, reflecting their ability to improve transduction by accelerating the expression of the transgene in host cells.

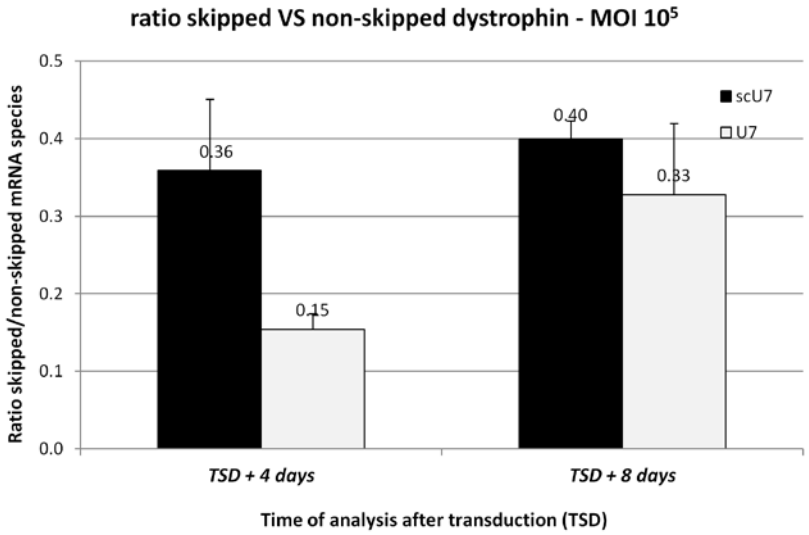

Figure 27: analysis of scAAV transgene expression rapidity reflected by induction of dystrophin exon 23 skipping

The graph represents the ratio between exon 22–24 junction and exon 22–23 junction in dystrophin mRNA species extracted from C2C12 cells transduced with AAV2/5-sc-U7DTex23 (scU7) or AAV2/5-U7DTex23 (U7) at MOI 10^5. The cells were harvested at 4 days or at 8 days post-transduction. Each condition represents an average of 3 independent experiments. No significant difference was observed between the various conditions, according to Student's t-test (P value > 0.05). Error bars are shown as mean ± SEM.

This result indicated that, at 8 days after transduction, the exon skipping efficiency obtained with the scAAV vector could be considered equivalent to the one obtained with the single-stranded AAV (ssAAV) according to the error bars, which reflected a comparable level of U7Dtex23 transgene expression in the host cells. However, at 4 days after transduction the scAAV vector importantly increased the ratio of skipped against non-skipped dystrophin transcripts, bringing the proportion of skipped species approximately two times higher than the one recorded with the regular AAV vector (left hand side histograms). Moreover, four days only after transduction the scAAV vector already generated a proportion of skipped mRNA species equivalent to the one obtained at 8 days post-transduction with the same vector or with the ssAAV vector, indicating that the levels of U7Dtex23 expressed in the transduced cells after 4 days was already at its highest. On the other hand, the ratio of corrected transcripts compared to full-length transcripts obtained from the conventional AAV vector only reached similar levels after 8 days of incubation, implying a slower rate of transgene expression. These data demonstrate that the rate of transgene production is accelerated with the use of scAAV vectors, inducing a high level of dystrophin exon 23 skipping early after transduction and much faster than that obtained form ssAAV vectors.

We further wanted to assess the efficiency of the self-complementary AAV vectors to induce an increase in transgene expression compared to conventional single stranded AAV vectors by comparing the exon skipping efficiency obtained from each vector at a lower dose. It is possible that the scAAV efficiency over the single-stranded vector is mostly noticeable at lower doses as the high amount of antisense sequence expressed at such high doses (MOI 10^5) could have reached a saturating level, which could trigger the antisense sequences to naturally associate two by two and form double-stranded structures spontaneously, hiding the potential effect of the self-complementary vector.

- Analysis of scAAV increase of transgene expression illustrated by dystrophin exon 23 skipping

To ascertain the potential beneficial effect of the self-complementary vector in increasing the amount of U7 snRNA transgene expression at a given vector dose, we transduced C2C12 muscular cells at a lower MOI, 10^3 viral genome per cell, to limit the number of available antisense sequences in each cells and prevent the spontaneous formation of double-stranded sequences. The cells were harvested and RNAs extracted at 4 days after transduction and a Taqman RQ-PCR quantitifying of the proportion of skipped transcripts against non-skipped dystrophin mRNA was realised (**Figure 47**).

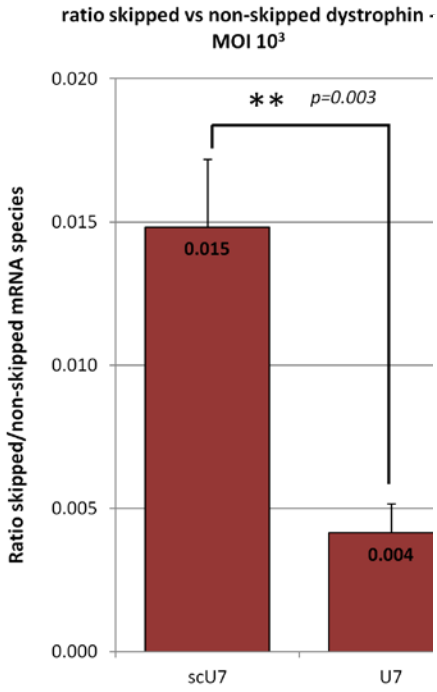

Figure 28: analysis of scAAV levels of transgene expression illustrated by induction of dystrophin exon 23 skipping

The graph represents the ratio between exon 22–24 junction and exon 22–23 junction in dystrophin mRNA species extracted from C2C12 cells transduced with AAV2/5-sc-U7DTex23 (scU7) or AAV2/5-U7DTex23 (U7) at MOI 10^3. The cells were harvested at 4 days post-transduction. Each condition represents an average of 5 independent experiments. A significant difference is observed as shown by the P value (P = 0.023), according to Student's t-test. Error bars are shown as mean ± SEM.

The delivery of U7Dtex23 cassettes by scAAV vectors induced a dramatic increase in the proportion of skipped dystrophin transcripts compared to the use of ssAAV vectors. This improvement was highlighted here by the use of lower doses of vectors and the difference between the samples treated with AAV-sc-U7DTex23 and the samples treated with AAV-U7DTex23 was statistically significant according to the Student's *t*-test (*P* value <0.05). This result illustrates the potential of scAAV vectors to increase transgene expression. We could observe that, for the same dose of vector, the scAAV is able to readily express higher amount of U7 snRNA cassette than the ssAAV, thereby inducing a three to four fold increase in the resulting proportion of exon 23 skipped dystrophin mRNA. This observation suggests that the use of self

complementary vectors could enable to decrease the dose necessary to obtain significant levels of exon skipping, which is of utmost interest in the context of DMD where high doses of vector are so far deemed necessary for therapeutic benefits. The self-complementary AAV vectors therefore constitute a very promising tool to express U7 snRNA cassettes for dystrophin exon skipping as they enable a faster and increased expression of chimeric snRNAs, resulting in a greater efficiency of exon 23 skipping in the mouse dystrophin mRNA in the C2C12 cell line.

I. 3. Analysing the origin of the extra-processed U7Dtex23 sub-products

On the Northern blots for U7Dtex23 transcripts detection after transduction of muscular cells we observed the generation of a second smaller transcript, about 20 nucleotides shorter than the expected U7Dtex23 size, that also reacted with the U7Dtex23-specific probe (**Figure 43**, bottom question mark). This surprising observation compelled us to investigate the origin and cause of this hyper-processing of the U7 snRNA cassette.

- Effect of the size of the transcript on U7 snRNA cassettes processing

Our first hypothesis was that this shortening could be triggered by the increased size of the U7 transcript when modifying its antisense sequence, as the U7Dtex23 is 86 nucleotides long, more than 20 nucleotides longer than the original mouse U7 snRNA transcript of 62 nucleotides. It is therefore possible that the intracellular processing machinery further process the modified U7 cassette to shorten the U7 transcript back to its natural size. To assess this hypothesis, the M23D(+07-18) antisense sequence targeting the donor splice site of dystrophin exon 23 [379] was inserted in the U7 snRNA cassette by fusion PCR to replace the Dtex23 antisense sequence in the scAAV vector (see Material and Methods). This U7-M23D cassette produces a U7 transcript that is 66bp long, close to the original U7 wt size, and therefore provides a good way to assess if the extra-shortening of chimeric U7 snRNAs is size-dependent. We transduced C2C12 myotubes with self-complementary AAV2/5 vectors expressing the U7M23D or the U7Dtex23 constructs at MOI 10^4 to 10^6 and harvested the cells at 7 days post-transduction. The extracted RNAs were used in a Northern blot assay using the Dtex23 probe and the M23D probe for U7M23D transcripts detection (**Figure 48**).

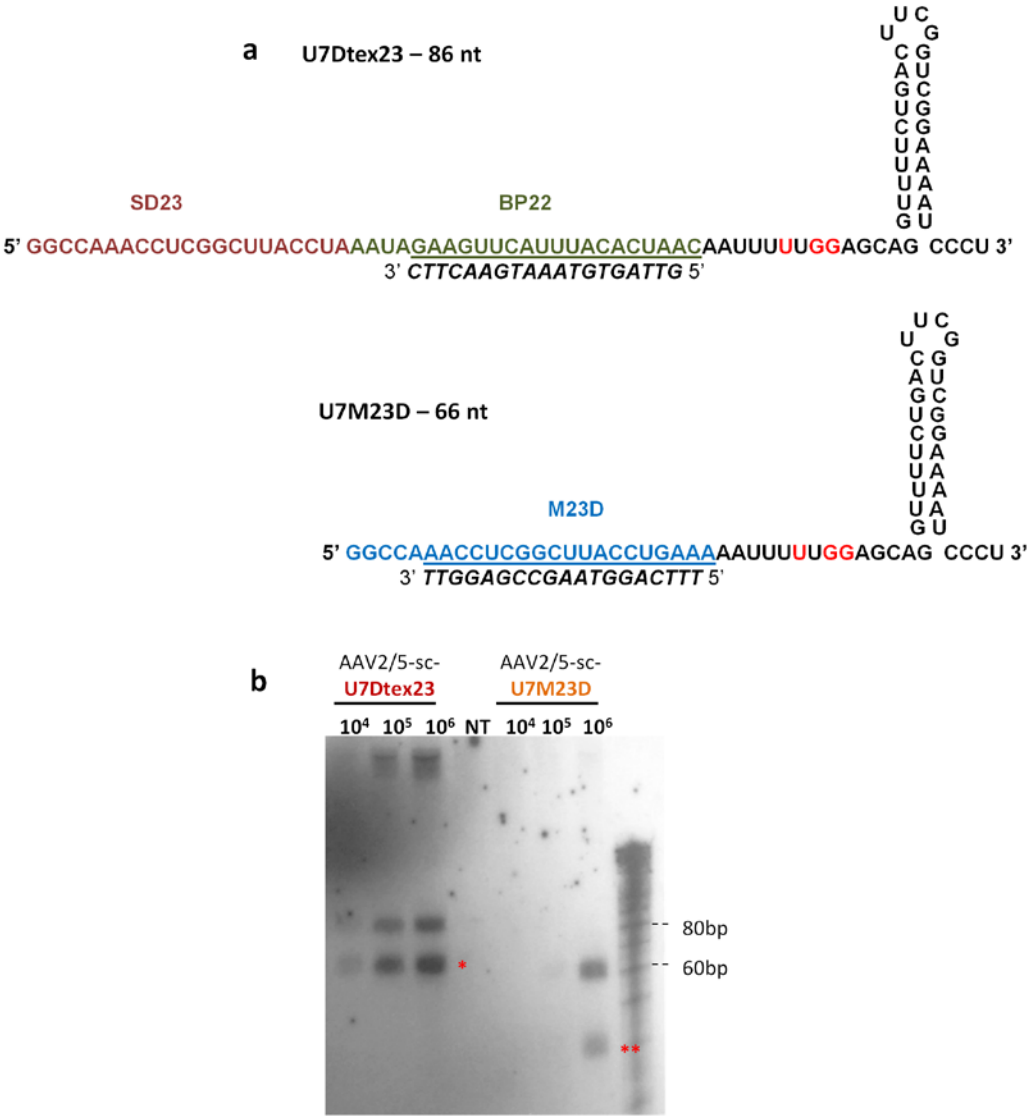

Figure 29: northern blot analysis of U7Dtex23 and U7M23D snRNAs expressed in transduced mouse C2C12 cells

(a) The 86 nt U7Dtex23 and the 66 nt U7M23D snRNA cassettes are represented, along with the position of the ^{32}P end-labeled Dtex23 and M23D probes for their respective detection. (b) The U7Dtex23 is detected at 86 nt and a second band of approximately 65 nt also reacts with the Dtex23 probe (red star). The U7M23D is detected at 66 nt and a second band approximately 20 nucleotides shorter is also observed (double red star). NT represents non-transduced cells. A 15% TBE-Urea gel was used.

Regarding the samples obtained from the transduction of C2C12 cells with AAV2/5-sc-U7Dtex23 vectors (lanes 1 to 3), the Northern blot indicated that the expected full-length U7Dtex23 transcript is produced in the transduced cells in a dose-dependent manner. The additional extra-processed transcripts of approximately 65 nt were detected in these samples, similarly to the former Northern blot experiment (**Figure 43**), and their production also depended on the dose of vector administered. Interestingly, in the AAV2/5-sc-U7M23 samples a second 20 nt-shorter band was also observed along with the 66 nt band corresponding to the U7M23D transcript, indicating that the U7M23D mRNA was also submitted to an additional processing event in the transduced cells. This observation proved that this additional processing is independent of the size of the U7snRNA cassette, thereby ruling out our first hypothesis. Since the probe used to detect the U7M23D transcript hybridizes on the 5' end of the snRNA, this experiment also indicated that the additional processing event occurs on the 3' side of the transcript, somewhere along the stem-loop structure, thereby shortening the 3' end of the transcript of approximately 20 nucleotides.

Further analysis were therefore needed to determine if this phenomenon also occurs on the endogenous U7 transcript or if it affects only the exogenous modified U7 expression cassettes.

- Analysis of endogenous U7 snRNA transcripts in human and mouse cell cultures

To detect the endogenous U7snRNA in human 293T cells and mouse C2C12 cells by Northern blot a probe specifically hybridizing in the natural antisense sequence present in the U7 wild-type mRNA was designed. The U7 snRNA is well-conserved between mice and humans and the antisense sequence against histone H1 is identical in both species (**Figure 49a)**, therefore a single probe (U7wt) could be designed to detect the U7 transcript in both the human and the mouse cell lines by Northern blot (**Figure 49b**).

a

```
                10        20        30        40        50        60
human CAGUGUUACAGCUCUUUUAGAAUUUGUCUAGUAGGCUUUCUGGCUUUUACCGGAAAGCCCCU
      :::::::::::::::::::::::::::::: ::: :::::: :::     : ::::: :::::
mouse AAGUGUUACAGCUCUUUUAGAAUUUGUCUAGCAGGUUUUCUGACUUCGGUC-GGAAAACCCCU
                10        20        30        40        50        60
```

b

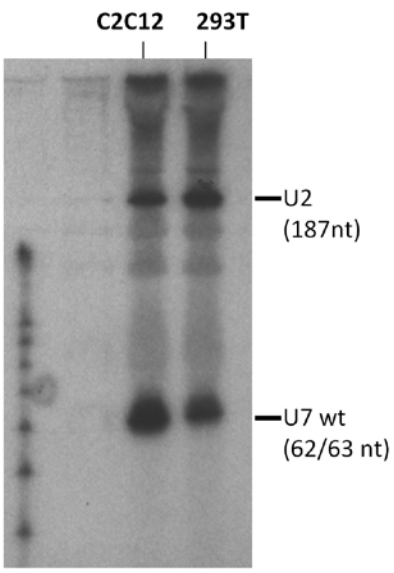

Figure 30: northern blot analysis of endogenous U7 snRNA transcripts in mouse C2C12 and human 293T cells

(a) Nucleotide alignment of human and mouse U7 snRNAs. The nucleotides labelled in red are non-conserved. The conserved histone H1 antisense sequence, corresponding to the position of the U7wt probe for Northern blot detection, is underlined. (b) Northern blot analysis of endogenous U7 snRNA transcripts in mouse C2C12 and human 293T cells. The murine and human transcripts are detected at 62 and 63 nt, respectively. The 187 nt U2 snRNA is used as a control. 15% TBE-Urea gel

We could detect the U7 snRNA in both cells lines at the expected 62 nt and 63 nt sizes for mouse and human transcripts, respectively. The additional 20 nt-shorter band previously observed on the Northern blot detection for U7Dtex23 and U7M23D transcripts could however not be observed, indicating that the wild-type U7 snRNAs are not naturally submitted to an extra-processing event. This observation suggested that the 3' shortening of U7 chimeras is a result of modifications realised on the U7 snRNAs to engineer it into an antisense shuttle. The cause of this phenomenon must be either the modifications of the Sm binding site into the consensus SmOPT sequence made to the U7 shuttles or could stem from the antisense sequence introduced in the cassette. Alternatively, the generation of these shortened species could be the result of the AAV-mediated delivery of U7 snRNAs and the transduction process.

- **Analysing the effect of the SmOPT sequence on the U7Dtex23 processing**

In order to analyse whether the SmOPT sequence present in the modified U7 snRNA cassettes plays a role in the additional processing of U7 transcripts, different variants of the U7snRNA cassette were analyzed. We first modified the Sm sequence of the wild-type U7 snRNA, which retains its natural 5' sequences targeting Histone H1 mRNA, into the consensus SmOPT site by site-directed mutagenesis. Alternatively, it was envisaged that the antisense sequence itself could trigger this processing event in a sequence specific manner. To assess this later hypothesis we also altered the U7Dtex23 cassette by reverting the SmOPT sequence to the Sm sequence present in U7 snRNAs. These constructs were transfected in human 293T cells and a Northern blot was realised on extracted RNAs (**Figure 50**).

The data indicate that an additional 3' shorten transcript was generated only in the presence of the SmOPT mutations (**Figure 50b**). The transfected wild-type U7 cassette was only detected at the expected 62 nt and no shorter transcripts reacted with the probe. Additionally, the U7Dtex23 cassette was detected at 86 nt and further shortened to around 65 nt only when containing the SmOPT sequence. In contrast, the U7Dtex23 construct containing the natural U7 Sm site was not processed into shorter transcripts but only detected at 86 nt, indicating that the antisense sequence itself does not contribute to this extra-processing event. This experiment also indicated that the generation of these shortened U7 species is not due to AAV gene transfer as we could also observe their production after cell transfection.

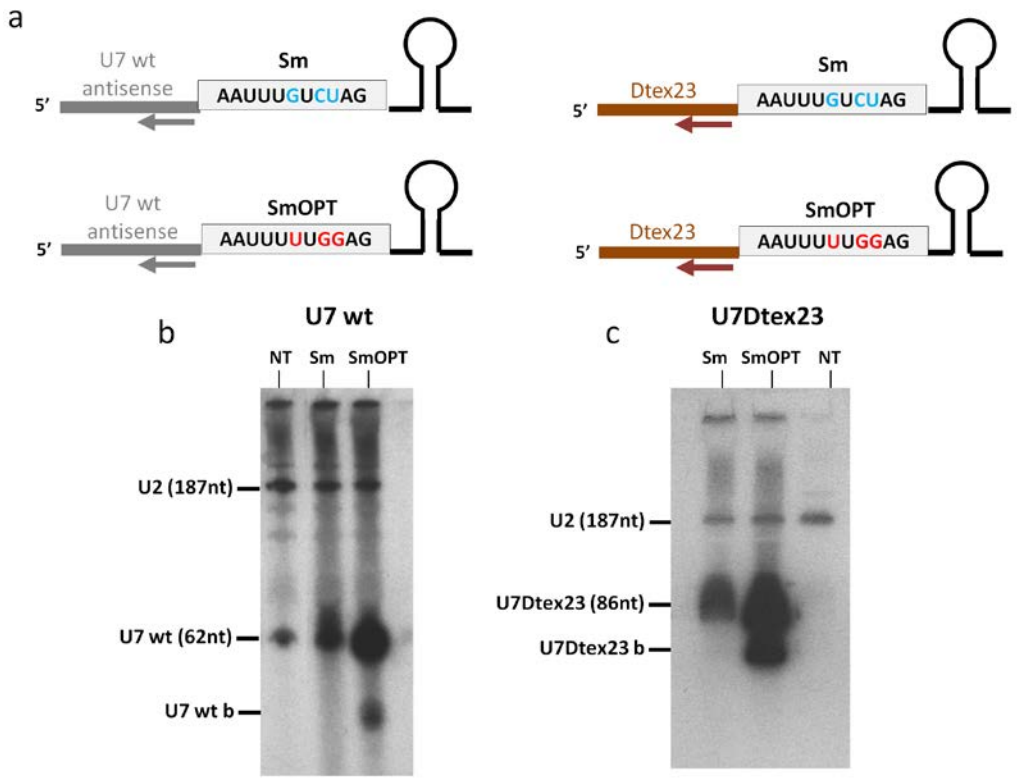

Figure 31: northern blot analysis of the effect of the SmOPT sequence on U7 snRNA processing

(a) Left panel: representation of the U7wt-Sm or -SmOPT constructs and the U7wt probe used in b. Right panel: the U7Dtex23-Sm or -SmOPT constructs and the Dtex23 probe used in c. (b) Northern blot detection of mouse U7 snRNAs transfected in 293T cells. The U7 snRNAs contain the wild-type antisense sequence against histone H1 mRNA (U7 wt) and either the U7 Sm site (Sm) or the optimised SmOPT sequence (SmOPT). U7 snRNA is detected at 62 nt using the U7wt probe. An additional band is detected in the sample containing the U7wt-SmOPT transcripts. The endogenous 63 nt human U7 snRNA is detected in the non-transfected 293T cells (NT). (c) Northern blot detection of U7Dtex23 transcripts transfected in 293T cells. U7Dtex23 contain the Dtex23 antisense sequence and either the U7 Sm site or the optimised SmOPT sequence. U7Dtex23 is detected at 86 nt using the Dtex23 probe. An additional band is detected in the sample containing the U7-SmOPT transcripts, but not in the U7-Sm transcripts. NT: non-transfected 293T cells. The 187 nt U2 snRNA is used as a control. 15% TBE-Urea gels were used.

Altogether, these data demonstrate that the optimised SmOPT sequence is responsible for an extra-processing of chimeric U7 snRNA cassettes on the 3' hairpin side of the transcripts, however the relevance of this phenomenon for the function of U7 cassettes as antisense carriers is unknown. It is not clear whether the 3' processed U7 species are also able to target pre-mRNAs for exon skipping and whether the SmOPT sequence is crucial or dispensable for this function.

- **The SmOPT sequences induce an additional processing of the U7snRNA cassette but increase chimeric U7snRNA expression in cultured myotubes**

U7 cassettes with or without the SmOPT mutations were compared in AAV2/5 transductions of C2C12 myoblasts. The U7Dtex23 expression levels obtained were determined by Northern blot, and the exon skipping efficiency obtained from each vector was assessed by nested PCR (**Figure 51**).

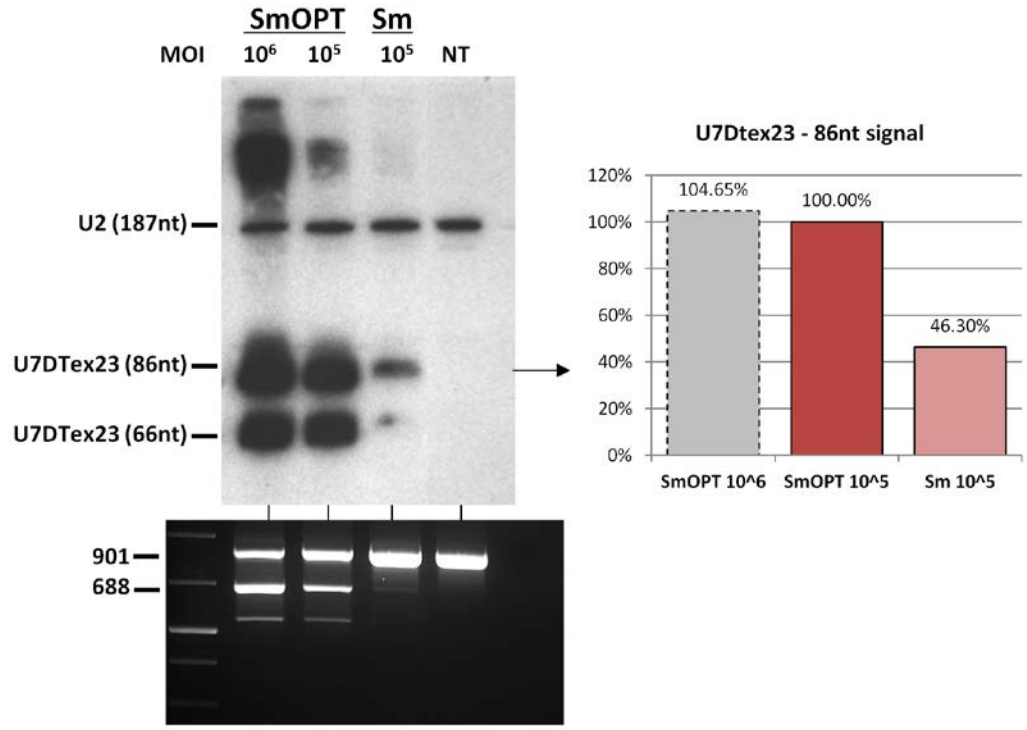

Figure 32: analysis of the effect of SmOPT mutations on the expression of chimeric U7snRNAs and on exon skipping efficiency in cultured myotubes

Northern blot analysis of U7DTex23 transcripts expressed in C2C12 cells transduced with AAV2/5 vectors (15% TBE-Urea gel). U7Dtex23 contains the Dtex23 antisense sequence for dystrophin exon skipping and either the natural U7 Sm binding site (Sm) or the SmOPT sequence (SmOPT). U7Dtex23 is detected at 86 nt using the Dtex23 probe. An additional band of approximately 65 nt is detected in the sample containing the U7Dtex23-SmOPT transcripts, but not in the U7Dtex23-Sm transcripts. NT: non-transduced C2C12 cells. The 187 nt U2 snRNA is used as a control. The right hand side panel represents the quantification by densitometry of the 86 nt band intensity for each sample, expressed as a percentage relative to the expression of U7Dtex23 (86 nt) obtained from the sample SmOPT 10^5. The lower panel presents the analysis of exon 23 skipping performed in the same RNA samples. The native dystrophin mRNA containing exon 23 is detected as a 901-bp fragment and the skipped mRNA yields a 688-bp product.

As expected, the presence of the SmOPT mutation in the U7Dtex23 cassette induced an additional 3' end shortening of the transcripts which was not detected in the U7Dtex23-Sm sample confirming that the SmOPT mutations determine this modification. The presence of the SmOPT sequence also yielded higher levels of expression of chimeric U7 cassettes, as the amounts of U7Dtex23 full-length transcripts (86 nt) obtained with the SmOPT construct were more than two fold higher than that obtained with the Sm-containing transcripts for the same dose of vector administered, according to the quantification by densitometry of the Northern blot. Consistent with this observation, analysis of dystrophin mRNA skipping resulting from the expression of each construct showed that the U7Dtex23-SmOPT cassette induced a strong exon 23 skipping whereas the same quantity of AAV expressing the U7Dtex23-Sm construct was almost inactive.

These data show that the shortened U7snRNA transcripts correspond to 3' processed species whose presence depends on the SmOPT modification and that this optimisation highly increases the U7snRNA transcript levels in vitro, leading to a several fold increase in dystrophin exon skipping in the transduced cells. It remains unclear whether the additionally-processed U7 species are also active in targeting dystrophin pre-mRNA for exon skipping, and the mechanism behind this 3' processing is not elucidated. However, the inclusion of the SmOPT mutation in chimeric U7 cassettes appears critical for the efficiency of U7 cassettes expression. Of note, we could again observe on the Northern blot the presence of high molecular weight RNA species that reacted with the Dtex23 probe and are produced in a dose-dependent manner. This observation could correspond to unprocessed RNA species accumulating at high doses of vector and are further investigated in section I. 5.

I. 4. Optimal levels of antisense sequences using a muscle and heart specific enhancer to drive the expression of U7 snRNA cassettes

I. 4. 1. Principle and objective

As an alternative optimization strategy, we tried to increase the level of expression of the U7 cassette in muscular and cardiac tissues using a transcriptional enhancer.

The non-conventional structure of snRNA genes and the coupling between their transcription and 3' processing via the Integrator complex imply that the presence of a snRNA-specific promoter is necessary for their proper expression. Replacing the snRNA promoter with an mRNA promoter, such as an entire muscle-specific promoter, would impair the expression of the snRNA cassette by preventing the correct 3' end processing of the transcript. It is however possible to stimulate an snRNA promoter through the use of an enhancer, as shown by Brun et al. who inserted an MCK enhancer element upstream of the U7 snRNA promoter [256].

The MCK promoter and its 206 bp enhancer have been thoroughly studied [380-381] and various synthetic versions of these elements have shown strong potential to drive the expression of transgenes in the skeletal muscle [382]. The MCK enhancer promotes tissue-specific expression in fast-twitch myofibers preferentially, but expression activity in the cardiac muscle is limited [382]. Yet, achieving robust transgene expression in the cardiac muscle is paramount for the treatment of DMD where heart failures are frequent. It is therefore important to identify enhancer/promoter elements capable of inducing high-level of transgene expression in both cardiac and skeletal muscle.

Such regulatory cassettes have been previously engineered by fusing enhancer and promoter elements from the α-Myosin Heavy Chain (α-MHC) and the Muscle Creatin Kinase (MCK) genes [383]. The best enhancer, MHCK7, is composed of a fusion between the α-MHC enhancer and an optimised MCK placed upstream of the 358 bp MCK promoter (**Figure 52**). The 188 bp α-Myosin Heavy Chain (α-MHC) enhancer present in this construct is known to promote strong cardiac muscle-specific expression of downstream genes, even in the context of a heterologous promoter [384]. As the α-MHC enhancer is a short element, its introduction in the MHCK7 cassette enabled to obtain high-level of transgene transcription in the heart muscle, in addition to the MCK-promoted activity in the skeletal muscle, without hampering the requirement for a small size enhancer cassette for introduction in rAAV vectors.

The MHCK7 element can be shortened by deleting 63 bp sequence in the MCK enhancer, thereby moving the right E-box closer to the MEF2 binding sites **(Figure 52**). This modification constitutes an important optimisation as various myogenic regulatory factors binding to the E-box, such as MyoD, are known to interact and to synergize with transcription factors associating with the MEF2 site of the MCK enhancer [385]. Bringing together the E-box and MEF2 sites through the modification in the MCK enhancer therefore facilitates interactions between the regulatory factors binding on both sites, which then cooperatively activate transcription in a more efficient fashion. The final 770 bp MHCK7 regulatory cassette induces a robust and mostly tissue-specific enhancement of gene expression in both skeletal and cardiac muscles when delivered via AAV vectors [383].

In order to drive high-level and tissue-specific expression of U7Dtex23 cassettes in skeletal and heart muscles we have assembled a similar enhancer, MHCK, by reproducing the synthetic fusion between the MCK and the α- MHC enhancers described in the MHCK7 hybrid cassette. The MCK promoter initially present in the cassette was omitted and the MHCK element was inserted directly upstream of the natural U7 promoter in the U7Dtex23 cassette (**Figure 53**).

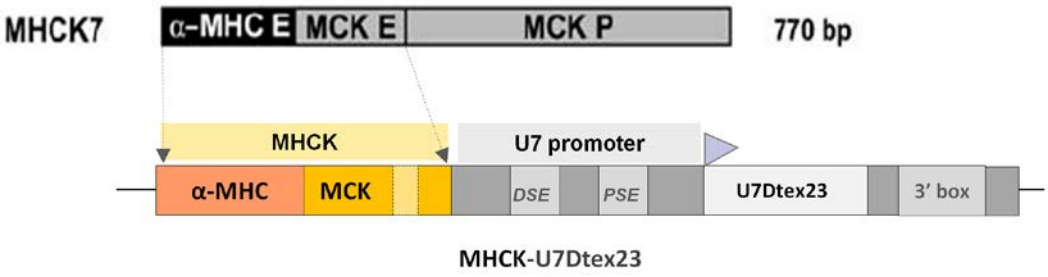

Figure 33: the MHCK-U7Dtex23 construct

The enhancer moiety from the MHCK7 cassette [383], containing the α- MHC enhancer fused to a 63 bp-deleted MCK enhancer, was inserted upstream of the U7 promoter in the U7Dtex23 cassette for mouse dystrophin exon 23 skipping.

AAV vectors containing the MHCK-U7Dtex23 transgene were tested in myoblast cultures and administered intramuscularly in *mdx* mice and their efficiency to promote dystrophin exon 23 skipping was compared to that of the conventional AAV vector without muscle-specific enhancer.

I. 4. 2. Design and validation of the MHCK enhancer sequence

We assembled a strong muscle-specific hybrid enhancer, MHCK, composed of elements from the α-Myosin Heavy Chain (α-MHC) and the Muscle Creatin Kinase (MCK) enhancers similar to the one described by Salva et al. [383] (see materials and methods). The ability of this enhancer to increase the expression of a transgene in muscular cells was first validated in murine C2C12 myotubes by assessing its activity when placed upstream of the human PGK promoter driving the expression of the green fluorescent protein (GFP) gene. AAV vectors carrying the PGK-GFP construct with or without MHCK enhancer were transduced in mouse myotubes and the resulting expression of GFP protein recorded by Western Blot showed that the MHCK enhancer induces an increase of GFP expression in myogenic cells, which was more than two fold higher than that obtained from the vector without enhancer (**Figure 54**).

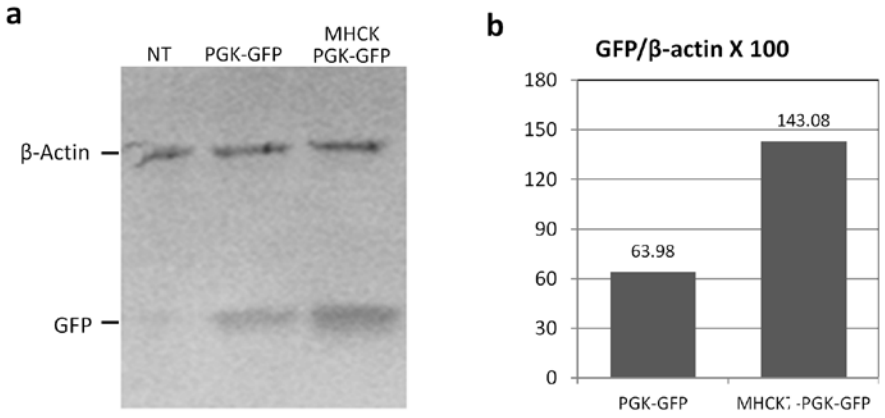

Figure 34: the MHCK enhancer increases GFP expression in myogenic cells.

(a) Western blot to detect GFP expression obtained from protein extracts of C2C12 cells transduced with self complementary AAV2/5 vectors expressing the PGK-GFP or the MHCK-PGK-GFP cassettes (MOI = 10^4 vg /cell). The non-transduced cells (NT) and the endogenous control β-actin are shown. (b) Representation of GFP expression levels relative to β-actin in the transduced cells. The quantification was performed from the western blot presented in (a). The membrane was converted to a numerical picture and band intensities were analysed by densitometry using the GeneTools analysis software (Syngene, UK). GFP levels are expressed as a percentage compared with β-actin levels.

I. 4. 3. Analysis of the expression of the MHCK-driven U7Dtex23 and on the resulting exon 23 skipping in muscular cultures

With the goal of testing the efficiency of the MHCK enhancer in the skeletal muscle of *mdx* mice, we then introduced it upstream of the U7DTex23 cassette designed for exon 23 skipping in the mouse dystrophin pre-mRNA (**Figure 55a**). AAV2/5 vectors carrying the U7Dtex23 cassette construct with or without the MHCK enhancer were produced and C2C12 myotubes were transduced at various MOI. Total RNAs extracted from the transduced cells were used in a Northern blot assay to detect the levels of U7Dtex23 expression resulting from each condition (**Figure 55b**).

The full-length 86 nt U7Dtex23 snRNA was detected in all samples, along with the 3' end processed species of around 66 nt, and expressed in a dose-dependent manner. We could observe that, from the same quantity of vector transduced, the expression of U7Dtex23 transcripts was increased in the presence of the MHCK enhancer (**Figure 55b**). This enhancement was mainly noticeable at the lower dose of vector (10^5 vg/ml), where the expression of chimeric U7 snRNAs was increased by approximately 1.5 fold when driven by the MHCK enhancer according to quantification by densitometry. At higher dose of vector the effect of the enhancer was almost negligible, potentially reflecting a saturation of transcript amounts in the cells when administering such high quantities of AAV vector. In agreement with this idea, the presence of high molecular weight RNA species reacting with the probe proportionally to the amount of vector further suggested an accumulation of unprocessed species when the quantity of U7 snRNAs produced in the cells reached a saturated level.

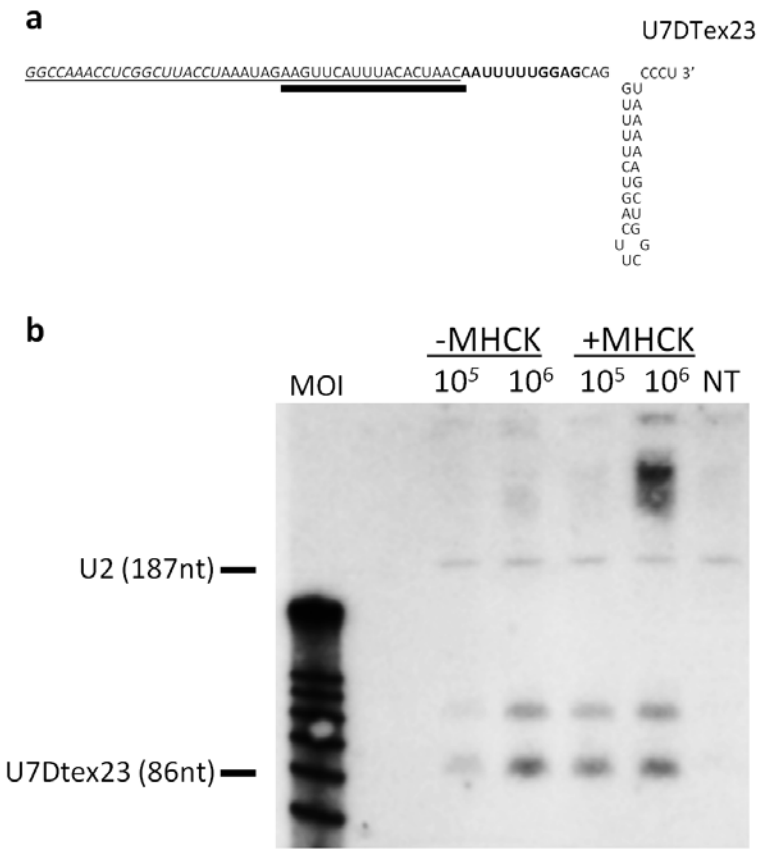

Figure 35: enhancement of U7Dtex23 snRNAs expression in C2C12 cells.

(a) The sequence of the U7DTex23 transcript is shown, including two antisense sequences (underlined) that encompass the donor splice site of exon 23 (SD23, shown in italics) and the branch point of intron 22 (BP22) on the murine DMD pre-mRNA. The optimized Sm binding sequence (SmOPT) is shown in bold. The probe used for the Northern blot analysis in (b) is underlined in bold.

(b) Northern blot analysis of U7DTex23 transcripts expressed in C2C12 cells transduced with AAV2/5-U7DTex23 (-MHCK) or AAV2/5-MHCK-U7DTex23 (+MHCK) at different MOI (10^5 or 10^6 viral genome per cell). NT: non-transduced cells. The blot was hybridized with two ^{32}P end-labeled probes recognizing the U2 snRNA and the U7DTex23 chimera (underlined in bold in (a)), respectively. U2 snRNA shows up as a 187 nt band. Two bands of 86nt and approximately 66 nt react with the U7DTex23 probe.

The RNAs from the same transduced myotubes were further used to analyse the dystrophin exon 23 skipping obtained from each condition by nested PCR and Taqman RQ-PCR (**Figure 56**).

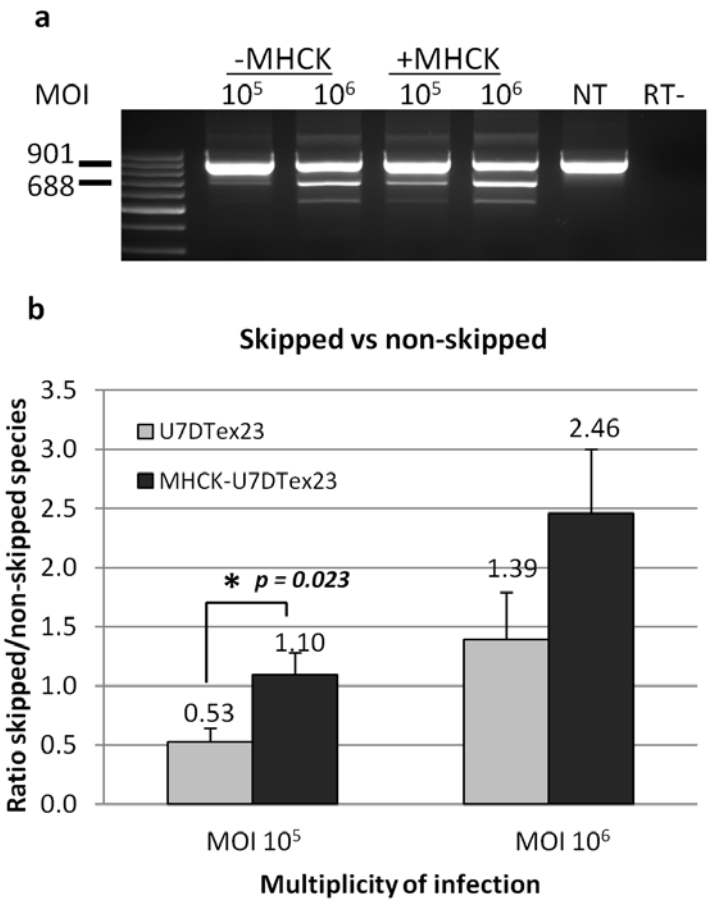

Figure 36: U7-mediated exon 23 skipping on the DMD pre-mRNA.

(a) Reverse Transcription and nested PCR detection of DMD exon 23 skipping after transduction of C2C12 cells with AAV2/5-U7DTex23 (-MHCK) or AAV2/5-MHCK-U7DTex23 (+MHCK) with different MOI (10^5 or 10^6 viral genome per cell). Controls include non-transduced cells (NT) and no reverse transcriptase (RT-). The native murine DMD mRNA containing exon 23 is detected as a 901bp fragment and the skipped mRNA yields a 688bp product. Results are representative of at least three independent transductions.

(b) Analysis of mouse exon 23 skipping by quantitative PCR. The graph represents the ratio between the exon 22-24 junction (skipped mRNA species) and the exon 22-23 junction (full-length mRNA species). Quantification was performed on mRNA extract from C2C12 cells transduced with AAV2/5-U7DTex23 or AAV2/5-MHCK-U7DTex23 at the indicated MOI. A significant difference is observed at MOI 10^5 as shown by the p value (p = 0,023), according to Student's t-test. Error bars are shown as mean ± SEM (n = 6).

Exon 23 skipping was achieved on the dystrophin pre-mRNA from all conditions and the proportion of skipped mRNA detected at 688 bp was consistent with the levels of U7Dtex23 transcripts detected on the Northern Blot in figure 55b. The proportion of skipped dystrophin transcripts compared to the non-skipped ones assessed by RQ-PCR

indicated that the MHCK enhancer induces an increase in dystrophin exon 23 skipping in muscular cells, especially at the lower dose of vector administered where the difference between the samples with or without the enhancer was statistically significant (p value < 0.05 in the Student's *t* test). Furthermore, we could notice that the proportion of skipped dystrophin mRNA obtained in presence of the enhancer was almost equivalent to the one achieved with a 10 fold higher dose of vector without enhancer. This observation is of importance as the main goal of this strategy is to enable to reduce the dose of AAV vector to be administered in order to restore therapeutic levels of dystrophin.

I. 4. 4. Evaluation of MHCK-U7Dtex23 in the skeletal muscle of mdx mice

AAV2/8 vectors carrying U7Dtex23 cassettes with or without the enhancer were produced and administered intra-muscularly in the tibialis anterior of *mdx* mice at doses of $5x10^9$, 10^{10} or 10^{11} vg per muscle. Total RNAs were extracted from the injected muscle sections one month after injection and were analysed by Northern Blot, nested PCR and Taqman RQ-PCR.

The nested PCR analysis indicated that a moderate increase in dystrophin exon 23 skipping was obtained in the presence of the MHCK enhancer, consistent with the increased expression of the U7Dtex23 cassette, approximately 1.5 fold, as seen on the Northern blot (**Figure 57a**). Of note, although hyper-processed U7 snRNA species were also present in each sample on the Northern blot, no high molecular weight RNAs could however be detected, even at the highest dose of AAV injected. The slight beneficial effect of the MHCK enhancer for dystrophin exon skipping *in vivo* could be observed by Taqman RQ-PCR quantifying of the proportion of skipped against non-skipped dystrophin transcripts. Although the highest dose of vector was too strong to clearly see an effect of the enhancer, the intermediate dose indicated an improvement in exon skipping efficiency (**Figure 57b**).

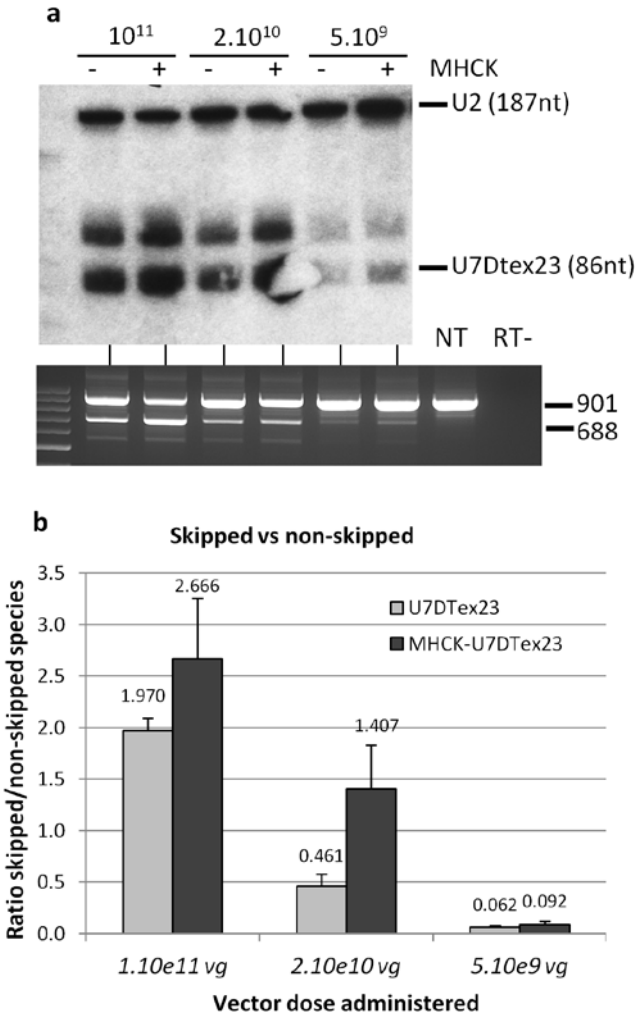

Figure 37: MHCK enhances U7DTex23-mediated exon skipping on the DMD pre-mRNA in the muscle of mdx mice.

(a) Northern blot analysis of U7DTex23 transcripts expressed in mdx mice muscles treated with AAV2/8-U7DTex23 (-MHCK) or AAV2/8-MHCK-U7DTex23 (+MHCK). Total injected vector doses in vector genomes are indicated on top. The blot was hybridized with two ^{32}P end-labeled probes recognizing the U2 snRNA and the U7DTex23 chimera, respectively. U2 snRNA shows up as a 187 nt band. Two bands of 86 nt and approximately 66 nt react with the U7DTex23 probe. The lower panel presents the analysis of exon 23 skipping performed in the same RNA samples. The native murine DMD mRNA containing exon 23 is detected as a 901bp fragment and the skipped mRNA yields a 688bp product. NT, non-transduced muscle. RT-, no reverse transcriptase. Results are representative of at least three independent transductions.

(b) Quantitative PCR analysis of exon 23 skipping in RNA samples prepared from injected muscles. The graph represents the ratio between the exon 22-24 junction (skipped mRNA species) and the exon 22-23 junction (full length mRNA species). No statistically significant difference could be observed between the groups (Student's t-test, $P > 0.05$). Error bars are shown as mean ± SEM (n = 3 per cohort).

This improvement was more visible when analysing Dystrophin protein rescue by Western blot in the same treated muscles. In the presence of the MHCK enhancer a three- to five-fold increase of *de novo* dystrophin protein production was obtained compared with the AAV-U7Dtex23 vector (**Figure 58**). At the highest dose of vector administered, the levels of dystrophin protein restored in treated muscles reached approximately 60% of the levels in wild-type tissue according to quantification by densitometry of the band intensities on the Western Blot. Furthermore, the amount of protein rescue obtained with a dose of 2.10^{10} vg per muscle was higher than the one obtained with a 5-fold higher dose of vector, again demonstrating that the MHCK enhancer can enable to reduce the vector dose needed for therapeutic exon skipping.

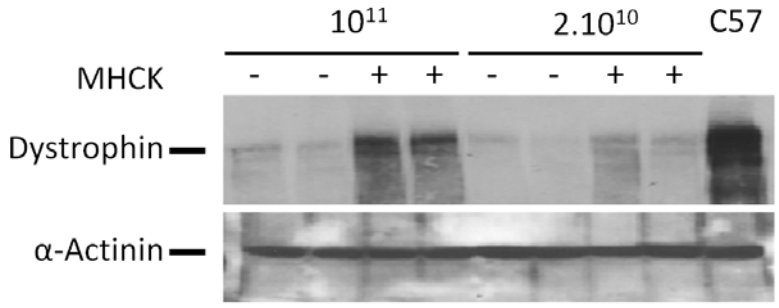

Figure 38: MHCK-mediated enhancement of DMD exon skipping in mdx mice leads to an increased rescue of Dystrophin protein in treated muscles.

Western blot analysis of Dystrophin levels in muscles of mdx mice treated with AAV2/8-U7DTex23 (-MHCK) or AAV2/8-MHCK-U7DTex23 (+MHCK), at doses of $2x10^{10}$ or $1x10^{11}$ vg / muscle. Samples from two different animals are analysed for each condition. The wild type C57 mouse Dystrophin is shown as a positive control and the endogenous α-Actinin is detected for normalization.

Altogether, these results consistently indicated that the MHCK sequence enhances U7DTex23-mediated exon skipping on the dystrophin pre-mRNA in the muscle of *mdx* mice, thereby leading to an increased rescue of Dystrophin protein in treated muscles. Furthermore, the efficiency of this strategy was also demonstrated in a human context, as a MHCK-driven chimeric U7snRNA containing antisense sequences for exons 51 and 52 skipping on the human dystrophin pre-mRNA and delivered by lentiviral vectors, induced an enhancement of dystrophin exon skipping in human CHQ myoblasts, readily visible at lower doses (**Figure 59**).

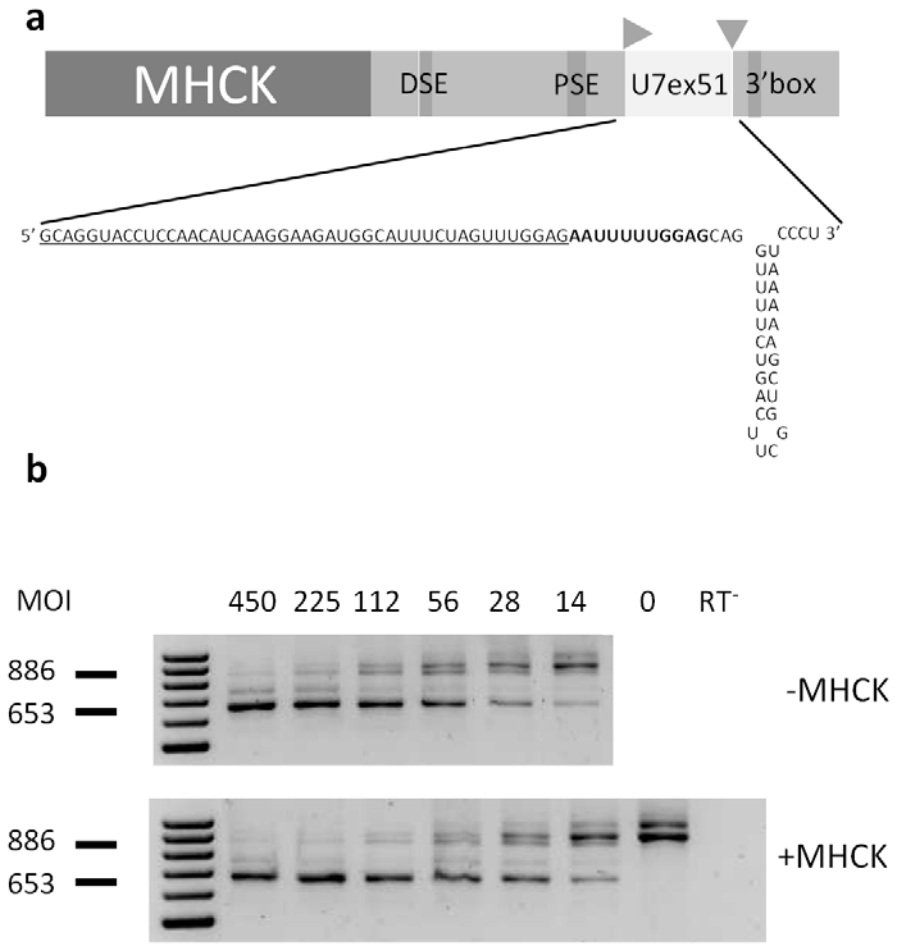

Figure 39: MHCK enhances U7ex51-mediated exon skipping on the DMD pre-mRNA in cultured human myotubes.

(a) U7ex51 expression is driven by the U7 promoter containing Distal and Proximal Sequence Elements (DSE and PSE respectively, shaded boxes). The MHCK enhancer [21] has been placed upstream. The rightward arrowhead indicates the site of transcription initiation and the downward arrowhead is the site of 3' processing, activated by recognition of the 3'box (shaded). The sequence of the U7ex51 transcript is shown, including an antisense sequence (underlined) that corresponds to an exonic splicing enhancer (ESE) in exon 51 of the human DMD pre-mRNA (h51AON1: 5'-UCAAGGAAGAUGGCAUUUCU-3') [22]. The optimized Sm binding sequence (SmOPT) is shown in bold.

(b) Reverse Transcription and nested PCR detection of DMD exon 51 skipping after transduction of human CHQ cells with lentiviral vectors containing the U7ex51 (upper panel) or the MHCK-U7ex51 (lower panel) cassettes, at increasing multiplicities of infection (MOI) Controls include non-transduced cells (0) and no reverse transcriptase (RT-). The native human DMD mRNA containing exon 51 is detected as a 886bp fragment and the skipped mRNA yields a 653bp product.

I. 5. Analysis of the high molecular weight species of the U7Dtex23 transcript.

We have observed a dose-dependent accumulation of high molecular weight (hMW) species during U7 snRNAs detection on Northern blots (Figures 3, 8, 11 and annex 2 Figure 2b). To further analyse these additional hMW species, a Northern blot membrane containing U7Dtex23 transcripts produced with or without the MHCK enhancer and at different doses, corresponding to figure 2b from annex 1, was stripped and hybridized with a new ^{32}P-end labelled probe that specifically recognises the 3' box of U7 transcripts localised at the processed 3' end, thereby only detecting the non-processed primary U7 transcripts (**Figure 60a**).

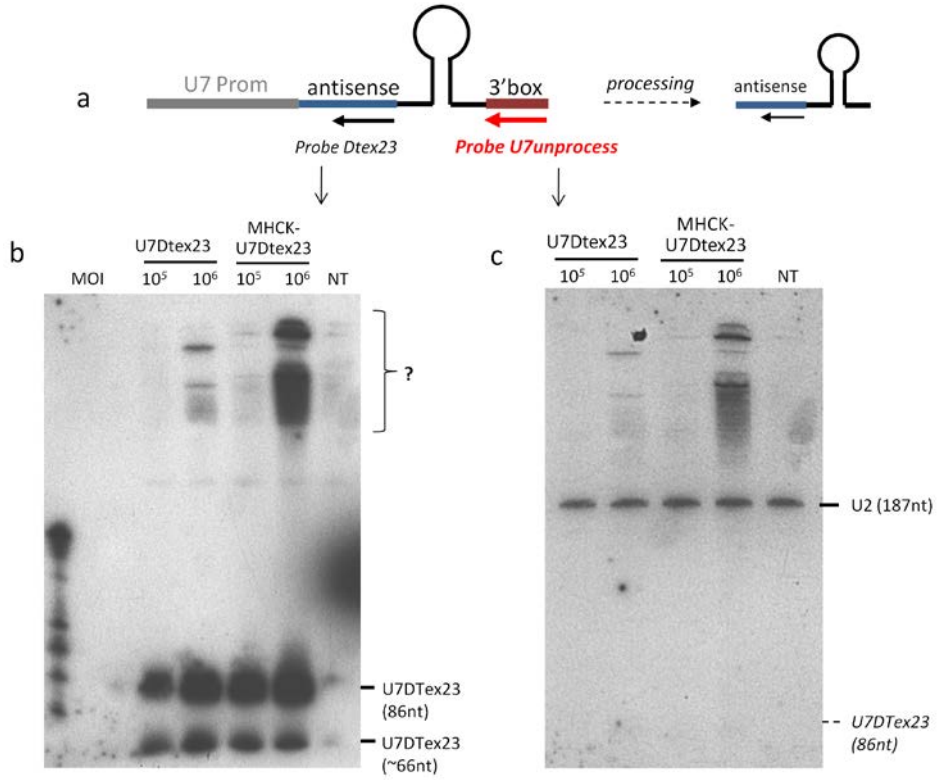

Figure 40: dose-dependent accumulation of readthrough U7 transcripts

(a) Representation of the primary U7 transcript and position of the Dtex23 probe (black arrow) and U7unprocess probe (red arrow) used for the Northern Blot detection in (b) and (c), respectively. (b) Northern blot analysis of U7DTex23 transcripts expressed from C2C12 cells transduced with AAV2/5-U7DTex23 or AAV2/5-MHCK-U7DTex23 at 10^5 or 10^6 vg per cell. The samples were detected using the Dtex23 probe that reacts with the full-length U7Dtex23 (86 nt), the 3' end processed U7Dtex23 (~66nt) and some HMW species. (c) The membrane from (b) was stripped and hybridized with a probe complementary to the U7 3' box for the detection of unprocessed primary U7 transcripts only. NT: non-transduced cells. 10% TBE-Urea gels were used.

The Northern Blot indicated that the hMW species detected earlier contain the U7 3' box that is normally removed during the processing of the mature U7 snRNA (**Figure 60b**). These species could therefore correspond to unprocessed primary U7 RNAs transcribed from tandem or circle copies of the AAV vector genome [294]. We also noticed a difference in size between the HMW species expressed from the AAV-U7Dtex23 and the ones expressed from the AAV-MHCK-U7Dtex23. This difference could correspond to the size of the MHCK enhancer itself present in the HMW species expressed from the later vector, and further analysis using a MHCK-specific probe indicated that the MHCK sequence was present in these HMW species (**Figure 61**).

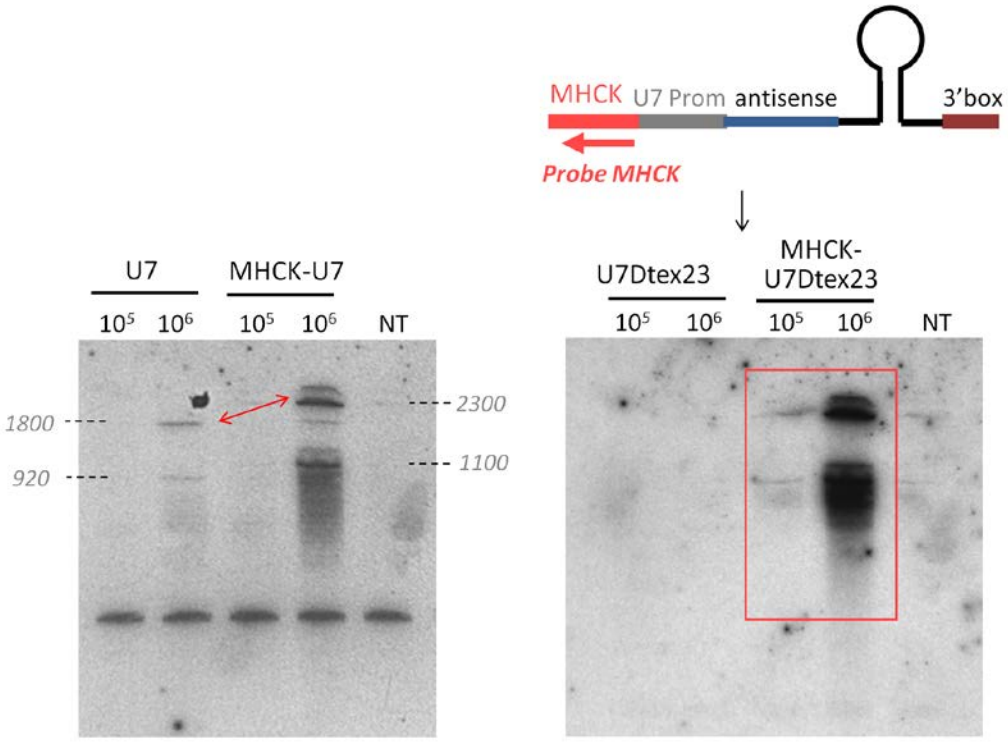

Figure 41: the readthrough U7 RNAs transcripts generated from MHCK-U7Dtex23 AAV vectors contain the MHCK sequence

The left panel corresponds to the Northern Blot from figure 12c where the U7unprocess probe for detection of U7 3' end was used. The approximate sizes of each band are indicated, calculated from the distances of migration. Right panel: Northern blot analysis of HMW U7 species expressed from C2C12 cells transduced with AAV2/5-U7DTex23 or AAV2/5-MHCK-U7DTex23 at 10^5 or 10^6 viral genomes per cell. The membrane used in figure 54c was stripped and a probe specific for MHCK sequences was used to detect MHCK-containing primary transcripts, as represented on the top panel. 10% TBE-Urea gels were used.

We calculated the expected sizes of the unprocessed transcripts, with or without enhancer, based on the idea that they were generated from circle copies of the AAV vector genome. Such transcripts should therefore contain: the U7 transcript, the 3' unprocessed extremity, the ITR region of the AAV vector, the U7 promoter region (with or without MHCK), and another U7 transcript and 3' box. We thereby estimated the size of the unprocessed U7 snRNA without enhancer to be around 920 nucleotides, and around 1100 nucleotides for the enhancer-containing transcript. The second main hMW bands observed in each sample, of higher size than the former, could correspond to dimers of unprocessed U7 transcripts produced from circle copies of AAV. In this case, the expected sizes of the bands should be approximately 1800 and 2300 nucleotides for the transcripts without or with enhancer, respectively. We further confirmed these measurements by estimating on a semi-logarythmic graph the approximate size of the hMW bands depending on their migration distances on the gel.

Altogether these data indicated that the high molecular weight RNA species observed during detection of U7 transcripts by Northern blot are consistent with readthrough primary transcripts, monomers and dimers generated from AAV vector genome concatemers or circles, containing both promoter and 3' untranslated regions and accumulating in a dose-dependent manner. This accumulation of unprocessed readthrough U7 species at high doses of AAV vector suggests that a saturation of the U7 3’ end processing machinery takes place: it is possible that the integrator-mediated formation of U7 snRNA 3’ end is saturated when the primary transcript is present in too many copies in the cells, thereby limiting the amount of mature U7 transcripts available that can be produced intracellularly.

I. 6. Coupling the effects of MHCK enhancer and self-complementary vectors

We next asked whether the benefits of the MHCK enhancer and scAAV could be used in synergy. We introduced the MHCK-U7Dtex23 cassette into a self-complementary AAV vector and AAV2/5-sc-MHCK-U7Dtex23 particles were produced and used to transduce C2C12 myotubes at 10^3 vg/cell. This low dose of AAV vector was chosen in order to avoid a potential saturation of the U7 3' end processing machinery when high amounts of U7 snRNA cassettes are expressed in the cells, which would prevent the comparison of exon-skipping efficiencies obtained from each vector. Total RNAs harvested at 4 days after transduction were analysed by Taqman RQ-PCR to determine the efficiency of dystrophin exon 23 skipping compared with the ones obtained following transduction with AAV-U7Dtex23, AAV-MHCK-U7Dtex23 and AAV-**sc**-U7Dtex23 vectors (**Figure 62**).

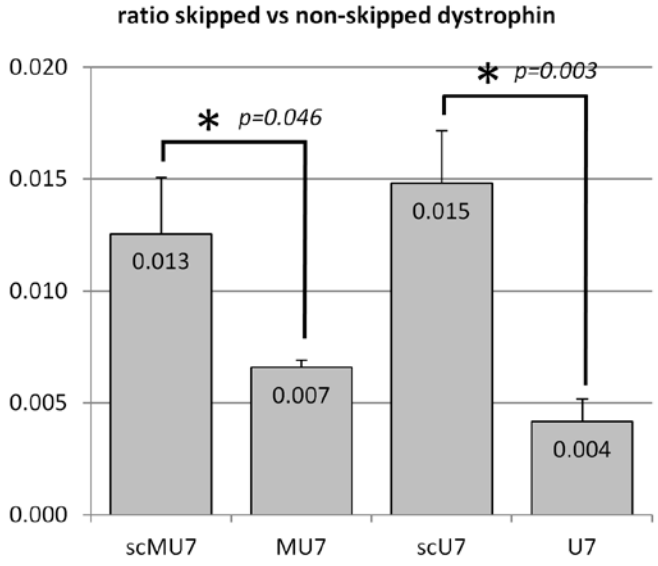

Figure 42: analysis of mouse dystrophin exon 23 skipping obtained from various U7Dtex23 vectors by RQ-PCR

*The graph represents the ratio between the exon 22–24 junction (skipped mRNA species) and the exon 22–23 junction (full-length mRNA species). Quantification was performed on mRNA extracts from C2C12 cells transduced with AAV2/5-**sc**-MHCK-U7Dtex23 (scMU7), AAV2/5-MHCK-U7Dtex23 (MU7), AAV2/5-**sc**-U7Dtex23 (scU7) or AAV2/5-U7Dtex23 (U7) at MOI 10^3 vg/cell and harvested at 4 days post-transduction. Each condition represents an average of 5 independent experiments. The data shown for the scU7 and U7 samples was already represented in figure 47. Significant differences are observed as shown by the P value ($P < 0.05$), according to the Student's t-test. Error bars are shown as mean ± SEM.*

Consistent with our previous experiments, the use of the MHCK enhancer to express U7Dtex23 cassettes, as well as their delivery by scAAV vectors, both increased the amounts of skipped dystrophin transcripts (compare MU7 or scU7 with U7). More importantly, we could observe that the use of a self-complementary AAV vector to deliver the MHCK-driven U7Dtex23 cassette enabled to increase the amount of U7 transcripts expressed in the transduced cells as a significant improvement in exon 23 skipping efficiency was recorded when using AAV-sc-MHCK-U7Dtex23 compared to single-stranded AAV-MHCK-U7Dtex23 vectors (compare scMU7 and MU7). Surprisingly, the proportion of skipped dystrophin obtained with scMU7 was not improved relative to scU7. The exon skipping efficiency obtained from both scAAV vectors was identical independently of the presence of an enhancer sequence, indicating that these two optimisation strategies do not act synergistically.

I. 7. Conclusion

In this chapter we tried to maximise the amount of antisense-expressing U7snRNA cassettes expressed in muscular cells for dystrophin exon skipping for a minimal dose of AAV vector administered. For this purpose two different optimisation strategies were tested. Here we show that the use of self-complementary AAV vectors to deliver U7 snRNA cassettes for mouse dystrophin exon skipping induce a faster and better transduction efficiency and enable to increase the proportion of exon 23 skipped dystrophin mRNA in mouse muscular cells compared to the same dose of single-stranded AAV vector. Furthermore, we show that using the heart and muscle-specific MHCK enhancer to promote the transcription of U7Dtex23 snRNA cassettes in muscular cells induces an increase in dystrophin exon 23 skipping both *in vitro* and *in vivo* and highly improves Dystrophin expression in transduced muscles of *mdx* mice.

Both the use of self-complementary AAV vectors or of muscular-specific enhancer elements therefore constitute promising tools to express U7 snRNA cassettes for dystrophin exon skipping as both strategies independently demonstrated a potential for decreasing the dose of vector necessary to obtain significant levels of dystrophin exon skipping, which is of utmost interest in the context of DMD where high doses of vector are so far deemed necessary for therapeutic benefits. Of note, we also observed the generation of high molecular weight (hMW) U7 RNA species accumulating at high doses of AAV vector and in a dose-dependent manner in transduced muscular cells. This accumulation of unprocessed readthrough U7 species at high doses of AAV vector pointed at a possible saturation of the U7 snRNA 3' end processing machinery when the primary U7 transcript is expressed in too many copies in the cells. This suggests that the potential for maximising the amounts of mature U7 transcripts that can be produced in muscular cells is limited by the cellular machinery capacity. Accordingly, further improvement of U7Dtex23 cassettes expression and dystrophin exon 23 skipping was not observed following the combination of both scAAV and MHCK strategies, probably owing to saturating levels of primary U7 transcripts produced.

II. Small nucleolar RNAs as new molecular platforms for the modification of pre-mRNA splicing

Another line of investigation explored in this study involved the creation new small RNA cassettes designed for the modulation of pre-mRNA processing. By using different carriers and methods for specific splicing alteration we hoped to identify more efficient and regulatable ways to modulate pre-mRNA processing.

We first explored the possibility of using small nucleolar RNAs (snoRNAs) as novel molecular platforms for the intracellular delivery of antisense sequences, aiming to alter dystrophin pre-mRNA splicing with potentially more efficiency than with the original U7snRNA shuttles. The natural presence of antisense elements in C/D box snoRNAs that present perfect complementarities to various cellular RNAs pointed at the feasibility of engineering snoRNAs as antisense carriers through the alteration of their natural antisense sequence. As snoRNAs are encoded within introns of RNA polymerase II transcripts, their concomitant synthesis would allow an antisense-expressing snoRNA to be transcribed from strong and regulatable promoters, allowing for tissue-specificity control and high levels of expression. We therefore derived a C/D box snoRNA into an antisense sequence carrier by replacing its original guide sequence with the Dtex23 antisense sequence for mouse dystrophin exon 23 skipping, and we tested the efficiency of this new system to modulate pre-mRNA splicing.

As an alternative strategy, snoRNAs were modified to specifically target the methylation of intronic adenosine branch points in order to inhibit their recognition by the splicing machinery. This pre-mRNA splicing modulation strategy still displays very low efficiency to-date and our aim was to reproduce and potentially improve this technique in order to obtain effective splicing inhibition in a human cell culture context. We re-directed a human C/D box snoRNA towards the branch point of an exon-skipping reporter construct, based on the splicing dependent activation of a luciferase gene, aiming to first validate this approach before ultimately trying to optimise it for dystrophin exon skipping application.

II. 1. Modified snoRNAs for the intracellular delivery of antisense sequences

II. 1. 1. The snoRNA MBII-52: characteristics and engineering into an antisense carrier

As mentioned earlier, some C/D box snoRNA are devoid of any complementarity to an rRNA or snRNA, suggesting a potential for functions other than 2'-O-methylation and for the targeting of other classes of RNAs. One of them in particular, the human HBII-52 C/D box snoRNA, naturally contains an 18-nt antisense element complementary to a brain-specific mRNA [203] and therefore constitutes a potential good tool for the delivery of antisense sequences designed to target pre-mRNAs.

The HBII-52 snoRNA is an 82 nt long brain-specific snoRNA located in a maternally imprinted region on human chromosome 15q11-13, and a loss of expression of this locus from the paternal allele has been associated with the Prader–Willi syndrome (PWS) [28, 386]. In addition to containing several other snoRNAs, this region contains 47 almost-identical copies of HBII-52 separated by poorly conserved non-protein coding exons [203]. The mouse RNA orthologue MBII-52, which is also exclusively brain-specific, is 79 nt long and is present in 130 tandemly repeated copies. HBII-52 as well as the murine MBII-52 contain a perfect antisense element complementary to the alternatively spliced exon Vb of the serotonin receptor 5-HT2CR pre-mRNA in a region subjected to site-specific A-to-I RNA editing [387]. It has been shown that HBII-52 regulates the 5-HT2C mRNA alternative splicing by base pairing through its 18nt antisense and can promote the inclusion of alternative exon Vb, potentially by masking an exonic silencer in 5-HT2C receptor pre-mRNA [105, 203].

In addition to the MBII-52 capacity to modulate the alternative splicing of the HT2CR mRNA, it has been suggested that the mouse MBII-52 gene could also generate shorter RNA species, exonuclease processing products of the full-length snoRNA termed psnoRNAs and involved in the splicing of other transcripts [106]. These shortened snoRNA species would not form conventional snoRNPs but could instead associate with proteins involved in splice site selection such as hnRNPs [106, 204], and their expression was suggested to modulate the alternative splicing of five additional endogenous pre-mRNAs by interfering with the splicing machinery [106]. These so-called psnoRNAs and the experimental procedure used for their identification have however been recently refuted in a study by Bortolin-Cavaillé et al. [388], but the role of

the full-length MBII-52 snoRNA in the regulation of HT2CR mRNA splicing has however not been questioned. The MBII-52 snoRNA is not an isolated case as another C/D snoRNA, HBII-180C, has also recently been implicated in the regulation of alternative splicing of various pre-mRNAs including FGFR3 [389].

It seems therefore possible to engineer snoRNAs into shuttles for the expression of antisense sequences through the alteration of their natural antisense sequence in order to target pre-mRNAs and modify their splicing pattern. In line with this hypothesis, a recent study showed that a new group of C/D box snoRNAs containing, in addition to the methylation guide sequence, a 21 nt sequence complementary to a pre-mRNA termed the M box could be engineered as a tool for the targeted knockdown of protein expression by replacing the natural M box sequence with a sequence complementary to a target mRNA [390]. This new vector technology called snoMEN (snoRNA modulator of gene expressioN) provides further evidence that natural functions of snoRNAs can be redirected towards a target mRNA in order to influence its expression.

Designing snoRNAs to transform them into antisense sequences carriers for the modulation of splicing presents the advantage that a snoRNA cassette encoded within introns of RNA polymerase II transcripts could be efficiently expressed from strong and regulatable promoters. This system may therefore allow more expression of antisense sequences than the original U7snRNA shuttles.

To test this snoRNA-based approach we replaced the original guide sequence of the mouse snoRNA MBII-52 with the Dtex23 antisense sequence able to induce the skipping of dystrophin exon 23 (**Figure 64**). We compared the efficiency of the new snoRNA-based antisense delivery system to the one obtained with U7snRNA by assessing the exon 23 skipping observed in C2C12 myotubes as a result of the Dtex23 antisense expression.

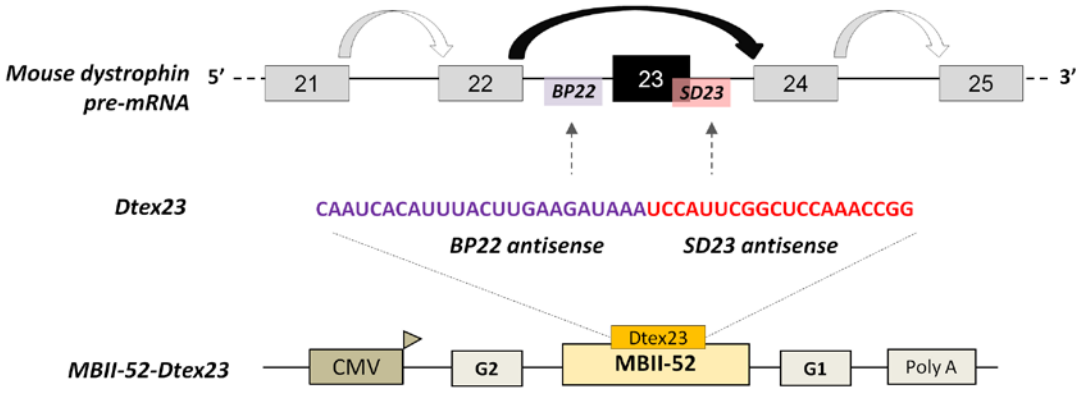

Figure 43: the chimeric MBII52-Dtex23 snoRNA

The original antisense sequence contained in the snoRNA MBII52 was replaced by the Dtex23 antisense for mouse dystrophin exon 23 skipping. The modified MBII52-Dtex23 cassette is included in the original MBII52 host gene, between the bipartite exon G (G1 and G2) of the ipw gene, and placed under the control of the CMV promoter

II. 1. 2. AAV vectors delivery of modified MBII-52-Dtex23 snoRNAs for mouse dystrophin exon 23 skipping

The original guide sequence of the mouse snoRNA MBII-52 directed against the serotonin receptor 5-HT2CR pre-mRNA was replaced with the Dtex23 antisense sequence for mouse dystrophin exon 23 skipping by fusion PCR (see material and methods). MBII52 snoRNA is naturally located in an intron of the *ipw* (imprinted Prader-Willi) gene, between the bipartite exon G (G1 and G2) [203], and this context was therefore conserved for the MBII52-Dtex23 construct in order to ensure correct synthesis and processing of the snoRNA cassette. The MBII52-Dtex23 product enclosed between exons G1 and G2 of *ipw* gene was then placed under the control of the CMV promoter and introduced into a scAAV vector. This vector was first assessed by transfection into 293T cells in order to ensure that the final construct could produce correctly processed MBII52-Dtex23 snoRNAs (**Figure 65**).

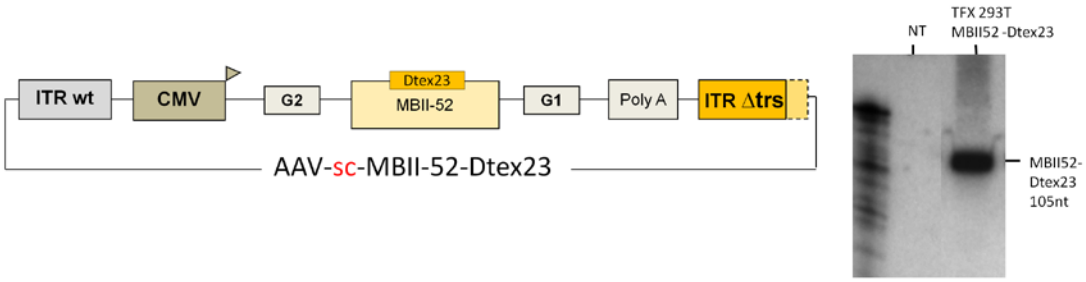

Figure 44: AAV-sc-MBII52-Dtex23 construct

The modified snoRNA MBII52 containing the Dtex23 antisense for mouse dystrophin exon 23 skipping is included in the original MBII52 context, between the bipartite exon G (G1 and G2) of the ipw gene, and placed under the control of the CMV promoter. The construct is inserted between the ITRs of a self-complementary vector where the 3' ITR is depleted of the terminal resolution site (Δtrs). Right panel: Northern blot to detect MBII52-Dtex23 transcripts in 293T cells transfected with the AAV-sc-MBII52-Dtex23 plasmid. The transcripts are detected at 105 nt using the ^{32}P-end labelled MBII52-Dtex23 probe (see figure 60a, right panel). A 15% TBE-urea gel was used.

The AAV-sc-MBII52-Dtex23 construct generated correctly processed chimeric snoRNAs of 105 nucleotides that could be specifically detected by Northern blot. We therefore produced AAV2/5-sc-MBII52-Dtex23 viral vectors and transduced mouse C2C12 myotubes at MOI 10^4 to 10^6 viral genomes per cell. The RNA extracts were first analysed by Northern blot to assess the levels of MBII52-Dtex23 transcripts actually genereated in the transduced cells, and a nested PCR was then realised to record the efficiency of dystrophin exon 23 skipping obtained with this construct (**Figure 66**).

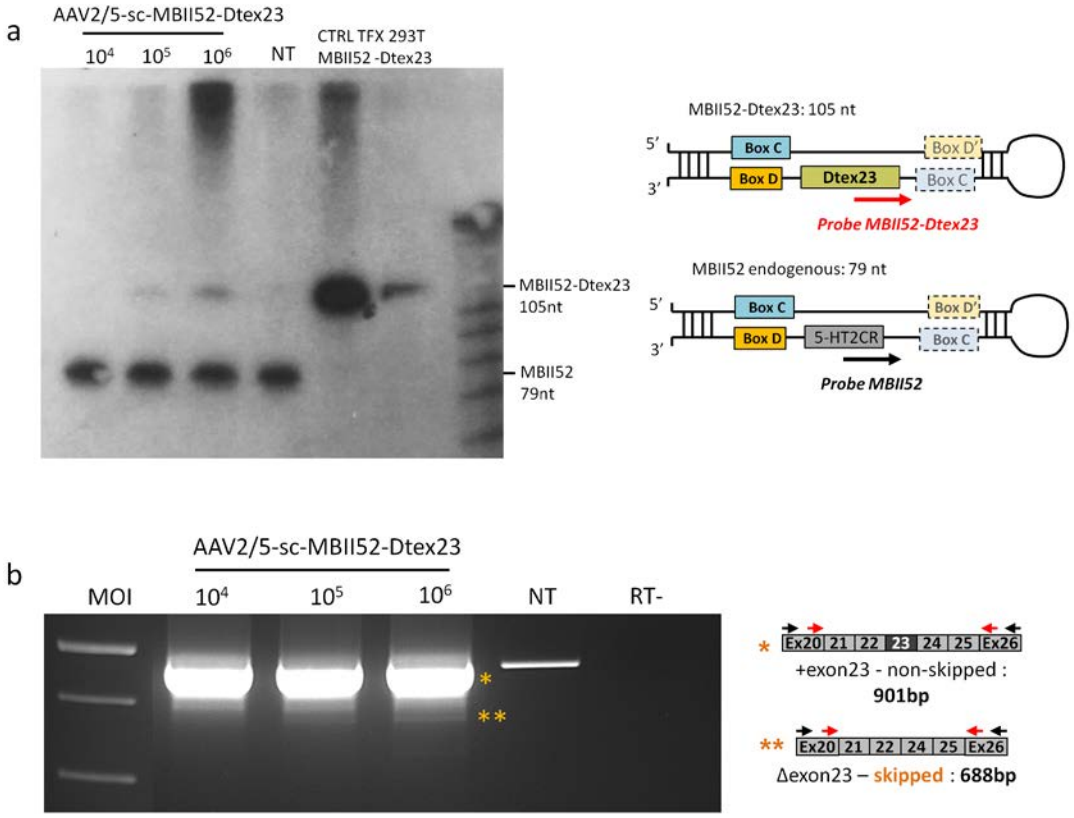

Figure 45: analysis of AAV-mediated snoRNA MBII-52-Dtex23 expression and of dystrophin exon 23 skipping in C2C12 myotubes

(a) Northern blot analysis of MBII52-Dtex23 transcripts expressed in C2C12 myotubes transduced with AAV2/5-sc-MBII52-Dtex23 vectors at MOI 10^4 to 10^6. A control of MBII52-Dtex23 transcripts expressed in transfected 293T cells is shown (ctrl TFX 293T). The MBII52-Dtex23 snoRNA is detected at 105 nt. The endogenous control MBII52 is detected at 79 nt. NT: non transduced cells. 15% TBE-Urea gel. Right panel: representation of MBII52-Dtex23 and MBII52 wt snoRNAs and the respective probes used for Northern Blot detection. (b) Reverse transcription and nested PCR detection of dystrophin exon 23 skipping in C2C12 cells transduced with AAV2/5-sc-MBII52-Dtex23 at various MOI. A schematic representation of the PCR primers and products is shown on the right hand side. NT represents non-transduced cells and RT- is a control without reverse transcriptase. The native dystrophin mRNA is detected at 901 bp (yellow star) and the skipped mRNA at 688 bp (double yellow star).

We could detect some 105 nt MBII52-Dtex23 transcripts on the Northern blot only in the cells transduced with the highest doses of AAV2/5 vectors (**Figure 66a**). The general level of expression of the modified snoRNAs was surprisingly low considering the amount of AAV vector administered and compared to the endogenous MBII52 snoRNA levels. The fact that we could detect the endogenous MBII52 transcript in C2C12 muscular cells is startling considering that its expression is supposed to be

exclusively brain-restricted in mice and humans [203]. This observation might be explained by the fact that the C2C12 cell line does not actually correspond to muscle tissue but correspond instead to transformed cells. We could also detect a strong signal at high molecular weight on the Northern blot that seemed to accumulate in a dose-dependent manner. This signal could represent non-processed MBII52-Dtex23 transcripts still surrounded by their host intron, thereby suggesting a potential problem in the processing of the pre-snoRNA from the debranched intron. Alternatively, it has been demonstrated that the MBII52 snoRNA is naturally processed into several shorter RNAs by exonuclease trimming [106]. It could be envisaged that the expected MBII52-Dtex23 form is under-represented in the RNA species produced from the MBII52 region and that, as most of the transcripts are further processed into short RNAs, we could only detect few full-length MBII52-Dtex23 species on the Northern blot.

Consistent with this observation, we could detect some dystrophin exon 23 skipping in the transduced myotubes but only at very low efficiency, reflecting the low amount of antisense cassettes correctly expressed and active in the cells (**Figure 66b**). Nonetheless, although this result proved relatively difficult to replicate, some dystrophin exon skipping events could be observed in different independent experiments, which demonstrated that the snoRNA-mediated delivery of antisense sequences is feasible, if imperfect.

In order to thoroughly compare the efficiency of the new snoRNA-based antisense delivery system to the one obtained with U7snRNA, C2C12 myotubes were transduced with AAV2/5 self-complementary vectors containing either the MBII52-Dtex23 cassette or the U7Dtex23 cassette at MOI 10^5 and 10^6 vg/cell. We assessed the amounts of Dtex23 antisense-containing cassettes expressed from each vector by Northern blot and analysed the resulting exon 23 skipping obtained on dystrophin mRNA by nested PCR and Taqman RQ-PCR (**Figure 67**).

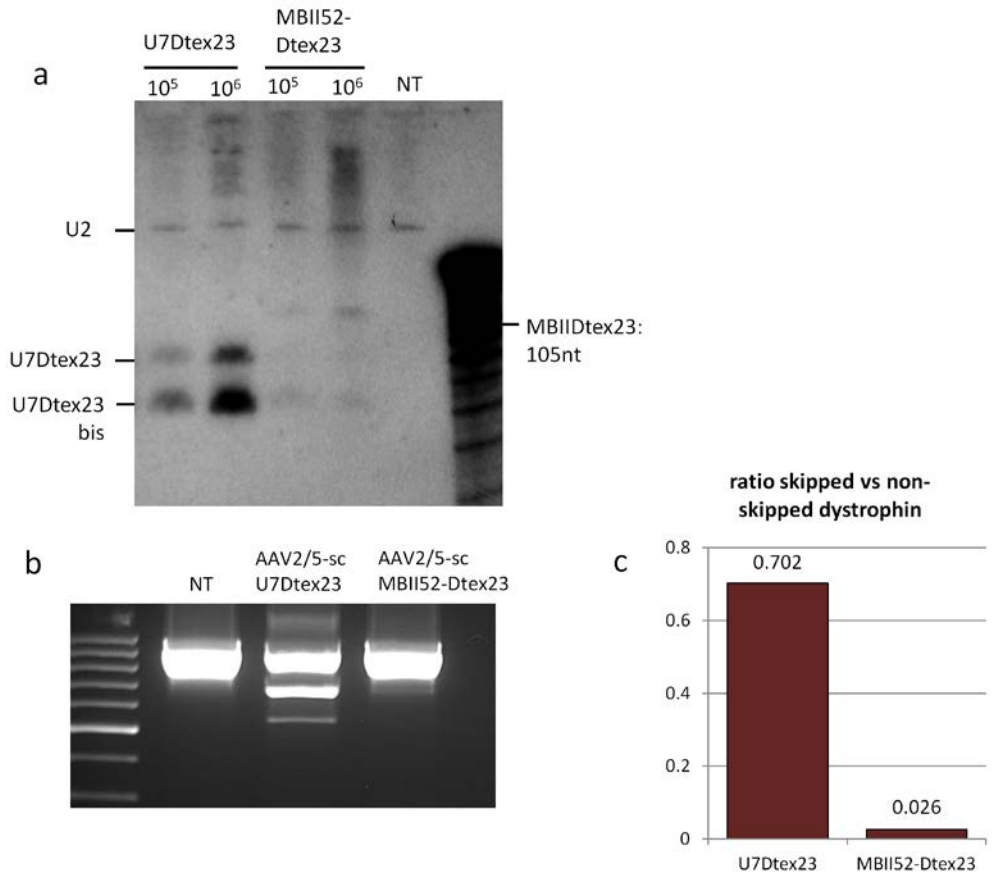

Figure 46: comparison of U7snRNA and MBII52 snoRNA cassettes for intracellular expression of antisense sequences and dystrophin exon skipping

(a) Northern blot analysis of U7Dtex23 and MBII52-Dtex23 transcripts expressed in C2C12 myotubes transduced with AAV2/5-sc-U7Dtex23 or AAV2/5-sc-MBII52-Dtex23 vectors at various MOI. The MBII52-Dtex23 and the U7Dtex23 transcripts are detected at 105 nt and 86 nt, respectively, using the Dtex23 probe. U2 snRNA is used as an endogenous control. NT: non transduced C2C12 cells. 15% TBE-Urea gel. (b) Reverse transcription and nested PCR detection of dystrophin exon 23 skipping after transduction of C2C12 cells with AAV2/5-sc-U7Dtex23 or AAV2/5-sc-MBII52-Dtex23 vectors at MOI 10^6. NT: non transduced C2C12 cells. (c) Taqman RQ-PCR to quantify the ratio between exon 22–24 junction and exon 22–23 junction in dystrophin mRNA species obtained from RNA samples visualised in (b).

It was clear from these experiments that, for a similar dose of AAV vector administered, the amount of MBII52-Dtex23 transcripts actually expressed in the cells are much lower than the levels of U7Dtex23 snRNAs (**Figure 67a**). Consistently, we observed that the snoRNA construct was much less efficient than the U7 cassette to inhibit the splicing of dystrophin exon 23, as seen on the nested PCR analysis and on the quantitative PCR assay (**Figure 67b and c**).

Overall these data indicated that the snoRNA-based delivery of antisense sequences to target pre-mRNAs is not an efficient system compared to the U7snRNA approach. The general shortage of MBII52-Dtex23 expression in AAV transduced cells could reflect a problem in the cassette expression or in the processing of the snoRNA from the transgene. In order to assess if the concatemerisation of AAV genomes in transduced cells could be a limiting factor for the correct synthesis of the MBII52 transcripts we decided to test this system in the context of lentiviral vectors as the integration of the vector DNA in the genome of host cells could potentially provide a better environment for the splicing of the transgene and the subsequent processing of the snoRNA cassette.

II. 1. 3. Lentiviral vectors delivery of snoRNA MBII-52-Dtex23 cassettes in mouse myotubes for dystrophin exon skipping

We wanted to assess if the snoRNA system for intracellular antisense sequences delivery could be improved through the use of lentiviral vectors, which could enable to express the snoRNA cassettes more steadily and therefore increase their efficiency for dystrophin exon skipping. We therefore subcloned the CMV-MBII52-Dtex23 sequence into the pRRL-PGK-GFP lentiviral vector, thereby obtaining a bicistronic vector that also expresses GFP as a control of transduction (**Figure 68a**). Lentiviral particles were produced by calcium phosphate transfection of HEK293T cells and C2C12 myoblasts were transduced with pRRL-MBII52-Dtex23-GFP viral particles at MOI of 30 or 40 infectious particles per cells. The cells were allowed to proliferate for 3 days and then induced to differentiate for 3 additional days before GFP observation and RNA extraction for analysis (see material and methods). A reverse transcription and nested PCR was realised on the RNA extracts to detect potential dystrophin exon 23 skipping events obtained with this vector (**Figure 68b**).

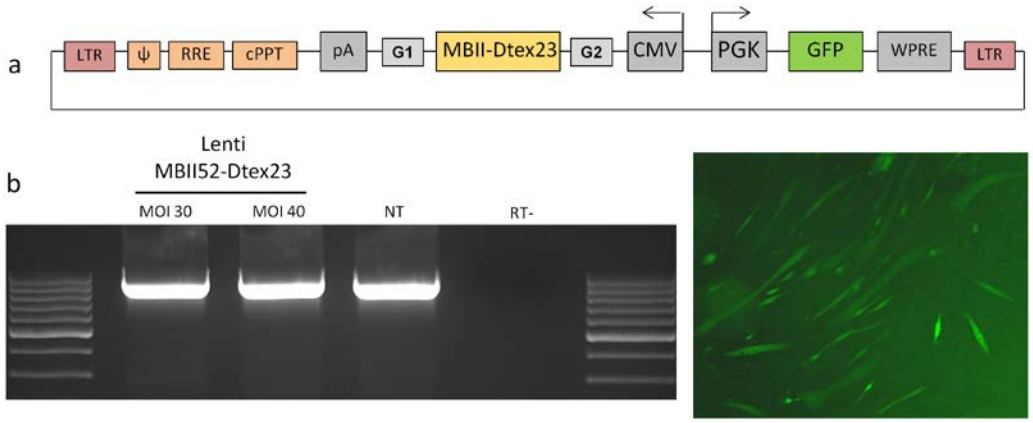

Figure 47: pRRL-MBII52-Dtex23-GFP construct and analysis of exon 23 skipping in transduced myotubes by nested PCR

(a) The modified snoRNA MBII52-Dtex23, containing the Dtex23 antisense for mouse dystrophin exon 23 skipping and surrounded by exons G1 and G2 of the ipw gene, is placed under the control of the CMV promoter and inserted between the LTR of a pRRL-PGK-GFP vector. (b) Reverse transcription and nested PCR detection of dystrophin exon 23 skipping after transduction of C2C12 cells with the pRRL-MBII52-Dtex23-GFP lentiviral vector at MOI 30 or 40 infectious particles per cell. NT: non transduced C2C12 cells. RT-: control without reverse transcriptase. Right panel: Fluorescent microscopy image of C2C12 myotubes transduced with pRRL-MBII52-Dtex23-GFP vector at MOI 40. The cells were observed after 3 days of induced differentiation, 6 days in total after transduction, at a magnitude of 10x.

From the nested PCR assay we could observe that, although a reasonable amount of vector seemed to have integrated in the transduced cells as reflected by the expression of GFP, no skipped dystrophin species could be obtained as only a 901 bp band corresponding to the full-length cDNA could be visualised on the gel (**Figure 68b**). This observation was further confirmed by RQ-PCR (data not shown). This somewhat surprising result might indicate that such doses of lentiviral vector did not allow the expression of enough copies of MBII52-Dtex23 transcript in the transduced cells to target the dystrophin pre-mRNA. Accordingly, on a Northern blot realised from the same samples we failed to observe a signal corresponding to the MBII52-Dtex23 transcripts (data not shown), suggesting that the modified snoRNAs were expressed in too low amounts, if expressed at all, from the pRRL-MBII52-Dtex23-GFP vector.

To increase the probability of observing an effect with the lentiviral construct we decided to sort and isolate the GFP-positive cells obtained after transduction of myoblasts with the pRRL-MBII52-Dtex23-GFP vector, in order to expand only the cells that integrated the lentiviral genome in their DNA. By doing that, every single cell should contain at least one MBII52-Dtex23 sequence stably integrated in their genome and readily available to express antisense sequences for dystrophin exon skipping after differentiation. The sorted GFP-expressing cells, termed C2C12-sno+, were induced to differentiate for 14 days before harvest and RNA extraction. We then analysed the proportion of dystrophin exon skipping obtained by nested PCR (**Figure 69**).

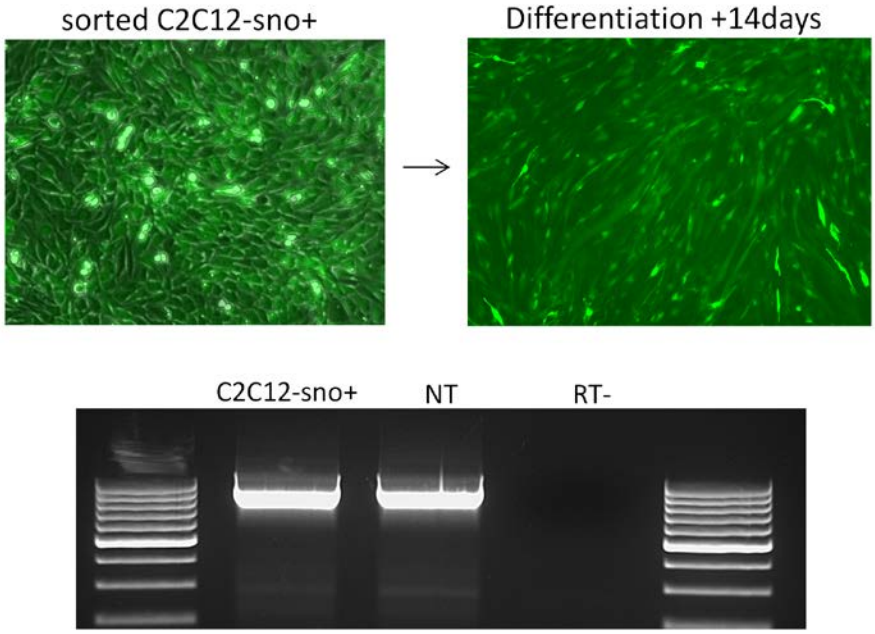

Figure 48: analysis of dystrophin exon 23 skipping in C2C12 cells stably expressing the MBII52-Dtex23 cassette by nested PCR

Top: Fluorescent microscopy pictures of C2C12-sno+ cells transduced with the pRRL-MBII52-Dtex23-GFP vector. The isolated GFP-positive myoblasts were expanded to reach confluence (top left, 20x) and induced to differentiate for 14 days (top right, 10x). Bottom: Reverse transcription and nested PCR detection of dystrophin exon 23 skipping in C2C12-sno+ sorted and differentiated cells. NT: non transduced cells. RT-: control without reverse transcriptase. Only a 901 bp band is detected, corresponding to the non-skipped dystrophin cDNA.

This experiment confirmed that, although all the analysed cells should contain the snoRNA construct stably integrated in their genome, no dystrophin exon skipping could be obtained using lentiviral vectors to deliver the MBII52-Dtex23 cassette. The most likely explanation for this lack of efficiency is a too low amount of antisense sequences expressed in the cells: it has been previously shown that to achieve some level of exon skipping a very high dose of antisense-expressing vector must be administered. This was achieved by using AAV vectors where the number of viral genome copies per cell used was up to 10^6 copies, leading to some exon skipping events. However in this experiment the number of antisense-expressing cassettes per C2C12-sno+ cell could be as little as one copy per cell, which was probably not enough to alter the dystrophin transcripts.

We therefore concluded that the delivery of snoRNA MBII52-Dtex23 construct by lentiviral vectors is inefficient to induce dystrophin exon 23 skipping in cultured muscular cells. Furthermore, although the use of AAV vectors seemed more suited for this snoRNA-antisense strategy as some skipping could actually be achieved through AAV delivery of high doses of transgene, the overall efficiency of this approach did not meet expectations. The use of snoRNAs as shuttles to express antisense sequences is therefore not a promising tool for dystrophin splicing modulation and the high efficiency of U7snRNA cassettes should be privileged for this purpose over the use of a snoRNA cassette. This strategy was thus not investigated further in this study.

II. 2. A human C/D box snoRNA to target the 2'-O- methylation of branch point adenosine as a new strategy for exon skipping

The engineering of C/D box snoRNAs to specifically target the 2'-O-ribose methylation of pre-mRNA branch points for exon skippimg has shown some moderate success in vertebrate systems previously, but the efficiency of this approach is still very limited. Here we tried to reproduce and improve this technique and to obtain effective splicing inhibition in cultured human cells, ultimately aiming to apply this approach to dystrophin exon skipping.

II. 2. 1. The U24 snoRNA: characteristics and modifications for directed 2'-O-ribose methylation of a pre-mRNA branch point

A modified U24snoRNA previously showed some efficiency in inducing the skipping of a downstream exon by targeted methylation of the branch point [205] and we therefore decided to derive the same human U24 snoRNA to induce the methylation of our target branch site and trigger exon skipping.

The human U24 snoRNA is a 76 nt long phylogenetically conserved C/D box snoRNA encoded in the second intron of the RPL7A (ribosomal protein L7a) gene and present at approximately 14 000 molecules per human cell. U24 contains two target recognition sequences of 12 nucleotides each situated upstream of the consensus D and D' boxes, which naturally guide the 2'-O-ribose methylation of two adjacent residues in 28S rRNA exactly 5 nucleotides upstream of the D/D' motif [391] (**Figure 70**).

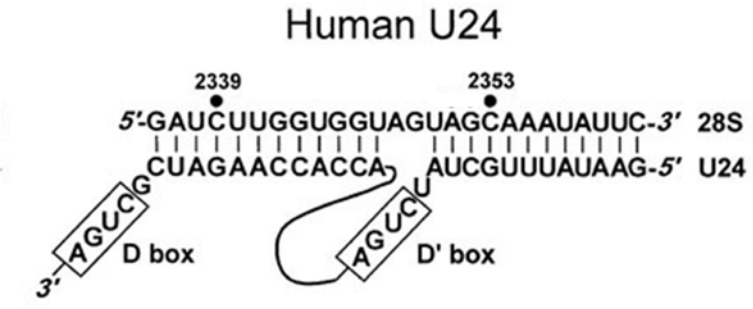

Figure 49: the human U24 snoRNA

The human U24 C/D box snoRNAs represented at the bottom contains two antisense sequences of 12 nt each situated upstream of the D and D' boxes (CUGA). These antisense sequences recognize the human 28s rRNA (top) to target the 2' O-methylation of nucleotides located 5 nt upstream of each CUGA motif (black dot).

The initial aim of this project was to modify the original guide sequences of human U24 snoRNA to perfectly match the branch point of intron 22 (BP22) of mouse dystrophin pre-mRNA, precisely positioning the adenosine for 2'-O-methylation, in order to target the exclusion of dystrophin exon 23. However, to first ensure that the BP methylation strategy would induce exon skipping the use of a reporter system rather than the endogenous dystrophin pre-mRNA seemed more informative and easily interpretable. The βglobin gene constitute a good candidate to use as a reporter system in this study as the splicing of this pre-mRNA has been extensively studied and the important splice signals such as the branching point are well-described [60, 392-398]. A reporter system based on the βglobin sequence and the splicing dependent activation of a luciferase gene was therefore created in order to facilitate the read-out in this study.

II. 2. 2. Design and validation of an exon skipping reporter construct

The reporter construct designed to test the branch point methylation strategy for pre-mRNA splicing modulation consists of the human βglobin sequence, composed of three exons and two introns, fused to an out-of-frame luciferase reporter gene and placed under the control of the CMV promoter. We first isolated the human βglobin gene by PCR and removed the termination codon at the 3' end of the third exon. The codon optimized X12-Luciferase gene [399], deleted of its ATG initiation codon, was then amplified and fused to the βglobin gene by fusion PCR, placing the X12-Luciferase in-frame with the βglobin gene to enable the production of a luciferase fusion protein through the ATG initiation codon of the βglobin mRNA (pCMV-βglobin**N**-Luciferase construct). Finally, a 2-base pair deletion in βglobin exon 3 was induced by mutagenesis in order to disrupt the reading frame of the mRNA and to generate a stop codon downstream of the mutation in βglobin exon 3, thereby placing the fused luciferase transcript out-of-frame and enable to be expressed. The skipping of βglobin exon 2 would however restore the correct reading frame of the construct and allow the production of Luciferase, and the resulting expression of luciferase recorded would therefore reflect the proportion of exon skipping obtained (**Figure 71**). This reporter plasmid, named pCMV-βglobin-Luciferase, was verified by sequencing.

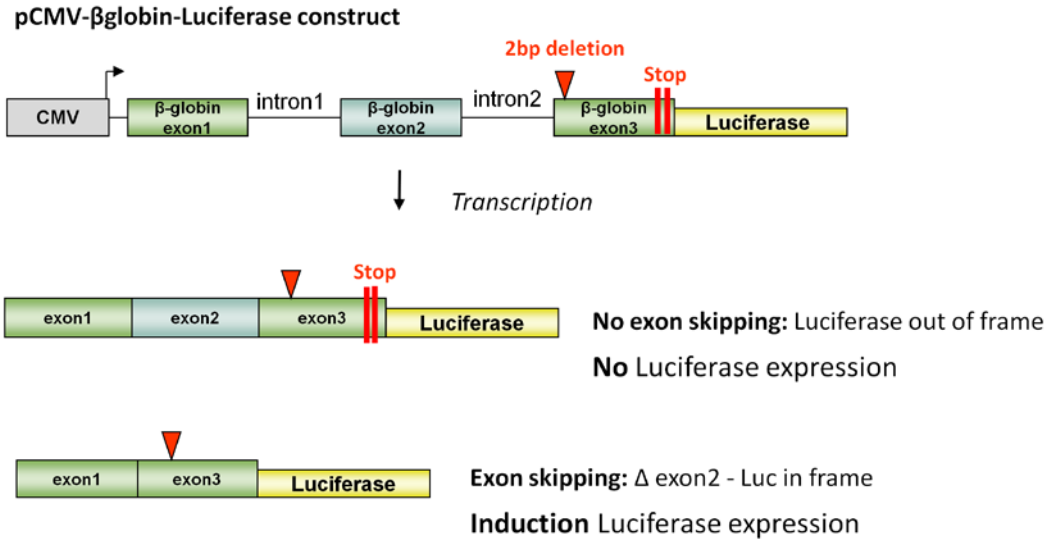

Figure 50: βglobin-Luciferase exon skipping reporter system

The pCMV-βglobin-Luciferase construct contains a GT dinucleotide deletion in βglobin exon2 and is fused to an out of frame luciferase gene. The luciferase mRNA becomes in frame and expressed only if exon 2 of the beta-globin gene is skipped.

To confirm that the construct is valid and does not spontaneously express the luciferase gene we first transfected the pCMV-βglobin-Luciferase construct into 293T cells and recorded the light signal obtained in the cell extracts after 48 hours by luciferase assay. The initial in-frame βglobin-Luciferase construct obtained before deletion of two base pairs in βglobin exon 3, named pCMV-βglobin**N**-Luciferase, should express high levels of luciferase and was therefore transfected in parallel as a positive control. The raw light units (RLU) obtained were normalised to the total amount of proteins present in each sample, which were quantified by BCA assay (see material and methods) (**Figure 72**).

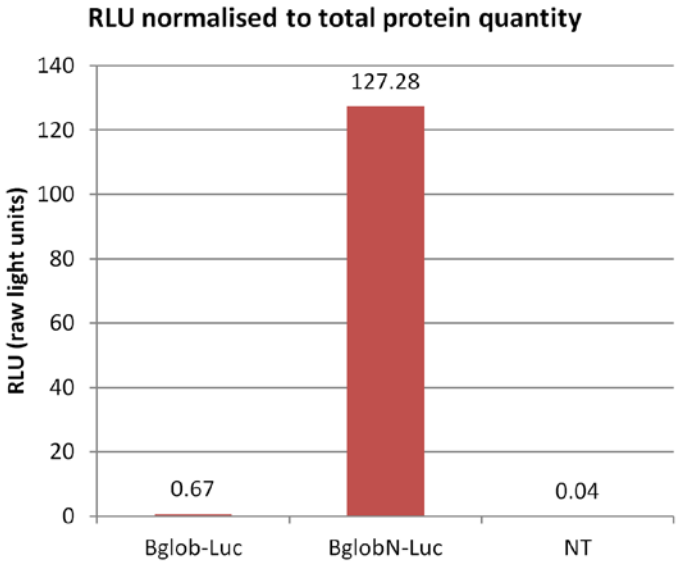

Figure 51: validation of the pCMV-βglobin-Luciferase reporter construct by luciferase assay

*Luciferase assay realised on cellular extracts from 293T transfected with the out-of-frame pCMV-βglobin-Luciferase (βglob-Luc) construct or the in-frame pCMV-βglobin**N**-Luciferase (βglobN-Luc). The average RLU recorded was normalised by the total amount of protein present in each sample. NT corresponds to non-transfected 293T cells.*

As expected, the pCMV-βglobin**N**-Luciferase (βglobN-Luc) control plasmid emitted a good light signal indicating that the βglobin-Luciferase fusion protein generated from this construct was successfully expressed in the transfected cells. On the other hand, the reporter construct pCMV-βglobin-Luciferase (βglob-Luc) was not able to generate any light signal in the transfected cells, which demonstrated that the luciferase sequence placed out of frame in the construct was not able to produce a luciferase protein spontaneously, thereby validating the system. We could therefore use the pCMV-βglobin-Luciferase reporter to test the branch point methylation strategy as a new approach for splicing alteration.

II. 2. 3. The chimeric U24met-βglobin for targeted methylation of the reporter construct and exon skipping

To target the 2'-O-methylation of the βglobin exon 2 adenosine branch point in the pCMV-βglobin-Luciferase reporter construct we engineered the human U24 snoRNA to specifically hybridize in intron 1 of βglobin pre-mRNA at the precise position to target the nucleotide to be modified. In order to insure the proper processing of the synthetic U24 snoRNA we decided to maintain it within its natural host introns context, between exons 2 and 3 of the human RPL7A gene. A RPL7A gene fragment from exon 2 to 3 was first isolated by PCR and inserted downstream of the CMV promoter to create the pCMV-U24 plasmid (**Figure 73a**). After verification of the construct by sequencing, we transfected it in NIH3T3 mouse fibroblasts to ensure that the human U24 snoRNA was correctly processed and expressed from this construct and we analysed the extracted RNAs by Northern blot (**Figure 73b**).

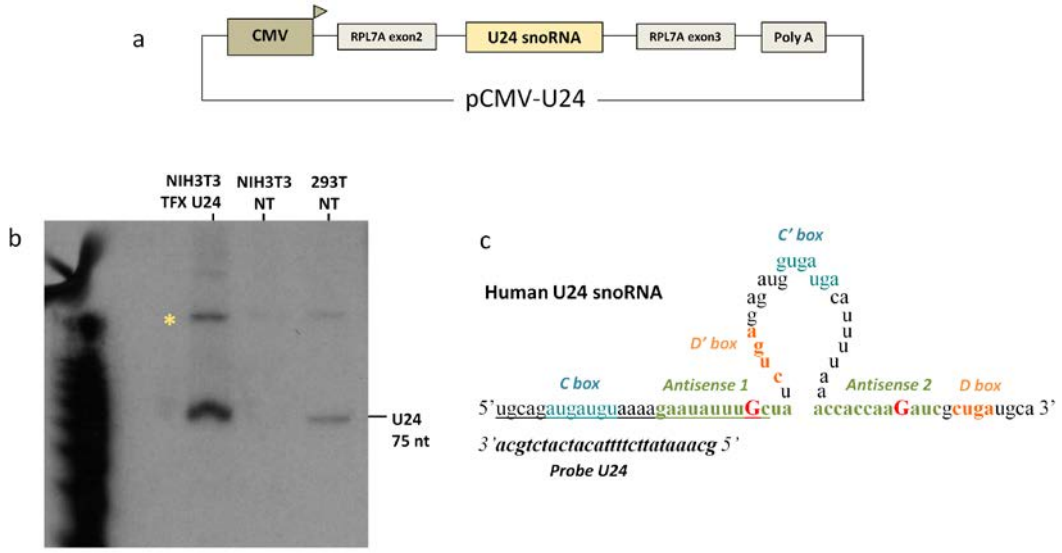

Figure 52: pCMV-U24 construct and detection by Northern blot

(a) Representation of the pCMV-U24 construct containing the human U24 snoRNA enclosed between exon 2 and 3 of the RPL7A gene, under the control of the CMV promoter. (b) Northern blot analysis of U24 snoRNAs in NIH3T3 cells transfected with the pCMV-U24 plasmid. The U24noRNA transcript is detected at 75 nt. The human-specific U24snoRNA probe hybridizes the transfected U24 in NIH3T3 mouse cells and the endogenous U24 in 293T cells. A second band of approximately 170 nt (yellow star) also reacts with the probe in the samples expressing U24 snoRNA. NT: non-transfected cells. 15% TBE-Urea gel. (c) Representation of the human U24 snoRNA containing two antisense elements against 28S rRNA (green). The red capital letters show the hybridization positions of the nucleotides to be methylated in the target RNAs. The probe used for U24 detection by Northern blot is indicated.

The U24 snoRNA was successfully detected at 79 nt on the Northern blot, both from transfected or endogenous source, indicating that its synthesis and processing was correctly achieved from the synthetic construct. We could also observe on the Northern blot the presence of a second band of higher molecular weight than the expected U24 in the samples expressing U24 snoRNA (**Figure 73b**, yellow star). This second transcript, of approximately 170 nucleotides, was present in both transfected 3T3 cells and non-transfected 293T cells but not detected in non-transfected mouse cells, suggesting that these RNA species are also present endogenously in cells naturally expressing the U24 snoRNA and could therefore represent premature U24 transcripts that are still attached to parts of their host introns before processing. As good amounts of correctly processed U24 snoRNAs were detected at the expected size after cell transfection, we therefore further used this construct to realise the modifications necessary to retarget U24 snoRNA to the chosen pre-mRNA.

We replaced the original U24 antisense sequences in the pCMV-U24 plasmid with sequences targeting βglobin intron 1 in the reporter construct, precisely positioning the branch point adenosine for 2'-O-methylation in order to trigger βglobin exon 2 skipping. The branch point adenosine residue to target has been well-described in the literature [393-394] and has been located 37 nucleotides upstream of the 3' splice site in βglobin intron 1, which we further confirmed using the software ESEfinder [56]. As the U24 snoRNA naturally contains two antisense sequences, both able to direct the 2'-O-methylation of target nucleotides, we decided to insert the branch point-specific sequence in the antisense box situated at the 3' side of the snoRNA. The 5' antisense sequence was designed to hybridize the βglobin pre-mRNA downstream of the first hybridization site, thereby strengthening the association between the modified U24 and its target (**Figure 74**). The pCMV-U24metβglob construct obtained was verified by sequencing.

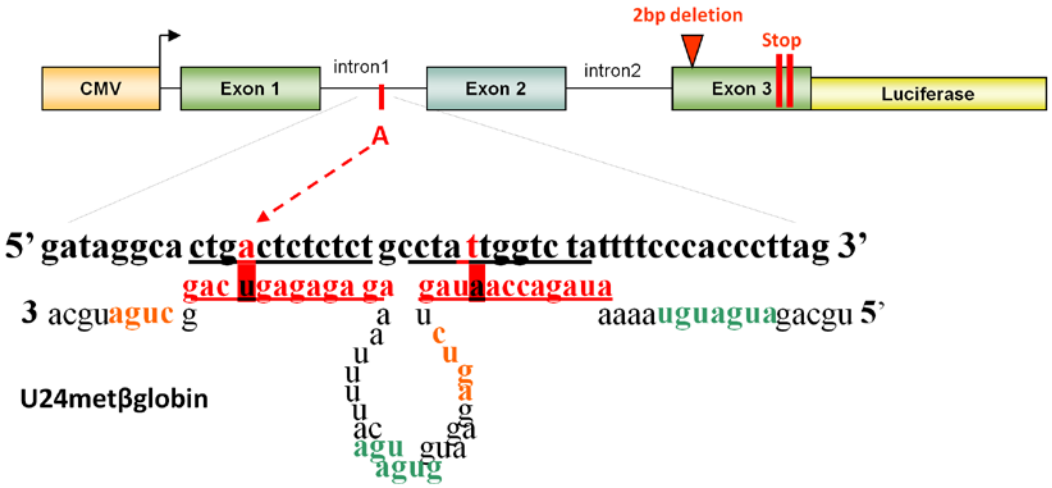

Figure 53: the modified U24metβglob snoRNA to target the adenosine branch point of βglobin intron 1 in the exon skipping reporter construct

The U24metβglob snoRNA contains two antisense sequences (red) directed against βglobin intron 1. The antisense sequence on the 3' side of the snoRNA specifically targets the branch point adenosine of the reporter construct and precisely positions it 5 nucleotides upstream of the D box for site-specific 2'-0-methylation. The C and D boxes of the snoRNA are represented in green and orange, respectively.

In order to test the efficiency of the pCMV-U24metβglob construct to trigger the skipping of βglobin exon 2 in the reporter construct by specifically targeting the branch point, the pCMV-U24metβglob plasmid was co-transfected with the luciferase reporter construct in 293T cells. To increase the probability of successfully targeting the βglobin sequence, the constructs were transfected following a 5:1 (U24:reporter) molecular ratio. The cell extracts were analysed 48 hours later by luciferase assay to assess the exon skipping potentially obtained in the transfected cells depending on the intensity of the light signal detected. Furthermore, to insure that the U24metβglob transcript was efficiently expressed in the transfected cells we realised a Northern blot on the RNA extracted from these samples (**Figure 75**).

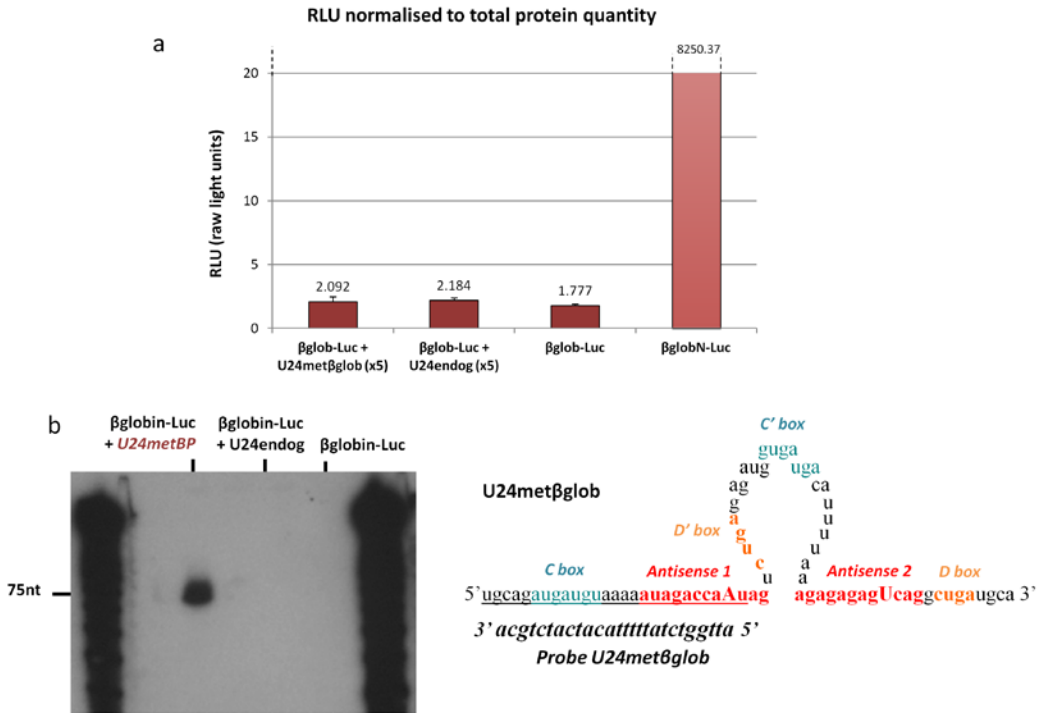

Figure 54: analysis of pCMV-U24metβglob expression and induction of exon skipping in the reporter construct by co-transfection in 293T cells

*(a) Luciferase assay realised on cellular extracts from 293T transfected with the pCMV-βglobin-Luciferase (βglob-Luc) construct and either the pCMV-U24metβglob or the pCMV-U24 (U24endog) constructs. The βglob-Luc transfected alone was used as a negative control and the βglob**N**-Luc constitutes the positive control. The average RLU recorded was normalised with the total amount of proteins. Each condition represents an average of 3 independent experiments. Error bars are shown as mean ± SEM. (b) Left: Northern blot analysis of U24metβglob snoRNAs expressed in transfected 293T cells, corresponding to samples in (a). The U24metβglob snoRNA is detected at 75 nt. 15% TBE-Urea gel. Right: representation of the U24metβglob containing the C/C' boxes (blue), the D/D' boxes (orange), and two antisense elements against βglobin intron 1 (red). The capital letters correspond to the position of the methylated nucleotides in the target RNA. The probe used for U24metβglob snoRNA detection by Northern blot is indicated.*

The luciferase assay showed that no significant increase of luciferase signal was recorded when the reporter construct was co-expressed with the U24metβglob snoRNA compared to the signal obtained with the reporter construct alone or co-transfected with the non-modified U24 snoRNA (U24 endog). This result indicated that no exon skipping could be obtained in this system, at least not at detectable levels. The inefficiency of the strategy was however not due to a lack of expression of the U24 cassette in the transfected cells as good amounts of U24metβglob transcripts could be observed at the right size on the Northern blot from these samples (**Figure 75b**).

These data suggested that this exon skipping strategy was greatly inefficient in this system, which could be due to various factors. It is possible that, even if a base-pairing occurs between the U24metβglob and the reporter construct at the right position, the modified U24 is not able to efficiently target the methylation of the intronic branch point of the reporter construct and cannot trigger the skipping of the downstream exon. Alternatively, by inactivating the major branch point of intron 1 of the reporter construct through targeted methylation, it is highly possible that a cryptic branch point could be activated to restore the normal splicing pattern [395]. To eliminate this possibility, the exon skipping reporter construct would need to be stringently examined in order to identify and mutate all the putative cryptic splice sites that could favour alternative splicing events or restore the splicing pattern even after successful alteration of the pre-mRNA. This strategy could therefore be further investigated and optimised.

Of note, in an attempt to validate the βglobin-luciferase reporter system used in this assay, we also tried to induce the skipping of exon 2 through the use of antisense oligonucleotides in order to assess if an expression of Luciferase could effectively be recovered. The second exon of βglobin pre-mRNA contains three well-characterised exonic splicing enhancers (ESE) that are activated by SR proteins, each of them independently able to induce the splicing of an heterologous pre-mRNA [60]. We therefore designed 2'-O-Methyl oligonucleotides directed against each of these ESE and tested them by co-transfection with the reporter construct. However, no expression of Luciferase could be detected as a result of AON expression, administered one by one or simultaneously (data not shown), demonstrating that antisense-mediated exon skipping was inefficient to modulate the splicing of this reporter construct. This observation could mean that the steric blocking of these ESE sequences was not sufficient to alter the splicing pattern and inhibit the splicing of this exon. Accordingly, it has been suggested that additional splicing enhancers sequences could be present in βglobin exon 2 [60], which could promote the splicing of this exon even when the three main characterised ESE have been inactivated.

The BP methylation strategy for splicing modulation, as well as the exon-skipping reporter construct designed in this study, therefore need to be further investigated and potentially optimised before considering further applications.

II. 3. Conclusion

Here we tried to design and improve different methods for splicing alteration by deriving small nucleolar RNAs to specifically target pre-mRNAs and induce exon skipping. We first tried to use snoRNAs as an alternative antisense delivery system for exon skipping application, and we also considered the property of C/D box snoRNAs to direct site-specific methylation of cellular RNAs as a mean to target the branch point of pre-mRNAs and inhibit the splicing of the downstream exon.

Overall these data indicated that the newly designed snoRNA-based delivery of antisense sequences to target mouse dystrophin exon 23 skipping is not an efficient system compared to the U7snRNA approach. Although some dystrophin exon skipping events could be observed in different experiments through AAV-mediated delivery of high doses of transgene, demonstrating that the modification of snoRNAs for antisense sequences delivery is feasible, the overall efficiency of this approach did not meet expectations.

Furthermore, the use of modified C/D box snoRNAs to direct the methylation of a branch point (BP) adenosine and inhibit the splicing of the downstream exon in a target pre-mRNA prove to be highly inefficient. No significant increase of luciferase signal was recorded when the modified U24 snoRNA was co-expressed with the luciferase reporter construct, indicating that no exon skipping could be obtained in the target pre-mRNA through this system.

Altogether, although both the snoRNA-based delivery of antisense sequences and the BP methylation strategy could benefit from further analysis and optimisations, the very poor exon skipping efficiencies obtained with these strategies suggest that the use of snoRNAs as new molecular platforms to modulate pre-mRNA processing presents little potential for an application to disease treatment by exon skipping.

DISCUSSION

Antisense oligonucleotide-mediated therapy for the treatment of splicing-associated diseases has been extensively studied over the past ten years and many efforts have been focused on correcting the mutated dystrophin transcript by exon skipping for the treatment of Duchenne muscular dystrophy [400]. However, the delivery of antisense sequences inserted into small nuclear RNA cassettes by viral vectors still faces some hurdles before being considered for clinical translation as large amounts of vector would still need to be administered to patients in order to stably restore therapeutic levels of dystrophin protein in all myofibers. In an attempt to maximise the therapeutic index of such vectors, different approaches were investigated in this study. Here we discuss the efficiency and limitations of the small RNA-antisense cassettes designed and optimised in this study, U7 snRNA- or snoRNA-based, and analyse their potential for clinical applications.

I. Limitations of the U7snRNA cassettes

I. 1. Cassette design and optimisation

During this study we have used the U7 snRNA to deliver antisense sequences for specific exon 23 skipping in the mouse dystrophin pre-mRNA. The choice of this snRNA over the other frequently used U1 snRNA was based on various observations. The endogenous U1 snRNP is detected in cells at much higher levels than the U7 snRNP due to its higher gene copy number and to a more efficient assembly into snRNP, which could lead us to favour its use over the U7 snRNA to shuttle antisense sequences. However, the modification of the Sm binding site into the consensus SmOPT sequence in the U7 gene leads to levels of expression equivalent to that of the U1 gene [254]. Accordingly, it has been demonstrated that modified U1 snRNA carrying antisense sequences were not more efficient than the engineered U7-SmOPT cassettes to alter the splicing of pre-mRNAs, and replacing the U7 promoter with the U1 counterpart to drive U7 snRNA transcription did not result in increased snRNA levels [209-210]. Based on these observations, the U7 snRNA seemed as good a candidate for the intracellular delivery of antisense sequences and was then chosen in our study to target the skipping of exon 23 in the murine dystrophin transcript.

Furthermore, a new regulatory role for U1 snRNA aside from its implication in pre-mRNA splicing has recently been brought to evidence, providing an explanation to the presence of such high levels of U1 transcripts in cells. The spliceosomal U1 snRNA has been shown to recognize and associate with cryptic polyadenylation signals that can occur at different positions along the transcript and to inhibit the premature cleavage and polyadenylation (PCPA) process, thereby ensuring the integrity of the mRNA [401]. This U1-mediated regulation of PCPA, termed telescripting, therefore protects pre-mRNAs from premature termination but also constitutes a mechanism to regulate gene expression. This novel function of U1 snRNA makes its use as an antisense carrier less attractive as it is not clear to what extent the overexpression of exogenous modified U1 could impair the PCPA-mediated regulation of gene expression.

I. 1. 1. scAAV vectors for U7 snRNA cassettes expression

The first optimisation tested in this study was the use of self-complementary AAV vectors to deliver the U7Dtex23 cassette in muscular cells, which have previously demonstrated enhanced transduction efficiencies over single-stranded counterparts. Our initial experiments did not show any significant increase of U7Dtex23 expression or of dystrophin exon 23 skipping compared to that obtained with conventional AAV vectors (**Figure 45**). This could be due to a spontaneous two-by-two association of small transgenes such as U7 snRNAs when high doses of ssAAV vector are administered. In such a situation, the snRNA cassettes would readily be present as dimer molecules in the transduced cells, thereby abolishing the need for double-strand synthesis of vector genome and mimicking the potential benefit obtained from self-complementary vectors. Accordingly, when lower doses were used, the delivery of U7Dtex23 cassettes by scAAV vectors induced a three-to four-folds increase in the proportion of exon 23-skipped dystrophin mRNAs compared to the same dose of ssAAV (**Figure 47**). Moreover, we observed that the rate of transgene synthesis was accelerated with the use of scAAV vectors, inducing a high level of dystrophin exon 23 skipping early after transduction and much faster than that obtained from ssAAV vectors (**Figure 46**). Altogether these data suggested that the use of self complementary vectors to express chimeric U7-antisense cassettes could enable to decrease the dose of vector necessary to obtain significant levels of dystrophin exon skipping, which presents an obvious interest for the treatment of DMD where high doses of vector are so far still needed.

These encouraging data are however to be taken precociously as various technical issues could have led to a mis-interpretation of the results. Firstly, it must be noted that the choice of Taqman probes and primers used for the quantification of dystrophin exon 23 skipping was probably not optimal. By only quantifying the transcripts containing exons 22-23 junctions and exons 22-24 junctions, we failed to consider the double exon skipping events leading to dystrophin transcripts lacking both exons 22 and 23 that are detected on agarose gel by nested PCR (**Figure 44**). We therefore did not determine the total amount of dystrophin transcripts expressed in the treated cells and the total percentage of exon 23-skipped mRNAs is only approximate. It would have been helpful to include in the assay a third pair of primers/probe in distant exons not affected by the antisense in order to quantify the total amount of dystrophin transcripts.

A second technical issue potentially hindering the correct interpretation of data concerns the quantitative PCR method used to assess the viral titer of an AAV preparation. This method is of limited reliability as the result obtained gives an estimation of the number of copies of viral genome present in the sample but does not relate to the number of capsid proteins and does not actually indicate the proportion of infectious viral particles. It is therefore possible that two AAV preparations quantified as containing comparable amounts of viral genomes could actually present important differences in transduction efficiencies. Furthermore, it has been recently demonstrated that the quantification of vector genomes in a self-complementary AAV preparation by qPCR can lead to an underestimation of the viral titer. The scAAV genomes have been shown to rapidly self anneal into the hairpin structure immediately after the PCR denaturation step, thereby inhibiting primer annealing and preventing the amplification of proximal amplicons [402]. This phenomenon often results in an underestimation of the copy number of vector genomes present in a scAAV preparation. The authors show that, depending on the position of the pair of primers used relative to the mutated ITR, the measured titer of an AAV preparation by conventional qPCR can be up to 10-fold lower than the actual viral genome content of the sample and could therefore lead to the administration of a 10-fold higher dose of scAAV vector than initially thought. To counter this effect, a method based on the cleavage of scAAV genomes by endonuclease digestion close to the terminal hairpin was designed in order to separate the two amplicons prior to the qPCR reaction and to enable an accurate quantification of vector genomes[402].

Considering this observation, it is possible that the improved transduction efficiency that we observed with the scAAV vectors was due to a higher dose of vector administered, leading to an overestimation of the benefit obtained. It would therefore be important to reiterate the titration of the scAAV preparations used in our study by realising the endonuclease digestion of terminal hairpins mentioned above [402]. Alternatively, the titration of vector genome by Southern blot or by dot blot assay could be accomplished as both techniques seem to be unaffected by the structure of the scAAV genome [372-373]

Nonetheless, our results suggested that the delivery of U7Dtex23 cassettes for dystrophin exon 23 skipping in mouse muscular cells was improved by the use of self-complementary AAV vectors. Accordingly, this benefit was recently demonstrated *in vivo* using the same U7Dtex23 cassette to correct the dystrophin transcript in the severe dKO mouse model [267]. In this study, a single systemic injection of 10^{13} vg of self-complementary AAV9 vector encoding the modified U7 snRNA cassette induced long-term restoration of Dystrophin protein in all muscles of the body, including the heart. Whilst this study did not directly compare the efficiency of scAAV over ssAAV vectors, the particularly high levels of dystrophin exon 23 skipping obtained following a single injection of scAAV9, ranging from 20% to 50% of total dystrophin transcripts depending on the muscle types, demonstrated that the delivery of U7Dtex23 cassettes by scAAV vectors was a very efficient strategy to transduce muscular tissues and stably correct the mouse dystrophin transcript.

I. 1. 2. The MHCK enhancer for increased U7Dtex23 expression

We further tried to optimise the expression of the U7Dtex23 cassette in cultured muscle cells by using muscle-specific enhancer elements, aiming to improve the skipping of mouse dystrophin exon 23 at a minimal dose of AAV vector administered. The MHCK enhancer used in our study was based on previous work where a synthetic fusion between the α-MHC enhancer and regions of the MCK enhancer and promoter induced high levels of transgene expression in muscular cells *in vitro* and *in vivo* [383]. We observed that the MHCK element placed upstream of the U7 promoter increased the levels of modified U7Dtex23 snRNAs expression by approximately 1.5-folds in cultured myotubes and resulted in a two-fold increase in the proportion of exon 23-skipped dystrophin transcripts, reaching equivalent levels than that achieved at a 10-fold higher dose of vector without enhancer (**Figure 55 and 56**). This improvement seems however weaker than the one reported in the original study, where the MHCK7 cassette, which consists of the same MHCK fusion enhancer coupled to a modified MCK promoter, induced a up to 10-fold increase in transgene activity in transfected mouse muscle cells compared to the native mouse MCK enhancer and promoter (CK1 cassette) [383]. The U7 snRNA promoter is however very active and of comparable strength to the spliceosomal snRNAs promoters [247, 254], which themselves are similar in activity to the CMV viral promoter [403]. It can therefore be expected that an enhancer element placed upstream of such an active promoter would only produce a moderate effect.

We could also obtain a beneficial effect *in vivo* when AAV vectors encoding the MHCK-driven U7Dtex23 cassette were administered in the tibilias anterior muscle of *mdx* mice. A moderate increase in the proportion of skipped dystrophin transcripts was observed in the presence of the MHCK enhancer, correlating with an increase of around 1.5-fold in U7DTex23 transcript expression (**Figure 57**). This enhancement in exon 23 skipping resulted in a three-to-five fold increase in Dystrophin protein expression and enabled to obtain levels of protein rescue equivalent to the ones obtained with a five-fold higher dose of vector without enhancer (**Figure 58**). Interestingly, we observed that the beneficial effect obtained with the MHCK cassette was mostly noticeable at intermediate doses of vector but that no significant improvement could be obtained at higher doses, which suggested that saturating levels of transcripts had been reached in the transduced cells. Nonetheless, our results consistently pointed at a positive effect of the MHCK sequence on U7Dtex23 expression, which could allow decreasing the

amount of vector administered to reach satisfactory amounts of dystrophin exon 23 skipping and protein rescue in treated muscular tissues.

We also observed the loss of tissue specificity of the MHCK-driven U7 snRNA cassette, demonstrated by the increase in U7Dtex23 expression in non-muscular cells in the presence of the enhancer (**Annex 1, Figure S4**). In contrast, the MHCK7 cassette originally reported allowed high-level of transgene expression in muscles whilst presenting an highly reduced activity in non-muscular tissues [383]. Yet, this MHCK7 element contains the MCK promoter which contributes to the muscle-specific activity of the cassette, whereas in our design the MHCK enhancer was fused to the ubiquitous U7 promoter. This is in line with the previously reported loss of tissue specificity of both MCK and MHC enhancers when coupled with constitutive promoters [404]. However, even if the expression of the U7Dtex23 cassette is augmented in all tissues in the situation of a systemic injection of vector, the tissue-specificity is dictated by the expression of the dystrophin transcript and therefore the U7Dtex23 can have no effect in non-muscular tissues that do not express dystrophin.

The MHCK enhancer-U7 hybrid promoter presents improved transcriptional activity of the antisense cassette in muscular tissues and therefore constitutes a promising tool to allow reducing the vector dose needed for therapeutic dystrophin exon skipping. It would now be interesting to assess the activity of this enhancer sequence in cardiac muscle after systemic injection of AAV vectors in *mdx* mice to assess if an increase in U7Dtex23 expression and in dystrophin exon 23 skipping could be obtained in this tissue, in agreement with the very high activity of the MHCK7 element reported in the heart muscle [383].

I. 2. Consequences of the engineering of U7 snRNA cassettes on their efficiency and processing.

I. 2. 1. Effects of the SmOPT modification

A surprising phenomenon was brought to light during our study as we observed the generation of smaller U7 snRNA species, about 20 nucleotides shorter than the expected U7Dtex23 size, following transfection or AAV transduction of mouse myocytes. Such shortened species issued from modified U7 snRNA cassettes have already been observed previously but the origin of this phenomenon was not further investigated [210]. This extra-processing of U7 transcripts was identified to occur on the 3' side of the snRNA, somewhere along the stem-loop structure, and we show that the formation of these 3' processed species is dependent on the presence of the SmOPT modification in the U7 snRNAs (**Figures 48-50**). Furthermore, we observed that the SmOPT sequence highly increases the expression levels of the U7Dtex23 transcripts *in vitro*, thereby greatly improving their efficiency as shuttles for antisense delivery and leading to a more efficient dystrophin exon skipping in the transduced cells (**Figure 51**). This observation has been described earlier by Grimm et al. who showed that modifying the U7-specific Sm sequence into the consensus SmOPT sequence redirected U7 transcripts into snRNPs containing the canonical Sm proteins, which led to a dramatic increase in U7 snRNA levels and a higher accumulation in the nucleus [254]. We concluded that, although the SmOPT mutation results in the formation of additional 3' processed U7 snRNA species, its presence is critical to achieve a maximum efficiency of antisense expression.

The mechanism involved in this SmOPT-dependent 3' processing of the U7 snRNA cassettes is however still unclear. By modifying the U7 binding site into the canonical SmOPT sequence the U7snRNA assembles into a particle whose behavior mimics the spliceosomal snRNPs. We could therefore speculate that the misprocessing is caused by an absence of correct cis-acting sequence on the U7snRNA that are important for the proper 3' processing of spliceosomal snRNAs. Yet, this explanation is not very likely as the U7snRNA-SmOPT cassette used in this study contains the native U7 3' box downstream of the mature 3' end and a snRNA-specific promoter, which are the critical signals for correct processing of a precursor snRNA [221, 225]. It was recently demonstrated that a cis-acting element present in the 3' stem-loop of Drosophila U7 snRNA is required for snRNA processing by the Integrator complex [223]. It is

therefore possible that the artificial packaging of U7-SmOPT into a spliceosomal snRNP with Sm proteins alters the secondary structure of the snRNA and exposes a cryptic nucleolytic site in the stem loop, which would trigger a partial cleavage of the snRNA during particle maturation and generate 3' shortened species.

It has also been shown that the SmOPT modification, in addition to increasing U7 transcript levels, also modifies the exonucleolitic trimming that takes place after assembly of the Sm particle and before re-entry into the nucleus [234, 253]. The endogenous U7snRNA transcript detected in the nucleus is 1 to 2 nucleotides shorter than the cytoplasmic transcript. On the other hand, the SmOPT-containing U7 transcript is shortened by 3 to 4 nt instead of 1 or 2, which indicates that the presence of the SmOPT binding site affects the processing of the U7 snRNA by promoting a 3' shortening reaction [254]. We could then hypothesize that the exonucleolitic trimming occuring before nuclear import could play a role in the generation of 20-nt shorter U7SmOPT transcripts.

More studies are therefore needed to understand this phenomenon as well as the factors responsible for this additional processing of the U7SmOPT transcript. It still remains unclear whether the shortened U7 species, which still carry the introduced antisense sequences, are also able to induce exon skipping or whether this 3' shortening along the stem-loop alters the secondary structure of the snRNAs and therefore destabilises and inactivates the transcripts. Nonetheless, the inclusion of the SmOPT mutation in chimeric U7 snRNAs appears critical for their efficiency and should therefore be conserved in the design of snRNA-based antisense cassettes.

I. 2. 2. The rate-limiting effect of the snRNA processing machinery

Another observation of interest in our study relates to the presence of high molecular weight (hMW) U7 RNA species accumulating at high doses of AAV vector and in a dose-dependent manner in transduced muscular cells. These hMW species may correspond to readthrough primary U7 transcripts containing both promoter and 3' untranslated regions, generated from tandem or circular copies of the vector genome known to form at high doses of AAV vector [294] (**Figure 60 and 61**).

We further noticed that by replacing the SmOPT sequence with the wild-type Sm sequence in the U7 snRNAs a diminution of the amount of non-processed hMW transcripts was detected (**Figure 51**). The proportion of hMW species observed could therefore be dependent on the presence of the SmOPT sequence. Considering that the SmOPT mutation enables the recruitment of Sm proteins to form a complex with the modified U7 snRNAs and to stabilize them after their 3' processing, it seems possible that the 3' extended transcripts containing the SmOPT consensus are also transported to the cytoplasm and stabilized by Sm proteins. This could explain the increased proportion of unprocessed transcripts observed in the presence of the SmOPT sequence as their association into canonical snRNPs would make them more stable than the ones containing the U7-specific Sm sequence.

We could also speculate that the presence of the SmOPT mutation directly affects the co-transcriptional 3' processing of U7 snRNAs by somehow preventing Integrator to cleave at the 3' end of the nascent transcript. This hypothesis is consistent with a previous study reporting that the formation of the correct 3'end of U1 snRNA is Sm site-dependent and that a mutation in this sequence causes misprocessing of the U1snRNA precursor, leading to accumulation of 3' extended species [235]. The SmOPT modification could have the same consequences and lead to a defect in 3' end formation. However, we also observed the generation of unprocessed hMW forms in the presence of the wild-type U7 Sm site, albeit in much lower quantities **(Figure 51),** suggesting that the misprocessing observed is not directly induced by the SmOPT sequence. Instead, it seems more likely that the accumulation of 3' unprocessed species depends on the amount of primary U7 transcript produced and reflects a saturation of transcript amounts in the cells when administering such high quantities of AAV vector.

Based on this idea, we could think that the SmOPT sequence directly activates the transcription and triggers the production of higher amounts of primary U7 transcript. In this situation, the Integrator complex that normally associates with the CTD of RNAPII during the transcription of U7 genes to induce their 3' end processing is then not present in sufficient amounts in the cells and is unable to associate on all the nascent transcripts, thereby leaving many U7 snRNAs in an unprocessed state. This accumulation of unprocessed readthrough U7 species at high doses of AAV vector therefore points at a saturation of the U7 3' end processing machinery when the primary transcript is produced in too many copies in the cells, which would limit the amount of mature U7 transcripts that can be produced intracellularly. Such accumulation of precursor U7 transcripts have been observed following injection of U7-SmOPT constructs in xenopus oocytes and have been suggested to result from a saturation of the processing machinery [247, 254]. Consistent with this hypothesis, a similar accumulation of primary U1 and U2 transcripts is observed after depletion of some sub-units of Integrator, suggesting that a shortage of Integrator complex available for the primary transcripts in the cells results in a defect in the processing of the snRNA 3' ends [224].

Further analysis are therefore needed to understand the origin of the high molecular weight RNA species generated when high levels of U7 snRNAs are transcribed in the cells. Understanding the molecular basis of this phenomenon could enable to modulate this effect and maybe increase the amount of mature U7 transcripts that can be produced in the cells. We could design an experiment where high doses of U7 snRNA-carrying AAV vectors are administered in the context of an overexpression of the Integrator complex's sub-unit Int11. If the amount of hMW species diminishes when Integrator is overexpressed, we could deduce that the appearance of unprocessed transcript is attributable to a saturation of the 3' end processing machinery and can be reversed by increasing the Integrator-to-primary transcripts ratio. Alternatively, we could use a previously described U7snRNA-GFP reporter system where the U7 snRNA gene, including its 3' processing signal, is fused to an in-frame GFP gene that is followed by a polyadenylation site [223]. In this system, aberrant 3' end formation of snRNA resulting in read-through transcription leads to a gain of GFP expression in the transduced cells. We could derive the same system with the mouse U7 snRNA gene containing the wild-type Sm or the SmOPT sequence and fused to the GFP gene, and deliver it to cultured cells with AAV vectors. If the U7 snRNA is properly processed in the cells no

expression of GFP should be detected. On the other hand, any detection of GFP expression would correspond to transcriptional readthrough past the U7 processing site and would therefore demonstrate an impairment of the processing machinery. By comparing the results obtained with or without the SmOPT sequence and at high or low, non-saturating doses of AAV vector administered, we could determine whether the misprocessing of U7 transcripts is function of the amount of primary U7 snRNAs produced, which would trigger increasing amounts of GFP being expressed with increasing doses of vector, or if the failure to process U7 snRNA 3' ends directly results from the presence of the SmOPT site itself, thereby generating a GFP signal even under non-saturating levels of vector.

This saturation hypothesis also explains the results obtained when trying to couple the effects of the self-complementary AAV vector delivery with the use of the muscle-specific enhancer MHCK. We have observed that the proportion of skipped dystrophin transcripts obtained with the scAAV vector was not improved by the enhancer (**Figure 62**). This is also consistent with the saturation hypothesis and it would be interesting to further establish this point in dose-response experiments, in which the minimal dose required to achieve comparable effects could be determined. It is however conceivable that these elements function sequentially, the scAAV vector providing early expression of high levels of U7 transcripts and the MHCK enhancer ensuring high levels of expression at later time points. A careful kinetic study is thefore also needed in order to assess the benefit of the self-complementary vector containing the MHCK- U7.

Altogether, these observations highlight the limitations faced when trying to maximise the expression of antisense-containing U7 cassettes in a tissue: the potential saturation of the cellular machinery implies that the processing of these transcripts can only occur at a certain rate and up to a certain level in the cells, leaving little room for optimisation of U7 snRNA shuttles expression. Of note, no high molecular weight RNA species was detected after local administration of U7Dtex23-containing AAV vectors in *mdx* mice muscles, even for the highest dose of vector injected. Differences in the cellular environment provided by myocytes in culture versus differentiated mdx muscle may account for this and it is therefore possible that saturation does not occur *in vivo*.

II. New strategies for exon skipping and the use of small nucleolar RNAs

II. 1. The snoRNA MBII-52 to express antisense sequences

II. 1. 1. Rationale for the study and choice of the MBII-52 snoRNA

We have tried to derive snoRNAs into antisense sequences carriers to target dystrophin exon 23 skipping, aiming to find new and potentially more efficient tools to modulate the splicing of this transcript. As snoRNAs are encoded within introns of RNA polymerase II transcripts, their concomitant synthesis allow the snoRNA to be transcribed by strong and regulatable mRNA promoters, which theoretically present an advantage over the U7 snRNA genes that specifically require snRNA-specific promoters for their transcription and processing. The natural presence of antisense elements in C/D box snoRNAs that present perfect complementarities to various cellular RNAs further deepened our argument for the feasibility of engineering snoRNAs as antisense carriers. Finally, it has been demonstrated that certain C/D box snoRNAs could be engineered as a tool for the targeted knockdown of protein expression by replacing the natural M box sequence with a sequence complementary to a target mRNA, providing further evidence that some snoRNAs can be redirected towards a target mRNA in order to influence its expression [390]. Following this line of reasoning, it seemed therefore possible to design a snoRNA to induce the skipping of dystrophin exon 23.

We chose to use the MBII-52 C/D box snoRNA to deliver antisense sequences for dystrophin splicing modulation as this snoRNA naturally associates with a messenger RNA, the HT2CR mRNA, to modulate its alternative splicing through its antisense element [203]. Furthermore, the several copies of the MBII-52 snoRNA present in the mouse genome display different sequence variations in their antisense sequences compared to the 18 nt consensus, and it has been recently suggested that these variants could associate with five additional mouse pre-mRNA targets and modulate their pattern of alternative splicing [106]. It seemed therefore fairly straightforward to derive such an mRNA-specific snoRNA to redirect it towards another pre-mRNA for splicing alteration.

II. 1. 2. low efficiency of the strategy and potential optimisations

The modification of the MBII-52 snoRNA into a Dtex23 antisense sequence delivery tool for mouse dystrophin exon 23 skipping did not prove to be an efficient system. The expression of MBII52-Dtex23 transcripts delivered by AAV vectors in muscular cells was detectable only at the highest dose of vector administered, and the overall mature transcript level was strikingly low for such high doses of vector (**Figure 66**). Accordingly, the skipping of dystrophin exon 23 occurred with very limited efficiency in transduced myotubes, and at levels much lower than that obtained with the same dose of U7Dtex23-containing AAV vectors (**Figure 67**). Furthermore, we observed that the delivery of MBII52-Dtex23 snoRNAs by lentiviral vectors was inefficient to induce dystrophin exon 23 skipping in cultured muscular cells (**Figures 68 and 69**). We concluded that, although the snoRNA-mediated delivery of antisense sequences to target pre-mRNAs is feasible in an AAV context and only at high doses, the general efficiency of this approach was very low and did not meet expectations. This lack of efficiency could stem from several issues related to the design of the system

- Problem of intracellular localisation

It is possible that the MBII52-Dtex23 transcript could not act on the dystrophin transcript because of different sub-nuclear localisations. The snoRNAs are predominantly localised in the nucleolus whilst mRNA splicing occurs in the nucleoplasm and, although it has been shown that endogenous MBII-52 snoRNAs could transiently localise in the nucleoplasm to modulate the alternative splicing of HT2CR pre-mRNA [204], this transit into the pre-mRNA compartment might be too transitory to allow splicing alteration of long dystrophin transcripts. Indeed, considering that the splicing of dystrophin pre-mRNA occurs co-transcriptionaly and that its transcription last approximately 16 hours, it is possible that the brief passage of the snoRNA in the nucleoplasm does not occur at all or at the right timing and for long enough to inhibit exon 23 splicing.

To assess this issue we could imagine a way to detect and localise the modified snoRNAs in live cells to gain information on their real-time sub-cellular trafficking as well as the abundance of mature molecules produced for a set amount of vector administered. The development of RNA-based probes that contain two aptamer

modules to specifically recognize small molecules and generate a fluorescent signal when bound to the targets could be envisaged for our purpose [405]. Such aptamer probes are composed of a detector moiety that bind to the small molecule to be detected and of a second aptamer domain that can bind to a dye such as malachite green. When the targeted ligand binds to the detector domain of the probe, a change of conformation of the probe occurs so that the dye that is added in the cell culture media can associate with the second aptamer moiety. By doing so, the malachite green that is not naturally fluorescent is submitted to a change of conformation that induces it to become fluorescent. The specificity of the technique is such that a fluorescent signal is emitted only when the targeted ligand binds to the detector module. As the probe is constituted entirely of RNA it can be easily produced through genetic encoding inside the cells and could be adapted to detect intracellular nucleic acids, such as snoRNAs [406].

It would be interesting to track the spatiotemporal localisation of transduced snoRNAs in the muscular cells using this type of approach to assess if they are able to localise in the nucleoplasm with the dystrophin pre-mRNA and if this potential co-localisation lasts long enough to act on the splicing of these transcripts. In order to discriminate multiple RNAs simultaneously., we could use an aptamer probe specific for the snoRNA and able to bind the malchite green dye, and a second aptamer probe for dystrophin transcript detection binding a different dye, such as the SRB dye described elsewhere [407]. In the case where the snoRNA is not able to transit through the nucleoplasm or at least not long enough, we could try to further modify the chimeric MBII52 molecule to force this localisation. It has been demonstrated that the conserved C/D box motif of the snoRNAs is necessary to induce nucleolar accumulation of the mature snoRNPs [201-202]. It could be envisaged to alter these C/D boxes in a way that would prevent the retention of the snoRNA in the nucleolus and would enable to redirect it towards the nucleoplasm, thereby promoting its co-localisation with target pre-mRNAs.

- Problem associated with the antisense sequence

The very low efficiency of the snoRNA cassette could be related to the size or the composition of the Dtex23 antisense sequence itself. The 44-nucleotide long Dtex23 antisense is more than twice larger than the original 18-nt HT2CR complementary sequence present in the endogenous MBII-52 snoRNA. The replacement of this natural antisense by the double-target Dtex23 sequence could therefore destabilise the transcript and lead to a disruption of the snoRNA secondary structure, thereby preventing the association of the transcript into a stable snoRNP and leading to its global inactivation and inability to associate with the target pre-mRNA. Instead of the long double-target Dtex23 antisense sequence we could therefore insert in the snoRNA construct a single-target antisense sequence that would be closer in size to the original MBII-52 antisense sequence. For example, a sequence directed against the 5' splice site of exon 23 alone and introduced into a U7snRNA cassette was reported to allow dystrophin exon 23 skipping in mouse *mdx* cells, albeit with less efficiency than the double-target construct [256]. This 24-nucleotide sequence inserted in the snoRNA cassette could be tested for its ability to induce exon 23 skipping, thereby deciphering whether the size of the antisense sequence plays a determining role in the efficiency of the snoRNA-based approach for dystrophin splicing modulation.

Alternatively, snoRNAs have also been derived as carriers to express ribozymes and target the cleavage of precise transcripts in *trans* [408]. In this study, a shortened version of the U3 C/D box snoRNA was modified by replacing its original 5' single stranded sequence involved in ribosomal RNA binding by the chosen hammerhead ribozyme sequence. This so-called "snorbozyme" was able to induce specific cleavage of the targeted transcript. To introduce the ribozyme into the U3 snoRNA construct, a 54-nucleotide segment of the 5' snoRNA extremity was replaced by the 46-nt catalytic portion of the ribozyme. This demonstrates that a large RNA sequence can be inserted into the U3 snoRNA without hampering its metabolic stability and whilst still maintaining the ability to direct the modification of cellular RNAs. We could therefore consider using this described mini-U3 snoRNA as a carrier to express the large 44-nt Dtex23 antisense sequence and assess if this approach enables to alter the splicing of mouse dystrophin transcript with more efficiency than the use of the MBII-52 snoRNA.

Another option would be to introduce the antisense into a snoRNA that naturally contains two complementary sequences directed against a cellular RNA. The U24 C/D

box snoRNA for example naturally contains two target recognition sequences of 12 nucleotides each to guide the methylation of two residues in the 28S rRNA [391]. We could therefore split the double-target Dtex23 sequence into the two sequences complementary to the branch point of intron 22 and the donor splice site of exon 23 of the mouse dystrophin pre-mRNA and insert each of these antisenses into each of the natural antisense boxes of U24. We could then assess if this strategy enables to form more stable snoRNPs and to induce the inhibition of exon 23 splicing more efficiently than the use of the MBII-52 snoRNA.

- Problem associated with the processing of the snoRNAs

The very low efficiency of the snoRNA-based delivery of antisense sequences for dystrophin exon skipping might also stem from inaccurate processing of the modified snoRNA. We observed in our experiments that only very low levels of mature MBII52-Dtex23 snoRNAs could be obtained in cells transduced with high doses of AAV vectors. Conversely, high quantities of high molecular weight species accumulated instead in a dose-dependent manner (**Figure 66**), probably corresponding to non-processed MBII52-Dtex23 transcripts still surrounded by their host pre-mRNA sequences and suggesting a problem in the processing of the pre-snoRNA from the debranched intron. As intronic snoRNA processing is physically and functionally linked to the splicing of the host pre-mRNA intron [194-195], it is possible that these hMW species correspond to full primary transcripts containing the host exons G2-G1 and the snoRNA embedded in the intronic sequence and could suggest that the splicing of exons G2-G1 is inefficient in the artificial construct.

It would therefore be important to test whether this transcript is actually being spliced and if the snoRNA-containing intron is properly debranched and released in the cells. We could for example realise a simple reverse transcription and PCR on RNA extracted from AAV-MBII52-Dtex23 transduced cells using primers placed on the surrounding exons. This would indicate whether exons G2 and G1 are properly joined in most of the transcripts or if an intronic sequence is still maintained in a proportion of them, thereby reflecting a default in the splicing of the artificial construct. In the latter case, optimisations of the chimeric construct could be proposed to force the splicing of the host pre-mRNA: it has been demonstrated that the position of the snoRNA in the host

intron relative to consensus splice sites is critical for both pre-mRNA splicing and snoRNA synthesis as the modification of the distance between the 3' end of the snoRNA and the branch point disrupts both processes and leads to a strong decrease in mature snoRNA accumulation. A distance of about 50 nucleotides between the snoRNA coding region and the branch point was then determined optimal for efficient snoRNA production [196-199]. In our construct the MBII52-Dtex23 snoRNA is located 60 nucleotides away from the intronic branch point sequence between exons G2 and G1, and it is possible that the splicing process is less effective in this construct as a result of non-optimal localisation of the snoRNA in the host intron. We could therefore try to optimise the position of the snoRNA in the intron by shortening the distance between its 3' end and the branch point sequence. Alternatively, we might consider expressing the snoRNA in another intronic context surrounded by exons that are known to be spliced with very high efficiency, such as in the βglobin pre-mRNA [393]. It has indeed been demonstrated that all the necessary signals for correct processing of a snoRNA reside within the snoRNA sequence itself and that snoRNAs artificially placed into introns of mRNAs that do not normally encode snoRNAs are accurately processed [191]. We could therefore insert the MBII52-Dtex23 in an optimal intronic context between strong splice signals in a constitutively spliced sequence and distant to the branch point by approximately 50 nucleotides. By doing this, we might be able to eliminate the issue of incomplete processing and hMW species generation and obtain a larger amount of antisense sequences expressed in the transduced myotubes.

Another processing-related issue is that the MBII52 snoRNA may be processed into several shorter fragments lacking the 5' and 3' ends by exonuclease trimming that finishes at the C and C' boxes [106]. In this case, it is possible that a good proportion of MBII52-Dtex23 transcripts would also be shortened into small RNA fragments, called psnoRNAs, and that the 105-nucleotide full-length form would be only a minor product of the MBII52-Dtex23 expression unit, explaining why only few 105-nt transcripts could be detected on the Northern blot assay. Furthermore, as the described shortened species often lack 5' and 3' sequences up to the C and C' boxes, such RNAs would not contain the Dtex23 sequence anymore and would not hybridize with the MBII52-Dtex23 radioactive probe, which could explain the absence of short fragments detected on the Northern blot. Although the existence of these MBII52-derived psnoRNAs is controversial and has since been questioned [388], studies have demonstrated the presence of snoRNA-derived short RNA species, usually shorter than 35 nucleotides,

stably accumulating and conserved in various cell types [389]. These small RNAs are processed from a proportion of C/D box or H/ACA box snoRNAs that are transported to the cytoplasm and many of them have been shown to act as regulatory RNAs and affect gene expression. Indeed, some of these short RNAs are able to associate with the RNA silencing machinery and to function like miRNAs in posttranscriptional gene silencing [409-412]. These findings support the fact that various snoRNAs can be processed into short RNA species and it is therefore possible to suggest that the same phenomenon could affect the modified MBII52 snoRNA, processing it into short RNA fragments undetected in our experiments and limiting the amount of full-length transcripts expressed in the transduced cells.

Interestingly, it has been demonstrated that MBII52 snoRNA could assemble into non-canonical RNPs by recruiting novel proteins distinct from known snoRNP-associated proteins [204]. In particular, the study from Kishore et al. suggests that shortened forms of the MBII52 RNA lacking the 5'-3' stem sequences are not able to associate with snoRNA-specific proteins such as fibrillarin but instead can assemble with hnRNP proteins known to be involved in pre-mRNA splicing [106]. The authors describe a model where the MBII52 snoRNP associates with its target RNA and is able to influence splice site selection through the action of the associated hnRNP proteins.

Based on this notion, we could imagine a strategy to force the association of the chimeric snoRNA with hnRNPs, thereby favouring its action on the splicing of the target pre-mRNA. This could be realised by eliminating the stem sequences at each end of the snoRNA coding region in our construct in order to artificially influence the generation of psnoRNAs that associate with hnRNPs. Although the presence of the terminal stem is thought to be necessary for correct processing of the snoRNA, many snoRNAs naturally lack a terminal helix and are yet still processed into mature and stable snoRNAs. It has been demonstrated that such snoRNAs are processed from the intron of host pre-mRNAs owing to the transient formation of external stems in the surrounding intronic sequences that enable to bring the C and D boxes close together and to allow the binding of snoRNA-associated proteins necessary for correct snoRNA processing [413]. Further studies of these flanking sequences and of the signals necessary to form an external stem structure could allow us to introduce a stem-less MBII52-Dtex23 snoRNA, truncated at the 5' and 3' ends, into intronic sequences able to form the necessary external stems. As these intronic structures would then be

degraded during snoRNA processing, this strategy could therefore allow the release of modified psnoRNAs able to associate with hnRNP proteins and to alter the splicing pattern of target pre-mRNAs with high efficiency.

II. 2. Targeted methylation of the branch point adenosine for exon skipping

II. 2. 1. Rationale for the study and choice of the U24 snoRNA

Our third line of investigation relates to the improvement of an alternative strategy to induce exon skipping in pre-mRNAs through the targeted methylation of the branch site by modified snoRNAs. Although this strategy had been previously validated for pre-mRNA splicing modulation in human cells, the low efficiency reported indicated the need for improvements [414]. In a recent study, artificial U24 C/D box snoRNAs targeting methylation at different consensus splice sites of a pre-mRNA altered the alternative splicing pattern by increasing the proportion of exon-depleted mRNA isoforms from 7% in non treated cells to up to 30% in snoRNA-transfected cells [415]. We decided to reproduce this approach by re-directing the human U24 snoRNA towards the branch point of an exon-skipping reporter construct, aiming to first validate this approach before ultimately trying to optimise it for dystrophin exon skipping application. The choice of the U24 snoRNA for our study stemmed from the demonstration in these earlier studies that this transcript can be derived to target pre-mRNAs and induce their site-specific methylation. Furthermore, this snoRNA presents the advantage of naturally containing two target recognition sequences that can hybridize on two adjacent sequences in the target RNA, thereby strengthening the association between the molecules.

II. 2. 2. low efficiency of the strategy and potential optimisations

Owing to the published results we were expecting to obtain low levels of exon skipping in our reporter system. However, even though the modified U24 snoRNA and the luciferase reporter construct were transfected at a 5:1 molecular ratio, providing an excess of U24 molecules to target the βglobin sequence in the reporter pre-mRNA, no significant increase of luciferase signal was recorded suggesting that no exon skipping could be obtained in the target pre-mRNA (**Figure 75**). Our data therefore implied that this exon skipping strategy was greatly inefficient in our system, which could be due to various factors.

- Problem associated with the modified snoRNA and its ability to direct the methylation of target nucleotides

We could consider that the inefficiency of the strategy stemmed from the same problems stated before associated with the design of the snoRNA. For example a non-effective splicing of the host gene or a problem of processing of the snoRNA from the debranched host intron could also occur in this system as described for the modified MBII52 snoRNA. However, such a processing issue is not likely as good amounts of modified U24 transcripts could be observed at the right size on the Northern blot from RNAs of cells transfected with both the reporter construct and the chimeric U24 plasmid. This indicated that U24 synthesis and processing was correctly achieved from the synthetic construct and was present in virtually sufficient amounts to direct the modification of target pre-mRNAs. A problem of sub-nuclear localisation is not very likely either as it has been shown that upon transfection of human cells a similarly modified U24 snoRNA was able to interact with the pre-mRNA target inside the nucleus to influence its splicing and that artificial snoRNAs could be detected both in the cytoplasm and nuclear fraction of transfected cells [415]. This suggested that no constraint forced the modified U24 to an exclusive nucleolar localisation but that it could instead transit in various sub-cellular compartments. It is therefore likely that the U24 transcript in our study was also able to reach different cellular compartments and that the absence of activity was not due to a failure to find its target in the nucleoplasm. It would however be important to experimentally verify the localisation and processing

of the modified U24 snoRNA in the transfected cells using the approach described above.

Another possibility is that the snoRNA was not able to efficiently direct the methylation of the pre-mRNA branch point and could not trigger the skipping of the downstream exon. Indeed, in the study realised by Stepanov et al., although a splicing impairment was obtained in the presence of the artificial U24 snoRNA it could not be clearly demonstrated that this increase in exon skipping was the result of site-specific 2'-O-methylation of the target pre-mRNA as the potential modification of the targeted nucleotide could not be verified [415]. It is therefore possible that the inhibition of splicing observed was instead the result of an antisense effect obtained by blocking the access of the splicing machinery to consensus splice sites via complementary hybridization of the artificial snoRNA. In this case, if the beneficial effect observed in their study is due to an antisense effect rather than the directed methylation of the splice sites, this could explain the absence of efficiency in our system: it is possible that the steric blocking of the branch sequence of βglobin pre-mRNA was not enough to inhibit the splicing of the downstream exon and that the antisense effect did not apply in this situation.

It would therefore be of key importance to test whether the 2' -O-methylation of the target branch point in βglobin pre-mRNA actually occurred in our assay, for example by realising a primer extension-based 2'-O-methylation assay at high or low dNTP concentrations as previously described [186]. In the case where no nucleotide modification could be detected, a likely explanation could be that the modification of U24 snoRNA antisense sequences prevented its correct assembly with methyltransferase proteins into a functional snoRNP, thereby precluding the 2'-O-methylation of the target RNA branch point and the skipping of the downstream exon. However, if the targeted branch point was accurately 2'-O-methylated whilst no exon skipping could be obtained, this could mean that the methylation of the branch point was not efficient enough to compete with the multiple known ESEs contained in βglobin exon 2 and to induce splicing inhibition of the downstream exon [60]. Finally, it is possible that the 2'-O-methylation occurred only on a small proportion of pre-mRNAs and left a majority of unmodified transcripts that present a normal splicing pattern, thereby precluding the detection of significant luciferase increase associated with exon skipped-transcripts.

As an alternative strategy to inhibit splicing we could try to target the 2'-O-methylation of nucleotides at different splice sites, such as the splice donor and splice acceptor sites at each extremity of the intron, as it has been shown that the targeted modification of these consensus splice sequences by C/D box snoRNAs induced the skipping of the downstream exon with a comparable efficiency than targeting the branch point [415]. As altering or masking different splice sequences in a transcript can trigger exon skipping with variable efficiencies depending on the pre-mRNA studied, it is possible that methylating the 5' or 3' splice site instead of the branch point in our βglobin-luciferase transcript could alter the pre-mRNA splicing with a better efficiency. It could also be envisaged to use the ability of U24 snoRNA to simultaneously target and methylate two different sequences on the same pre-mRNA: we could for example combine two antisense sequences that would concurrently hybridize the branch point and the 3' splice site of the pre-mRNA and induce the 2' -O-methylation at both sites, thereby increasing the chances of altering the splicing of the 3' exon.

It has also been recently demonstrated that H/ACA box snoRNAs that naturally catalyze the post-transcriptional isomerization of uridine to pseudouridine at specific site within target RNAs [416] could be redirected towards various RNAs, including mRNAs, by modifying their guide sequence [417-418]. In these studies, artificial H/ACA snoRNAs were able to successfully induce the pseudouridylation of a premature termination codon within the target pre-mRNA, thus allowing translational readthrough and the recovery of a full-length protein, albeit with a limited efficiency (5 - 10%). Yet, this strategy is not confined to nonsense suppression applications but could potentially be derived for splicing modulation purposes. As pseudouridylation of uridines induces structural and biochemical changes in the transcript, we hypothesize that the isomerisation of a uridine residue situated in a splice site could affect its function in the splicing process. Considering that pseudouridylation occurs only on uridine residues it is impossible in this case to target the modification of an adenosine branch point with this process. We could engineer an H/ACA snoRNA to direct the pseudouridylation of the uridine present in the critical GU dinucleotide at the donor splice site of our βglobin-reporter transcript in order to inhibit the splicing process.

- **Problem associated with the luciferase exon skipping reporter construct**

It is also plausible that the apparent absence of exon skipping obtained by the snoRNA-induced BP methylation strategy was linked to a wrong design of the reporter construct. The inactivation of the major branch point of intron 1 in the βglobin-luciferase transcript through targeted methylation could very possibly have triggered the activation of one or several cryptic branch points to restore the normal splicing pattern. Consistent with this hypothesis, it has previously been demonstrated that the inactivation by 2'-O-methylation of an authentic branch point in an intron of the adenovirus pre-mRNA led to a correct splicing pattern through the use of cryptic branch points [206]. Moreover, it has been shown that the intron 1 of human βglobin pre-mRNA contains several cryptic branch points, always adenosine residues situated 22 to 37 nucleotides upstream of the 3' splice site of exon 2, which can become activated following mutation of the authentic branch site to enable the accurate splicing of the pre-mRNA [395]. Providing the methylation and inactivation of βglobin intron 1 branch point did actually occur in the reporter construct, it is therefore very likely that the splicing machinery selected one or several of the cryptic adenosines identified by Ruskin et al. to recover the correct splicing pattern of βglobin pre-mRNA and thereby maintain the fused luciferase transcript out-of-frame. Therefore, even if the targeted 2'-O-methylation did occur at the authentic branch point and efficiently inactivated it, this potential success could have been concealed by the presence of cryptic branch points and would have led us to conclude on the inefficiency of the strategy. It would then be important to identify and mutate all the putative cryptic branch sites present in βglobin intron 1 of the reporter construct in order to decidedly assess if the snoRNA-directed BP methylation strategy could present some efficiency for the skipping of a downstream exon. This demonstrates a weakness of this approach as each target branch site in a specific pre-mRNA will have to be optimized relative to the presence of cryptic branch points in the surrounding intronic sequences, which could be very fastidious.

III. Relevance for treatment of diseases by exon skipping

In this study several exon skipping strategies have been investigated in the aim to identify suitable treatments for genetic diseases that could benefit from the skipping of one or several exons in a mutated pre-mRNA. The use of modified C/D box snoRNAs to direct the methylation of a branch point adenosine and inhibit the splicing of the downstream exon in a target pre-mRNA was investigated. However, although this strategy could benefit from further analysis, the overall low exon skipping efficiency suggests that it presents little potential for an application to disease treatment by exon skipping. The use of antisense sequences for the steric blocking of splice signals to inhibit the splicing of an exon is to date a much more promising strategy and is more likely to continue being the focus of research for the development of clinically applicable exon skipping strategies. We therefore concentrate here on these antisense-based approaches and analyse their potential for a possible translation towards clinical application.

For the intracellular delivery of antisense sequences we investigated and compared two different carrier molecules, snRNA or snoRNA cassettes, to express steady levels of antisense sequences able to trigger exon skipping. These exon skipping strategies were tested in the context of the dystrophin pre-mRNA for the treatment of Duchenne muscular dystrophy. Here we consider the difficulties associated with this choice of disease model to assess various exon skipping approaches. Furthermore, we expose the arguments in favour of the vectorisation of antisense sequences for long-term treatment. Finally, we discuss the likelihood of each antisense delivery approaches, snoRNA-based or U7 snRNA-based, to realistically be translated into clinical treatments of DMD.

III. 1. The model: dystrophin pre-mRNA exon skipping

Targeting the dystrophin pre-mRNA can be challenging as various factors can hamper the action of antisense sequences. As mentioned before, the availability of the dystrophin transcript in muscular cells varies depending on the myogenic stage of the culture (present study and [303, 362-363]). By expressing antisense sequences too early in myoblasts, where the dystrophin transcript has not yet been synthesised, or too late when the splicing of most of the transcripts already occurred, could impair the efficiency of exon skipping and could lead us to mistakenly judge a strategy as inefficient. It is therefore critical to carefully choose the appropriate time to administer antisense sequences to modulate the splicing of the dystrophin transcript in order to obtain the best exon skipping efficiency possible. In addition, it has been shown that the dystrophin transcript folds into transient secondary structures during its transcription [419], and this is the time when antisense sequences must hybridize on the transcript as the splicing process occurs co-transcriptionally. This pre-mRNA folding can therefore impair the accessibility of antisense sequences to their target sequences and therefore prevent exon skipping depending on the sequence targeted [419]. It is therefore important to verify the potential engagement of the targeted sequence into a secondary structure during dystrophin pre-mRNA transcription to ensure the efficiency of the exon skipping strategy.

Furthermore, there is a high variability in exon-skipping efficiency among muscular cell types. In our study, we evaluated the different antisense cassettes for dystrophin exon skipping in wild-type mouse muscular cells. However, it has been shown that the administration of antisense sequences in dystrophic muscles, such as *mdx* mouse muscles, induces superior exon skipping efficiencies than in healthy muscles [314, 420]. Testing an antisense-based strategy in wild-type cell cultures could therefore lead to an underestimation of efficiency. This observation however applies mainly for synthetic AONs and is primarily the result of an enhanced penetration of the oligonucleotides in dystrophic muscle fibres that present increased permeability. It is less likely to be a variability factor in our study where antisense oligonucleotides are delivered via viral vectors, which should enter healthy or pathologic cells with the same efficiency. Another important point to consider when testing exon skipping strategies on wild-type cells is the probability of disrupting the reading frame of the normal dystrophin transcript when inducing exon skipping. In the case of the skipping of dystrophin exon

23, the reading frame of the whole transcript is maintained even when altering wild-type pre-mRNAs. However, in the context of human dystrophin pre-mRNA splicing, there are some cases where the skipping of an exon would only restore the reading frame of a mutated transcript but would alter that of a wild type pre-mRNA. This is the case of human dystrophin exon 51 skipping that has been designed to restore the reading frame of exon 52-deleted transcripts: the reading frame is actually recovered upon skipping of the disease-associated transcript but is disrupted in wild-type transcripts that still contain exon 52, triggering the apparition of premature stop codons. In this case, wild-type mRNAs with disrupted reading frames would be susceptible to degradation by nonsense-mediated decay, which would impair the detection of skipping events and lead to an underestimation of exon skipping efficiencies [339].

In addition, it is to be considered that the situation observed *in vitro* might be fairly different than the one encountered *in vivo*. For example, amongst a pathologic muscle some muscle fibres are able to respond better than other to exon skipping treatment, which often leads to non-uniformly expressed dystrophin in a whole treated muscle, and the overall responsiveness to treatment also fluctuates from muscle to muscle in animal models [421]. The amount of dystrophin transcripts synthesised in mouse muscles *in vivo* also varies compared to that produced in cultured myotubes: it has been suggested that the level of dystrophin pre-mRNA observed *in vitro* is approximately 10 fold less than the one detected in adult mouse muscle, and a similar observation was made in human myotubes [364]. This could suggest that the amount of antisense sequences needed to be administered in order to correct a good proportion of dystrophin transcripts need to be adapted and increased for *in vivo* treatment.

Finally, it is now clear that very high doses of antisense sequences expression are necessary to correct enough dystrophin transcripts and restore therapeutical levels of dystrophin protein expression [422]. This need for high levels of antisense expression could lead us to overlook some strategies that display poor efficiencies in the context of DMD but that might prove sufficient for applications in other diseases.

Altogether, these peculiar characteristics associated with dystrophin expression patterns and variability amongst muscular cells point to the difficulty of trying to adapt an exon skipping strategy for the treatment of DMD and suggest that this disease might not constitute the best model when trying to test and compare new exon skipping strategies. For such purposes, it is therefore more adapted to develop exon skipping reporter constructs that can provide an easy read-out of exon skipping efficiency. In this study however our major aim was to identify strategies that enable to increase the amount of antisense sequences that can be expressed in muscular cells specifically for the treatment of DMD by exon skipping. In that prospect, the newly designed snoRNA-based system showed strong limitations in its ability to deliver high levels of antisense sequences and displayed very limited dystrophin exon skipping efficiency, even when high doses of vector were administered. This need for high levels of expression of antisense sequences therefore led us to dismiss this snoRNA-based strategy for an application to dystrophin pre-mRNA splicing modulation. The already well-studied U7 snRNA cassette still constitutes the most likely strategy to reach further stages towards potential clinical trials for DMD therapy and we will now focus for the remainder of the discussion on this system and the optimisations realised in this study.

III. 2. The pros and cons of the vectorisation of antisense sequences

In this study we have chosen to deliver antisense sequences to muscular tissues and cells through the use of viral vectors. The modified U7 snRNA can be used as a carrier and expression cassette for antisense sequences and can be easily introduced in a recombinant AAV vector for transduction of muscular cells. The use of viral vectors presents several advantages over the delivery of synthetic oligonucleotides but still faces several challenging barriers that have precluded clinical trial so far.

Firstly, it is important to note that some antisense sequences that are active and efficient for inducing exon skipping when used as synthetic oligonucleotides can be inactive when introduced in expression vectors. One example is the M23D sequence that targets the donor splice site of mouse dystrophin exon 23, which is able to induce good levels of exon skipping when used as an oligonucleotide [423-424] but is not efficient when vectorised into a U7 snRNA cassette delivered by AAV vectors (data not shown). A plausible explanation for this observation is that a particular sequence introduced into the U7 snRNA might trigger the U7 snRNA to form an abnormal secondary structure and prevent its association into a functional snRNP, thereby abolishing its potential effect to modulate splicing. In addition, the efficient expression of antisense sequences introduced into small RNA cassettes can be dependent on the choice of viral vector used for intracellular delivery and in some cases, small RNA-antisense cassettes that are active when delivered by AAV vectors can display a complete absence of activity in a lentiviral background. Nonetheless, once an antisense sequence has proven active when introduced into a small RNA cassette in a chosen viral vector, the use of recombinant viruses to deliver antisense sequences has proven to be an advantageous strategy on many levels.

- Administration and biodistribution

The first advantage of using viral vectors relates to the mode of administration and the biodistribution of antisense sequences. In order to target all the muscular tissues in the body, including the heart, systemic administration of the sequences is necessary. However, although AONs have proven very efficient when administered locally in isolated muscles, their widespread distribution after systemic injection in blood circulation is still problematic. Conventional AONs administered intravenously, PMO as well as 2'-O-Me AONs, present a poor cellular uptake in some muscles particularly in the cardiac muscle, leading to an absence of cardiac dystrophin restoration [425]. Conjugating PMO with cell-penetrating peptides to improve delivery of AONs in the heart has enabled good levels of exon skipping in cardiac muscle of animal models [426], but these peptide-conjugates can induce a strong dose-dependent toxicity. Therefore, the determination of a dose that would be efficient for bodywide dystrophin restoration whilst remaining safe for the patient still constitute a challenge [427].

In contrast, the delivery of antisense-carrying U7 cassettes by AAV vectors offers the ability to transduce all muscle fibres efficiently, including the heart, when a specific AAV serotype is used. AAV serotypes 6, 8 and 9 display strong tropisms for muscle tissues and have demonstrated high gene transfer efficiencies in cardiac muscles following systemic administration in animal models [318, 358-359, 422, 428]. It is therefore possible to greatly improve muscle and heart delivery of U7 snRNA cassettes carrying antisense sequences using such AAV serotypes. The potential for an immune reaction to the viral capsid in the host when using AAV vectors is however non negligible and would often result to a diminution of transduction efficiencies as well as constitute an obstacle to readministration of the vector. Yet, AAV-mediated gene therapy in the presence of mild immunosuppressive treatment have shown some success in recent clinical trials [429] and new strategies aiming at avoiding the immune system are starting to emerge. For example the recent creation of a chimeric AAV capsid variant, termed AAV2.5, enabled efficient muscle transduction without eliciting any immune response even in the presence of preexisting antibodies [430]. The delivery of antisense sequences by viral vectors therefore seems to be the favoured method for systemic administration applications.

- Stability, long-term expression and safety

The vectorisation of antisense sequences in U7 snRNA cassettes delivered by viral vectors also enables to stabilise and protect the antisense sequence against intracellular degradation by incorporating them into stable snRNPs. Furthermore, the use of U7 as a carrier of antisense sequences promote their accumulation into the nucleus, where the splicing reaction takes place, and the SmOPT modifications present on U7 shuttles enable to redirect the modified U7 snRNP to the spliceosome, thereby facilitating its action on the splicing of the target pre-mRNA.

Yet, one of the most interesting point associated with the use of viral vectors for antisense delivery is that, compared to AON-based therapies that necessitate repeated administrations of high doses of oligonucleotides to obtain a prolonged effect, the use of U7 snRNA as an expression vector implies that the antisense sequence is continuously expressed from the strong U7 promoter, thereby providing a stable and long-term production of antisense sequences in transduced tissues and avoiding, or strongly reducing, the need for readministration. Although it is possible that a readministration of U7-carrying AAV vectors might become necessary over time in treated DMD patients, it has been shown that a single administration of AAV9 containing U7 snRNA cassettes for dystrophin exon 23 skipping was able to rescue the phenotype of a severe mouse model of DMD and strongly increased its lifespan, demonstrating the long-lasting effect of the treatment [422].

The development of gene expression profiling methods to evaluate the efficiency and safety of various antisense-based therapies in the *mdx* mouse model enabled to highlight the strong potential of AAV-mediated treatment. Following intramuscular AAV delivery of modified U7 cassettes specific for dystrophin pre-mRNA, the DMD signature genes that correlate with the severity of the disease displayed a highly significant shift towards wild-type profiles and a strong reduction in genes implicated in inflammation and fibrosis, characteristic of *mdx* muscle fibres [431]. Additionally, no disruption of other unrelated genes was associated with the overexpression of the modified U7 snRNAs, except a slight increase in expression of some immunological markers. Such remarkable results were on the other hand not obtained with synthetic AONs complexed with various carriers, which displayed only mild exon skipping efficiencies, no significant change of expression profiles towards the wild-type situation and in some cases an aggravation of the inflammation characteristic of muscular dystrophies.

- Potential for multiple exon skipping

Finally, owing to the small size of U7 snRNA cassettes, the use of viral vectors also provides the possibility to express several modified U7 snRNAs containing different antisense sequences in the same vector, thereby allowing the simultaneous delivery of various cassettes into the same cell [269]. This strategy is very attractive for the prospect of multiple exons skipping, such as the skipping of exons 45 to 55 of human dystrophin pre-mRNA that could constitute a therapy for over 60 % of DMD patients. On the other hand, multiexon skipping is still considered challenging using AONs as it is difficult to control the simultaneous introduction of many different oligonucleotides in the same nucleus, although very encouraging results are starting to emerge [432]. Furthermore, as several antisense sequences would then be contained into a single viral vector, the development of this construct as a drug would be greatly simplified compared to the need to validate and obtain an authorisation for each antisense sequence designed for one mutation for personalised treatment.

Altogether, this points at a strong advantage of the AAV-mediated delivery of antisense sequences over the administration of synthetic AONs for the correction of dystrophin pre-mRNA by exon skipping. Providing the immune reaction towards viral capsids can be controlled and diminished, the vectorisation of antisense sequence via AAV vectors encoding modified U7 snRNA cassettes therefore seems a very promising strategy for the treatment of DMD by exon skipping, allowing a long-term, more targeted and stronger effect than the use of AONs. As the immunological reactions against AAV viruses appear to be dependent on the dose administered, the importance of identifying strategies that would enable the reduction of the dose of vector needed is evident, and this is the goal we tried to achieve in this study by increasing the amount of antisense sequence that can be expressed from a viral vector.

III. 3. Application to DMD - Low efficiency of the strategies versus the requirement for strong therapies

The clinical application of AAV-mediated dystrophin exon skipping for the treatment of DMD patients is proving more and more realistic but would still require the administration of high doses of viral vectors in order to obtain satisfying therapeutic effects, which is currently challenging in terms of safety and vector production capacity. In the severe mouse model of DMD, the dKO mouse, it has been demonstrated that a minimal dose of 5 x 10^{12} vg needed to be injected in order to obtain an improvement of phenotype and lifespan, and that doses below this threshold were not sufficient to restore therapeutic levels of dystrophin expression when administered systemically [422].

Based on the exon skipping efficiencies obtained, we first concluded that both snoRNA-based strategies developed here for dystrophin splicing modulation did not show much potential for a prospective application to DMD patients. However, the results obtained from the various optimisations of U7 snRNA cassettes were more encouraging. We observed that the use of self-complementary AAV vectors to express U7 snRNA cassettes specific for dystrophin exon skipping induce a faster and better transduction efficiency and enable to increase the proportion of exon 23 skipped dystrophin mRNA by more than three folds compared to the same dose of conventional AAV vector. Additionally, we showed that using the heart and muscle-specific MHCK enhancer to promote the transcription of U7Dtex23 snRNA cassettes in muscular cells induced an increase in dystrophin exon 23 skipping both *in vitro* and *in vivo* and resulted in a three-to-five fold increase in Dystrophin expression in transduced muscles of *mdx* mice. Both strategies independently demonstrated a potential for decreasing the dose of vector necessary to obtain significant levels of dystrophin exon skipping, as exon skipping efficiencies or protein rescue levels obtained by each strategy reached equivalent levels than that achieved at a five-to-ten fold higher dose of non-optimised vector.

However, our data also pointed at a possible saturation of the U7 snRNA 3' end processing machinery when the primary U7 transcript is expressed in too many copies in the cells at high doses of vectors, which suggest that the potential for maximising the amounts of mature U7 transcripts that can be produced in muscular cells is limited by the cellular machinery capacity. Accordingly, further improvement of U7Dtex23

cassettes expression and dystrophin exon 23 skipping was not observed following the combination of both scAAV and MHCK strategies, probably owing to saturating levels of primary U7 transcripts produced.

Nonetheless, the improvements on exon skipping efficiency and Dystrophin protein rescue observed suggest that these optimisation strategies, by allowing a small reduction of the vector dose, could be interesting to pursue towards clinical translation. Interestingly, we could not observe the accumulation of primary unprocessed U7 transcripts in vivo in treated *mdx* muscles even for the highest dose of AAV vector administered, which could suggest that this issue of processing machinery saturation might not apply *in vivo*, at least at the doses used and with these vectors. It is therefore possible that the combination of scAAV vectors and MHCK enhancer for the expression of U7Dtex23 cassettes would display further enhancement of dystrophin exon skipping and protein rescue in the *mdx* mouse model, which would then be very encouraging for the clinical application of AAV-mediated delivery of antisense sequences for DMD treatment.

III. 4. Further optimisations of exon skipping-based therapy for DMD and remaining challenges

Further optimisation strategies could yet be envisaged and used alongside the AAV-based delivery of U7 snRNA cassettes to improve the effect of the treatment for Duchenne muscular dystrophy. For example, it could be interesting to inhibit miR-31, a microRNA that has been shown to reduce Dystrophin protein synthesis, by designing an antagomir that would sequester miR-31 and would then allow the increase of protein rescue levels, as has been demonstrated before [433]. Alternatively, myostatin is a factor that inhibits muscle growth and the inhibition of its expression can be achieved by expressing the myostatin propeptide (MRPO) gene, a natural inhibitor, or by exon-skipping aimed at disrupting the reading frame of the myostatin transcript, thereby promoting an increase in muscle mass [434-435]. Finally, the recent discovery that miR-29 expression is disrupted in DMD and that restoring its expression in pathologic muscles improves the dystrophic phenotype could represent a new strategy for DMD treatment [436]. These various approaches could then be coupled with the U7-mediated exon skipping of dystrophin pre-mRNA by introducing either a miR-31 inhibitor, the

miR-29 microRNA, the 288-bp MRPO gene or a cassette expressing an antisense sequence for myostatin exon skipping in a multicistronic vector along with the U7 cassette for dystrophin splicing modulation. Recent studies focus on combining strategies to improve muscle mass and function together with the selective skipping of dystrophin pre-mRNA and the reported results have so far been very encouraging [433, 437-438].

The path towards clinical application of gene therapy for DMD seems now clearer and many hopes lie on antisense-based strategies for dystrophin pre-mRNA splicing modulation. Yet, it is important to remember that exon skipping approaches, even in the eventuality of major successes in the future, would not constitute a cure for the disease as exon-skipping of the dystrophin transcript only induce a conversion of DMD fibres into BMD-like fibres, which is still not equivalent to the normal situation. Furthermore, the muscle tissue lost during the progression of the disease would not be repaired as it is replaced by fibrotic and adipose cells that do not express dystrophin. Finally, exon skipping approaches is only applicable for patients that present frameshifting deletions or non-sense mutations and the identification of other strategies for the treatment of patients that cannot benefit from exon skipping is therefore critical.

IV. Conclusion

The work described in this thesis focused on the development of improved vector that enable to maximise the amount of antisense sequences delivered to a target tissue for the therapeutic modulation of pre-mRNA splicing. We describe various optimisations strategies designed for the antisense-mediated exon skipping of dystrophin pre-mRNA in the context of Duchenne muscular dystrophy and show that some modest improvement can be obtained on exon skipping efficiency and restoration of protein expression. This work highlights the fact that exon skipping approaches for pre-mRNA splicing modulation can still benefit from further improvements [166].

It is therefore clear that the use of RNA as a therapeutic molecule for intervention at the transcript level is an appealing and very promising strategy for the treatment of many RNA-based disorders that are otherwise virtually impossible to correct. The highly intricate network of regulatory non-coding RNAs expressed intracellularly is only just beginning to be discovered and understood and, although this adds a high level of complexity for the fine control of RNA molecules to modulate transcripts, it also provides a virtually endless array of possible tools that could be derived to develop RNA-based therapeutics.

MATERIALS AND METHODS

I. Polymerase Chain Reaction (PCR) and fusion PCR

I. 1. PCR

For all the PCR reactions the PCR Master Mix from Promega was used following the manufacturer's instructions. One reaction mix of 50 µl is composed of 25 µl of Master Mix 2X (containing dNTPs, PCR buffer and Taq polymerase), 1 µl each of 20 µM Forward and Reverse primers, and 50 ng of template DNA. The PCR reactions were realised under the cycle conditions: a) 94°C for 5 minutes, b) 94°C for 30 seconds, c) TM (melting temperature, adapted for each PCR reaction) for 1 minute, d) 72°C for 1 minute/kb, repeating from step b) for 29 additional cycles.

I. 2. Fusion PCR

To realise a fusion PCR, which enables to introduce or replace an entire stretch of nucleotides inside a specific DNA sequence, two pairs of primers are used **(Figure 76a)**:

- an internal pair (Fint and Rint) that enables to insert the chosen modification via modified 5'-ends. Each internal primer is composed of a 3' part complementary to the template sequence adjacent to the modification, and of a 5' unpaired tail containing a fragment of the sequence to be inserted. The 5' ends of both internal primers are designed to contain overlapping sequences.

- an external pair of primers (Fext and Rext) to amplify the template DNA sequence surrounding the region to be modified.

Two PCR reactions are realised in parallel from the template DNA (**Figure 76b**): the first one (PCRI) with the primers Fext and Rint, and the second one (PCRII) using Fint and Rext. As the modified 5' tails of the internal primers contain complementary sequences, the products of PCRI and II therefore present overlapping extremities. A final amplification with the external primers using PCRI and II products as templates therefore enables to fuse both products into the final construct containing the modification (PCRIII) **(Figure 76c)**.

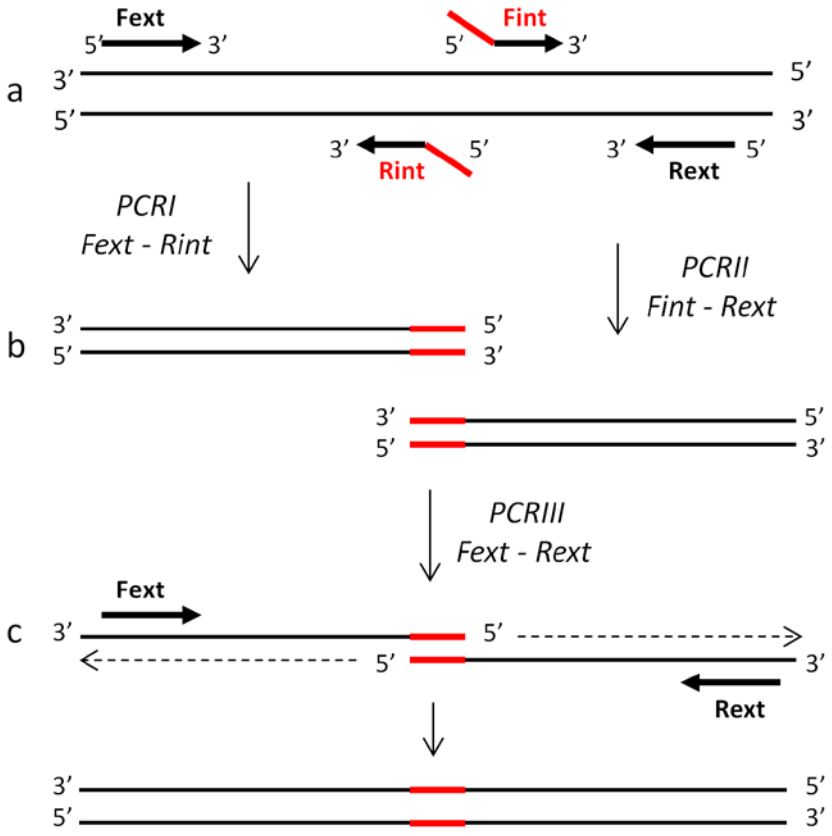

Figure 55: mechanism of fusion PCR

The final products of amplification are purified on a 1% agarose gel and extracted using the QIAGEN gel extraction kit. Restriction sites that produce compatible cohesive ends are added at each extremity of the products to allow their cloning into plasmid vectors.

The sets of primers used for the amplification of each gene are listed in **Table 2.**

PCR PRODUCTS	*PRIMERS FORWARD*	*PRIMERS REVERSE*	*SOURCE*
U7Dtex23 snRNA for mouse dystrophin exon 23 skipping	5’TTTTGCTAGCTAACAACATA GGAGCTGTG 3’	5’TTTTACTAGTCACATACGCGTT TCCTAGG 3’	AAV-pSMD2-U7Dtex23 [266]
MHCK synthetic enhancer Fusion of α-MHC and MCK enhancers	**Fext** 5'TTTACTAGTCCCTTCAGATTA AAAATAACTGA 3'	**Rext** 5'CCCTCTAGAGATCCACCAGGG ACAGGGTTAT 3'	Mouse genomic DNA from C2C12 cells
	Fint 5'CCCTGCTGTCCCACTACGGG TCTAGGCTGC 3'	**Rint** 5'CCCGTAGTGGGACAGCAGGGC CCCAAGGTT 3'	
U7M23D snRNA (fusion PCR)	**Fext** 5'TTTTGCTAGCTAACAACATA GGAGCTGTGATTGGC 3'	**Rext** 5'ACCUCGGCUUACCUGAAAAAT TTTTGGAGCAGGTT 3'	AAV-pSMD2-U7Dtex23 [266]
	Fint 5'ACCUCGGCUUACCUGAAAA ATTTTTGGAGCAGGTT 3'	**Rint** 5'CGAGGTTTGGCCGCGGAAGTG CG 3'	
MBIIDtex23 (fusion PCR)	**Fext** 5’TTTTCTCGAGCTGCTTAGGT AAGACATTCTTCC 3’	**Rext** 5’TCTATTTAGGTAAGCCGAGGT TTGGCCGAATTTTATGTCATCAC 3’	pCMV-MBII-52 [203]
	Fint 5’CCTAAATAGAAGTTCATTTA CACTAACCTGAGGCCCAACCA GGA 3’	**Rint** 5’TTTTTCTAGATGACCACTTACC TGATTATCTGTGTTTCT 3’	
U24 snoRNA (RPL7A gene)	5’TTAGCTCTCGAGTTCTGCCC AGCCGAAAGGAAAGAAG 3’	5’TGAACTTCTAGAAGAACCTCA CCTGTTTGGCGGTCCA 3’	Human genomic DNA from 293T cells
human β-globin	5’TTTTGCTAGCACATTTGCTT CTGACACAAC 3’	5’TTTTGCTAGCGTGATACTTGTG GGCCA 3’	Human genomic DNA from 293T
X12-Luciferase	5’TTTGCTAGCGAAGACGCCAA AAACATAA 3’	5’TTTTACTAGTTTACAATTTGGA CTTTCCGCCC3’	pET16b-lucX12 [399] Gift from M.Pule

Table 2: list of primers used for the PCR reactions

II. Mutagenesis

The various site-specific mutagenesis reactions, used to modify few nucleotides inside a particular sequence, were realized using the QuikChange II Site-Directed Mutagenesis Kit from Stratagene following the manufacturer's instructions.

The sets of primers used for each site-specific modification are listed in **Table 3**.

MHCK synthetic enhancer: Deletion 63 bp in the MCK part of MHCK [21]	**Fwd:** 5'CCCCAACACCTGCTGCCTGCTAAAAATAACCCTGTCCCTGGTGG 3'
	Rev: 5'CCACCAGGGACAGGGTTATTTTTAGCAGGCAGCAGGTGTTGGGG 3'
U7Dtex23-Sm: Modification SmOPT site of the U7Dtex23 into canonical Sm site	**Fwd:** 5'GCTTACCTAAATAGAAGTTCATTTACACTAACAATTTGTCTAGCAGGTTTTCTGACTTCG 3'
	Rev: 5'CGAAGTCAGAAAACCTGCTAGACAAATTGTTAGTGTAAATGAACTTCTATTTAGGTAAGC 3'
U7wt-SmOPT: Modification Dtex23 antisense of the U7Dtex23 into original U7snRNA antisense	**Fwd:** 5'CAGACGCACTTCCGCAAGTGTTACAGCTCTTTTAGAATTTTTGGAGCAGGTTT 3'
	Rev: 5'AAACCTGCTCCAAAAATTCTAAAAGAGCTGTAACACTTGCGGAAGTGCGTCTG 3'
U7wt-Sm: Modification SmOPT site of the U7wt-SmOPT into canonical Sm site	**Fwd:** 5'ACTTCCGCAAGTGTTACAGCTCTTTTAGAATTTGTCTAGCAGGTTTTCTGACTT 3'
	Rev: 5'AAGTCAGAAAACCTGCTAGACAAATTCTAAAAGAGCTGTAACACTTGCGGAAG 3'
pCMV-βglobin-X12Luc: Deletion 2bp in ß-globin exon3	**Fwd:** 5' GCTCCTGGGCAACGCTGGTCTGTGTG3'
	Rev: 5' CACACAGACCAGCGTTGCCCAGGAGC3'
pCMV-U24metβglob: Modification **antisense 1** in pCMV-U24snoRNA	**Fwd:** 5'TGCAGATGATGTAAAAATAGACCAATAGTCTGAGAGATGGTGATG 3'
	Rev: 5'CATCACCATCTCTCAGACTATTGGTCTATTTTTACATCATCTGCA 3'
pCMV-U24metβglob: Modification **antisense 2** in pCMV-U24snoRNA	**Fwd:** 5'TGGTGATGACATTTTAAAGAGAGAGTCAGGCTGATGCACCAACAC 3'
	Rev: 5'GTGTTGGTGCATCAGCCTGACTCTCTCTTTAAAATGTCATCACCA 3'

Table 3: list of primers used for the PCR-directed mutagenesis reactions

III. Molecular cloning and plasmid preparation

The various PCR fragments produced were digested by New England Biolabs (NEB) restriction enzymes in order to create cohesive ends. The backbone vectors were digested with matching restriction enzymes. The ligation reactions were done using the Quick ligation kit from NEB following the manufacturer's instructions and were transformed in NEB 5α competent cells. The bacteria were plated on agar plates containing ampicillin and incubated at 37°C overnight. The colonies obtained were cultured in LB medium supplemented with ampicillin and incubated in a shaker at 37°C for 8h. The plasmid DNA obtained was purified using the Qiagen Miniprep kit. The constructs were verified by specific enzyme digestion and sequencing and, when correct, were amplified and purified using the Qiagen MegaPrep and PureLink HiPure Plasmid DNA GigaPrep (Invitrogen) kits.
The different vectors constructed are listed thereafter.

- AAV-sc-U7Dtex23

The self-complementary AAV-U7Dtex23 vector was constructed by subcloning the U7Dtex23 PCR product amplified from the AAV-U7DTex23 vector (previously described as AAV-U7-SD23/BP22 [266]) into the self-complementary AAV backbone scAAV-LP1-hFIX [378], from which the LP1-hFIX sequence was removed (**Figure 77**). The polyA fragment was conserved in the plasmid to facilitate the quantification by RQ-PCR of the copy number of viral genome present in the self-complementary AAV vectors produced.

The self-complementary AAV backbone contains a mutation in one of the AAV inverted terminal repeats (ITRs): one of the ITRs is deleted of the D-sequence (the packaging signal) and the adjacent terminal resolution site (trs), allowing the AAV vector to package an inverted repeat genome that can fold directly into double stranded DNA.

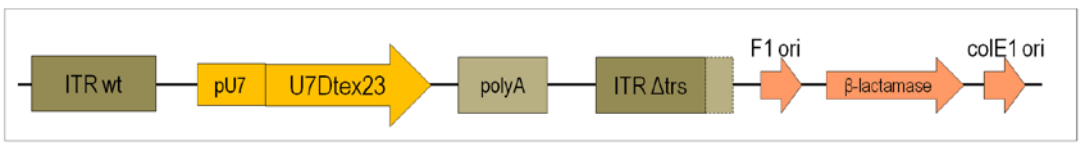

Figure 56: plasmid AAV-sc-U7Dtex23

- AAV-sc-U7M23D

The plasmid AAV-sc-U7M23D was constructed by subcloning the U7M23D cassette into the self-complementary AAV backbone scAAV-LP1-hFIX [378] (**Figure 78**). The U7M23D was produced by fusion PCR to replace the Dtex23 double target antisense sequence in the U7Dtex23 snRNA by the M23D(+07-18) antisense targeting the donor splice site of dystrophin exon 23 [379].

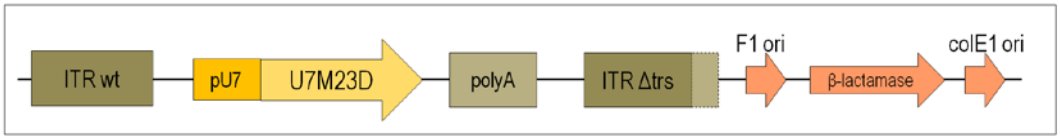

Figure 57: plasmid AAV-sc-U7M23D

- AAVsc-U7DTex23-Sm, AAVsc-U7wt-SmOPT and AAVsc-U7wt-Sm

The SmOPT site of the U7DTex23 sequence (AAUUUUUGGAG) was changed into the native U7 Sm sequence (AAUUUGUCUAG) in the AAV-sc-U7DTex23 vector by site-specific mutagenesis to obtain the AAV-sc-U7Dtex23-Sm vector.

The AAV-sc-U7wt-SmOPT was obtained by modifying the Dtex23 antisense sequence of the AAV-sc-U7Dtex23 plasmid into the original U7snRNA antisense sequence (AAGTGTTACAGCTCTTTTAG).

Finally the AAV-sc-U7wt-Sm vector was obtained by changing the SmOPT sequence of the AAV-sc-U7wt-SmOPT into the canonical Sm sequence by site-specific mutagenesis (**Figure 79**).

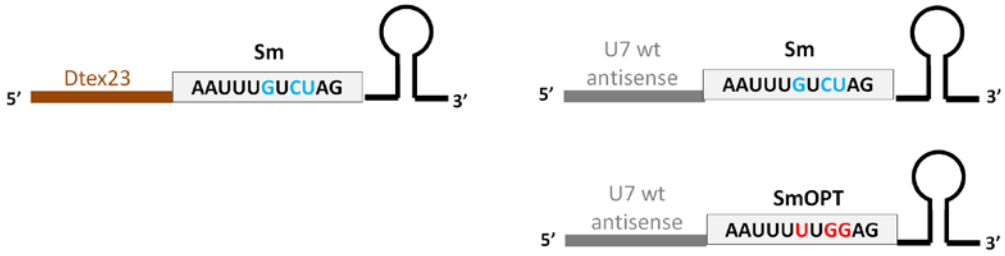

Figure 58: representation of AAV-sc-U7DTex23-Sm, AAV-sc-U7wt-SmOPT and AAV-sc-U7wt-Sm constructs

- AAV-MHCK-U7Dtex23 and AAV-sc-MHCK-U7Dtex23

The AAV-MHCK-U7Dtex23 vector was obtained by introducing the MHCK PCR product upstream of the U7snRNA cassette in the AAV-U7DTex23 vector [266]. The MHCK enhancer was obtained by fusing the 188 base pairs (bp) long murine α- MHC enhancer, and the 207 bp murine MCK enhancer. To faithfully reproduce the MHCK cassette as was described by Salva et al. [383], a 63 base pairs deletion in the MCK part of the synthetic enhancer was realized by mutagenesis (**Figure 80**).

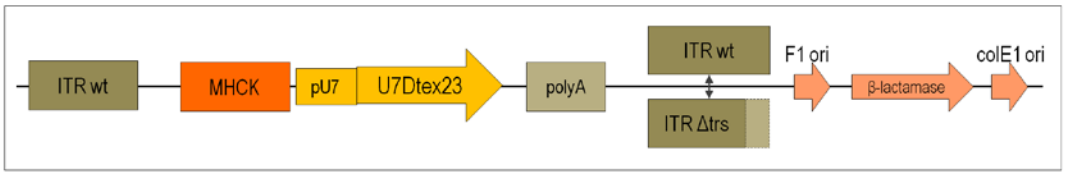

Figure 59: AAV-MHCK-U7Dtex23 and AAV-sc-MHCK-U7Dtex23

When verified by sequencing, the final AAV-MHCK-U7DTex23 construct showed a deletion of three cytosines between the left and right E-boxes of the MCK enhancer compared to the published NCBI sequence (GenBank: AF188002.1). However, this deletion was also observed in C2C12 genomic DNA and was considered a polymorphism.

The vector AAV-sc-MHCK-U7 was constructed by subcloning the MHCK-U7Dtex23 fragment from the aforementioned vector into the self-complementary AAV vector scAAV-LP1-hFIX.

- AAV-sc-PGK-GFP and AAV-sc-MHCK-PGK-GFP vectors.

The AAV-sc-PGK-GFP plasmid was constructed by subcloning the PGK-GFP digested fragment from the pRRLSIN-cPPT-PGK-GFP-WPRE lentiviral vector [439] into the scAAV backbone vector. The MHCK enhancer isolated by restriction digest from the AAV-MHCK-U7Dtex23 plasmid was then inserted upstream of the PGK promoter to obtain the AAV-sc-MHCK-PGK-GFP vector (**Figure 81**).

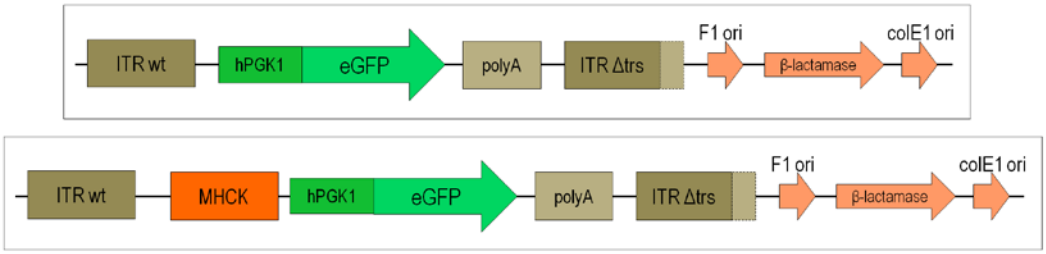

Figure 60: plasmids AAV-sc-PGK-GFP and AAV-sc-MHCK-PGK-GFP

- AAVsc-CMV-MBIIDtex23 and pRRL-CMV-MBIIDtex23-PGK-GFP

The MBII52-Dtex23 cassette was produced by fusion PCR to replace the natural antisense sequence present in the snoRNA MBII-52 by the Dtex23 antisense for dystrophin exon 23 skipping, using the plasmid pCMV-MBII-52 as a template [203]. The plasmid pCMV-MBII52-Dtex23 was then constructed by subcloning the MBII52-Dtex23 PCR product into the pcDNA3.1(+) (Invitrogen) backbone under the control of the CMV promoter. Finally, the AAV-sc-CMV-MBII52-Dtex23 vector was obtained by subcloning the CMV-MBII52-Dtex23 digested fragment into the self-complementary AAV backbone [378] (**Figure 82**).

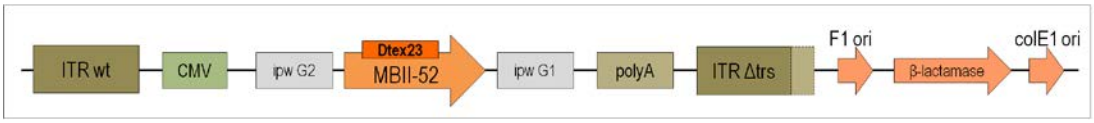

Figure 61: plasmid AAV-sc-CMV- MBII52-Dtex23

The lentiviral vector pRRL-CMV-MBII52-Dtex23-PGK-GFP was obtained by inserting the CMV-MBIIDtex23-polyA restriction fragment from the pCMV-MBII52-Dtex23 vector into the pRRLSIN-cPPT-PGK-GFP-WPRE vector [439] (**Figure 83**).

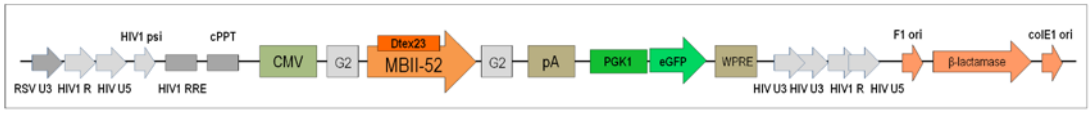

Figure 62: plasmid pRRL-CMV-MBIIDtex23-PGK-GFP

- pCMV-βglobin-Luciferase

The plasmid pCMV-βglobin was constructed by subcloning the β-globin PCR fragment into the pcDNA3.1 (+) vector (Invitrogen) under the control of the CMV promoter. The pCMV-βglobin-Luciferase was obtained by inserting the X12-Luciferase PCR fragment downstream of the βglobin gene in the pCMV-βglobin plasmid. The X12-Luciferase is a highly thermostable and pH-tolerant mutant of the wild-type firefly luciferase [399]. The X12-luciferase product, without its ATG initiation codon, was inserted in the same open reading frame as the βglobin.

The pCMV-βglobin-Luciferase construct was further modified by deleting 2 base pairs in exon 3 of the βglobin sequence to generate an out of frame construct. The 2 base pair deletion (GT) was realised by site-specific mutagenesis (**Figure 84**).

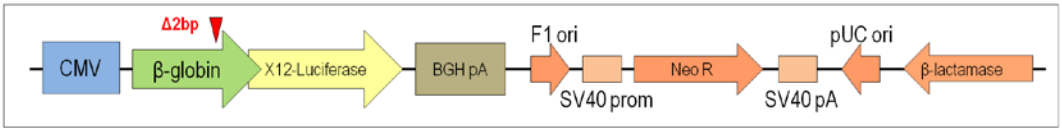

Figure 63: pCMV-βglobin-Luciferase plasmid

- pCMV-U24metβglob

The plasmid pCMV-U24 was constructed by subcloning the U24 snoRNA PCR fragment, surrounded by the host exons 2 and 3 of the RPL7A gene, into the pcDNA3.1(+) backbone under the control of the CMV promoter. The pCMV-U24metβglob was obtained by replacing the U24 snoRNA natural antisense sequences with sequences complementary to the branch point of βglobin intron 1 in the pCMV-U24 vector (**Figure 85**). Two successive site-specific mutagenesis reactions were necessary to modify both antisense sequences.

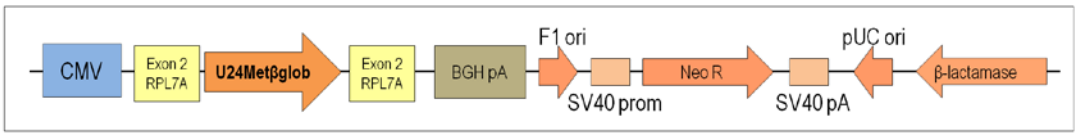

Figure 64: pCMV-U24metβglob plasmid

IV. AAV virus preparations

To produce recombinant AAV (rAAV) vectors, cells are transiently co-transfected with three plasmids: the plasmid containing the trangene flanked by the ITRs of AAV2, a plasmid helper that provides the excised AAV-specific genes rep and cap in *trans*, and a plasmid containing the adenovirus helper functions necessary for virion formation (**Figure 86**).

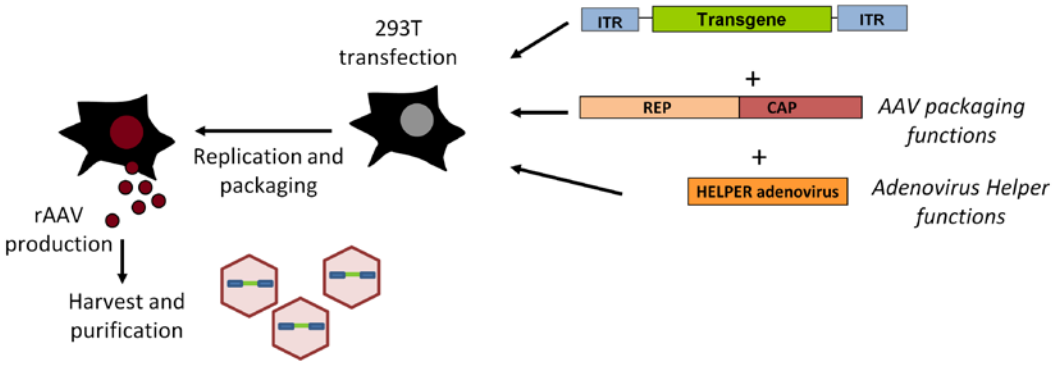

Figure 65: production of recombinant AAV vectors

AAV vectors were produced by calcium phosphate triple transfection of HEK293T cells. For each AAV production, 40 plates of 70-90% confluent 293T cells were transfected with 400 µg of transgene plasmid, 400 µg of capsid plasmid AAV2/8 (LTAAVhelp2-8 [440]) or AAV2/5 (Napoli 2/5 [441]), and 1200 µg of the plasmid for adenovirus helper function (pHGTI-Adeno1 [442]). The plasmids were added to 5 mL of 2.5M CaCl2 in a total volume of 50mL. Then, 50 mL of 2x HEPES buffered solution (HBS: 273.9mM NaCl, 9.9mM KCl, 1.5mM NaH2PO4H2O, 11.1mM Dextrose, 41.9mM HEPES – pH 7.05) were added drop by drop in the plasmid mix whilst vortexing, and 100 mL of D10 medium was finally added to stop the reaction. The total volume was divided between 40 plates, which were then incubated for 72 hours at 37°C.

Cells were harvested by scraping and centrifugating the media and cells at 1800rpm for 10 minutes at 4°C. The supernatant was removed and the pellets resuspended in 40 ml of TD buffer (140.4mM NaCl, 4.9mM KCl, 0.7mM K2HPO4, 3.4mM MgCl2, 24.7mM Tris - pH 7.5). A second centrifugation, aspiration of the supernatant, and resuspension in 40ml of TD buffer was performed. The harvested cells underwent five freeze/thaw

cycles to lyse the cells, where each cycle consists of 30 minutes at -70°C followed by 30 minutes at 37°C. After the last thaw 2000U of benzonase were added, the suspension was incubated for 30 minutes at 37°C and centrifuged for 20 minutes at 3,000g. The resulting supernatant was filtered using a 0.45 μm cellulose acetate filter and the AAV2/5 or AAV2/8 virus preparations were purified on an AVB-sepharose column.

V. rAAV virus purification with AVB Sepharose Column

The prepared virus sample was purified by affinity chromatography using the ÄKTAexplorer™ chromatography system (GE Healthcare) and a prepacked AVB sepharose affinity column (HiTrap™ AVB Sepharose High Performance, GE Healthcare – 5ml). After equilibration of the column with 1 column volume (CV - 1 CV = 5.228ml) with filtered PBS at pH 7, the virus sample is loaded into the system at a rate of 2ml/min. The unbound sample is then washed with 4 CV of filtered PBS. The sample is eluted with 7 CV of 50mM Glycine at pH2.7 and collected into fractionating tubes containing 30μl each of Tris 0.5M pH 8.8 to neutralise the glycine elution buffer. The chromatogram shows the graphs of the absorbance at 260nm and at 280nm, and a peak on these graphs represents the fraction containing the eluted viral sample. The corresponding fractions are pooled and injected into a dialysis cassette (Slide-A-Lyser Dialysis cassette 10000MWCO – Thermo scientific) and the sample is allowed to dialyse in 2L of filtered PBS at 4°C, stirring for 12-16 hours. Finally the dialysed virus preparation is sterilised by filtration through a 0.2μm cellulose acetate filter.

VI. Lentiviral vectors preparation

The production of SIN lentiviral vectors involves transient transfection of a cell line with four different plasmids: a plasmid containing the gag and pol genes, one with the envelope gene - most commonly the VSV-G gene, a third one containing the rev gene, and last the SIN vector containing the transgene of interest (**Figure 87**).

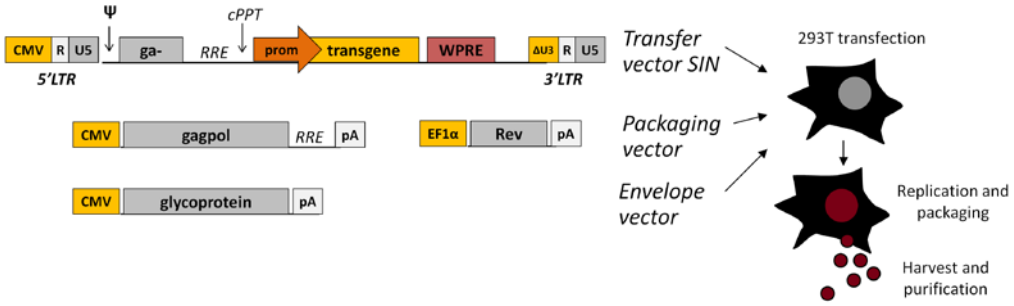

Figure 66: lentiviral vectors production

Lentiviral virions were produced by calcium phosphate transfection of HEK293T cells. The day before transfection, 6.10^6 293T cells were seeded per 150 mm culture dishes. The following day, 7.9 µg of the envelope plasmid pMD.G (VSV-G, [443]), 1 µg each of the packaging plasmids pHDM15 gpm2 (codon-optimized HIV-1 Gag-Pol) and pRC/CMV-REV1B (REV) (both gifts from Dr Jeng-Shin Lee, Harvard Medical School), and 22.5 µg of the plasmid pRRLSIN-transgene were cotransfected by the calcium phosphate method.. Viral supernatants were harvested 20 hours, 28 hours, 36 hours and 44 hours post-transfection, pooled and centrifuged 5 minutes at 1500 rpm, filtered through a 0.45 µm filter, and ultracentrifuged 2 hours at 19500 rpm at 12°C. Pellets were resuspended in PBS 1%BSA and stored at −80 °C.

VII. Cell culture and transduction

All cells were cultivated at 37°C and 5% CO2. HEK 293T cells were grown in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% fetal calf serum (Invitrogen). C2C12 mouse myoblast cells were grown in proliferation medium containing DMEM supplemented with 20% fetal calf serum. Differentiation of C2C12 was induced in DMEM supplemented with 2% Horse serum (Sigma).

For AAV vector transduction, 100 000 C2C12 cells were seeded in each well of a 12-wells plate. The following day proliferation medium was changed into differentiation medium and the cells were left to differentiate into myotubes for a week. The cells were then transduced with the AAV preparation at the chosen multiplicity of infection (MOI) in a total volume of 300 µl of medium without serum. Each condition is realized in triplicate. After 6 hours of incubation, 1 ml of differentiation medium was added per well. Cells were incubated for seven additional days in differentiation medium before harvesting for analysis.

For lentiviral vector transduction, 30000 C2C12 cells were seeded in each well of a 12-wells plate. 24 hours later, the lentiviral vector preparation was administered to the cells in a total volume of 250 µL of proliferation medium at the chosen multiplicity of infection (MOI). After 6 hours incubation, 1 ml of D20 medium was added per well. 72 hours post-transduction, the medium was replaced with differentiation medium and cells were incubated for 72 additional hours. Cells were then harvested for analysis.

VIII. Cell transfections

293T cells at an average of 80% of confluence are transfected with X µg of transgene plasmid (scAAV-U7Dtex23, scAAV-U7Dtex23-Sm, scAAV-U7wt-Sm, scAAV-U7wt-SmOPT or scAAV-MBII52 Dtex23) in 100mm plates using Lipofectamine 2000 from Invitrogen following the manufacturer's protocol. The transfected cells are incubated for 48h at 37°C before RNA extraction and Northern blot analysis. The reporter plasmid pcDNA-βglobin-X12Luc and the pCMV-U24metβglob constructs were co-transfected into 80% confluent 293T cells in a 6-well plate using Lipofectamine 2000, in order to obtain a 1:5 molar ratio. Each transfection condition was realized in triplicate. The cells were left to incubate at 37ºC for 48h before harvesting.

IX. Western blot

IX. 1. GFP detection

C_2C_{12} protein extracts were obtained by incubation 30 minutes at 4°C of cells in lysis buffer (1× complete protease inhibitor cocktail (Roche, Diagnostics), 0.5% NP-40, 20 mM Tris pH 7.5, 10% glycerol, 100 mM $(NH_4)_2SO_4$). Equal amounts of protein (determined by the Bradford kit (Bio-Rad)) were mixed with 3X loading buffer (30% glycerol, 150 mM Tris pH 6.8 , 6% SDS, 0.03% bromophenol blue, 15% β-mercaptoethanol) denaturated 5 minutes at 95°C, separated on a 12% polyacrylamide gel 1 hour at 110 V and transferred to a PVDF membrane (Amersham) 90 minutes at 100 V. The membrane was blocked with PBS 5% milk 0,1% Tween 20 and probed with the rabbit polyclonal anti-GFP antibody (sc-8334; Santa Cruz Biotechnology) used at a 1:500 dilution and the mouse anti-actin (A-5441; Sigma) used at a 1:15000 dilution, and then with secondary goat anti-mouse antibody and goat anti-rabbit antibody, both conjugated with peroxidase (Sigma) used at a 1:2000 dilution. Signals were detected with the ECL kit (Amersham).

IX. 2. Dystrophin detection

Total protein was extracted from muscle samples with Newcastle buffer (3.8% SDS, 75 mM Tris-HCl pH 6.7, 4 M urea, 10% β-mercaptoethanol, 10% glycerol, 0.001% bromophenol blue) and quantified using the bicinchoninic acid (BCA) protein assay kit, according to the manufacturer's instructions (Perbio Science, UK). Samples were denatured at 95°C for 5 minutes before 100 μg of protein was loaded in a 5% polyacrylamide gel with a 4% stacking gel. Gels were electrophoresed for 4-5 hours at 100 V and blotted to a PVDF membrane overnight at 50 V. Blots were blocked for 1 hour with 10% non-fat milk in PBS–Tween (PBST) buffer. Dystrophin and α-Actinin proteins were detected by probing the membrane with 1:100 dilution of NCL-DYS1 primary antibody (monoclonal antibody to dystrophin R8 repeat; NovoCastra) and 1:200 dilution of α-actinin primary antibody (Santa Cruz Biotechnology) respectively. Incubation with mouse horseradish peroxidase-conjugated secondary antibody (1:2000) or goat horseradish peroxidase-conjugated secondary antibody (1:160,000) allowed visualisation using ECL Analysis System (GE Healthcare). Membranes were converted to numerical pictures and band intensities were analysed using ImageJ 1.33a software.

X. RNA isolation and Nested RT-PCR analysis

Total RNAs from C2C12 cells were extracted using the miRvana miRNA Isolation Kit from Ambion according to the manufacturer's instructions. RNAs from mice muscles were isolated using TriZol Reagent according to the manufacturer's instructions (Invitrogen).

The reverse transcription-polymerase chain reaction (RT-PCR) reaction was carried out on 300 ng of total RNA as the starting material using the Access RT-PCR System (Promega), which enables to realise the reverse transcription and the PCR in one single step. The cDNA synthesis was carried out at 45°C for 45 min, directly followed by the primary PCR of 30 cycles of 94°C (30 s), 55°C (1min) and 72°C (2 min). The forward and reverse primers used, amplifying from exons 20 and 26 respectively, were as follow:
Ex20Fext: 5'-CAGAATTCTGCCAATTGCTGAG-3';
Ex26Rext: 5'-TTCTTCAGCTTGTGTCATCC-3'.

Two microlitres of the RT-PCR product was used as template in a 50 µl secondary nested PCR that was carried out for 22 cycles under cycling conditions identical to the primary amplification and using the PCR master mix kit (Promega). The internal primers used were:
Ex20Fint: 5'-CCCAGTCTACCACCCTATCAGAGC-3';
Ex26Rint: 5'-CCTGCCTTTAAGGCTTCCTT-3'.
The products were analysed on 2% agarose gel.

XI. Real-time quantitative (RQ) PCR

XI. 1. Titration of AAV vectors by RQ-PCR

The titers of the AAV virus preparations were assessed by real-time PCR quantifying of vector genome copies using the SYBR Green master mix (Applied Biosystems). A mix for one reaction contained 1 µL each of 2.5 µM Forward and 2.5 µM Reverse primers, 12.5 µL SYBR Green master mix and 0.5 µL H2O. Each AAV sample was diluted 1 in 100, 1 in 1000, or 1 in 10000, and 10 µL of either dilution was added in each well. Each condition was realised in triplicate and the qPCR reaction was realised under the cycle conditions: a) 95°C for 2 minutes, b) 95°C for 15 seconds, c) 60°C for 15 seconds, repeating from step b) for 40 cycles. The primers used for each AAV vector titration are listed in **Table 4**.

AAV VECTORS	*PRIMERS FORWARD*	*PRIMERS REVERSE*
All AAV vectors containing the **Dtex23 sequence**	**F-Dtex23** 5'GGCCAAACCTCGGCTTA 3'	**R-Dtex23** 5'GTTAGTGTAAATGAACTTC 3'
scAAV-U7M23D scAAV-U7wt-SmOPT scAAV-U7wt-Sm	**F-polyA** 5'ATTTTATGTTTCAGGTTCAGGGGGAGGTG 3'	**R-polyA** 5'GCGCAGAGAGGGAGTGGACTAGT 3'
scAAV-PGK-GFP scAAV-MHCK-PGK-GFP	**F-GFP** 5'GACGGCAACATCCTGGGGCACAAG 3'	**R-GFP** 5'CGGCGGCGGTCACGAACTC 3'

Table 4: list of primers used for AAV vectors qPCR titrations

XI. 2. Titration of lentiviral vectors by RQ-PCR

After transduction of 293T cells with different dilutions of the lentiviral preparation, total DNA was extracted from the cells with the Qiagen DNeasy Blood and Tissue Kit. The infectious titers were determined by quantitative PCR using the SYBR Green PCR Master Mix (Applied Biosystems). A mix for one reaction contained 0.5 µL each of 10 µM Forward and 10 µM Reverse primers, 12.5 µL SYBR Green master mix, 1 µL H2O and 10 µL of each DNA sample at 10 ng/µl. Each condition was realised in triplicate and the qPCR reaction was realised under the cycle conditions: a) 95°C for 2 minutes, b) 95°C for 15 seconds, c) 60°C for 1minute, repeating from step b) for 40 cycles. Primers specific for the lentiviral genome (HIV-LTR) allowed the quantification of viral genome

copies integrated in the DNA, and primers for the albumin endogenous gene were used in parallel for normalization. The primers used for lentiviral vector titrations are listed in **Table 5**.

	PRIMERS FORWARD	***PRIMERS REVERSE***
HIV-LTR primers	5'AGCTTGCCTTGAGTGCTTCAA 3'	5'AGGGTCTGAGGGATCTCTAGTTACC 3'
Albumin primers	5'GCTGTCATCTCTTGTGGGCTGT 3'	5'ACTCATGGGAGCTGCTGGTTC 3'

Table 5: primers used for lentiviral vectors titrations

XI. 3. Taqman assay for Dystrophin levels quantification

The Taqman RQ-PCR assay relies on the 5´–3´ exonuclease activity of Taq polymerase to cleave a dual-labeled probe designed to hybridize within the complementary target sequence [444-445]. After the reverse transcription step successful quantification requires the annealing of three oligonucleotides to the DNA. The probe is covalently linked to a fluorophore at the 5' end and to a quencher at the 3'-end, which inhibits the fluorescence emitted by the fluorophore via Fluorescence Resonance Energy Transfer (FRET) as long as the molecules are in close proximity [446]. During the amplification of the target sequence through the use of a specific set of primers surrounding the Taqman probe, the 5'-nuclease activity of the Taq polymerase displaces and hydrolyses the labelled probe, which releases the fluorescent dye from its proximity to the quencher and allows the emission of a fluorescent signal that is directly proportional to the amount of product formed at each PCR cycle [445] (**Figure 88**). The Ct (cycle threshold) value is defined as the number of PCR cycles required for the fluorescent signal to cross the threshold level and is therefore a relative measure of the concentration of target DNA in the PCR reaction. In relative quantification we can then compare the proportion of two target DNAs by subtracting the Ct value of one target gene to another, which is called the delta Ct value (ΔCT). As the PCR reactions are exponential functions, the fold difference in the amounts of two target DNAs is therefore calculated by the formula $2^{\Delta CT}$.

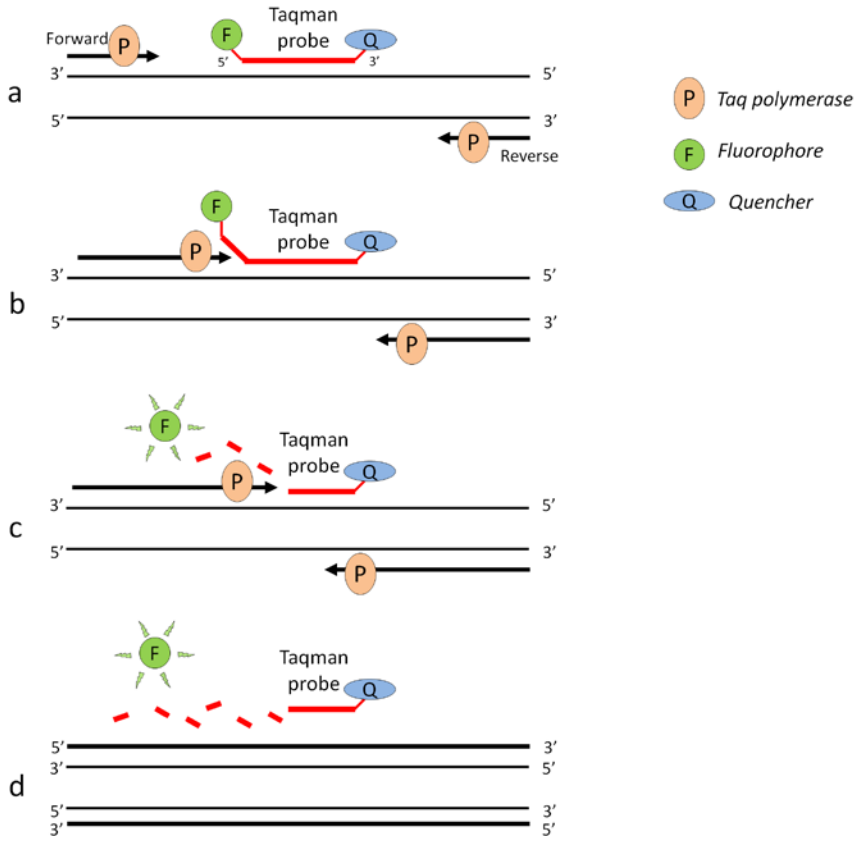

Figure 67: Taqman assay

a) Probe and primers annealing on denaturated DNA. The fluorophore on 5' of the probe is inhibited by the proximity of the quencher on 3'.
b) Polymerisation starts and the Taq polymerase displaces the probe.
c) The polymerase hydrolyses the probe, the fluorescent dye is released from its proximity to the quencher and fluorescence is emitted.
d) The fluorescent signal is proportional to the number of molecules at the end of the cycle

Total RNAs from C2C12 cells were extracted using the miRvana miRNA Isolation Kit from Ambion according to the manufacturer's instructions. For each sample 1 µg of total RNA was used for reverse transcription using the SuperScript II RT system by Invitrogen. Reverse transcription was carried out according to the manufacturer's instructions, using random hexamers as the priming nucleotides. For qPCR analysis, comparative CT quantitative analysis of dystrophin exon 22-23 junction (Taqman Assay Mm01216934_m1, Applied Biosystems) and 18S rRNA (Eukaryotic 18s rRNA Control, Applied Biosystems) was performed on each RNA sample. A mix for one reaction contained 1 µL of 20X TaqMan Gene Expression Assay containing the primers and the Taqman probe, 10 µL of 2X TaqMan Gene Expression Master Mix, 10 ng of the cDNA sample and water up to 20 µL. Reactions were performed in triplicate using the cycle

conditions: 95°C for 10 minutes, and 40 cycles at 95°C for 15 seconds and 60°C for 1 minute.

Data Analysis: For each sample the CTs for the 22-23 assay (dystrophin) are normalised 18S Ct to generate a delta-CT value. This involves the subtraction of the 18S Ct from the experimental Ct. The delta Ct can be used to calculate the relative amounts of dystrophin expressed relative to 18S. All results are expressed as mean values ± SEM.

XI. 4. Taqman assay for Dystrophin exon skipping

Total RNAs from transduced C2C12 cells were extracted using the miRvana miRNA Isolation Kit from Ambion according to the manufacturer's instructions. For each sample 1 μg of total RNA was used for reverse transcription using the SuperScript II RT system according to the manufacturer's instructions and using random hexamers as the priming nucleotides. For qPCR analysis, 10 ng of the resulting cDNA was used per well. Comparative CT quantitative analysis of the exon 22-23 junction (Taqman Assay Mm01216934_m1, Applied Biosystems) and the exon 22-24 junction (Fw: 5'-CTGAATATGAAATAATGGAGGAGAGACTCG-3', Rev: 5'-CTTCAGCCATCCATTTCTGTAAGGT-3', probe: FAM-ATGTGATTCTGTAATTTCC-NFQ) was performed on each RNA sample. Reactions were performed in triplicate with the appropriate non-template controls (NTCs) using the cycle conditions: 95°C for 10 minutes, and 40 cycles at 95°C for 15 seconds and 60°C for 1 minute.

Data Analysis: For each sample the CTs for the 22-24 assay (skipped dystrophin) are normalised to the CT for the 22-23 assay (non-skipped dystrophin) to generate a delta-CT value. This involves the subtraction of the 22-23 CT from the 22-24 Ct. The delta CT can be used to calculate the relative abundance in skipped dystrophin RNA species relative to full length species.

An unpaired Student's t-test was used to assess statistical significance of differences between the samples treated with AAV-U7Dtex23 and the samples treated with AAV2-MHCK7-U7Dtex23. A P-value below 0.05 was considered statistically significant. All results are expressed as mean values ± SEM.

XII. Northern blot assay

Northern blots were realized on 7 µg of total RNA. Samples of total RNA were mixed with an equal volume of loading buffer (8M Urea, 2xTBE, Bromephenol Blue and Xylene Cyanol) and heated at 70°C for 15 minutes. RNAs were loaded on a 15% Novex TBE-Urea precast gel (Invitrogen) along with 1 µl of γ-^{32}P-labeled RNA Decade Markers (Ambion). Gels were electrophoresed for 1 hour at 180 V in 2,5X TBE buffer, and transfered to a Hybond-N+ membrane (Amersham) for 1 hour at constant current (390 mA) in 2X TBE buffer. The membrane was then UV-crosslinked with 100 mJ of energy and prehybridized in 10 ml of Rapid Hyb buffer (Amersham) containing salmon sperm DNA (50 µg/ml) at 42°C. For radiolabeling of the complementary DNA probe, 50 pmol of oligonucleotide was end-labeled using γ-^{32}P-ATP (Amersham, 6000 Ci/mmol) and T4 PolyNucleotide Kinase (BioLabs) for 1 hour at 37°C. The reaction is stopped by adding 1 µl of 0,2M EDTA. The probe was purified from unincorporated label using G-25 MicroSpin Columns (Amersham), and heated for 1 minute at 95°C before addition to the membrane in the hybridization solution and incubation overnight at 42°C. The following day 3 successive washes of the membrane are realized for 15 min at 42°C: a first wash with 50ml of wash solution A (2xSSC, 0,2% SDS), the second with 50 ml of wash solution B (1xSSC, 0,2% SDS), and the third with 50 ml of Stringent Wash solution (0,1xSSC, 0.2%SDS). The blots are enveloped in Saran wrap, placed in a cassette with an autoradiogram. The films were exposed for 72 hours.
The probes used for detection of the different transcripts are listed in **Table 6**.

TRANSCRIPTS	***PROBES***
U7Dtex23 **MBII-52-Dtex23**	**Dtex23:** 5' ^{32}P- GTTAGTGTAAATGAACTTC 3'
Mouse U2 snRNA	**U2snRNA:** 5'-^{32}P-CTGATAAGAACAGATAC-3'
Mouse and human U7 snRNA	**U7wt:** 5'-^{32}P-CTAAAAGAGCTGTAACACTT-3'
Unprocessed U7 snRNA	**U7unprocess:** 5'-^{32}P-CATTGTAGACCAGTGAAATT-3'
U7M23D	**probeM23D:** 5'-^{32}P-TTTCAGGTAAGCCGAGGTT-3'
Unprocessed MHCK-U7	**ProbeMHCK:** 5'-^{32}P-GATCCACCAGGGACAGGGTTA-3'
Human U24 snoRNA	**U24human:** 5'-^{32}P-GCAAATATTCTTTTACATCATCTGCA-3'
Chimeric U24MetBP-βglob	**ProbeU24mod:** 5'-^{32}P-ATTGGTCTATTTTTACATCATCTGCA-3'

Table 6: list of probes used for Northern blot detection

XIII. Luciferase assay

The Luciferase Assay System from Promega was used to assess the expression of luciferase obtained after transfection of the βglobin-Luciferase reporter construct in 293T cells. The cells were harvested two days after transfection by adding 150 µL of room temperature CCLR in each well and incubating for 2min. The collected samples were centrifuged to pellet the debris and 20µL of supernatant was mixed with 20µL of room temperature Luciferase Assay Reagent containing the substrate luciferin. We immediately measure the light produced by the cell lysates with a luminometer at 600nm.

The bioluminescence data were normalised to the total amount of proteins expressed for each sample. The Pierce BCA Protein Assay Kit (Thermo Scientific) was used according to the manufacturer's instructions for the colorimetric detection and quantitation of total protein per sample, using a standard curve reference of bovine serum albumin (BSA) protein.

XIV. MDX mice injection

All animal experiments were performed according to the guidelines and protocols approved by the Home Office. Five month-old *mdx* mice were anesthetized under isoflurane and injected in the tibialis anterior with 5.10^9, 2.10^{10} or 10^{11} viral genome copies (vg) of AAV2/8-U7DTex23 or AAV2/8-MHCK-U7DTex23 (n=3 for each condition). For each mouse, contra-lateral muscle was injected with PBS as a negative control of injection. Mice were sacrificed 4 weeks after injection and tibialis anterior muscles were isolated and snap frozen in liquid nitrogen-cooled isopentane.

REFERENCES

1. Crick, F, Central dogma of molecular biology. *Nature*, 1970. **227**(5258): p. 561-563.
2. Jacob, F and Monod, J, Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 1961. **3**(3): p. 318-356.
3. Crick, FH, The origin of the genetic code. *Journal of Molecular Biology*, 1968. **38**(3): p. 367-379.
4. Orgel, LE, Evolution of the genetic apparatus. *Journal of Molecular Biology*, 1968. **38**(3): p. 381-393.
5. Woese, CR, The genetic code: The molecular basis for genetic expression. *Harper & Row*, 1967: p. p. 186.
6. Cech, TR, A model for the RNA-catalyzed replication of RNA. *Proceedings of the National Academy of Sciences*, 1986. **83**(12): p. 4360-4363.
7. Gilbert, W, Origin of life: The RNA world. *Nature*, 1986. **319**(6055): p. 618-618.
8. Cech, TR, The RNA Worlds in Context. *Cold Spring Harbor Perspectives in Biology*, 2011.
9. Wang, E, Sandberg, R, Luo, S, Khrebtukova, I, Zhang, L, Mayr, C, Kingsmore, S, Schroth, G and Burge, C, Alternative isoform regulation in human tissue transcriptomes. *Nature*, 2008. **456**(7221): p. 470-476.
10. Sharp, PA, The Centrality of RNA. *Cell*, 2009. **136**(4): p. 577-580.
11. Xu, H, Wang, P, Fu, Y, Zheng, Y, Tang, Q, Si, L, You, J, Zhang, Z, Zhu, Y, Zhou, L, Wei, Z, Lin, B, Hu, L and Kong, X, Length of the ORF, position of the first AUG and the Kozak motif are important factors in potential dual-coding transcripts. *Cell Res*, 2010. **20**(4): p. 445-457.
12. Berretta, J and Morillon, A, Pervasive transcription constitutes a new level of eukaryotic genome regulation. *EMBO Rep*, 2009. **10**(9): p. 973-982.
13. Ulveling, D, Francastel, C and Hubé, F, When one is better than two: RNA with dual functions. *Biochimie*, 2011. **93**(4): p. 633-644.
14. Dinger, ME, Gascoigne, DK and Mattick, JS, The evolution of RNAs with multiple functions. *Biochimie*, 2011. **93**(11): p. 2013-2018.
15. Brosnan, C and Voinnet, O, The long and the short of noncoding RNAs. *Current Opinion in Cell Biology*, 2009. **21**(3): p. 416-425.
16. Gibb, E, Brown, C and Lam, W, The functional role of long non-coding RNA in human carcinomas. *Molecular Cancer*, 2011. **10**(1): p. 38.
17. Brown, CJ, Ballabio, A, Rupert, JL, Lafreniere, RG, Grompe, M, Tonlorenzi, R and Willard, HF, A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, 1991. **349**(6304): p. 38-44.
18. Tuck, AC and Tollervey, D, RNA in pieces. *Trends in genetics*, 2011. **27**(10): p. 422-432.

19. Cooper, TA, Wan, L and Dreyfuss, G, RNA and Disease. *Cell*, 2009. **136**(4): p. 777-793.
20. Wilkie, SE, Vaclavik, V, Wu, H, Bujakowska, K, Chakarova, CF, Bhattacharya, SS, Warren, MJ and Hunt, DM, Disease mechanism for retinitis pigmentosa (RP11) caused by missense mutations in the splicing factor gene PRPF31. *Mol Vis*, 2008. **14**: p. 683-690.
21. Tazi, J, Bakkour, N and Stamm, S, Alternative splicing and disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 2009. **1792**(1): p. 14-26.
22. Andreadis, A, Tau gene alternative splicing: expression patterns, regulation and modulation of function in normal brain and neurodegenerative diseases. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 2005. **1739**(2–3): p. 91-103.
23. Srebrow, A and Kornblihtt, AR, The connection between splicing and cancer. *Journal of Cell Science*, 2006. **119**(13): p. 2635-2641.
24. Esteller, M, Non-coding RNAs in human disease. *Nat Rev Genet*, 2011. **12**(12): p. 861-874.
25. Clop, A, Marcq, F, Takeda, H, Pirottin, D, Tordoir, X, Bibe, B, Bouix, J, Caiment, F, Elsen, J-M, Eychenne, F, Larzul, C, Laville, E, Meish, F, Milenkovic, D, Tobin, J, Charlier, C and Georges, M, A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet*, 2006. **38**(7): p. 813-818.
26. Abelson, J, Kwan, K, O'Roak, B, Baek, D, Stillman, A, Morgan, T, Mathews, C, Pauls, D, Rasin, M-R, Gunel, M, Davis, N, Ercan Sencicek, AG, Guez, D, Spertus, J, Leckman, J, Dure, L, Kurlan, R, Singer, H, Gilbert, D, Farhi, A, Louvi, A, Lifton, R, Sestan, N and State, M, Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science*, 2005. **310**(5746): p. 317-320.
27. Calin, GA, Dumitru, CD, Shimizu, M, Bichi, R, Zupo, S, Noch, E, Aldler, H, Rattan, S, Keating, M, Rai, K, Rassenti, L, Kipps, T, Negrini, M, Bullrich, F and Croce, CM, Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences*, 2002. **99**(24): p. 15524-15529.
28. Sahoo, T, del Gaudio, D, German, JR, Shinawi, M, Peters, SU, Person, RE, Garnica, A, Cheung, SW and Beaudet, AL, Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet*, 2008. **40**(6): p. 719-721.
29. Kole, R, Krainer, A and Altman, S, RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nature reviews Drug discovery*, 2012. **11**(2): p. 125-140.

30. Haugen, P, Simon, DM and Bhattacharya, D, The natural history of group I introns. *Trends in genetics*, 2005. **21**(2): p. 111-119.
31. Toro, N, Jimnez-Zurdo, J and Garca-Rodrguez, F, Bacterial group II introns: not just splicing. *FEMS microbiology reviews*, 2007. **31**(3): p. 342-358.
32. Randau, L and Sll, D, Transfer RNA genes in pieces. *EMBO reports*, 2008. **9**(7): p. 623-628.
33. Buratti, E, Baralle, M and Baralle, FE, Defective splicing, disease and therapy: searching for master checkpoints in exon definition. *Nucleic Acids Research*, 2006. **34**(12): p. 3494-3510.
34. Cartegni, L, Chew, SL and Krainer, AR, Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet*, 2002. **3**(4): p. 285-298.
35. Zhang, MQ, Statistical Features of Human Exons and Their Flanking Regions. *Human Molecular Genetics*, 1998. **7**(5): p. 919-932.
36. Reed, R and Maniatis, T, A role for exon sequences and splice-site proximity in splice-site selection. *Cell*, 1986. **46**(5): p. 681-690.
37. Cunningham, SA, Else, AJ, Potter, BV and Eperon, IC, Influences of separation and adjacent sequences on the use of alternative 5' splice sites. *Journal of Molecular Biology*, 1991. **217**(2): p. 265-281.
38. Reed, R, The organization of 3' splice-site sequences in mammalian introns. *Genes & Development*, 1989. **3**(12B): p. 2113-2123.
39. Hui, J, Regulation of mammalian pre-mRNA splicing. *Science in China Series C: Life Sciences*, 2009. **52**(3): p. 253-260.
40. Sharp, PA and Burge, CB, Classification of Introns: U2-Type or U12-Type. *Cell*, 1997. **91**(7): p. 875-879.
41. Hall, SL and Padgett, RA, Conserved Sequences in a Class of Rare Eukaryotic Nuclear Introns with Non-consensus Splice Sites. *Journal of Molecular Biology*, 1994. **239**(3): p. 357-365.
42. Robberson, BL, Cote, GJ and Berget, SM, Exon definition may facilitate splice site selection in RNAs with multiple exons. *Molecular and Cellular Biology*, 1990. **10**(1): p. 84-94.
43. Brow, D, Allosteric cascade of spliceosome activation. *Annual Review of Genetics*, 2002. **36**: p. 333-360.
44. Staley, JP and Guthrie, C, Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell*, 1998. **92**(3): p. 315-326.

45. Query, CC, Moore, MJ and Sharp, PA, Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes & Development*, 1994. **8**(5): p. 587-597.
46. Horowitz, DS, The mechanism of the second step of pre-mRNA splicing. *Wiley Interdisciplinary Reviews: RNA*, 2012. **3**(3): p. 331-350.
47. Valadkhan, S, snRNAs as the catalysts of pre-mRNA splicing. *Current Opinion in Chemical Biology*, 2005. **9**(6): p. 603-608.
48. Valadkhan, S and Manley, J, Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA*, 2003. **9**(7): p. 892-904.
49. Valadkhan, S, Mohammadi, A, Wachtel, C and Manley, J, Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing. *RNA*, 2007. **13**(12): p. 2300-2311.
50. Lim, LP and Burge, CB, A computational analysis of sequence features involved in recognition of short introns. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. **98**(20): p. 11193-11198.
51. Wang, Z and Burge, C, Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA*, 2008. **14**(5): p. 802-813.
52. Hiller, M, Zhang, Z, Backofen, R and Stamm, S, Pre-mRNA secondary structures influence exon recognition. *PLOS Genetics*, 2007. **3**(11): p. e204-e204.
53. Liu, H-X, Zhang, M and Krainer, AR, Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes & Development*, 1998. **12**(13): p. 1998-2012.
54. Fairbrother, W, Yeh, R-F, Sharp, P and Burge, C, Predictive identification of exonic splicing enhancers in human genes. *Science*, 2002. **297**(5583): p. 1007-1013.
55. Zhang, XH-F and Chasin, LA, Computational definition of sequence motifs governing constitutive exon splicing. *Genes & Development*, 2004. **18**(11): p. 1241-1250.
56. Cartegni, L, Wang, J, Zhu, Z, Zhang, MQ and Krainer, AR, ESEfinder: a web resource to identify exonic splicing enhancers. *Nucl Acids Res*, 2003. **31**(13): p. 3568-3571.
57. Brudno, M, Gelfand, MS, Spengler, S, Zorn, M, Dubchak, I and Conboy, JG, Computational analysis of candidate intron regulatory elements for tissue-specific alternative pre-mRNA splicing. *Nucleic Acids Research*, 2001. **29**: p. 2338-2348.
58. Ladd, A and Cooper, T, Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biology*, 2002. **3**(11): p. 0008.0001 - 0008.0016.

59. Mardon, HJ, Sebastio, G and Baralle, FE, A role for exon sequences in alternative splicing of the human fibronectin gene. *Nucleic Acids Research*, 1987. **15**(19): p. 7725-7733.
60. Schaal, TD and Maniatis, T, Multiple Distinct Splicing Enhancers in the Protein-Coding Sequences of a Constitutively Spliced Pre-mRNA. *Mol Cell Biol*, 1999. **19**(1): p. 261-273.
61. Coulter, LR, Landree, MA and Cooper, TA, Identification of a new class of exonic splicing enhancers by in vivo selection. *Molecular and Cellular Biology*, 1997. **17**(4): p. 2143-2150.
62. Graveley, BR and Maniatis, T, Arginine/Serine-Rich Domains of SR Proteins Can Function as Activators of Pre-mRNA Splicing. *Molecular Cell*, 1998. **1**(5): p. 765-771.
63. Graveley, BR, Sorting out the complexity of SR protein functions. *RNA*, 2000. **6**(9): p. 1197-1211.
64. Kan, JL and Green, MR, Pre-mRNA splicing of IgM exons M1 and M2 is directed by a juxtaposed splicing enhancer and inhibitor. *Genes & Development*, 1999. **13**(4): p. 462-471.
65. Zheng, ZM, Huynen, M and Baker, CC, A pyrimidine-rich exonic splicing suppressor binds multiple RNA splicing factors and inhibits spliceosome assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 1998. **95**(24): p. 14088-14093.
66. Del Gatto-Konczak, F, Olive, M, Gesnel, MC and Breathnach, R, hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer. *Molecular and Cellular Biology*, 1999. **19**(1): p. 251-260.
67. Wagner, EJ and Garcia Blanco, MA, Polypyrimidine tract binding protein antagonizes exon definition. *Molecular and Cellular Biology*, 2001. **21**(10): p. 3281-3288.
68. Blanchette, M and Chabot, B, Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO J*, 1999. **18**(7): p. 1939-1952.
69. Zhu, J, Mayeda, A and Krainer, AR, Exon Identity Established through Differential Antagonism between Exonic Splicing Silencer-Bound hnRNP A1 and Enhancer-Bound SR Proteins. *Molecular Cell*, 2001. **8**(6): p. 1351-1361.
70. McCullough, AJ and Berget, SM, G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. *Molecular and Cellular Biology*, 1997. **17**(8): p. 4562-4571.

71. McCullough, AJ and Berget, SM, An intronic splicing enhancer binds U1 snRNPs to enhance splicing and select 5' splice sites. *Molecular and Cellular Biology*, 2000. **20**(24): p. 9225-9235.
72. Hui, J, Stangl, K, Lane, W and Bindereif, A, HnRNP L stimulates splicing of the eNOS gene by binding to variable-length CA repeats. *Nature structural biology*, 2003. **10**(1): p. 33-37.
73. Chen, CD, Kobayashi, R and Helfman, DM, Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat β-tropomyosin gene. *Genes & Development*, 1999. **13**(5): p. 593-606.
74. Dominski, Z and Kole, R, Identification of exon sequences involved in splice site selection. *Journal of Biological Chemistry*, 1994. **269**(38): p. 23590-23596.
75. Smith, CWJ and Valcárcel, J, Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends in Biochemical Sciences*, 2000. **25**(8): p. 381-388.
76. Tazi, J, Durand, S and Jeanteur, P, The spliceosome: a novel multi-faceted target for therapy. *Trends in Biochemical Sciences*, 2005. **30**(8): p. 469-478.
77. Maniatis, T and Reed, R, An extensive network of coupling among gene expression machines. *Nature*, 2002. **416**(6880): p. 499-506.
78. Tennyson, CN, Klamut, HJ and Worton, RG, The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nat Genet*, 1995. **9**(2): p. 184-190.
79. Neugebauer, KM, On the importance of being co-transcriptional. *Journal of Cell Science*, 2002. **115**(20): p. 3865-3871.
80. Proudfoot, NJ, Furger, A and Dye, MJ, Integrating mRNA Processing with Transcription. *Cell*, 2002. **108**(4): p. 501-512.
81. de la Mata, M, Alonso, CR, Kadener, S, Fededa, JP, Blaustein, Ma, Pelisch, F, Cramer, P, Bentley, D and Kornblihtt, AR, A Slow RNA Polymerase II Affects Alternative Splicing In Vivo. *Molecular Cell*, 2003. **12**(2): p. 525-532.
82. Hicks, MJ, Yang, C-R, Kotlajich, MV and Hertel, KJ, Linking Splicing to Pol II Transcription Stabilizes Pre-mRNAs and Influences Splicing Patterns. *PLoS Biol*, 2006. **4**(6): p. e147.
83. de la Mata, M and Kornblihtt, AR, RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nat Struct Mol Biol*, 2006. **13**(11): p. 973-980.
84. Das, R, Yu, J, Zhang, Z, Gygi, MP, Krainer, AR, Gygi, SP and Reed, R, SR Proteins Function in Coupling RNAP II Transcription to Pre-mRNA Splicing. *Molecular Cell*, 2007. **26**(6): p. 867-881.

85. Fong, YW and Zhou, Q, Stimulatory effect of splicing factors on transcriptional elongation. *Nature*, 2001. **414**(6866): p. 929-933.
86. Kornblihtt, A, de la Mata, M, Fededa, J, Munoz, M and Nogues, G, Multiple links between transcription and splicing. *RNA*, 2004. **10**(10): p. 1489-1498.
87. Martinson, HG, An active role for splicing in 3′-end formation. *Wiley Interdisciplinary Reviews: RNA*, 2011. **2**(4): p. 459-470.
88. Niwa, M, Rose, SD and Berget, SM, In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes & Development*, 1990. **4**(9): p. 1552-1559.
89. Niwa, M, MacDonald, CC and Berget, SM, Are vertebrate exons scanned during splice-site selection? *Nature*, 1992. **360**(6401): p. 277-280.
90. Stamm, S, Ben Ari, S, Rafalska, I, Tang, Y, Zhang, Z, Toiber, D, Thanaraj, TA and Soreq, H, Function of alternative splicing. *Gene*, 2005. **344**: p. 1-20.
91. Graveley, BR, Alternative splicing: increasing diversity in the proteomic world. *Trends in genetics*, 2001. **17**(2): p. 100-107.
92. Black, D, Mechanisms of alternative pre-messenger RNA splicing. *Annual review of biochemistry*, 2003. **72**: p. 291-336.
93. Chasin, LA, Searching for splicing motifs. *Adv Exp Med Biol*, 2007. **623**: p. 85-106.
94. Luco, RF and Misteli, T, More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. *Current Opinion in Genetics & Development*, 2011. **21**(4): p. 366-372.
95. Schwartz, S, Meshorer, E and Ast, G, Chromatin organization marks exon-intron structure. *Nature structural & molecular biology*, 2009. **16**(9): p. 990-995.
96. Kolasinska Zwierz, P, Down, T, Latorre, I, Liu, T, Liu, XS and Ahringer, J, Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature Genetics*, 2009. **41**(3): p. 376-381.
97. Luco, RF, Pan, Q, Tominaga, K, Blencowe, BJ, Pereira-Smith, OM and Misteli, T, Regulation of Alternative Splicing by Histone Modifications. *Science*, 2010. **327**(5968): p. 996-1000.
98. Luco, RF, Allo, M, Schor, IE, Kornblihtt, AR and Misteli, T, Epigenetics in Alternative Pre-mRNA Splicing. *Cell*, 2011. **144**(1): p. 16-26.
99. Muro, AF, Caputi, M, Pariyarath, R, Pagani, F, Buratti, E and Baralle, FE, Regulation of Fibronectin EDA Exon Alternative Splicing: Possible Role of RNA Secondary Structure for Enhancer Display. *Molecular and Cellular Biology*, 1999. **19**(4): p. 2657-2671.
100. Buratti, E and Baralle, FE, Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process. *Molecular and Cellular Biology*, 2004. **24**(24): p. 10505-10514.

101. Shin, C and Manley, J, Cell signalling and the control of pre-mRNA splicing. *Nature reviews Molecular cell biology*, 2004. **5**(9): p. 727-738.

102. Weg-Remers, S, Ponta, H, Herrlich, P and Konig, H, Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J*, 2001. **20**(15): p. 4194-4203.

103. Stamm, S, Regulation of Alternative Splicing by Reversible Protein Phosphorylation. *Journal of Biological Chemistry*, 2008. **283**(3): p. 1223-1227.

104. Boutz, PL, Chawla, G, Stoilov, P and Black, DL, MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development. *Genes & Development*, 2007. **21**(1): p. 71-84.

105. Kishore, S and Stamm, S, The snoRNA HBII-52 Regulates Alternative Splicing of the Serotonin Receptor 2C. *Science*, 2006. **311**(5758): p. 230-232.

106. Kishore, S, Khanna, A, Zhang, Z, Hui, J, Balwierz, PJ, Stefan, M, Beach, C, Nicholls, RD, Zavolan, M and Stamm, S, The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. *Hum Mol Genet*, 2010: p. ddp585.

107. Pagani, F and Baralle, FE, Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet*, 2004. **5**(5): p. 389-396.

108. Pagani, F, Raponi, M and Baralle, F, Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 2005. **102**(18): p. 6368-6372.

109. Krawczak, M, Reiss, J and Cooper, DN, The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Human genetics*, 1992. **90**(1-2): p. 41-54.

110. Matlin, A, Clark, F and Smith, CWJ, Understanding alternative splicing: towards a cellular code. *Nature reviews Molecular cell biology*, 2005. **6**(5): p. 386-398.

111. Hilleren, P and Parker, R, Mechanisms of mRNA surveillance in eukaryotes. *Annu Rev Genet*, 1999. **33**: p. 229-260.

112. Nissim Rafinia, M and Kerem, B, Splicing regulation as a potential genetic modifier. *Trends in genetics*, 2002. **18**(3): p. 123-127.

113. Ibrahim, E, Hims, M, Shomron, N, Burge, C, Slaugenhaupt, S and Reed, R, Weak definition of IKBKAP exon 20 leads to aberrant splicing in familial dysautonomia. *Human Mutation*, 2007. **28**(1): p. 41-53.

114. Roca, X, Sachidanandam, R and Krainer, A, Intrinsic differences between authentic and cryptic 5' splice sites. *Nucleic Acids Research*, 2003. **31**(21): p. 6321-6333.

115. Wieringa, B, Meyer, F, Reiser, J and Weissmann, C, Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit beta-globin gene. *Nature*, 1983. **301**(5895): p. 38-43.

116. Atweh, GF, Anagnou, NP, Shearin, J, Forget, BG and Kaufman, RE, Beta-thalassemia resulting from a single nucleotide substitution in an acceptor splice site. *Nucleic Acids Research*, 1985. **13**(3): p. 777-790.

117. Péquignot, MO, Dey, R, Zeviani, M, Tiranti, V, Godinot, C, Poyau, A, Sue, C, Di Mauro, S, Abitbol, M and Marsac, C, Mutations in the SURF1 gene associated with Leigh syndrome and cytochrome C oxidase deficiency. *Human Mutation*, 2001. **17**(5): p. 374-381.

118. Sun, H and Chasin, LA, Multiple splicing defects in an intronic false exon. *Molecular and Cellular Biology*, 2000. **20**(17): p. 6414-6425.

119. Tuffery Giraud, S, Saquet, C, Chambert, S and Claustres, M, Pseudoexon activation in the DMD gene as a novel mechanism for Becker muscular dystrophy. *Human Mutation*, 2003. **21**(6): p. 608-614.

120. Cartegni, L and Krainer, A, Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nature Genetics*, 2002. **30**(4): p. 377-384.

121. Kashima, T and Manley, J, A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nature Genetics*, 2003. **34**(4): p. 460-463.

122. Cartegni, L, Hastings, M, Calarco, J, de Stanchina, E and Krainer, A, Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *American journal of human genetics*, 2006. **78**(1): p. 63-77.

123. Spillantini, MG, Van Swieten, JC and Goedert, M, Tau gene mutations in frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17). *neurogenetics*, 2000. **2**(4): p. 193-205.

124. Kanadia, R, Johnstone, K, Mankodi, A, Lungu, C, Thornton, C, Esson, D, Timmers, A, Hauswirth, W and Swanson, M, A muscleblind knockout model for myotonic dystrophy. *Science*, 2003. **302**(5652): p. 1978-1980.

125. McKie, AB, McHale, JC, Keen, TJ, Tarttelin, EE, Goliath, R, van Lith-Verhoeven, JJ, Greenberg, J, Ramesar, RS, Hoyng, CB, Cremers, FP, Mackey, DA, Bhattacharya, SS, Bird, AC, Markham, AF and Inglehearn, CF, Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Human Molecular Genetics*, 2001. **10**(15): p. 1555-1562.

126. Neumann, M, Sampathu, D, Kwong, L, Truax, A, Micsenyi, M, Chou, T, Bruce, J, Schuck, T, Grossman, M, Clark, C, McCluskey, L, Miller, B, Masliah, E, Mackenzie, I, Feldman, H, Feiden, W, Kretzschmar, H, Trojanowski, J and Lee, VMY, Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*, 2006. **314**(5796): p. 130-133.

127. Ghigna, C, Giordano, S, Shen, H, Benvenuto, F, Castiglioni, F, Comoglio, P, Green, M, Riva, S and Biamonti, G, Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene. *Molecular Cell*, 2005. **20**(6): p. 881-890.

128. Christofk, H, Vander Heiden, M, Harris, M, Ramanathan, A, Gerszten, R, Wei, R, Fleming, M, Schreiber, S and Cantley, L, The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature*, 2008. **452**(7184): p. 230-233.

129. David, C, Chen, M, Assanah, M, Canoll, P and Manley, J, HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature*, 2010. **463**(7279): p. 364-368.

130. Clower, C, Chatterjee, D, Wang, Z, Cantley, L, Vander Heiden, M and Krainer, A, The alternative splicing repressors hnRNP A1/A2 and PTB influence pyruvate kinase isoform expression and cell metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 2010. **107**(5): p. 1894-1899.

131. Grosso, A, Martins, S and Carmo Fonseca, M, The emerging role of splicing factors in cancer. *EMBO reports*, 2008. **9**(11): p. 1087-1093.

132. Mazoyer, S, Puget, N, Perrin-Vidoz, L, Lynch, HT, Serova-Sinilnikova, OM and Lenoir, GM, A BRCA1 Nonsense Mutation Causes Exon Skipping. *American journal of human genetics*, 1998. **62**(3): p. 713-715.

133. Sumanasekera, C, Watt, D and Stamm, S, Substances that can change alternative splice-site selection. *Biochemical Society Transactions*, 2008. **36**(3): p. 483-490.

134. Muraki, M, Ohkawara, B, Hosoya, T, Onogi, H, Koizumi, J, Koizumi, T, Sumi, K, Yomoda, J-i, Murray, M, Kimura, H, Furuichi, K, Shibuya, H, Krainer, A, Suzuki, M and Hagiwara, M, Manipulation of alternative splicing by a newly developed inhibitor of Clks. *Journal of Biological Chemistry*, 2004. **279**(23): p. 24246-24254.

135. Novoyatleva, T, Heinrich, B, Tang, Y, Benderska, N, Butchbach, MER, Lorson, CL, Lorson, MA, Ben-Dov, C, Fehlbaum, P, Bracco, L, Burghes, AHM, Bollen, M and Stamm, S, Protein phosphatase 1 binds to the RNA recognition motif of several splicing factors and regulates alternative pre-mRNA processing. *Human Molecular Genetics*, 2008. **17**(1): p. 52-70.

136. Chang, JG, Hsieh Li, HM, Jong, YJ, Wang, NM, Tsai, CH and Li, H, Treatment of spinal muscular atrophy by sodium butyrate. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. **98**(17): p. 9808-9813.
137. Brichta, L, Hofmann, Y, Hahnen, E, Siebzehnrubl, FA, Raschke, H, Blumcke, I, Eyupoglu, IY and Wirth, B, Valproic acid increases the SMN2 protein level: a well-known drug as a potential therapy for spinal muscular atrophy. *Human Molecular Genetics*, 2003. **12**(19): p. 2481-2489.
138. Zingman, LV, Park, S, Olson, TM, Alekseev, AE and Terzic, A, Aminoglycoside-induced translational read-through in disease: overcoming nonsense mutations by pharmacogenetic therapy. *Clinical pharmacology & therapeutics*, 2007. **81**(1): p. 99-103.
139. Hainrichson, M, Nudelman, I and Baasov, T, Designer aminoglycosides: the race to develop improved antibiotics and compounds for the treatment of human genetic diseases. *Organic & biomolecular chemistry*, 2008. **6**(2): p. 227-239.
140. Hammond, S and Wood, MJA, Genetic therapies for RNA mis-splicing diseases. *Trends in genetics*, 2011. **27**(5): p. 196-205.
141. Sazani, P and Kole, R, Therapeutic potential of antisense oligonucleotides as modulators of alternative splicing. *J Clin Invest*, 2003. **112**(4): p. 481-486.
142. Dominski, Z, Kole, R., Restoration of correct splicing in thalassemic pre-mRNA by antisens oligonucleotides. *Proc Natl Acad Sci USA*, 1993. **vol.90**: p. pp.8673-8677.
143. Du, L and Gatti, RA, Progress toward therapy with antisense-mediated splicing modulation. *Current opinion in molecular therapeutics*, 2009. **11**(2): p. 116-123.
144. Branch, AD, A good antisense molecule is hard to find. *Trends in Biochemical Sciences*, 1998. **23**(2): p. 45-50.
145. Chan, JHP, Lim, S and Wong, WSF, Antisense Oligonucleotides: from design to therapeutic application. *Clinical and Experimental Pharmacology and Physiology*, 2006. **33**(5-6): p. 533-540.
146. Zuker, M, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 2003. **31**(13): p. 3406-3415.
147. Vickers, TA, Wyatt, JR and Freier, SM, Effects of RNA secondary structure on cellular antisense activity. *Nucleic Acids Research*, 2000. **28**(6): p. 1340-1347.
148. Aartsma Rus, A, van Vliet, L, Hirschi, M, Janson, AAM, Heemskerk, H, de Winter, C, de Kimpe, S, van Deutekom, JCT, t Hoen, PAC and van Ommen, G-J, Guidelines for antisense oligonucleotide design and insight into splice-modulating mechanisms. *Molecular therapy*, 2009. **17**(3): p. 548-553.
149. Mathews, DH, Burkard, ME, Freier, SM, Wyatt, JR and Turner, DH, Predicting oligonucleotide affinity to nucleic acid targets. *RNA*, 1999. **5**(11): p. 1458-1469.

150. Crooke, S, Progress in antisense technology. *Annual Review of Medicine*, 2004. **55**: p. 61-95.
151. Kurreck, J, Antisense technologies. Improvement through novel chemical modifications. *European Journal of Biochemistry*, 2003. **270**(8): p. 1628-1644.
152. Amantana, A and Iversen, P, Pharmacokinetics and biodistribution of phosphorodiamidate morpholino antisense oligomers. *Current Opinion in Pharmacology*, 2005. **5**(5): p. 550-555.
153. Lundin, K, Good, L, Strömberg, R, Gräslund, A and Smith, CIE, Biological activity and biotechnological aspects of peptide nucleic acid. *Adv Genet*, 2006. **56**: p. 1-51.
154. Pande, V and Nilsson, L, Insights into structure, dynamics and hydration of locked nucleic acid (LNA) strand-based duplexes from molecular dynamics simulations. *Nucleic Acids Research*, 2008. **36**(5): p. 1508-1516.
155. Du, L, Pollard, JM and Gatti, RA, Correction of prototypic ATM splicing mutations and aberrant ATM function with antisense morpholino oligonucleotides. *Proceedings of the National Academy of Sciences*, 2007. **104**(14): p. 6007-6012.
156. Sierakowska, H, Sambade, MJ, Agrawal, S and Kole, R, Repair of thalassemic human beta-globin mRNA in mammalian cells by antisense oligonucleotides. *Proc Natl Acad Sci U S A*, 1996. **93**(23): p. 12840-12844.
157. Friedman, KJ, Kole, J, Cohn, JA, Knowles, MR, Silverman, LM and Kole, R, Correction of Aberrant Splicing of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Gene by Antisense Oligonucleotides. *Journal of Biological Chemistry*, 1999. **274**(51): p. 36193-36199.
158. Rodrguez-Pascau, L, Coll, M, Vilageliu, L and Grinberg, D, Antisense oligonucleotide treatment for a pseudoexon-generating mutation in the NPC1 gene causing Niemann-Pick type C disease. *Human Mutation*, 2009. **30**(11): p. E993-E1001.
159. Collin, RWJ, den Hollander, AI, van der Velde-Visser, SD, Bennicelli, J, Bennett, J and Cremers, FPM, Antisense Oligonucleotide (AON)-based Therapy for Leber Congenital Amaurosis Caused by a Frequent Mutation in CEP290. *Mol Ther Nucleic Acids*, 2012. **1**: p. e14.
160. Gerard, X, Perrault, I, Hanein, S, Silva, E, Bigot, K, Defoort-Delhemmes, S, Rio, M, Munnich, A, Scherman, D, Kaplan, J, Kichler, A and Rozet, J-M, AON-mediated Exon Skipping Restores Ciliation in Fibroblasts Harboring the Common Leber Congenital Amaurosis CEP290 Mutation. *Mol Ther Nucleic Acids*, 2012. **1**: p. e29.
161. Hua, Y, Vickers, T, Baker, B, Bennett, CF and Krainer, A, Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon. *PLoS Biology*, 2007. **5**(4): p. e73-e73.

162. Hua, Y, Vickers, T, Okunola, H, Bennett, CF and Krainer, A, Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *American journal of human genetics*, 2008. **82**(4): p. 834-848.

163. Hua, Y, Sahashi, K, Hung, G, Rigo, F, Passini, M, Bennett, CF and Krainer, A, Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes & Development*, 2010. **24**(15): p. 1634-1644.

164. Skordis, L, Dunckley, M, Yue, B, Eperon, I and Muntoni, F, Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*, 2003. **100**(7): p. 4114-4119.

165. Cartegni, L and Krainer, A, Correction of disease-associated exon skipping by synthetic exon-specific activators. *Nature structural biology*, 2003. **10**(2): p. 120-125.

166. Aartsma-Rus, A and Van Ommen, GB, Antisense-mediated exon skipping: A versatile tool with therapeutic and research applications. *RNA*, 2007. **13**(10): p. 1609-1624.

167. Aartsma Rus, A, Antisense-mediated modulation of splicing: therapeutic implications for Duchenne muscular dystrophy. *RNA Biology*, 2010. **7**(4): p. 453-461.

168. Khoo, B, Roca, X, Chew, S and Krainer, A, Antisense oligonucleotide-induced alternative splicing of the APOB mRNA generates a novel isoform of APOB. *BMC molecular biology*, 2007. **8**: p. 3-3.

169. Villemaire, J, Dion, I, Elela, S and Chabot, B, Reprogramming alternative pre-messenger RNA splicing through the use of protein-binding antisense oligonucleotides. *Journal of Biological Chemistry*, 2003. **278**(50): p. 50031-50039.

170. Sutton, RE and Boothroyd, JC, Evidence for trans splicing in trypanosomes. *Cell*, 1986. **47**(4): p. 527-535.

171. Krause, M and Hirsh, D, A trans-spliced leader sequence on actin mRNA in C. elegans. *Cell*, 1987. **49**(6): p. 753-761.

172. Puttaraju, M, Jamison, SF, Mansfield, SG, Garcia Blanco, MA and Mitchell, LG, Spliceosome-mediated RNA trans-splicing as a tool for gene therapy. *Nature biotechnology*, 1999. **17**(3): p. 246-252.

173. Coady, T, Baughan, T, Shababi, M, Passini, M and Lorson, C, Development of a single vector system that enhances trans-splicing of SMN2 transcripts. *PLoS ONE*, 2008. **3**(10): p. e3468-e3468.

174. Coady, T and Lorson, C, Trans-splicing-mediated improvement in a severe mouse model of spinal muscular atrophy. *The Journal of neuroscience*, 2010. **30**(1): p. 126-130.

175. Liu, X, Jiang, Q, Mansfield, SG, Puttaraju, M, Zhang, Y, Zhou, W, Cohn, J, Garcia Blanco, M, Mitchell, L and Engelhardt, J, Partial correction of endogenous DeltaF508 CFTR in human cystic fibrosis airway epithelia by spliceosome-mediated RNA trans-splicing. *Nature biotechnology*, 2002. **20**(1): p. 47-52.

176. Bachellerie, J-P, Cavaillé, J and Hüttenhofer, A, The expanding snoRNA world. *Biochimie*, 2002. **84**(8): p. 775-790.

177. Kiss, T, Small Nucleolar RNAs: An Abundant Group of Noncoding RNAs with Diverse Cellular Functions. *Cell*, 2002. **109**(2): p. 145-148.

178. Reichow, SL, Hamma, T, Ferre-D'Amare, AR and Varani, G, The structure and function of small nucleolar ribonucleoproteins. *Nucl Acids Res*, 2007. **35**(5): p. 1452-1464.

179. Lestrade, L and Weber, MJ, snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Research*, 2005. **34**(suppl 1): p. D158-D162.

180. Fedorov, A, Stombaugh, J, Harr, MW, Yu, S, Nasalean, L and Shepelev, V, Computer identification of snoRNA genes using a Mammalian Orthologous Intron Database. *Nucleic Acids Research*, 2005. **33**(14): p. 4578-4583.

181. Washietl, S, Hofacker, IL, Lukasser, M, Huttenhofer, A and Stadler, PF, Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. *Nat Biotech*, 2005. **23**(11): p. 1383-1390.

182. Yang, J-H, Zhang, X-C, Huang, Z-P, Zhou, H, Huang, M-B, Zhang, S, Chen, Y-Q and Qu, L-H, snoSeeker: an advanced computational package for screening of guide and orphan snoRNA genes in the human genome. *Nucleic Acids Research*, 2006. **34**(18): p. 5112-5123.

183. Kiss-Laszlo, Z, Henry, Y and Kiss, T, Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *The EMBO journal*, 1998. **17**(3): p. 797-807.

184. Cavaille, J, Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature*, 1996. **383**(6602): p. 732-735.

185. Kiss-Laszlo, Z, Henry, Y, Bachellerie, JP, Caizergues Ferrer, M and Kiss, T, Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell*, 1996. **85**(7): p. 1077-1088.

186. Zhao, X and Yu, Y-T, Targeted pre-mRNA modification for gene silencing and regulation. *Nat Meth*, 2008. **5**(1): p. 95-100.

187. Tollervey, D and Kiss, T, Function and synthesis of small nucleolar RNAs. *Current Opinion in Cell Biology*, 1997. **9**(3): p. 337-342.

188. Tycowski, KT, Shu, MD and Steitz, JA, A mammalian gene with introns instead of exons generating stable RNA products. *Nature*, 1996. **379**(6564): p. 464-466.

189. Tanaka, R, Satoh, H, Moriyama, M, Satoh, K, Morishita, Y, Yoshida, S, Watanabe, T, Nakamura, Y and Mori, S, Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. *Genes to cells*, 2000. **5**(4): p. 277-287.

190. Filipowicz, W and Pogacic, V, Biogenesis of small nucleolar ribonucleoproteins. *Current Opinion in Cell Biology*, 2002. **14**(3): p. 319-327.

191. Kiss, T and Filipowicz, W, Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes & Development*, 1995. **9**(11): p. 1411-1424.

192. Caffarelli, E, Fatica, A, Prislei, S, De Gregorio, E, Fragapane, P and Bozzoni, I, Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *EMBO journal*, 1996. **15**(5): p. 1121-1131.

193. Cavaillé, J and Bachellerie, JP, Processing of fibrillarin-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure. *Biochimie*, 1996. **78**(6): p. 443-456.

194. Richard, P and Kiss, T, Integrating snoRNP assembly with mRNA biogenesis. *EMBO Rep*, 2006. **7**(6): p. 590-592.

195. Kiss, T, SnoRNP biogenesis meets Pre-mRNA splicing. *Molecular Cell*, 2006. **23**(6): p. 775-776.

196. Hirose, T, Shu, M-D and Steitz, JA, Splicing-Dependent and -Independent Modes of Assembly for Intron-Encoded Box C/D snoRNPs in Mammalian Cells. *Molecular Cell*, 2003. **12**(1): p. 113-123.

197. Hirose, T and Steitz, JA, Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proceedings of the National Academy of Sciences*, 2001. **98**(23): p. 12914-12919.

198. Vincenti, S, De Chiara, V, Bozzoni, I and Presutti, C, The position of yeast snoRNA-coding regions within host introns is essential for their biosynthesis and for efficient splicing of the host pre-mRNA. *RNA*, 2007. **13**(1): p. 138-150.

199. Hirose, T, Ideue, T, Nagai, M, Hagiwara, M, Shu, M-D and Steitz, J, A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Molecular Cell*, 2006. **23**(5): p. 673-684.

200. Weinstein, LB and Steitz, JA, Guided tours: from precursor snoRNA to functional snoRNP. *Current Opinion in Cell Biology*, 1999. **11**(3): p. 378-384.

201. Lange, TS, Borovjagin, A, Maxwell, ES and Gerbi, SA, Conserved boxes C and D are essential nucleolar localization elements of U14 and U8 snoRNAs. *EMBO journal*, 1998. **17**(11): p. 3176-3187.

202. Samarsky, DA, Fournier, MJ, Singer, RH and Bertrand, E, The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO journal*, 1998. **17**(13): p. 3747-3757.

203. Cavaille, J, Buiting, K, Kiefmann, M, Lalande, M, Brannan, CI, Horsthemke, B, Bachellerie, J-P, Brosius, Jr and Hüttenhofer, A, Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proceedings of the National Academy of Sciences of the United States of America*, 2000. **97**(26): p. 14311-14316.

204. Soeno, Y, Taya, Y, Stasyk, T, Huber, LA, Aoba, T and Hüttenhofer, A, Identification of novel ribonucleo-protein complexes from the brain-specific snoRNA MBII-52. *RNA*, 2010. **16**(7): p. 1293-1300.

205. Semenov, DV, Vratskih, OV, Kuligina, EV and Richter, VA, Splicing by Exon Exclusion Impaired by Artificial Box C/D RNA Targeted to Branch-Point Adenosine. *Annals of the New York Academy of Sciences*, 2008. **1137**: p. 119-124.

206. Ge, J, Liu, H and Yu, Y-T, Regulation of pre-mRNA splicing in Xenopus oocytes by targeted 2′-O-methylation. *RNA*, 2010. **16**(5): p. 1078-1085.

207. Jason, TLH, Koropatnick, J and Berg, R, Toxicology of antisense therapeutics. *Toxicology and applied pharmacology*, 2004. **201**(1): p. 66-83.

208. Gorman, L, Suter, D, Emerick, V, Schümperli, D and Kole, R, Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, 1998. **95**(9): p. 4929-4934.

209. Gorman, L, Mercatante, DR and Kole, R, Restoration of Correct Splicing of Thalassemic beta -Globin Pre-mRNA by Modified U1 snRNAs. *J Biol Chem*, 2000. **275**(46): p. 35914-35919.

210. De Angelis, F, Sthandier, O, Berarducci, B, Toso, S, Galluzzi, G, Ricci, E, Cossu, G and Bozzoni, I, Chimeric snRNA molecules carrying antisense sequences against the splice junctions of exon 51 of the dystrophin pre-mRNA induce exon skipping and restoration of a dystrophin synthesis in Delta 48-50 DMD cells. *Proceedings of the National Academy of Sciences of the United States of America*, 2002. **99**(14): p. 9456-9461.

211. Matera, AG, Terns, RM and Terns, MP, Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol*, 2007. **8**(3): p. 209-220.

212. Westin, G, Zabielski, J, Hammarstrm, K, Monstein, HJ, Bark, C and Pettersson, U, Clustered genes for human U2 RNA. *Proceedings of the National Academy of Sciences of the United States of America*, 1984. **81**(12): p. 3811-3815.

213. Hernandez, N, Small Nuclear RNA Genes: a Model System to Study Fundamental Mechanisms of Transcription. *Journal of Biological Chemistry*, 2001. **276**(29): p. 26733-26736.

214. Murphy, S, Yoon, JB, Gerster, T and Roeder, RG, Oct-1 and Oct-2 potentiate functional interactions of a transcription factor with the proximal sequence element of small nuclear RNA genes. *Molecular and Cellular Biology*, 1992. **12**(7): p. 3247-3261.

215. Yoon, JB, Murphy, S, Bai, L, Wang, Z and Roeder, RG, Proximal sequence element-binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. *Molecular and Cellular Biology*, 1995. **15**(4): p. 2019-2027.

216. Ma, B and Hernandez, N, A Map of Protein-Protein Contacts within the Small Nuclear RNA-activating Protein Complex SNAPc. *Journal of Biological Chemistry*, 2001. **276**(7): p. 5027-5035.

217. Kuhlman, TC, Cho, H, Reinberg, D and Hernandez, N, The General Transcription Factors IIA, IIB, IIF, and IIE Are Required for RNA Polymerase II Transcription from the Human U1 Small Nuclear RNA Promoter. *Molecular and Cellular Biology*, 1999. **19**(3): p. 2130-2141.

218. Egloff, S, O'Reilly, D and Murphy, S, Expression of human snRNA genes from beginning to end. *Biochemical Society Transactions*, 2008. **36**(4): p. 590–594.

219. Cuello, P, Boyd, DC, Dye, MJ, Proudfoot, NJ and Murphy, S, Transcription of the human U2 snRNA genes continues beyond the 3[prime] box in vivo. *EMBO J*, 1999. **18**(10): p. 2867-2877.

220. Hernandez, N, Formation of the 3' end of U1 snRNA is directed by a conserved sequence located downstream of the coding region. *EMBO journal*, 1985. **4**(7): p. 1827-1837.

221. Hernandez, N and Weiner, AM, Formation of the 3′ end of U1 snRNA requires compatible snRNA promoter elements. *Cell*, 1986. **47**(2): p. 249-258.

222. de Vegvar, HE, Lund, E and Dahlberg, JE, 3' end formation of U1 snRNA precursors is coupled to transcription from snRNA promoters. *Cell*, 1986. **47**(2): p. 259-266.

223. Ezzeddine, N, Chen, J, Waltenspiel, B, Burch, B, Albrecht, T, Zhuo, M, Warren, WD, Marzluff, WF and Wagner, EJ, A Subset of Drosophila Integrator Proteins Is Essential for Efficient U7 snRNA and Spliceosomal snRNA 3'-End Formation. *Mol Cell Biol*, 2011. **31**(2): p. 328-341.

224. Baillat, D, Hakimi, M-A, Näär, AM, Shilatifard, A, Cooch, N and Shiekhattar, R, Integrator, a Multiprotein Mediator of Small Nuclear RNA Processing, Associates with the C-Terminal Repeat of RNA Polymerase II. *Cell*, 2005. **123**(2): p. 265-276.

225. Chen, J and Wagner, EJ, snRNA 3'end formation: the dawn of the Integrator complex. *Biochemical Society Transactions*, 2010. **38**(4): p. 1082–1087.

226. Uguen, P and Murphy, S, The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing. *EMBO J*, 2003. **22**(17): p. 4544-4554.

227. Medlin, JE, Uguen, P, Taylor, A, Bentley, DL and Murphy, S, The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3[prime] processing of U2 snRNA. *EMBO J*, 2003. **22**(4): p. 925-934.

228. Egloff, S, O'Reilly, D, Chapman, RD, Taylor, A, Tanzhaus, K, Pitts, L, Eick, D and Murphy, S, Serine-7 of the RNA Polymerase II CTD Is Specifically Required for snRNA Gene Expression. *Science*, 2007. **318**(5857): p. 1777-1779.

229. Ohno, M, Segref, A, Bachi, A, Wilm, M and Mattaj, IW, PHAX, a Mediator of U snRNA Nuclear Export Whose Activity Is Regulated by Phosphorylation. *Cell*, 2000. **101**(2): p. 187-198.

230. Segref, A, Mattaj, IW and Ohno, M, The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export. *RNA*, 2001. **7**(3): p. 351-360.

231. Kolb, SJ, Battle, DJ and Dreyfuss, G, Molecular Functions of the SMN Complex. *Journal of Child Neurology*, 2007. **22**(8): p. 990-994.

232. Patel, SB and Bellini, M, The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Research*, 2008. **36**(20): p. 6482-6493.

233. Mouaikel, J, Narayanan, U, Verheggen, C, Matera, AG, Bertrand, E, Tazi, J and Bordonne, R, Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep*, 2003. **4**(6): p. 616-622.

234. Stefanovic, B, Hackl, W, Luhrmann, R and Schumperll, D, Assemble, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into Xenopus oocytes. *Nucl Acids Res*, 1995. **23**(16): p. 3141-3151.

235. Seipelt, RL, Zheng, B, Asuru, A and Rymond, BC, U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway. *Nucleic Acids Research*, 1999. **27**(2): p. 587-595.

236. Huang, Q and Pederson, T, A human U2 RNA mutant stalled in 3′ end processing is impaired in nuclear import. *Nucleic Acids Research*, 1999. **27**(4): p. 1025-1031.

237. Jady, BE, Darzacq, X, Tucker, KE, Matera, AG, Bertrand, E and Kiss, T, Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO J*, 2003. **22**(8): p. 1878-1888.

238. Nesic, D, Tanackovic, G and Krämer, A, A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *Journal of Cell Science*, 2004. **117**(19): p. 4423-4433.

239. Denti, MA, Rosa, A, D'Antona, G, Sthandier, O, De Angelis, FG, Nicoletti, C, Allocca, M, Pansarasa, O, Parente, V, Musarò, A, Auricchio, A, Bottinelli, R and Bozzoni, I, Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model. *Proceedings of the National Academy of Sciences of the United States of America*, 2006. **103**(10): p. 3758-3763.

240. Incitti, T, De Angelis, FG, Cazzella, V, Sthandier, O, Pinnaro, C, Legnini, I and Bozzoni, I, Exon Skipping and Duchenne Muscular Dystrophy Therapy: Selection of the Most Active U1 snRNA Antisense Able to Induce Dystrophin Exon 51 Skipping. *Mol Ther*, 2010. **18**(9): p. 1675-1682.

241. Pinotti, M, Rizzotto, L, Balestra, D, Lewandowska, MA, Cavallari, N, Marchetti, G, Bernardi, F and Pagani, F, U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency. *Blood*, 2008. **111**(5): p. 2681-2684.

242. Pinotti, M, Balestra, D, Rizzotto, L, Maestri, I, Pagani, F and Bernardi, F, Rescue of coagulation factor VII function by the U1+5A snRNA. *Blood*, 2009. **113**(25): p. 6461-6464.

243. Schmid, F, Glaus, E, Barthelmes, D, Fliegauf, M, Gaspar, H, Nürnberg, G, Nürnberg, P, Omran, H, Berger, W and Neidhardt, J, U1 snRNA-mediated gene therapeutic correction of splice defects caused by an exceptionally mild BBS mutation. *Human Mutation*, 2011. **32**(7): p. 815-824.

244. Fernandez Alanis, E, Pinotti, M, Dal Mas, A, Balestra, D, Cavallari, N, Rogalska, M, Bernardi, F and Pagani, F, An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. *Human Molecular Genetics*, 2012. **21**(11): p. 2389-2398.

245. Müller, B and Schümperli, D, The U7 snRNP and the hairpin binding protein: Key players in histone mRNA metabolism. *Seminars in Cell & Developmental Biology*, 1997. **8**(6): p. 567-576.

246. Dominski, Z and Marzluff, WF, Formation of the 3′ end of histone mRNA. *Gene*, 1999. **239**(1): p. 1-14.

247. Phillips, SC and Turner, PC, A transcriptional analysis of the gene encoding mouse U7 small nuclear RNA. *Gene*, 1992. **116**(2): p. 181-186.

248. Southgate, C and Busslinger, M, In vivo and in vitro expression of U7 snRNA genes: cis- and trans-acting elements required for RNA polymerase II-directed transcription. *The European Molecular Biology Organization Journal, ,* 1989. **8**: p. 539-549.

249. Dominski, Z, Erkmann, J, Yang, X, SÃ nchez, R and Marzluff, W, A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3'-end processing. *Genes & Development*, 2002. **16**(1): p. 58-71.

250. Pillai, RS, Will, CL, Luhrmann, R, Schumperli, D and Muller, B, Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J*, 2001. **20**(19): p. 5470-5479.

251. Pillai, RS, Grimmler, M, Meister, G, Will, CL, Lührmann, R, Fischer, U and Schümperli, D, Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes & Development*, 2003. **17**(18): p. 2321-2333.

252. Schümperli, D and Pillai, RS, The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cellular and Molecular Life Sciences*, 2004. **61**(19): p. 2560-2570.

253. Kolev, NG and Steitz, JA, In vivo assembly of functional U7 snRNP requires RNA backbone flexibility within the Sm-binding site. *Nat Struct Mol Biol*, 2006. **13**(4): p. 347-353.

254. Grimm, C, Stefanovic, B and Schümperli, D, The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO journal*, 1993. **12**(3): p. 1229-1238.

255. Goyenvalle, A, Engineering U7snRNA gene to reframe transcripts. *Methods Mol Biol*, 2012. **867**: p. 259-271.

256. Brun, C, Suter, D, Pauli, C, Dunant, P, Lochmller, H, Burgunder, JM, Schümperli, D and Weis, J, U7 snRNAs induce correction of mutated dystrophin pre-mRNA by exon skipping. *Cellular and molecular life sciences*, 2003. **60**(3): p. 557-566.

257. Suter, D, Tomasini, R, Reber, U, Gorman, L, Kole, R and Schümperli, D, Double-target antisense U7 snRNAs promote efficient skipping of an aberrant exon in three human beta-thalassemic mutations. *Hum Mol Genet*, 1999. **8**(13): p. 2415-2423.

258. Madocsai, C, Lim, SR, Geib, T, Lam, BJ and Hertel, KJ, Correction of SMN2 Pre-mRNA Splicing by Antisense U7 Small Nuclear RNAs. *Mol Ther*, 2005. **12**(6): p. 1013-1022.

259. Meyer, K, Marquis, J, Trüb, J, Nlend Nlend, R, Verp, S, Ruepp, M-D, Imboden, H, Barde, I, Trono, D and Schümperli, D, Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation. *Human Molecular Genetics*, 2009. **18**(3): p. 546-555.

260. Marquis, J, Meyer, K, Angehrn, L, Kampfer, SS, Rothen-Rutishauser, B and Schumperli, D, Spinal Muscular Atrophy: SMN2 Pre-mRNA Splicing Corrected by a U7 snRNA Derivative Carrying a Splicing Enhancer Sequence. *Mol Ther*, 2007. **15**(8): p. 1479-1486.

261. Voigt, T, Meyer, K, Baum, O and Schümperli, D, Ultrastructural changes in diaphragm neuromuscular junctions in a severe mouse model for Spinal Muscular Atrophy and their prevention by bifunctional U7 snRNA correcting SMN2 splicing. *Neuromuscular Disorders*, 2010. **20**(11): p. 744-752.

262. Geib, T and Hertel, K, Restoration of full-length SMN promoted by adenoviral vectors expressing RNA antisense oligonucleotides embedded in U7 snRNAs. *PLoS ONE*, 2009. **4**(12): p. e8204-e8204.

263. Asparuhova, M, Marti, G, Liu, S, Serhan, F, Trono, D and Schümperli, D, Inhibition of HIV-1 multiplication by a modified U7 snRNA inducing Tat and Rev exon skipping. *The Journal of Gene Medicine*, 2007. **9**(5): p. 323-334.

264. Liu, S, Asparuhova, M, Brondani, V, Ziekau, I, Klimkait, T and Schümperli, D, Inhibition of HIV-1 multiplication by antisense U7 snRNAs and siRNAs targeting cyclophilin A. *Nucleic Acids Research*, 2004. **32**(12): p. 3752-3759.

265. Vacek, MM, Ma, H, Gemignani, F, Lacerra, G, Kafri, T and Kole, R, High-level expression of hemoglobin A in human thalassemic erythroid progenitor cells following lentiviral vector delivery of an antisense snRNA. *Blood*, 2003. **101**(1): p. 104-111.

266. Goyenvalle, A, Vulin, A, Fougerousse, F, Leturcq, F, Kaplan, J-C, Garcia, L and Danos, O, Rescue of Dystrophic Muscle Through U7 snRNA-Mediated Exon Skipping. *Science*, 2004. **306**(5702): p. 1796-1799.

267. Goyenvalle, A, Babbs, A, Wright, J, Wilkins, V, Powell, D, Garcia, L and Davies, K, Rescue of severely affected dystrophin/utrophin-deficient mice through scAAV-U7snRNA-mediated exon skipping. *Human Molecular Genetics*, 2012. **21**(11): p. 2559-2571.

268. Goyenvalle, A, Babbs, A, van Ommen, G-J, Garcia, L and Davies, K, Enhanced exon-skipping induced by U7 snRNA carrying a splicing silencer sequence: Promising tool for DMD therapy. *Molecular therapy*, 2009. **17**(7): p. 1234-1240.

269. Goyenvalle, A, Wright, J, Babbs, A, Wilkins, V, Garcia, L and Davies, K, Engineering multiple U7snRNA constructs to induce single and multiexon-skipping for Duchenne muscular dystrophy. *Molecular therapy*, 2012. **20**(6): p. 1212-1221.

270. Bish, LT, Sleeper, MM, Forbes, SC, Wang, B, Reynolds, C, Singletary, GE, Trafny, D, Morine, KJ, Sanmiguel, J, Cecchini, S, Virag, T, Vulin, A, Beley, C, Bogan, J, Wilson, JM, Vandenborne, K, Kornegay, JN, Walter, GA, Kotin, RM, Garcia, L and Sweeney, HL, Long-term Restoration of Cardiac Dystrophin Expression in Golden Retriever Muscular Dystrophy Following rAAV6-mediated Exon Skipping. *Mol Ther*, 2012. **20**(3): p. 580-589.

271. Barbash, IM, Cecchini, S, Faranesh, AZ, Virag, T, Li, L, Yang, Y, Hoyt, RF, Kornegay, JN, Bogan, JR, Garcia, L, Lederman, RJ and Kotin, RM, MRI roadmap-guided transendocardial delivery of exon-skipping recombinant adeno-associated virus restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Gene Ther*, 2012.

272. Voigt, T, Meyer, K, Baum, O and Schmperli, D, Ultrastructural changes in diaphragm neuromuscular junctions in a severe mouse model for Spinal Muscular Atrophy and their prevention by bifunctional U7 snRNA correcting SMN2 splicing. *Neuromuscular Disorders*, 2010. **20**(11): p. 744-752.

273. Wein, N, Avril, A, Bartoli, M, Beley, C, Chaouch, S, Laforêt, P, Behin, A, Butler Browne, G, Mouly, V, Krahn, M, Garcia, L and Lévy, N, Efficient bypass of mutations in dysferlin deficient patient cells by antisense-induced exon skipping. *Human Mutation*, 2010. **31**(2): p. 136-142.

274. Uchikawa, H, Fujii, K, Kohno, Y, Katsumata, N, Nagao, K, Yamada, M and Miyashita, T, U7 snRNA-mediated correction of aberrant splicing caused by activation of cryptic splice sites. *J Hum Genet*, 2007. **52**(11): p. 891-897.

275. Francois, V, Klein, A, Beley, C, Jollet, A, Lemercier, C, Garcia, L and Furling, D, Selective silencing of mutated mRNAs in DM1 by using modified hU7-snRNAs. *Nature structural & molecular biology*, 2011. **18**(1): p. 85-87.

276. Danos, O, AAV vectors for RNA-based modulation of gene expression. *Gene Ther*, 2008. **15**(11): p. 864-869.

277. Hacein-Bey-Abina, S, von Kalle, C, Schmidt, M, Le Deist, F, Wulffraat, N, McIntyre, E, Radford, I, Villeval, J-L, Fraser, CC, Cavazzana-Calvo, M and Fischer, A, A Serious Adverse Event after Successful Gene Therapy for X-Linked Severe Combined Immunodeficiency. *New England Journal of Medicine*, 2003. **348**(3): p. 255-256.

278. Hacein-Bey-Abina, S, Von Kalle, C, Schmidt, M, McCormack, MP, Wulffraat, N, Leboulch, P, Lim, A, Osborne, CS, Pawliuk, R, Morillon, E, Sorensen, R, Forster, A, Fraser, P, Cohen, JI, de Saint Basile, G, Alexander, I, Wintergerst, U, Frebourg, T, Aurias, A, Stoppa-Lyonnet, D, Romana, S, Radford-Weiss, I, Gross, F, Valensi, F, Delabesse, E, Macintyre, E, Sigaux, F, Soulier, J, Leiva, LE, Wissler, M, Prinz, C, Rabbitts, TH, Le Deist, F, Fischer, A and Cavazzana-Calvo, M, LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*, 2003. **302**(5644): p. 415-419.

279. Deichmann, A, Hacein-Bey-Abina, S, Schmidt, M, Garrigue, A, Brugman, MH, Hu, J, Glimm, H, Gyapay, G, Prum, B, Fraser, CC, Fischer, N, Schwarzwaelder, K, Siegler, ML, de Ridder, D, Pike-Overzet, K, Howe, SJ, Thrasher, AJ, Wagemaker, G, Abel, U, Staal, FJ, Delabesse, E, Villeval, JL, Aronow, B, Hue, C, Prinz, C, Wissler, M, Klanke, C, Weissenbach, J, Alexander, I, Fischer, A, von Kalle, C and Cavazzana-Calvo, M, Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy. *J Clin Invest*, 2007. **117**(8): p. 2225-2232.

280. Raper, SE, Chirmule, N, Lee, FS, Wivel, NA, Bagg, A, Gao, G-p, Wilson, JM and Batshaw, ML, Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Molecular Genetics and Metabolism*, 2003. **80**(1–2): p. 148-158.

281. King, JA, Dubielzig, R, Grimm, D and Kleinschmidt, JA, DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J*, 2001. **20**(12): p. 3282-3291.

282. Dubielzig, R, King, JA, Weger, S, Kern, A and Kleinschmidt, JA, Adeno-Associated Virus Type 2 Protein Interactions: Formation of Pre-Encapsidation Complexes. *Journal of Virology*, 1999. **73**(11): p. 8989-8998.

283. Surosky, RT, Urabe, M, Godwin, SG, McQuiston, SA, Kurtzman, GJ, Ozawa, K and Natsoulis, G, Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. *Journal of Virology*, 1997. **71**(10): p. 7951-7959.

284. Kronenberg, S, Kleinschmidt, JA and Bottcher, B, Electron cryo-microscopy and image reconstruction of adeno-associated virus type 2 empty capsids. *EMBO Rep*, 2001. **2**(11): p. 997-1002.

285. Trempe, JP and Carter, BJ, Alternate mRNA splicing is required for synthesis of adeno-associated virus VP1 capsid protein. *Journal of Virology*, 1988. **62**(9): p. 3356-3363.

286. Sonntag, F, Schmidt, K and Kleinschmidt, J, A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proceedings of the National Academy of Sciences of the United States of America*, 2010. **107**(22): p. 10220-10225.

287. Snyder, RO, Samulski, RJ and Muzyczka, N, In vitro resolution of covalently joined AAV chromosome ends. *Cell*, 1990. **60**(1): p. 105-113.

288. Snyder, RO, Im, DS, Ni, T, Xiao, X, Samulski, RJ and Muzyczka, N, Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. *Journal of Virology*, 1993. **67**(10): p. 6096-6104.

289. Henckaerts, E and Linden, RM, Adeno-associated virus: a key to the human genome? *Future virology*, 2010. **5**(5): p. 555-574.

290. Kotin, RM, Linden, RM and Berns, KI, Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *EMBO J*, 1992. **11**(13): p. 5071-5078.

291. Linden, RM, Ward, P, Giraud, C, Winocour, E and Berns, KI, Site-specific integration by adeno-associated virus. *Proceedings of the National Academy of Sciences*, 1996. **93**(21): p. 11288-11294.

292. Duan, D, Sharma, P, Yang, J, Yue, Y, Dudus, L, Zhang, Y, Fisher, KJ and Engelhardt, JF, Circular Intermediates of Recombinant Adeno-Associated Virus Have Defined Structural Characteristics Responsible for Long-Term Episomal Persistence in Muscle Tissue. *Journal of Virology*, 1998. **72**(11): p. 8568-8577.

293. Nakai, H, Yant, SR, Storm, TA, Fuess, S, Meuse, L and Kay, MA, Extrachromosomal Recombinant Adeno-Associated Virus Vector Genomes Are Primarily Responsible for Stable Liver Transduction In Vivo. *Journal of Virology*, 2001. **75**(15): p. 6969-6976.

294. Vincent-Lacaze, N, Snyder, RO, Gluzman, R, Bohl, D, Lagarde, C and Danos, O, Structure of adeno-associated virus vector DNA following transduction of the skeletal muscle. *Journal of Virology*, 1999. **73**(3): p. 1949-1955.

295. Weitzman, M and Linden, RM, Adeno-associated virus biology. *Methods in Molecular Biology*, 2011. **807**: p. 1-23.

296. Halbert, CL, Rutledge, EA, Allen, JM, Russell, DW and Miller, AD, Repeat Transduction in the Mouse Lung by Using Adeno-Associated Virus Vectors with Different Serotypes. *Journal of Virology*, 2000. **74**(3): p. 1524-1532.

297. Daya, S and Berns, KI, Gene Therapy Using Adeno-Associated Virus Vectors. *Clinical Microbiology Reviews*, 2008. **21**(4): p. 583-593.

298. Vigna, E and Naldini, L, Lentiviral vectors: excellent tools for experimental gene transfer and promising candidates for gene therapy. *The Journal of Gene Medicine*, 2000. **2**(5): p. 308-316.

299. Dull, T, Zufferey, R, Kelly, M, Mandel, RJ, Nguyen, M, Trono, D and Naldini, L, A Third-Generation Lentivirus Vector with a Conditional Packaging System. *Journal of Virology*, 1998. **72**(11): p. 8463-8471.

300. Zufferey, R, Dull, T, Mandel, RJ, Bukovsky, A, Quiroz, D, Naldini, L and Trono, D, Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery. *J Virol*, 1998. **72**(12): p. 9873-9880.

301. O'Brien, KF and Kunkel, LM, Dystrophin and Muscular Dystrophy: Past, Present, and Future. *Molecular Genetics and Metabolism*, 2001. **74**(1–2): p. 75-88.

302. Love, DR, C., BB, M., TJ, J., BD and E., DK, Dystrophin and dystrophin-related proteins: a review of protein and RNA studies. *Neuromuscular Disorders*, 1993. **3**(1): p. 5-21.

303. Klamut, HJ, Zubrzycka-Gaarn, EE, Bulman, DE, Malhotra, SB, Bodrug, SE, Worton, RG and Ray, PN, Myogenic regulation of dystrophin gene expression. *Br Med Bull*, 1989. **45**(3): p. 681-702.

304. Nowak, KJ and Davies, KE, Duchenne muscular dystrophy and dystrophin: pathogenesis and opportunities for treatment. *EMBO Rep*, 2004. **5**(9): p. 872-876.

305. Davison, MD and Critchley, DR, Alpha-actinins and the DMD protein contain spectrin-like repeats. *Cell*, 1988. **52**(2): p. 159-160.

306. Bonilla, E, Samitt, CE, Miranda, AF, Hays, AP, Salviati, G, DiMauro, S, Kunkel, LM, Hoffman, EP and Rowland, LP, Duchenne muscular dystrophy: Deficiency of dystrophin at the muscle cell surface. *Cell*, 1988. **54**(4): p. 447-452.

307. Lapidos, KA, Kakkar, R and McNally, EM, The Dystrophin Glycoprotein Complex. *Circulation Research*, 2004. **94**(8): p. 1023-1031.

308. Kunkel, LM and co, a, Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. *Nature*, 1986. **322**(6074): p. 73-77.

309. Muntoni, F, Torelli, S and Ferlini, A, Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *The Lancet Neurology*, 2003. **2**(12): p. 731-740.

310. Anthony, K, Cirak, S, Torelli, S, Tasca, G, Feng, L, Arechavala-Gomeza, V, Armaroli, A, Guglieri, M, Straathof, CS, Verschuuren, JJ, Aartsma-Rus, A, Helderman-van den Enden, P, Bushby, K, Straub, V, Sewry, C, Ferlini, A, Ricci, E, Morgan, JE and Muntoni, F, Dystrophin quantification and clinical correlations in Becker muscular dystrophy: implications for clinical trials. *Brain*, 2011. **134**(12): p. 3547-3559.

311. Sicinski, P, Geng, Y, Ryder Cook, AS, Barnard, EA, Darlison, MG and Barnard, PJ, The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. *Science*, 1989. **244**(4912): p. 1578-1580.

312. Grady, RM, Teng, H, Nichol, MC, Cunningham, JC, Wilkinson, RS and Sanes, JR, Skeletal and cardiac myopathies in mice lacking utrophin and dystrophin: a model for Duchenne muscular dystrophy. *Cell*, 1997. **90**(4): p. 729-738.

313. Deconinck, AE, Rafael, JA, Skinner, JA, Brown, SC, Potter, AC, Metzinger, L, Watt, DJ, Dickson, JG, Tinsley, JM and Davies, KE, Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. *Cell*, 1997. **90**(4): p. 717-727.

314. Bremmer-Bout, M, Aartsma-Rus, A, de Meijer, EJ, Kaman, WE, Janson, AAM, Vossen, RHAM, van Ommen, G-JB, den Dunnen, JT and van Deutekom, JCT, Targeted Exon Skipping in Transgenic hDMD Mice: A Model for Direct Preclinical Screening of Human-Specific Antisense Oligonucleotides. *Mol Ther*, 2004. **10**(2): p. 232-240.

315. Howell, JM, Fletcher, S, Kakulas, BA, O'Hara, M, Lochmuller, H and Karpati, G, Use of the dog model for Duchenne muscular dystrophy in gene therapy trials. *Neuromuscular Disorders*, 1997. **7**(5): p. 325-328.

316. England, SB, Nicholson, LVB, Johnson, MA, Forrest, SM, Love, DR, Zubrzycka-Gaarn, EE, Bulman, DE, Harris, JB and Davies, KE, Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature*, 1990. **343**(6254): p. 180-182.

317. Wang, B, Li, J and Xiao, X, Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model. *Proceedings of the National Academy of Sciences*, 2000. **97**(25): p. 13714-13719.

318. Gregorevic, P, Allen, JM, Minami, E, Blankinship, MJ, Haraguchi, M, Meuse, L, Finn, E, Adams, ME, Froehner, SC, Murry, CE and Chamberlain, JS, rAAV6-microdystrophin preserves muscle function and extends lifespan in severely dystrophic mice. *Nat Med*, 2006. **12**(7): p. 787-789.

319. Wang, B, Li, J, Fu, F and Xiao, X, Systemic human minidystrophin gene transfer improves functions and life span of dystrophin and dystrophin/utrophin-deficient mice. *Journal of Orthopaedic Research*, 2009. **27**(4): p. 421-426.

320. Wang, Z, Allen, J, Riddell, S, Gregorevic, P, Storb, R, Tapscott, S, Chamberlain, J and Kuhr, C, Immunity to adeno-associated virus-mediated gene transfer in a random-bred canine model of Duchenne muscular dystrophy. *Human gene therapy*, 2007. **18**(1): p. 18-26.

321. Wang, Z, Storb, R, Lee, D, Kushmerick, MJ, Chu, B, Berger, C, Arnett, A, Allen, J, Chamberlain, JS, Riddell, SR and Tapscott, SJ, Immune Responses to AAV in Canine Muscle Monitored by Cellular Assays and Noninvasive Imaging. *Mol Ther*, 2010. **18**(3): p. 617-624.

322. Mingozzi, F, Hasbrouck, NC, Basner-Tschakarjan, E, Edmonson, SA, Hui, DJ, Sabatino, DE, Zhou, S, Wright, JF, Jiang, H, Pierce, GF, Arruda, VR and High, KA, Modulation of tolerance to the transgene product in a nonhuman primate model of AAV-mediated gene transfer to liver. *Blood*, 2007. **110**(7): p. 2334-2341.

323. Mingozzi, F, Meulenberg, JJ, Hui, DJ, Basner-Tschakarjan, E, Hasbrouck, NC, Edmonson, SA, Hutnick, NA, Betts, MR, Kastelein, JJ, Stroes, ES and High, KA, AAV-1–mediated gene transfer to skeletal muscle in humans results in dose-dependent activation of capsid-specific T cells. *Blood*, 2009. **114**(10): p. 2077-2086.

324. Mingozzi, F and High, K, Immune responses to AAV in clinical trials. *Current gene therapy*, 2011. **11**(4): p. 321-330.

325. Wang, Z, Kuhr, C, Allen, J, Blankinship, M, Gregorevic, P, Chamberlain, J, Tapscott, S and Storb, R, Sustained AAV-mediated dystrophin expression in a canine model of Duchenne muscular dystrophy with a brief course of immunosuppression. *Molecular therapy*, 2007. **15**(6): p. 1160-1166.

326. Jiang, H, Couto, LB, Patarroyo-White, S, Liu, T, Nagy, D, Vargas, JA, Zhou, S, Scallan, CD, Sommer, J, Vijay, S, Mingozzi, F, High, KA and Pierce, GF, Effects of transient immunosuppression on adenoassociated, virus-mediated, liver-directed gene transfer in rhesus macaques and implications for human gene therapy. *Blood*, 2006. **108**(10): p. 3321-3328.

327. van Deutekom, JCT and van Ommen, G-J, Advances in Duchenne muscular dystrophy gene therapy. *Nature reviews Genetics*, 2003. **4**(10): p. 774-783.

328. Mendell, J, Campbell, K, Rodino Klapac, L, Sahenk, Z, Shilling, C, Lewis, S, Bowles, D, Gray, S, Li, C, Galloway, G, Malik, V, Coley, B, Clark, KR, Li, J, Xiao, X, Samulski, J, McPhee, S, Samulski, RJ and Walker, C, Dystrophin immunity in Duchenne's muscular dystrophy. *The New England journal of medicine*, 2010. **363**(15): p. 1429-1437.

329. Nicholson, LVB, The “rescue” of dystrophin synthesis in boys with Duchenne muscular dystrophy. *Neuromuscular Disorders*, 1993. **3**(5–6): p. 525-531.

330. Lu, QL, Morris, GE, Wilton, SD, Ly, T, Artem'yeva, OV, Strong, P and Partridge, TA, Massive Idiosyncratic Exon Skipping Corrects the Nonsense Mutation in Dystrophic Mouse Muscle and Produces Functional Revertant Fibers by Clonal Expansion. *The Journal of Cell Biology*, 2000. **148**(5): p. 985-996.

331. Thanh, LT, Nguyen, TM, Helliwell, TR and Morris, GE, Characterization of revertant muscle fibers in Duchenne muscular dystrophy, using exon-specific monoclonal antibodies against dystrophin. *American journal of human genetics*, 1995. **56**(3): p. 725-731.

332. Goyenvalle, A, Seto, JT, Davies, KE and Chamberlain, J, Therapeutic approaches to muscular dystrophy. *Human Molecular Genetics*, 2011. **20**(R1): p. R69-R78.

333. Ahmad, A, Brinson, M, Hodges, BL, Chamberlain, JS and Amalfitano, A, Mdx mice inducibly expressing dystrophin provide insights into the potential of gene therapy for Duchenne muscular dystrophy. *Human Molecular Genetics*, 2000. **9**(17): p. 2507-2515.

334. Lu, QL, Rabinowitz, A, Chen, YC, Yokota, T, Yin, H, Alter, J, Jadoon, A, Bou-Gharios, G and Partridge, T, Systemic delivery of antisense oligoribonucleotide restores dystrophin expression in body-wide skeletal muscles. *Proceedings of the National Academy of Sciences of the United States of America*, 2005. **102**(1): p. 198-203.

335. Fletcher, S, Honeyman, K, Fall, AM, Harding, PL, Johnsen, RD, Steinhaus, JP, Moulton, HM, Iversen, PL and Wilton, SD, Morpholino Oligomer-Mediated Exon Skipping Averts the Onset of Dystrophic Pathology in the mdx Mouse. *Mol Ther*, 2007. **15**(9): p. 1587-1592.

336. Fletcher, S, Honeyman, K, Fall, AM, Harding, PL, Johnsen, RD and Wilton, SD, Dystrophin expression in the *mdx* mouse after localised and systemic administration of a morpholino antisense oligonucleotide. *The Journal of Gene Medicine*, 2006. **8**(2): p. 207-216.

337. Yokota, T, Lu, Q-l, Partridge, T, Kobayashi, M, Nakamura, A, Takeda, S and Hoffman, E, Efficacy of systemic morpholino exon-skipping in duchenne dystrophy dogs. *Annals of Neurology*, 2009. **65**(6): p. 667-676.

338. Yokota, T, Hoffman, E and Takeda, Si, Antisense oligo-mediated multiple exon skipping in a dog model of duchenne muscular dystrophy. *Methods in Molecular Biology*, 2011. **709**: p. 299-312.

339. van Deutekom, JCT, Bremmer-Bout, M, Janson, AAM, Ginjaar, IB, Baas, F, den Dunnen, JT and van Ommen, G-JB, Antisense-induced exon skipping restores dystrophin expression in DMD patient derived muscle cells. *Hum Mol Genet*, 2001. **10**(15): p. 1547-1554.

340. Aartsma Rus, A, Janson, AAM, Heemskerk, JA, De Winter, CL, Van Ommen, GJB and Van Deutekom, JCT, Therapeutic modulation of DMD splicing by blocking exonic splicing enhancer sites with antisense oligonucleotides. *Annals of the New York Academy of Sciences*, 2006. **1082**: p. 74-76.

341. Aartsma Rus, A, Houlleberghs, H, van Deutekom, JCT, van Ommen, G-J and t Hoen, PAC, Exonic sequences provide better targets for antisense oligonucleotides than splice site sequences in the modulation of Duchenne muscular dystrophy splicing. *Oligonucleotides*, 2010. **20**(2): p. 69-77.

342. Aartsma Rus, A, Kaman, W, Weij, R, den Dunnen, J, van Ommen, G-J and van Deutekom, JCT, Exploring the frontiers of therapeutic exon skipping for Duchenne muscular dystrophy by double targeting within one or multiple exons. *Molecular therapy*, 2006. **14**(3): p. 401-407.

343. Adams, A, Harding, P, Iversen, P, Coleman, C, Fletcher, S and Wilton, S, Antisense oligonucleotide induced exon skipping and the dystrophin gene transcript: cocktails and chemistries. *BMC molecular biology*, 2007. **8**: p. 57-57.

344. Mitrpant, C, Fletcher, S and Wilton, S, Personalised genetic intervention for Duchenne muscular dystrophy: antisense oligomers and exon skipping. *Current Molecular Pharmacology*, 2009. **2**(1): p. 110-121.

345. Béroud, C, Tuffery Giraud, S, Matsuo, M, Hamroun, D, Humbertclaude, V, Monnier, N, Moizard, M-P, Voelckel, M-A, Calemard, L, Boisseau, P, Blayau, M, Philippe, C, Cosse, M, Pags, M, Rivier, F, Danos, O, Garcia, L and Claustres, M, Multiexon skipping leading to an artificial DMD protein lacking amino acids from exons 45 through 55 could rescue up to 63% of patients with Duchenne muscular dystrophy. *Human Mutation*, 2007. **28**(2): p. 196-202.

346. Ferreiro, V, Giliberto, F, Muñiz, GMN, Francipane, L, Marzese, DM, Mampel, A, Roqué, M, Frechtel, GD and Szijan, I, Asymptomatic Becker muscular dystrophy in a family with a multiexon deletion. *Muscle & Nerve*, 2009. **39**(2): p. 239-243.

347. van Vliet, L, de Winter, C, van Deutekom, JCT, van Ommen, G-J and Aartsma Rus, A, Assessment of the feasibility of exon 45-55 multiexon skipping for Duchenne muscular dystrophy. *BMC medical genetics*, 2008. **9**: p. 105-105.

348. Aartsma-Rus, A, Fokkema, I, Verschuuren, J, Ginjaar, I, van Deutekom, J, van Ommen, G-J and den Dunnen, JT, Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Human Mutation*, 2009. **30**(3): p. 293-299.

349. van Deutekom, JC, Janson, AA, Ginjaar, IB, Frankhuizen, WS, Aartsma-Rus, A, Bremmer-Bout, M, den Dunnen, JT, Koop, K, van der Kooi, AJ, Goemans, NM, de Kimpe, SJ, Ekhart, PF, Venneker, EH, Platenburg, GJ, Verschuuren, JJ and van Ommen, G-JB, Local Dystrophin Restoration with Antisense Oligonucleotide PRO051. *N Engl J Med*, 2007. **357**(26): p. 2677-2686.

350. Goemans, NM, Tulinius, M, van den Akker, JT, Burm, BE, Ekhart, PF, Heuvelmans, N, Holling, T, Janson, AA, Platenburg, GJ, Sipkens, JA, Sitsen, JMA, Aartsma-Rus, A, van Ommen, G-JB, Buyse, G, Darin, N, Verschuuren, JJ, Campion, GV, de Kimpe, SJ and van Deutekom, JC, Systemic Administration of PRO051 in Duchenne's Muscular Dystrophy. *New England Journal of Medicine*, 2011. **364**(16): p. 1513-1522.

351. Kinali, M, Arechavala-Gomeza, V, Feng, L, Cirak, S, Hunt, D, Adkin, C, Guglieri, M, Ashton, E, Abbs, S, Nihoyannopoulos, P, Garralda, ME, Rutherford, M, McCulley, C, Popplewell, L, Graham, IR, Dickson, G, Wood, MJA, Wells, DJ, Wilton, SD, Kole, R, Straub, V, Bushby, K, Sewry, C, Morgan, JE and Muntoni, F, Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *The Lancet Neurology*, 2009. **8**(10): p. 918-928.

352. Cirak, S, Arechavala Gomeza, V, Guglieri, M, Feng, L, Torelli, S, Anthony, K, Abbs, S, Garralda, M, Bourke, J, Wells, D, Dickson, G, Wood, MJA, Wilton, S, Straub, V, Kole, R, Shrewsbury, S, Sewry, C, Morgan, J, Bushby, K and Muntoni, F, Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *Lancet*, 2011. **378**(9791): p. 595-605.

353. Denti, MA, Rosa, A, D'Antona, G, Sthandier, O, Angelis, FGD, Nicoletti, C, Allocca, M, Pansarasa, O, Parente, V, Musarò, A, Auricchio, A, Bottinelli, R and Bozzoni, I, Chimeric Adeno-Associated Virus/Antisense U1 Small Nuclear RNA Effectively Rescues Dystrophin Synthesis and Muscle Function by Local Treatment of mdx Mice. *Human Gene Therapy*, 2006. **17**(5): p. 565-574.

354. Denti, MA, Incitti, T, Sthandier, O, Nicoletti, C, Angelis, FGD, Rizzuto, E, Auricchio, A, Musaro, A and Bozzoni, I, Long-Term Benefit of Adeno-Associated Virus/Antisense-Mediated Exon Skipping in Dystrophic Mice. *Human gene therapy*, 2008. **19**(6): p. 601-608.

355. Vaillend, C, Perronnet, C, Ros, C, Gruszczynski, C, Goyenvalle, A, Laroche, S, Danos, O, Garcia, L and Peltekian, E, Rescue of a Dystrophin-like Protein by Exon Skipping In Vivo Restores GABAA-receptor Clustering in the Hippocampus of the mdx Mouse. *Mol Ther*, 2010. **18**(9): p. 1683-1688.

356. Mitrpant, C, Adams, AM, Meloni, PL, Muntoni, F, Fletcher, S and Wilton, SD, Rational Design of Antisense Oligomers to Induce Dystrophin Exon Skipping. *Mol Ther*, 2009. **17**(8): p. 1418-1426.

357. Aartsma-Rus, A, Bremmer-Bout, M, Janson, AAM, den Dunnen, JT, van Ommen, G-JB and van Deutekom, JCT, Targeted exon skipping as a potential gene correction therapy for Duchenne muscular dystrophy. *Neuromuscular disorders : NMD*, 2002. **12**: p. S71-S77.

358. Wang, Z, Zhu, T, Qiao, C, Zhou, L, Wang, B, Zhang, J, Chen, C, Li, J and Xiao, X, Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. *Nat Biotech*, 2005. **23**(3): p. 321-328.

359. Inagaki, K, Fuess, S, Storm, T, Gibson, G, McTiernan, C, Kay, M and Nakai, H, Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Molecular therapy*, 2006. **14**(1): p. 45-53.

360. Tang, Y, Cummins, J, Huard, J and Wang, B, AAV-directed muscular dystrophy gene therapy. *Expert Opinion on Biological Therapy*, 2010. **10**(3): p. 395-408.

361. Eckenfelder, A, Tordo, J, Babbs, A, Davies, KE, Goyenvalle, A and Danos, O, The Cellular Processing Capacity Limits the Amounts of Chimeric U7 snRNA Available for Antisense Delivery. *Mol Ther Nucleic Acids*, 2012. **1**: p. e31.

362. Nudel, U, Robzyk, K and Yaffe, D, Expression of the putative Duchenne muscular dystrophy gene in differentiated myogenic cell cultures and in the brain. *Nature*, 1988. **331**(6157): p. 635-638.

363. Lev, AA, Feener, CC, Kunkel, LM and Brown, RH, Expression of the Duchenne's muscular dystrophy gene in cultured muscle cells. *Journal of Biological Chemistry*, 1987. **262**(33): p. 15817-15820.

364. Chelly, J, Didier, M, Christian, P, Yoheved, B-N, Jean-Claude, K and Axel, K, Quantitative estimation of minor mRNAs by cDNA-polymerase chain reaction. *European Journal of Biochemistry*, 1990. **187**(3): p. 691-698.

365. Paule, MR and White, RJ, Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Research*, 2000. **28**(6): p. 1283-1298.

366. Aerts, JL, Gonzales, MI and Topalian, SL, Selection of appropriate control genes to assess expression of tumor antigens using real-time RT-PCR. *BioTechniques*, 2004. **36**(1): p. 84-91.

367. Goidin, D, Mamessier, A, Staquet, M-J, Schmitt, D and Berthier-Vergnes, O, Ribosomal 18S RNA Prevails over Glyceraldehyde-3-Phosphate Dehydrogenase and β-Actin Genes as Internal Standard for Quantitative Comparison of mRNA Levels in Invasive and Noninvasive Human Melanoma Cell Subpopulations. *Analytical biochemistry*, 2001. **295**(1): p. 17-21.

368. Fisher, KJ, Gao, GP, Weitzman, MD, DeMatteo, R, Burda, JF and Wilson, JM, Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *Journal of Virology*, 1996. **70**(1): p. 520-532.

369. Nakai, H, Storm, TA and Kay, MA, Recruitment of Single-Stranded Recombinant Adeno-Associated Virus Vector Genomes and Intermolecular Recombination Are Responsible for Stable Transduction of Liver In Vivo. *Journal of Virology*, 2000. **74**(20): p. 9451-9463.

370. McCarty, DM, Monahan, PE and Samulski, RJ, Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Therapy*, 2001. **8**(16): p. 1248-1254.

371. Le Bec, C and Douar, AM, Gene Therapy Progress and Prospects - Vectorology: design and production of expression cassettes in AAV vectors. *Gene Ther*, 2006. **13**(10): p. 805-813.

372. McCarty, DM, Fu, H, Monahan, PE, Toulson, CE, Naik, P and Samulski, RJ, Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther*, 2003. **10**(26): p. 2112-2118.

373. Wang, Z, Ma, HI, Li, J, Sun, L, Zhang, J and Xiao, X, Rapid and highly efficient transduction by double-stranded adeno-associated virus vectors in vitro and in vivo. *Gene Ther*, 2003. **10**(26): p. 2105-2111.

374. McCarty, DM, Self-complementary AAV Vectors; Advances and Applications. *Mol Ther*, 2008. **16**(10): p. 1648-1656.

375. Zhong, S, Sun, S and Teng, B-B, The recombinant adeno-associated virus vector (rAAV2)-mediated apolipoprotein B mRNA-specific hammerhead ribozyme: a self-complementary AAV2 vector improves the gene expression. *Genetic Vaccines and Therapy*, 2004. **2**(1): p. 5-5.

376. Xu, D, McCarty, D, Fernandes, A, Fisher, M, Samulski, RJ and Juliano, RL, Delivery of MDR1 small interfering RNA by self-complementary recombinant adeno-associated virus vector. *Molecular therapy*, 2005. **11**(4): p. 523-530.

377. Nathwani, AC, Gray, JT, Ng, CYC, Zhou, J, Spence, Y, Waddington, SN, Tuddenham, EGD, Kemball-Cook, G, McIntosh, J, Boon-Spijker, M, Mertens, K and Davidoff, AM, Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. *Blood*, 2006. **107**(7): p. 2653-2661.

378. Nathwani, AC, Gray, JT, McIntosh, J, Ng, CYC, Zhou, J, Spence, Y, Cochrane, M, Gray, E, Tuddenham, EGD and Davidoff, AM, Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates. *Blood*, 2007. **109**(4): p. 1414-1421.

379. Fall, A, Johnsen, R, Honeyman, K, Iversen, P, Fletcher, S and Wilton, S, Induction of revertant fibres in the mdx mouse using antisense oligonucleotides. *Genetic Vaccines and Therapy*, 2006. **4**(1): p. 3.

380. Jaynes, JB, Johnson, JE, Buskin, JN, Gartside, CL and Hauschka, SD, The muscle creatine kinase gene is regulated by multiple upstream elements, including a muscle-specific enhancer. *Molecular and Cellular Biology*, 1988. **8**(1): p. 62-70.

381. Amacher, SL, Buskin, JN and Hauschka, SD, Multiple regulatory elements contribute differentially to muscle creatine kinase enhancer activity in skeletal and cardiac muscle. *Molecular and Cellular Biology*, 1993. **13**(5): p. 2753-2764.

382. Wang, B, Li, J, Fu, FH, Chen, C, Zhu, X, Zhou, L, Jiang, X and Xiao, X, Construction and analysis of compact muscle-specific promoters for AAV vectors. *Gene Ther*, 2008. **15**(22): p. 1489-1499.

383. Salva, MZ, Himeda, CL, Tai, PWL, Nishiuchi, E, Gregorevic, P, Allen, JM, Finn, EE, Nguyen, QG, Blankinship, MJ, Meuse, L, Chamberlain, JS and Hauschka, SD, Design of Tissue-specific Regulatory Cassettes for High-level rAAV-mediated Expression in Skeletal and Cardiac Muscle. *Mol Ther*, 2006. **15**(2): p. 320-329.

384. Molkentin, JD, Jobe, SM and Markham, BE, [alpha]-myosin Heavy Chain Gene Regulation: Delineation and Characterization of the Cardiac Muscle-specific Enhancer and Muscle-specific Promoter. *Journal of Molecular and Cellular Cardiology*, 1996. **28**(6): p. 1211-1225.

385. Black, BL, Molkentin, JD and Olson, EN, Multiple Roles for the MyoD Basic Region in Transmission of Transcriptional Activation Signals and Interaction with MEF2. *Molecular and Cellular Biology*, 1998. **18**(1): p. 69-77.

386. Royo, H and Cavaillé, J, Non-coding RNAs in imprinted gene clusters. *Biology of the Cell*, 2008. **100**(3): p. 149-166.

387. Burns, CM, Chu, H, Rueter, SM, Hutchinson, LK, Canton, H, Sanders-Bush, E and Emeson, RB, Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature*, 1997. **387**(6630): p. 303-308.

388. Bortolin-Cavaillé, M-L and Cavaillé, J, The SNORD115 (H/MBII-52) and SNORD116 (H/MBII-85) gene clusters at the imprinted Prader–Willi locus generate canonical box C/D snoRNAs. *Nucleic Acids Research*, 2012. **40**(14): p. 6800-6807.

389. Scott, MS, Ono, M, Yamada, K, Endo, A, Barton, GJ and Lamond, AI, Human box C/D snoRNA processing conservation across multiple cell types. *Nucleic Acids Research*, 2012. **40**(8): p. 3676-3688.

390. Ono, M, Yamada, K, Avolio, F, Scott, MS, van Koningsbruggen, S, Barton, GJ and Lamond, AI, Analysis of Human Small Nucleolar RNAs (snoRNA) and the Development of snoRNA Modulator of Gene Expression Vectors. *Molecular Biology of the Cell*, 2010. **21**(9): p. 1569-1584.

391. Qu, LH and Bachellerie, U24 a novel intron-encoded small nucleolar RNA with two 12 nt long phylogenetically conserved complementarities to 28S rRNA. *Nucleic Acids Research*, 1995. **23**(14): p. 2669-2676.

392. Krainer, AR, Maniatis, T, Ruskin, B and Green, MR, Normal and mutant human β-globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell*, 1984. **36**(4): p. 993-1005.

393. Ruskin, B, Krainer, AR, Maniatis, T and Green, MR, Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell*, 1984. **38**(1): p. 317-331.

394. Hall, KB, Green, MR and Redfield, AG, Structure of a pre-mRNA branch point/3' splice site region. *Proceedings of the National Academy of Sciences of the United States of America*, 1988. **85**(3): p. 704-708.

395. Ruskin, B, Greene, JM and Green, MR, Cryptic branch point activation allows accurate in vitro splicing of human β-globin intron mutants. *Cell*, 1985. **41**(3): p. 833-844.

396. Reed, R and Maniatis, T, Intron sequences involved in lariat formation during pre-mRNA splicing. *Cell*, 1985. **41**(1): p. 95-105.

397. Orkin, SH and Kazazian, HH, The Mutation and Polymorphism of the Human beta-Globin Gene and its Surrounding DNA. *Annual Review of Genetics*, 2003. **18**(1): p. 131-171.

398. Dominski, Z and Kole, R, Effects of exon sequences on splicing of model pre-mRNA substrates in vitro. *Acta Biochimica Polonica*, 1996. **43**(1): p. 161-173.

399. Jathoul, A, Law, E, Gandelman, O, Pule, M, Tisi, L and Murray, J, Development of a pH-Tolerant Thermostable Photinus pyralis Luciferase for Brighter In Vivo Imaging. *Bioluminescence - Recent Advances in Oceanic Measurements and Laboratory Applications*, 2012.

400. Foster, H, Popplewell, L and Dickson, G, Genetic therapeutic approaches for duchenne muscular dystrophy. *Human gene therapy*, 2012. **23**(7): p. 676-687.

401. Berg, Michael G, Singh, Larry N, Younis, I, Liu, Q, Pinto, Anna M, Kaida, D, Zhang, Z, Cho, S, Sherrill-Mix, S, Wan, L and Dreyfuss, G, U1 snRNP Determines mRNA Length and Regulates Isoform Expression. *Cell*, 2012. **150**(1): p. 53-64.

402. Fagone, P, Wright, JF, Nathwani, A, Nienhuis, A, Davidoff, A and Gray, J, Systemic errors in quantitative polymerase chain reaction titration of self-complementary adeno-associated viral vectors and improved alternative methods. *Human gene therapy Part B Methods*, 2012. **23**(1): p. 1-7.

403. Bartlett, JS, Sethna, M, Ramamurthy, L, Gowen, SA, Samulski, RJ and Marzluff, WF, Efficient expression of protein coding genes from the murine U1 small nuclear RNA promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 1996. **93**(17): p. 8852-8857.

404. Bojak, A, Hammer, D, Wolf, H and Wagner, R, Muscle specific versus ubiquitous expression of Gag based HIV-1 DNA vaccines: a comparative analysis. *Vaccine*, 2002. **20**(15): p. 1975-1979.

405. Stojanovic, M and Kolpashchikov, D, Modular aptameric sensors. *Journal of the American Chemical Society*, 2004. **126**(30): p. 9266-9270.

406. Famulok, M, Chemical biology: Green fluorescent RNA. *Nature*, 2004. **430**(7003): p. 976-977.

407. Babendure, JR, Adams, SR and Tsien, RY, Aptamers Switch on Fluorescence of Triphenylmethane Dyes. *Journal of the American Chemical Society*, 2003. **125**(48): p. 14716-14717.

408. Samarsky, DA, Ferbeyre, G, Bertrand, E, Singer, RH, Cedergren, R and Fournier, MJ, A small nucleolar RNA:ribozyme hybrid cleaves a nucleolar RNA target in vivo with near-perfect efficiency. *Proceedings of the National Academy of Sciences*, 1999. **96**(12): p. 6609-6614.

409. Ono, M, Scott, MS, Yamada, K, Avolio, F, Barton, GJ and Lamond, AI, Identification of human miRNA precursors that resemble box C/D snoRNAs. *Nucleic Acids Research*, 2011. **39**(9): p. 3879-3891.

410. Brameier, M, Herwig, A, Reinhardt, R, Walter, L and Gruber, J, Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Research*, 2011. **39**(2): p. 675-686.

411. Ender, C, Krek, A, Friedländer, MR, Beitzinger, M, Weinmann, L, Chen, W, Pfeffer, S, Rajewsky, N and Meister, G, A Human snoRNA with MicroRNA-Like Functions. 2008. **32**(4): p. 519-528.

412. Taft, R, Glazov, E, Lassmann, T, Hayashizaki, Y, Carninci, P and Mattick, J, Small RNAs derived from snoRNAs. *RNA*, 2009. **15**(7): p. 1233-1240.

413. Darzacq, X and Kiss, T, Processing of Intron-Encoded Box C/D Small Nucleolar RNAs Lacking a 5',3'-Terminal Stem Structure. *Mol Cell Biol*, 2000. **20**(13): p. 4522-4531.

414. Semenov, DV, Vratskih, OV, Kuligina, EV and Richter, VA, Splicing by Exon Exclusion Impaired by Artificial Box C/D RNA Targeted to Branch-Point Adenosine. *Annals of the New York Academy of Sciences*, 2008. **1137**(Circulating Nucleic Acids in Plasma and Serum V): p. 119-124.

415. Stepanov, GA, Semenov, DV, Kuligina, EV, Koval, OA, Rabinov, IV, Kit, YY and Richter, VA, Analogues of Artificial Human Box C/D Small Nucleolar RNA As Regulators of Alternative Splicing of a pre-mRNA Target. *Acta Naturae*, 2012. **4**(1): p. 32-41.

416. Ganot, P, Bortolin, M-L and Kiss, T, Site-Specific Pseudouridine Formation in Preribosomal RNA Is Guided by Small Nucleolar RNAs. *Cell*, 1997. **89**(5): p. 799-809.

417. Karijolich, J and Yu, Y-T, Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature*, 2011. **474**(7351): p. 395-398.

418. Huang, C, Wu, G and Yu, Y-T, Inducing nonsense suppression by targeted pseudouridylation. *Nat Protocols*, 2012. **7**(4): p. 789-800.

419. Wee, KB, Pramono, ZAD, Wang, JL, MacDorman, KF, Lai, PS and Yee, WC, Dynamics of *Co-Transcriptional* Pre-mRNA Folding Influences the Induction of Dystrophin Exon Skipping by Antisense Oligonucleotides. *PLoS ONE*, 2008. **3**(3): p. e1844.

420. Heemskerk, H, de Winter, C, van Kuik, P, Heuvelmans, N, Sabatelli, P, Rimessi, P, Braghetta, P, van Ommen, G-J, de Kimpe, S, Ferlini, A, Aartsma Rus, A and van Deutekom, JCT, Preclinical PK and PD studies on 2'-O-methyl-phosphorothioate RNA antisense oligonucleotides in the mdx mouse model. *Molecular therapy*, 2010. **18**(6): p. 1210-1217.

421. Heemskerk, HA, de Winter, CL, de Kimpe, SJ, van Kuik-Romeijn, P, Heuvelmans, N, Platenburg, GJ, van Ommen, G-JB, van Deutekom, JCT and Aartsma-Rus, A, In vivo comparison of 2'O-methyl phosphorothioate and morpholino antisense oligonucleotides for Duchenne muscular dystrophy exon skipping. *The Journal of Gene Medicine*, 2009. **11**(3): p. 257-266.

422. Goyenvalle, A, Babbs, A, Wright, J, Wilkins, V, Powell, D, Garcia, L and Davies, KE, Rescue of severely affected Dystrophin/Utrophin deficient mice through scAAV-U7snRNA mediated exon skipping. *Human Molecular Genetics*, 2012.

423. Gebski, BL, Mann, CJ, Fletcher, S and Wilton, SD, Morpholino antisense oligonucleotide induced dystrophin exon 23 skipping in mdx mouse muscle. *Hum Mol Genet*, 2003. **12**(15): p. 1801-1811.

424. Goyenvalle, A, Babbs, A, Powell, D, Kole, R, Fletcher, S, Wilton, SD and Davies, KE, Prevention of Dystrophic Pathology in Severely Affected Dystrophin/Utrophin-deficient Mice by Morpholino-oligomer-mediated Exon-skipping. *Mol Ther*, 2010. **18**(1): p. 198-205.

425. Malerba, A, Boldrin, L and Dickson, G, Long-term systemic administration of unconjugated morpholino oligomers for therapeutic expression of dystrophin by exon skipping in skeletal muscle: implications for cardiac muscle integrity. *Nucleic Acid Therapeutics*, 2011. **21**(4): p. 293-298.

426. Yin, H, Saleh, A, Betts, C, Camelliti, P, Seow, Y, Ashraf, S, Arzumanov, A, Hammond, S, Merritt, T, Gait, M and Wood, M, Pip5 transduction peptides direct high efficiency oligonucleotide-mediated dystrophin exon skipping in heart and phenotypic correction in mdx mice. *Molecular therapy*, 2011. **19**(7): p. 1295-1303.

427. Moulton, H and Moulton, J, Morpholinos and their peptide conjugates: therapeutic promise and challenge for Duchenne muscular dystrophy. *Biochimica et biophysica acta*, 2010. **1798**(12): p. 2296-2303.

428. Zincarelli, C, Soltys, S, Rengo, G and Rabinowitz, J, Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Molecular therapy*, 2008. **16**(6): p. 1073-1080.

429. Nathwani, A, Tuddenham, EGD, Rangarajan, S, Rosales, C, McIntosh, J, Linch, D, Chowdary, P, Riddell, A, Pie, A, Harrington, C, O'Beirne, J, Smith, K, Pasi, J, Glader, B, Rustagi, P, Ng, CYC, Kay, M, Zhou, J, Spence, Y, Morton, C, Allay, J, Coleman, J, Sleep, S, Cunningham, J, Srivastava, D, Basner Tschakarjan, E, Mingozzi, F, High, K, Gray, J, Reiss, U, Nienhuis, A and Davidoff, A, Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *The New England journal of medicine*, 2011. **365**(25): p. 2357-2365.

430. Bowles, D, McPhee, SWJ, Li, C, Gray, S, Samulski, J, Camp, A, Li, J, Wang, B, Monahan, P, Rabinowitz, J, Grieger, J, Govindasamy, L, Agbandje McKenna, M, Xiao, X and Samulski, RJ, Phase 1 gene therapy for Duchenne muscular dystrophy using a translational optimized AAV vector. *Molecular therapy*, 2012. **20**(2): p. 443-455.

431. t Hoen, PAC, van der Wees, CGC, Aartsma-Rus, A, Turk, R, Goyenvalle, A, Danos, O, Garcia, L, van Ommen, G-JB, den Dunnen, JT and van Deutekom, JCT, Gene expression profiling to monitor therapeutic and adverse effects of antisense therapies for Duchenne muscular dystrophy. *Pharmacogenomics*, 2006. **7**(3): p. 281-297.

432. Aoki, Y, Yokota, T, Nagata, T, Nakamura, A, Tanihata, J, Saito, T, Duguez, SMR, Nagaraju, K, Hoffman, EP, Partridge, T and Takeda, Si, Bodywide skipping of exons 45–55 in dystrophic mdx52 mice by systemic antisense delivery. *Proceedings of the National Academy of Sciences*, 2012.

433. Cacchiarelli, D, Incitti, T, Martone, J, Cesana, M, Cazzella, V, Santini, T, Sthandier, O and Bozzoni, I, miR-31 modulates dystrophin expression: new implications for Duchenne muscular dystrophy therapy. *EMBO Rep*, 2011. **12**(2): p. 136-141.

434. Qiao, C, Li, J, Jiang, J, Zhu, X, Wang, B and Xiao, X, Myostatin propeptide gene delivery by adeno-associated virus serotype 8 vectors enhances muscle growth and ameliorates dystrophic phenotypes in mdx mice. *Human gene therapy*, 2008. **19**(3): p. 241-254.

435. Kang, J, Malerba, A, Popplewell, L, Foster, K and Dickson, G, Antisense-induced myostatin exon skipping leads to muscle hypertrophy in mice following octa-guanidine morpholino oligomer treatment. *Molecular therapy*, 2011. **19**(1): p. 159-164.

436. Wang, L, Zhou, L, Jiang, P, Lu, L, Chen, X, Lan, H, Guttridge, D, Sun, H and Wang, H, Loss of miR-29 in myoblasts contributes to dystrophic muscle pathogenesis. *Molecular therapy*, 2012. **20**(6): p. 1222-1233.

437. Kemaladewi, DU, Hoogaars, WM, van Heiningen, SH, Terlouw, S, de Gorter, DJ, den Dunnen, JT, van Ommen, GJ, Aartsma-Rus, A, ten Dijke, P and t Hoen, PA, Dual exon skipping in myostatin and dystrophin for Duchenne muscular dystrophy. *BMC Med Genomics*, 2011. **4**: p. 36.

438. Hoogaars, W, Mouisel, E, Pasternack, A, Hulmi, JJ, Relizani, K, Schuelke, M, Schirwis, E, Garcia, L, Ritvos, O, Ferry, A, t Hoen, PA and Amthor, H, Combined effect of AAV-U7-induced dystrophin exon skipping and soluble activin type IIB receptor in mdx mice. *Hum Gene Ther*, 2012.

439. VandenDriessche, T, Naldini, L, Collen, D and Chuah, MKL, *Oncoretroviral and lentiviral vector-mediated gene therapy*, in *Methods in Enzymology*, MI Phillips, Editor. 2002, Academic Press. p. 573-589.

440. Allay, J, Sleep, S, Long, S, Tillman, D, Clark, R, Carney, G, Fagone, P, McIntosh, J, Nienhuis, A, Davidoff, A, Nathwani, A and Gray, J (2011) Good manufacturing practice production of self-complementary serotype 8 adeno-associated viral vector for a hemophilia B clinical trial. *Hum Gene Ther* **22**, 595-604.

441. Allocca, M, Doria, M, Petrillo, M, Colella, P, Garcia Hoyos, M, Gibbs, D, Kim, S, Maguire, A, Rex, T, Di Vicino, U, Cutillo, L, Sparrow, J, Williams, D, Bennett, J and Auricchio, A (2008) Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J Clin Invest* **118**, 1955-1964.

442. Herson, S, Hentati, F, Rigolet, A, Behin, A, Romero, NB, Leturcq, F, Laforêt, P, Maisonobe, T, Amouri, R, Haddad, H, Audit, M, Montus, M, Masurier, C, Gjata, B, Georger, C, Cheraï, M, Carlier, P, Hogrel, J-Y, Herson, A, Allenbach, Y, Lemoine, FM, Klatzmann, D, Sweeney, HL, Mulligan, RC, Eymard, B, Caizergues, D, Voït, T and Benveniste, O (2012) A phase I trial of adeno-associated virus serotype 1-γ-sarcoglycan gene therapy for limb girdle muscular dystrophy type 2C. *Brain* **135**, 483-492.

443. Naldini, L, Blömer, U, Gallay, P, Ory, D, Mulligan, R, Gage, FH, Verma, IM and Trono, D, In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. *Science*, 1996. **272**(5259): p. 263-267.

444. Holland, PM, Abramson, RD, Watson, R and Gelfand, DH, Detection of specific polymerase chain reaction product by utilizing the 5'- 3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proceedings of the National Academy of Sciences*, 1991. **88**(16): p. 7276-7280.

445. Bustin, S, Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology*, 2000. **25**(2): p. 169-193.

446. Kutyavin, IV, Afonina, IA, Mills, A, Gorn, VV, Lukhtanov, EA, Belousov, ES, Singer, MJ, Walburger, DK, Lokhov, SG, Gall, AA, Dempcy, R, Reed, MW, Meyer, RB and Hedgpeth, J, 3'-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. *Nucleic Acids Research*, 2000. **28**(2): p. 655-661.

ANNEXES

Citation: Molecular Therapy–Nucleic Acids (2012) 1, e31; doi:10.1038/mtna.2012.24

www.nature.com/mtna

The Cellular Processing Capacity Limits the Amounts of Chimeric U7 snRNA Available for Antisense Delivery

Agathe Eckenfelder[1], Julie Tordo[1,2], Arran Babbs[3], Kay E Davies[3], Aurélie Goyenvalle[3,4] and Olivier Danos[1,2,4]

Many genetic diseases are induced by mutations disturbing the maturation of pre-mRNAs, often affecting splicing. Antisense oligoribonucleotides (AONs) have been used to modulate splicing thereby circumventing the deleterious effects of mutations. Stable delivery of antisense sequences is achieved by linking them to small nuclear RNA (snRNAs) delivered by viral vectors, as illustrated by studies where therapeutic exon skipping was obtained in animal models of Duchenne muscular dystrophy (*DMD*). Yet, clinical translation of these approaches is limited by the amounts of vector to be administered. In this respect, maximizing the amount of snRNA antisense shuttle delivered by the vector is essential. Here, we have used a muscle- and heart-specific enhancer (MHCK) to drive the expression of U7 snRNA shuttles carrying antisense sequences against the human or murine DMD pre-mRNAs. Although antisense delivery and subsequent exon skipping were improved both in tissue culture and in vivo, we observed the formation of additional U7 snRNA by-products following gene transfer. These included aberrantly 3′ processed as well as unprocessed species that may arise because of the saturation of the cellular processing capacity. Future efforts to increase the amounts of functional U7 shuttles delivered into a cell will have to take this limitation into account.

Molecular Therapy–Nucleic Acids (2012) **1**, e31; doi:10.1038/mtna.2012.24; published online 26 June 2012

INTRODUCTION

A variety of mRNAs with different exon composition, stability or subcellular localization, as well as small RNAs with regulatory roles can be produced from a single transcription unit, thus increasing the complexity of genome information.[1] A large proportion of disease-associated mutations in humans affect the processing of pre-mRNAs[2] often by disturbing the delicate balance of alternative splicing, which is observed in over 90% of transcription units.[3] The modulation of pre-mRNA splicing has been proposed as an approach to compensate the deleterious effect of mutations.[4] It can be achieved with a high-specificity, using antisense oligoribonucleotides (AON) that mask key splicing signals through Watson-Crick pairing with the pre-mRNA. The successful AON-mediated restoration of correct splicing has been reported in a number of cellular and animal models of genetic diseases.[5] In particular cases, AONs can be exploited to induce exon skipping in such a way that it removes deleterious mutations or re-establish a reading frame disrupted by a deletion. This approach is actively pursued in clinical trials involving patients with Duchenne muscular dystrophy (DMD).[6–9]

A limitation of AON-based therapies for treating chronic genetic diseases such as DMD is the need for repeated administration of high doses of oligonucleotides (up to several mg/kg). A sustained intracellular production of antisense sequences can be obtained using modified small nuclear (sn) RNAs such as U1, U2, or U7 delivered by viral vectors.[10–12] We and others have demonstrated the efficacy of chimeric snRNAs for sustained therapeutic exon skipping in animal models of DMD, following local or systemic gene delivery with AAV vectors targeting skeletal muscles, heart, and brain.[13–16] Yet, clinical applications to DMD patients by systemic administration of AAV vectors carrying a modified snRNA will require very high doses that are potentially toxic and represent a challenge for the current capacity of vector manufacturing. Reducing the injected dose is paramount to the success of current translational studies that aim at implementing vector mediated exon skipping into the clinic. It may be achieved in part by using the appropriate AAV serotype for muscle and heart delivery,[14] but also by optimizing the amount of chimeric snRNAs transcribed and assembled into active snRNPs.

The Sm-class snRNA genes are transcribed by RNA polymerase II in association with a specialized complex called Integrator. The polymerase initiate transcription on short bipartite promoters composed of proximal and distal sequence elements (PSE and DSE, respectively), while the Integrator complex recognizes signals in the internal stem-loop structures of the coding sequence, as well as in a conserved 3′ box and mediates co-transcriptional 3′ processing.[17] The snRNAs are then exported to the cytoplasm where they undergo further modifications and assemble with Sm proteins before re-entry into the nucleus. Proper assembly requires a consensus Sm binding site.[18] U7 is a non-spliceosomal snRNA that functions as a processing factor of the non-polyadenylated replication dependent histone mRNA. It is the shortest and simplest of

The first two authors contributed equally to this work.

[1]Inserm U845, Hôpital Necker-Enfants Malades, Université Paris Descartes, Paris, France; [2]Cancer Institute, University College London, London, UK; [3]MRC Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, UK; [4]Biothérapies des Maladies Neuromusculaires Unité Mixte: Um76 UPMC—UMR 7215 CNRS—U974 Inserm—Institut de Myologie—Faculté de Médecine Pierre et Marie Curie, Paris, France. Correspondence: Olivier Danos UCL Gene Therapy Consortium, UCL Cancer Institute, Paul O'Gorman Building, University College London, 72 Huntley Street, London WC1E 6BT, UK. E-mail: o.danos@ucl.ac.uk

Keywords: antisense; dystrophin; exon skipping; U7snRNA

Received 2 April 2012; accepted 21 May 2012; advance online publication 26 June 2012. doi:10.1038/mtna.2012.24

snRNAs and it includes a 5′ antisense sequence recognizing the histone pre-mRNA, a non-consensus Sm-binding site and a stem-loop structure containing determinants for the co-transcriptional 3′ endonucleolytic cleavage. The low abundance of the U7 small nuclear ribonucleoproteins (snRNP) is reversed when a consensus spliceosomal Sm sequence is introduced (SmOPT) and substitutions of the original 5′ sequences result in abundant and stable shuttles for ectopic antisense delivery.[10,19]

Here, we have added a strong muscle- and heart-specific transcriptional enhancer upstream of the DSE in chimeric *U7* genes carrying antisense sequences for exon skipping on the human or murine *DMD* pre-mRNAs. We have observed that although the amount of transcripts produced by the U7 promoter is readily enhanced, many of them are not processed and may not be assembled in snRNPs with exon-skipping activity on the *DMD* gene pre-mRNA. In addition, the SmOPT sequence required for stabilizing the U7 transcript induces the generation of 3′-processed species missing 20 nucleotides (nt), whose stability or activity may be affected. We conclude that the potential for enhancing exon-skipping activity of chimeric U7 cassettes is limited by the snRNA processing capacity of the host cell.

RESULTS

Design and validation of an enhanced U7 cassette

Brun *et al*. have shown that transcription from the U7 promoter can be increased by placing an enhancer upstream of the Distal Sequence Element.[20] Based on this observation, we have constructed a U7 cassette in which a strong muscle-specific transcriptional enhancer was inserted (**Figure 1a**). Salva *et al*. have designed such an enhancer by fusing elements from the α-myosin heavy chain (α-MHC) and the muscle creatin kinase (MCK) enhancers, and shown that, in combination with the CK promoter it results in a very strong and mostly tissue-specific enhancement of gene expression when used in AAV vectors.[21] We have assembled a similar enhancer [muscle- and heart-specific enhancer (MHCK), see Materials and Methods] and its activity was first validated in murine C2C12 myoblasts using an AAV vector carrying green fluorescent protein (GFP) under the control of the human PGK promoter (**Supplementary Figure S1**). The MHCK enhancer was then introduced upstream of a U7snRNA cassette containing an antisense sequence able to induce skipping of exon 51 on the human *DMD* pre-mRNA by masking an exonic splicing enhancer (U7ex51)[22] (**Figure 1a**). Lentiviral vectors carrying U7ex51 with or without the enhancer were constructed and human CHQ myoblasts were transduced with increasing multiplicities of infection (MOI, see Materials and Methods). Three days post-transduction, cells were differentiated into myotubes for 3 days and exon-skipping efficiency was analyzed by reverse transcription (RT)-PCR. The efficiency of skipping was evaluated by estimating the relative intensities of the skipped versus nonskipped diagnostic bands appearing after electrophoresis of RT-PCR products on agarose gels (**Figure 1b**). The results indicate a dose-dependent exon skipping for both vectors, with a higher activity for the one containing the MHCK enhancer, detectable at lower doses.

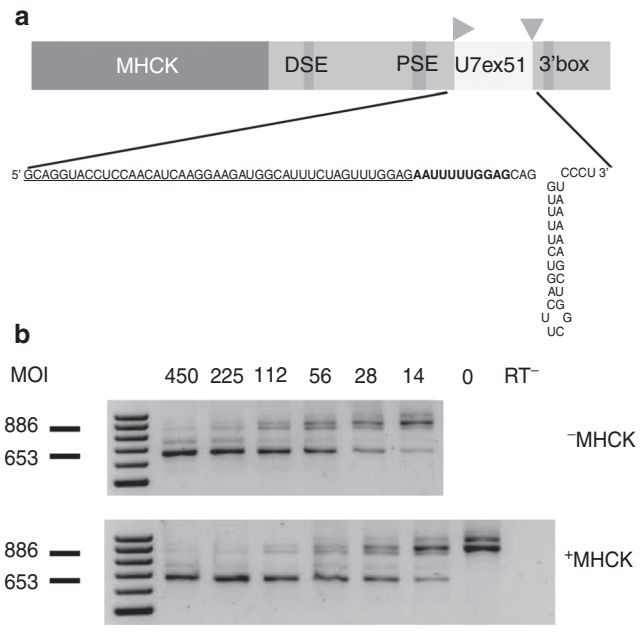

Figure 1 Muscle- and heart-specific enhancer (MHCK) enhances U7ex51-mediated exon skipping on the Duchenne muscular dystrophy (*DMD*) pre-mRNA in cultured human myotubes. (**a**) U7ex51 expression is driven by the U7 promoter containing distal and proximal sequence elements (DSE and PSE, respectively, shaded boxes). The MHCK enhancer[21] has been placed upstream. The rightward arrowhead indicates the site of transcription initiation and the downward arrowhead is the site of 3′ processing, activated by recognition of the 3′box (shaded). The sequence of the U7ex51 transcript is shown, including an antisense sequence (underlined) that corresponds to an exonic splicing enhancer (ESE) in exon 51 of the human *DMD* pre-mRNA (h51AON1: 5′-UCAAGGAAGAUGGCAUUUCU-3′).[22] The optimized Sm binding sequence (SmOPT) is shown in bold. (**b**) Reverse transcription and nested PCR detection of *DMD* exon 51 skipping after transduction of human CHQ cells with lentiviral vectors containing the U7ex51 (upper panel) or the MHCK-U7ex51 (lower panel) cassettes, at increasing multiplicities of infection (MOI) Controls include nontransduced cells (0) and no reverse transcriptase (RT–). The native human *DMD* mRNA containing exon 51 is detected as a 886-bp fragment and the skipped mRNA yields a 653-bp product.

In order to test the efficiency of the MHCK enhancer in the skeletal muscle of mdx mice, we introduced it into an AAV vector containing the U7DTex23 cassette that carries antisense sequences able to induce skipping of exon 23 in the murine *Dmd* pre-mRNA[13] (**Figure 2a**). AAV 2/5 vectors were produced with or without enhancer and used to transduce murine C2C12 myoblasts (see Materials and Methods). The U7DTex23 transcripts were analyzed by northern blot on total RNA isolated from transduced cells and normalized to the quantity of U2 snRNA (**Figure 2b**). Unexpectedly, two small RNA species reacted with a probe in the *Dmd* antisense region (**Figure 2a** and Materials and Methods). The predicted 86 nt U7DTex23 transcript was present along with equivalent quantities of a shorter species of ~66 nt. Altogether, the amounts of chimeric U7 RNA were enhanced in the presence of the MHCK sequences. This effect was clearly visible at the lower vector dose, presumably because nonsaturating levels of transcripts are made under these conditions. Of note, high-molecular weight RNA recognized by

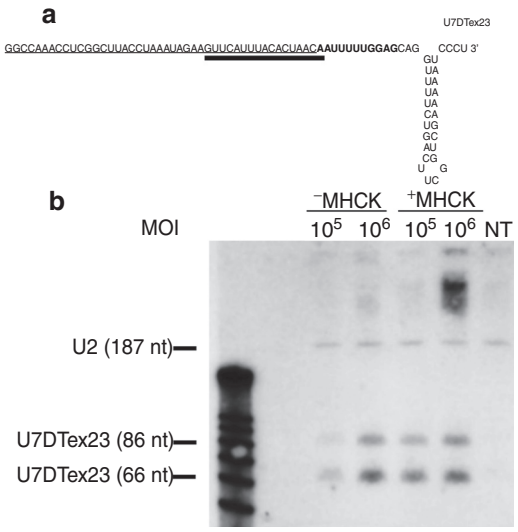

Figure 2 Enhancement of U7DTex23 small nuclear RNAs (snRNAs) expression in murine C2C12 cells. (**a**) The sequence of the U7DTex23 transcript is shown, including two antisense sequences (underlined) that encompass the donor splice site of exon 23 (SD23, shown in italics) and the branch point region of intron 22 (BP22) on the murine Duchenne muscular dystrophy (*Dmd*) pre-mRNA.[13] The optimized Sm-binding sequence (SmOPT) is shown in bold. The probe used for the northern blot analysis in (**b**) is underlined in bold. (**b**) Northern blot analysis of U7DTex23 transcripts expressed in C2C12 cells transduced with AAV2/5-U7DTex23 (–MHCK) or AAV2/5-MHCK-U7DTex23 (+MHCK) at different multiplicities of infection (MOI) (10^5 or 10^6 viral genome per cell). NT, nontransduced cells. The blot was hybridized with two 32P end-labeled probes recognizing the U2 snRNA and the U7DTex23 chimera (underlined in bold in (a)), respectively. U2 snRNA shows up as a 187-nt band. Two bands of 86 nt and ~66-nt react with the U7DTex23 probe.

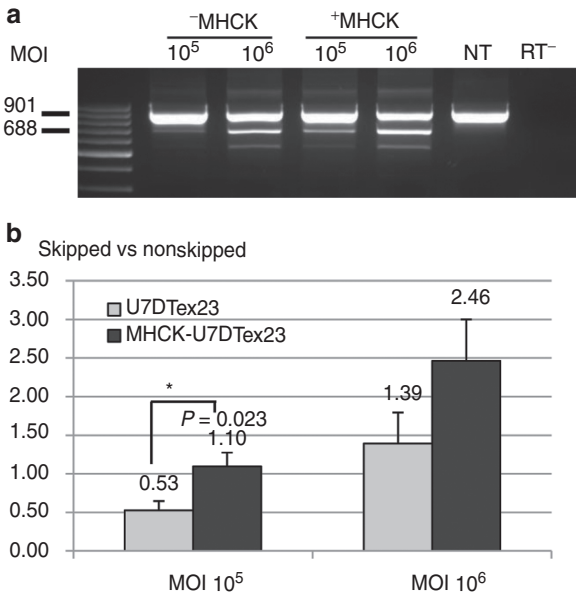

Figure 3 U7-mediated exon 23 skipping on the Duchenne muscular dystrophy (*Dmd*) pre-mRNA. (**a**) Reverse transcription and nested PCR detection of *Dmd* exon 23 skipping after transduction of C2C12 cells with AAV2/5-U7DTex23 (–MHCK) or AAV2/5-MHCK-U7DTex23 (+MHCK) with different multiplicities of infection (MOI) (10^5 or 10^6 viral genome per cell). Controls include nontransduced cells (NT) and no reverse transcriptase (RT–). The native *Dmd* mRNA containing exon 23 is detected as a 901-bp fragment and the skipped mRNA yields a 688-bp product. Results are representative of at least three independent transductions. (**b**) Analysis of mouse exon 23 skipping by quantitative PCR. The graph represents the ratio between the exon 22–24 junction (skipped mRNA species) and the exon 22–23 junction (full-length mRNA species). Quantification was performed on mRNA extract from C2C12 cells transduced with AAV2/5-U7DTex23 or AAV2/5-MHCK-U7DTex23 at the indicated MOI. A significant difference is observed at MOI 10^5 as shown by the *P* value ($P = 0.023$), according to Student's t-test. Error bars are shown as mean ± SEM ($n = 6$). MHCK, muscle- and heart-specific enhancer.

the probe accumulated at the high dose and in the presence of the enhancer. Further analysis indicated that it contained sequences 3′ of the U7-coding sequence and therefore represented unprocessed primary U7 RNA, possibly transcribed from tandem or circle copies of the AAV vector genome[23] (**Supplementary Figure S2**). It suggests that the integrator-mediated formation of the U7 snRNA 3′end may be limiting when the transcript is too abundant.

RNAs from transduced myoblasts were also analyzed by RT-PCR and quantitative RT-PCR (**Figure 3**). Consistently, the results indicate that the MHCK sequences increase exon 23 skipping, particularly at the lower dose of AAV.

The SmOPT sequences induce an additional processing of the U7snRNA cassette

To investigate the origin of the short RNA species detected with the antisense probe, variants of the U7snRNA cassette were analyzed. Human 293T cells were transfected with plasmids expressing an unmodified U7 which retains its natural 5′ sequences targeting Histone H1 mRNA or U7SmOPT, the parent construct for all U7 antisense shuttles.[19] An additional shorten transcript was generated only in the presence of the SmOPT mutations, suggesting that they were responsible for the extra processing (**Supplementary Figure S3**). Since probes at the 5′ extremity of the U7snRNA were used, the experiment also indicated that the processing removes the 3′ extremity. U7DTex23 cassettes with or without the SmOPT mutations were compared in AAV2/5 transductions of C2C12 myoblasts (**Figure 4**). The inclusion of the SmOPT mutations resulted in higher levels of chimeric U7[19] and the 3′ deleted U7 transcripts were enhanced accordingly. The higher molecular weight species were also proportionally increased in the presence of the SmOPT mutations. The cassette containing the original Sm sequences yielded lower levels of transcripts, which were not further processed in 3′, confirming that the SmOPT mutations determine this modification.

The amounts of the different U7-derived species (86 nt, 66 nt, and over 200 nt) were evaluated by densitometry of the northern blot showed in **Figure 4**. The 86 nt species represented 30% and 40% of all transcripts in lanes 1 and 2, respectively. The difference was due to the lower amounts of high-molecular weight unprocessed species at the lower MOI. In contrast, when the original Sm sequence was present, the 86-nt band accounted for >85% of total signal.

RT-PCR analysis of the *Dmd* pre-mRNA skipping was consistent with the U7snRNA expression analysis. The SmOPT construct induced a strong exon 23 skipping whereas the Sm construct was almost inactive (**Figure 4**, lower panel).

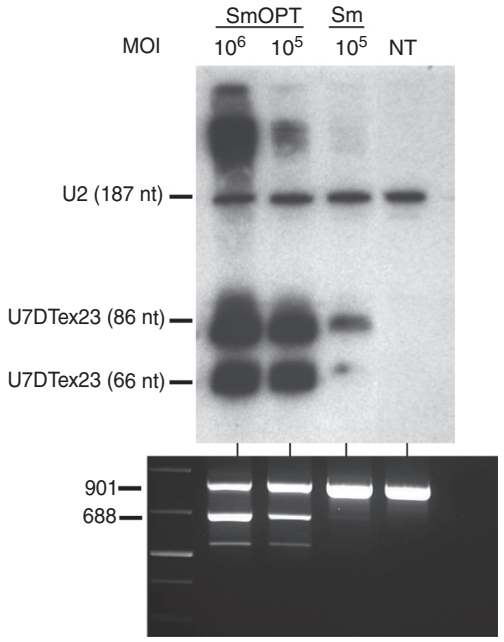

Figure 4 The SmOPT sequence induces an additional 3′ processing of the U7 small nuclear RNA (snRNA). Northern blot analysis of U7DTex23 transcripts expressed in C2C12 cells transduced with AAV2/5-U7DTex23 containing the SmOPT sequence (AAUUUUUGGAG, **Figure 1a**), or the native U7 Sm sequence (AAUUUGUCUAG). NT, nontransduced cells. The blot was hybridized with two 32P end-labeled probes recognizing the U2 snRNA and the U7DTex23 chimera, respectively. U2 snRNA shows up as a 187-nt band. Two bands of 86 nt and ~66-nt react with the U7DTex23 probe. The lower panel presents the analysis of exon 23 skipping performed in the same RNA samples. The native *Dmd* mRNA containing exon 23 is detected as a 901-bp fragment and the skipped mRNA yields a 688-bp product.

MHCK enhances dystrophin rescue in the skeletal muscle of mdx mice

In order to determine the efficiency of the MHCK containing construct *in vivo*, AAV2/8 vectors with or without the MHCK enhancer were produced. Mdx mice received doses of 10^{11}, 2×10^{10}, or 5×10^{9} viral genomes into the tibialis anterior. One month after injection, mice were sacrificed and protein and RNA were isolated from the injected muscles. U7DTex23 and U2snRNA were analyzed by northern blot, and exon 23 skipping efficiency was evaluated by RT-PCR and quantitative RT-PCR (**Figure 5a,b**). The northern blot analysis indicated that the processed U7 species was present and that the MHCK resulted in a modest increase of U7 levels in the muscle. This increase resulted in slightly higher efficiencies of skipping, although the differences analyzed by quantitative PCR were not significant. A stronger evidence of the MHCK enhancement was obtained by analyzing Dystrophin rescue in the injected tissues by western blot. At each dose analyzed, animals receiving the MHCK vector had three- to fivefold increased levels of rescued Dystrophin (**Figure 6**).

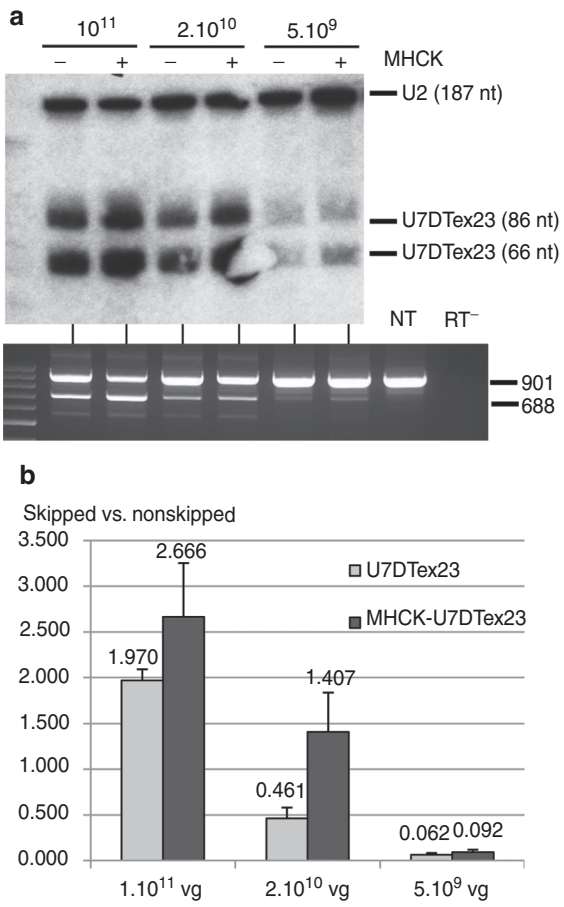

Figure 5 Muscle- and heart-specific enhancer (MHCK) enhances U7DTex23-mediated exon skipping on the Duchenne muscular dystrophy (*Dmd*) pre-mRNA in the muscle of mdx mice. (**a**) Northern blot analysis of U7DTex23 transcripts expressed in mdx mice muscles treated with AAV2/8-U7DTex23 (⁻MHCK) or AAV2/8-MHCK-U7DTex23 (⁺MHCK). Total injected vector doses in vector genomes are indicated on top. The blot was hybridized with two 32P end-labeled probes recognizing the U2 small nuclear RNA (snRNA) and the U7DTex23 chimera, respectively. U2 snRNA shows up as a 187-nt band. Two bands of 86 nt and ~66-nt react with the U7DTex23 probe. The lower panel presents the analysis of exon 23 skipping performed in the same RNA samples. The native *Dmd* mRNA containing exon 23 is detected as a 901-bp fragment and the skipped mRNA yields a 688-bp product. NT, nontransduced muscle; RT–, no reverse transcriptase. Results are representative of at least three independent transductions. (**b**) Quantitative PCR analysis of exon 23 skipping in RNA samples prepared from injected muscles. The graph represents the ratio between the exon 22–24 junction (skipped mRNA species) and the exon 22–23 junction (full-length mRNA species). No statistically significant difference could be observed between the groups (Student's t-test, $P > 0.05$). Error bars are shown as mean ± SEM ($n = 3$ per cohort).

DISCUSSION

The successful clinical translation of AAV-based approaches for exon skipping in diseases such as DMD will depend on our ability to improve the therapeutic index of the vectors. The effective dose should be sufficient to reach extended skeletal muscle territories and to introduce enough transcriptionally active copies of the vector genome in each cell nucleus, without compromising safety. Minimizing the dose will therefore be a combination of using an AAV vector with the appropriate tropism[14] and optimizing the snRNA transcription unit.

The promoters of *snRNA* genes are compact—around 250 base pairs (bp) in length—and very active. For example, the U1 promoter was shown to be equivalent to the immediate

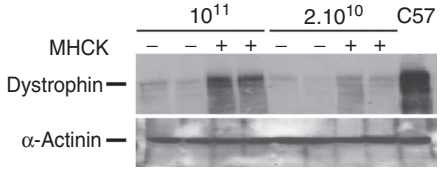

Figure 6 Muscle- and heart-specific enhancer (MHCK)-mediated enhancement of Duchenne muscular dystrophy (*Dmd*) exon skipping in mdx mice leads to an increased rescue of Dystrophin protein in treated muscles. Western blot analysis of dystrophin levels in muscles of mdx mice treated with AAV2/8-U7DTex23 ($^-$MHCK) or AAV2/8-MHCK-U7DTex23 ($^+$MHCK), at doses of 2×10^{10} or 1×10^{11} vg/muscle. Samples from two different animals are analyzed for each condition. The wild type C57 mouse Dystrophin is shown as a positive control and the endogenous α-actinin is detected for normalization.

early CMV enhancer/promoter for expressing a *lacZ* reporter gene.[24] Although the related U7 promoter is of comparable strength,[12,19,25] Brun *et al.* have shown that it can be boosted by a muscle-specific MCK enhancer.[20] Here, we have introduced a strong dual muscle- and heart-specific transcriptional enhancer (MHCK) upstream of a U7 shuttle for antisense sequences in order to evaluate the improvement of exon-skipping efficiency associated with a higher number of chimeric snRNA molecules in cultured myogenic cells and in the mouse skeletal muscle.

The effect of the enhancer on the amounts of snRNA produced in cultured cells or in vivo is modest relative to the one reported in association with the CK promoter and 5′ untranslated region.[21] This is likely due to the high constitutive activity of the U7 promoter. We have observed that transcription enhancement by the MHCK sequences also occurs in 293 cells, indicating that the ubiquitous nature of the U7 promoter overrides the tissue specificity of the enhancer (**Supplementary Figure S4**). This is consistent with reports showing that both MCK and MHC enhancers lose their tissue specificity when combined with constitutive promoters.[26]

Recombinant AAV genomes are found as either linear or circular monomers and concatemers. High doses of AAV vectors result in the formation of concatemers of the vector genome.[23,27] With increasing doses of vector applied to myogenic cells we observe the formation of high-molecular weight RNA species containing internal U7 sequences as well as sequences in the 3′ box. These are consistent with unprocessed read through transcripts generated from AAV concatemers. A similar dose-dependent accumulation of unprocessed U7 transcripts has been reported following injection of U7-based constructs in Xenopus oocytes and it was attributed to a saturation of the U7 3′ end processing machinery.[19,25] The fact that high-molecular weight U7 transcripts were not observed in the injected muscle is probably due to the lower MOI reached in vivo, but could also reflect a higher U7 processing capacity in muscle fibers compared to cultured myotubes.

Our data also indicate the presence in transfected or AAV transduced cells of a U7 species lacking 20 nt at the 3′ end. Such 3′-deleted forms of U7 chimeras have been observed previously[12] and we show here that they are a consequence of the SmOPT mutations. U7 is normally assembled into non-spliceosomal snRNP where a specific core of Sm proteins interacts with the Sm sequence.[28] Changing this sequence to SmOPT redirects U7 transcripts into snRNPs containing the canonical Sm proteins, which accumulate to a much higher levels than unmodified U7 transcripts.[19] As shown in **Figure 4**, the efficiency of U7 shuttles for antisense delivery critically depends on the presence of an SmOPT sequence.[10,29] The shortened U7 transcripts could be generated by a misprocessing of the 3′ end associated with mutations in the Sm sequences, as observed for U1.[30] So far however, only mutations in the 3′ box or in the U7 stem loop are known to affect processing.[31] More likely, the packaging of U7SmOPT into a spliceosomal snRNP may expose a cryptic nucleolytic site in the stem loop and trigger a partial cleavage or degradation during particle maturation.[18] It is unlikely that the misprocessed U7SmOPT species remain properly assembled into functional snRNPs and therefore functional for antisense delivery. Moreover, the shortened U7 RNAs are probably destabilized since the 3′ stem loop of snRNAs acts as a protection against exonucleases. Therefore, the steady-state ratio of full-length versus misprocessed U7 RNA observed by northern blot may in fact underestimate the proportion of transcripts that undergo removal of the terminal 20 nt.

Our study indicates that potentially inactive by-products of U7SmOPT shuttles may be produced following gene transfer. Further understanding of the formation of these molecules may allow improving the potency of U7-based chimera for antisense delivery.

MATERIALS AND METHODS

Primer oligonucleotides

F1	5′-TTTACTAGTCCCTTCAGATTAAAAATAACTGA-3′
R1	5′-CCCGTAGTGGGACAGCAGGGCCCCAAGGTT-3′
F2	5′-CCCTGCTGTCCCACTACGGGTCTAGGCTGC-3′
R2	5′-CCCTCTAGAGATCCACCAGGGACAGGGTTAT-3′
FMCKdel63	5′-CCCCAACACCTGCTGCCTGCTAAAAATAACCCTG TCCCTGGTGG-3′
RMCKdel63	5′-CCACCAGGGACAGGGTTATTTTTAGCAGGCAGCA GGTGTTGGGG-3′
FDTex23Sm	5′gcttacctaaatagaagttcatttacactaacaatttgtctagcaggttttc tgacttcg-3′
RDTex23Sm	5′cgaagtcagaaaacctgctagacaaattgttagtgtaaatgaacttctat ttaggtaagc-3′
HIV_LTR_F	5′-AGCTTGCCTTGAGTGCTTCAA-3′
HIV_LTR_R	5′-AGGGTCTGAGGGATCTCTAGTTACC-3′
DTex23F	5′-GGCCAAACCTCGGCTTA-3′
DTex23R	5′-GTTAGTGTAAATGAACTTC-3′
PolyA-F	5′-ATTTTATGTTTCAGGTTCAGGGGGAGGTG-3′
PolyA-R	5′-GCGCAGAGAGGGAGTGGACTAGT-3′
Ex46Fo	5′-AGGAAGCAGATAACATTGCT-3′
Ex53Ro	5′-TTTCATTCAACTGTTGCCTC-3′
Ex47Fi	5′-TTACTGGTGGAAGAGTTGCC-3′
Ex52Ri	5′-TGATTGTTCTAGCCTCTTGA-3′
Ex20Fo	5′-CAGAATTCTGCCAATTGCTGAG-3′
Ex26Ro	5′-TTCTTCAGCTTGTGTCATCC-3′
Ex20Fi	5′-CCCAGTCTACCACCCTATCAGAGC-3′
Ex26Ri	5′-CCTGCCTTTAAGGCTTCCTT-3′

Vector constructions. The MHCK enhancer was obtained by fusing the 188 bp long-murine α-MHC enhancer, and the 207 bp murine MCK enhancer, as described by Salva *et al*.[21] The sequences were amplified from C2C12 genomic DNA using the PCR Master Mix (Promega, Madison, WI) following the manufacturer's instructions, and primers F1 and R1 (α-MHC enhancer) or F2 and R2 (MCK enhancer). The two amplicons were fused by PCR using primers F1 and R2 and introduced upstream of the U7snRNA cassette in the AAV-U7DTex23 vector, previously described as AAV-U7-SD23/BP22.[13] The α-MHC-MCK fusion was further modified by deleting a 63 bp region as described,[21] using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) and the primers FMCKdel63 and RMCKdel63.

A deletion of three cytosines between the left and right E-boxes of the MCK enhancer was observed in the final AAV-MHCK-U7DTex23 construct, compared to the published NCBI sequence (GenBank: AF188002.1). This deletion was also observed in C2C12 genomic DNA and was considered a polymorphism.

U7ex51 was derived from U7DTex23 with the mouse dystrophin antisense sequences replaced by a sequence targeting a splicing enhancer (underlined) in human dystrophin exon 51 (5′-GCAGGTACCTCCAACATCAAGGAAGATGGC ATTTCTAGTTTGGAG-3′).[22] U7ex51 was introduced in the pRRLSIN-cPPT-PGK-GFP-WPRE lentiviral vector[32] to create pRRLSIN-U7ex51. The MHCK enhancer was introduced upstream of the U7snRNA cassette to generate pRRLSIN-MHCK-U7ex51.

The SmOPT site of the U7DTex23 sequence was changed into the canonical Sm sequence in the AAV-U7DTex23 vector by site-specific mutagenesis using the primers FDTex23Sm and RDTex23Sm.

The scAAV-PGK-GFP plasmid was constructed by introducing the PGK-GFP fragment from pRRLSIN-cPPT-PGK-GFP-WPRE into the self-complementary vector scAAV-LP1-hFIX,[33] from which the LP1-hFIX sequence was removed. The MHCK enhancer was inserted upstream of the PGK promoter to obtain scAAV-MHCK-PGK-GFP.

Lentiviral vectors preparation. Lentiviral virions were produced by calcium phosphate transfection of HEK293T cells with the envelope plasmid pMD.G (VSV-G[34]), the packaging plasmids pHDM-gpm2 (codon-optimized HIV-1 Gag-Pol) and pRC/CMV-REV1B (REV) (both gifts from Dr Jeng-Shin Lee, Harvard Medical School), and the vector plasmid pRRLSIN-U7ex51. Viral supernatants were harvested 20, 28, and 36 hours post-transfection, centrifuged 5 minutes at 1,500 rpm, filtered through a 0.45-µm filter, and ultracentrifuged 2 hours at 19,500 rpm at 12 °C. Pellets were resuspended in phosphate-buffered saline (PBS) 1% bovine serum albumin and stored at –80 °C. The infectious titers were determined by quantitative PCR using the SYBR Green PCR Master Mix (Applied Biosystems, Life Technologies, Carlsbad, CA) and primers specific for the lentiviral genome HIV_LTR_F and HIV_LTR_R. The albumin endogenous gene was used for normalization.

AAV vectors preparations. AAV vectors were produced by calcium phosphate transfection of HEK293T cells with the transgene plasmid, the capsid plasmid AAV2/8 (LTAAVhelp2-8[35]) or AAV2/5 (Napoli 2/5[36]), and the plasmid for adenovirus helper function (pHGTI-Adeno1[37]). After 72 hours cells and media were harvested and centrifuged at 1,800 rpm for 10 minutes at 4 °C and the pellets resuspended in 40 ml of TD buffer (140.4 mmol/l NaCl, 4.9 mmol/l KCl, 0.7 mmol/l K_2HPO_4, 3.4 mmol/l $MgCl_2$, 24.7 mmol/l Tris—pH 7.5). After five freeze/thaw cycles of the suspension (30 minutes at –80 °C followed by 30 minutes at 37 °C) 2,000 U of benzonase were added, the suspension was incubated for 30 minutes at 37 °C and centrifuged for 20 minutes at 3,000*g*. The supernatant was filtered through a 0.45-µm filter and the AAV preparations were purified by affinity chromatography using the AKTA explorer chromatography system (GE Healthcare, Waukesha, WI) and a prepacked AVB sepharose affinity column (HiTrap; GE Healthcare). The collected fractions of AAV vectors sample were pooled and dialysed in PBS at 4 °C for 16 hours. The viral genome titers were assessed by real-time PCR using the SYBR Green PCR Master Mix (Applied Biosystems). For titration of the AAV-U7DTex23 and AAV-MHCK-U7DTex23 vectors, primers specific of the DTex23 antisense sequence were used: DTex23F and DTex23R. For the scAAV-PGK-GFP and scAAV-MHCK-PGK-GFP vectors, primers in the polyA sequence were used: PolyA-F and PolyA-R.

Cell culture and transduction. All the cells were cultivated at 37 °C and 5% CO_2. 293T cells were grown in Dulbecco's modified Eagle's medium (DMEM; Invitrogen, Carlsbad, CA) supplemented with 10% fetal calf serum (Invitrogen). CHQ cells were grown in F-10 medium (Invitrogen) supplemented with 20% fetal calf serum (Invitrogen). Differentiation of CHQ cells was induced in DMEM without pyruvate (Invitrogen) supplemented with 100 µg/ml of human apolipotransferrin (Sigma, St Louis, MI) and 10 µg/ml of bovin insulin (Sigma). C2C12 cells were grown in DMEM supplemented with 20% fetal calf serum (Invitrogen). Differentiation of C2C12 was induced in DMEM supplemented with 2% horse serum (Sigma).

For lentiviral vector transduction, 30,000 CHQ cells were seeded in each well of a 12-wells plate. After 24 hours, the lentiviral vector preparation was administered in a total volume of 250 µl of proliferation medium at the chosen MOI. After 6-hour incubation 1 ml of proliferation medium was added per well. Seventy two hours post-transduction the medium was replaced with differentiation medium for 72 additional hours. Cells were then harvested for analysis.

For AAV vector transduction, 100,000 C2C12 cells were seeded in each well of a 12-wells plate. The following day proliferation medium was changed to differentiation medium and the cells were left to differentiate into myotubes for a week. The cells were then transduced with the AAV preparation at the chosen MOI in a total volume of 300 µl of medium without serum. After 6 hours of incubation 1 ml of differentiation medium was added per well. Cells were incubated for 7 days in differentiation medium before harvesting for analysis.

Western blot. For dystrophin detection total protein was extracted from muscle samples with Newcastle buffer (3.8% SDS, 75 mmol/l Tris–HCl pH 6.7, 4 mol/l urea, 10% βmercaptoethanol, 10% glycerol, 0.001% bromophenol blue) and quantified using the bicinchoninic acid protein assay kit, according to the manufacturer's instructions (Perbio Science, Brebieres, France). Samples were denatured at

95 °C for 5 minutes before 100 µg of protein was loaded in a 5% polyacrylamide gel with a 4% stacking gel. Gels were electrophoresed for 4–5 hours at 100 V and blotted to a PVDF membrane overnight at 50 V. Blots were blocked for 1 hour with 10% nonfat milk in PBS–Tween (PBST) buffer. Dystrophin and α-actinin proteins were detected by probing the membrane with 1:100 dilution of NCL-DYS1 primary antibody (monoclonal antibody to dystrophin R8 repeat; NovoCastra) and 1:200 dilution of α-actinin primary antibody (Santa Cruz Biotechnology), respectively. An incubation with a mouse horseradish peroxidase-conjugated secondary antibody (1:2,000) or goat horseradish peroxidase-conjugated secondary antibody (1:160,000) allowed visualization using ECL Analysis System (GE Healthcare). Membranes were converted to numerical pictures by scanning.

RNA isolation and nested RT-PCR analysis. RNA were isolated from cells or muscle sections using TriZol Reagent according to the manufacturer's instructions (Invitrogen). RT-PCR was carried out on 300 ng of total RNA using the Access RT-PCR System (Promega). For human sequences analysis external primers Ex46Fo and Ex53Ro were used. The cDNA synthesis was carried out at 45 °C for 45 minutes, directly followed by the primary PCR of 30 cycles of 94 °C (30 s), 55 °C (1 minute) and 72 °C (2 minutes). Two microlitres of the RT-PCR product was used for the nested PCR with the internal primers Ex47Fi and Ex52Ri for 22 cycles under identical cycling conditions. The same nested PCR conditions were used for mouse sequences analysis, with external primers Ex20Fo and Ex26Ro, and internal primers Ex20Fi and Ex26Ri. Products were analyzed on 2% agarose gel.

Quantitative PCR. For each sample, 1 µg of total RNA was used for RT using the SuperScript II RT system (Invitrogen) and random hexamers as the priming nucleotides, according to the manufacturer's instructions. For quantitative PCR analysis, 10 ng of the resulting cDNA was used per well. Comparative CT quantitative analysis of the exon 22-23 junction (Taqman Assay Mm01216934_m1; Applied Biosystems) and the exon 22-24 junction (Fw: CTGAATATGAAATAATGGAGGAGAGACTCG, Rev: CTTCAGCCATCCATTTCTGTAAGGT, probe: FAM-ATGTGATTCTGTAATTTCC-NFQ) was performed on each RNA sample. Reactions were performed in triplicate with the appropriate non-template controls using the cycle conditions: 95 °C for 10 minutes, and 40 cycles at 95 °C for 15 seconds and 60 °C for 1 minute.

Data analysis. For each sample, the CTs for the 22–24 assay (skipped dystrophin) are normalized to the CT for the 22–23 assay (nonskipped dystrophin) to generate a delta-CT value. This involves the subtraction of the 22–23 CT from the 22–24 Ct. The delta CT can be used to calculate the relative abundance in skipped dystrophin RNA species relative to full-length species.

An unpaired Student's *t*-test was used to assess the statistical significance of differences between the samples treated with AAV-U7DTex23 and the samples treated with AAV2-MHCK-U7DTex23. A P value <0.05 was considered statistically significant. All results are expressed as mean values ± SEM.

Northern blot. Northern blots were realized on 7 µg of total RNA. Samples of total RNA were denaturated at 70 °C for 15 minutes and loaded on a 15% Novex TBE-Urea precast gel (Invitrogen) along with 1 µl of γ-32P-labeled RNA Decade Markers (Ambion, Austin, TX). Gels were electrophoresed for 1 hour at 180 V in 2.5× TBE buffer, and transferred to a Hybond-N+ membrane (Amersham, UK) for 2 hours at 390 mA in 2× TBE buffer. The membrane was UV-crosslinked with 100 mJ of energy and prehybridized in 10 ml of Rapid Hyb buffer (Amersham) containing salmon sperm DNA (50 µg/ml) at 42 °C. For radiolabeling of the complementary DNA probe 50 pmol of oligonucleotide, U7DTex23 (5′-32P-GTTAGTGTAAATGAACTTC-3′) or endogenous U2snRNA (5′-32P-CTGATAAGAACAGATAC-3′), was end-labeled using γ-32P-ATP (Amersham, 6 000 Ci/mmol) and T4 PolyNucleotide Kinase (New England BioLabs, Ipswich, MA) for 1 hour at 37 °C. The probe was purified from unincorporated label using G-25 MicroSpin Columns (Amersham), and heated for 1 minute at 95 °C before addition to the hybridization solution and incubation overnight at 42 °C. The following day 3 successive washes of increasing stringency were realized for 15 minutes at 42 °C before placing the blots in a cassette with an autoradiogram. The films were exposed for 72 hours.

MDX mice injection. Five-month-old mdx mice were anesthetized with isoflurane and injected in the tibialis anterior with 5×10^9, 2×10^{10}, or 10^{11} viral genome copies of AAV2/8-U7DTex23 or AAV2/8-MHCK-U7DTex23 ($n = 3$ for each condition). For each mouse, contra-lateral muscle was injected with PBS as a negative control. Mice were sacrificed 4 weeks after injection and tibialis anterior muscles were isolated and snap frozen in liquid nitrogen-cooled isopentane. All animal experiments were carried out in Biomedical Science Building, University of Oxford, Oxford, UK and performed according to the guidelines and protocols approved by the Home Office.

Acknowledgments. This work was supported by grants from the Agence Nationale de la Recherche (Genesplice program), the Association Française contre les Myopathies, the Monegasque Association against Muscular Dystrophy and Duchenne Parent Project France (ICE program), The Muscular Dystrophy Campaign the Alexander Patrick trust and the Medical Research Council. We thank Dr Jeng-Shin Lee for providing reagents and Arnold Munnich (Inserm U785) for support. The authors declared no conflict of interest.

Supplementary Material

Figure S1. The MHCK enhancer increases GFP expression in myogenic cells.
Figure S2. Dose-dependent accumulation of read through U7 RNAs transcripts generated from AAV vector genome concatemers.
Figure S3. Northern blot analysis of (**a**) U7DTex23 and (**b**) U7 transcripts expressed in 293T cells transfected with AAV constructs containing either the SmOPT sequence (AAUUUUGGAG, **Figure 1a**), or the native U7 Sm sequence (AAUUUGUCUAG).
Figure S4. The MHCK enhancer loses its tissue specificity when coupled with a ubiquitous promoter.

REFERENCES

1. Tuck, AC and Tollervey, D (2011). RNA in pieces. *Trends Genet* 27: 422–432.
2. Buratti, E, Baralle, M and Baralle, FE (2006). Defective splicing, disease and therapy: searching for master checkpoints in exon definition. *Nucleic Acids Res* 34: 3494–3510.
3. Wang, ET, Sandberg, R, Luo, S, Khrebtukova, I, Zhang, L, Mayr, C *et al.* (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456: 470–476.
4. Tazi, J, Durand, S and Jeanteur, P (2005). The spliceosome: a novel multi-faceted target for therapy. *Trends Biochem Sci* 30: 469–478.
5. Hammond, SM and Wood, MJ (2011). Genetic therapies for RNA mis-splicing diseases. *Trends Genet* 27: 196–205.
6. Kinali, M, Arechavala-Gomeza, V, Feng, L, Cirak, S, Hunt, D, Adkin, C *et al.* (2009). Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *Lancet Neurol* 8: 918–928.
7. van Deutekom, JC, Janson, AA, Ginjaar, IB, Frankhuizen, WS, Aartsma-Rus, A, Bremmer-Bout, M *et al.* (2007). Local dystrophin restoration with antisense oligonucleotide PRO051. *N Engl J Med* 357: 2677–2686.
8. Goemans, NM, Tulinius, M, van den Akker, JT, Burm, BE, Ekhart, PF, Heuvelmans, N *et al.* (2011). Systemic administration of PRO051 in Duchenne's muscular dystrophy. *N Engl J Med* 364: 1513–1522.
9. Cirak, S, Arechavala-Gomeza, V, Guglieri, M, Feng, L, Torelli, S, Anthony, K *et al.* (2011). Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *Lancet* 378: 595–605.
10. Gorman, L, Suter, D, Emerick, V, Schümperli, D and Kole, R (1998). Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. *Proc Natl Acad Sci USA* 95: 4929–4934.
11. Gorman, L, Mercatante, DR and Kole, R (2000). Restoration of correct splicing of thalassemic beta-globin pre-mRNA by modified U1 snRNAs. *J Biol Chem* 275: 35914–35919.
12. De Angelis, FG, Sthandier, O, Berarducci, B, Toso, S, Galluzzi, G, Ricci, E *et al.* (2002). Chimeric snRNA molecules carrying antisense sequences against the splice junctions of exon 51 of the dystrophin pre-mRNA induce exon skipping and restoration of a dystrophin synthesis in Delta 48-50 DMD cells. *Proc Natl Acad Sci USA* 99: 9456–9461.
13. Goyenvalle, A, Vulin, A, Fougerousse, F, Leturcq, F, Kaplan, JC, Garcia, L *et al.* (2004). Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. *Science* 306: 1796–1799.
14. Goyenvalle, A, Babbs, A, Wright, J, Wilkins, V, Powell, D, Garcia, L *et al.* (2012). Rescue of severely affected dystrophin/utrophin-deficient mice through scAAV-U7snRNA-mediated exon skipping. *Hum Mol Genet* 21: 2559–2571.
15. Bish, LT, Sleeper, MM, Forbes, SC, Wang, B, Reynolds, C, Singletary, GE *et al.* (2012). Long-term restoration of cardiac dystrophin expression in golden retriever muscular dystrophy following rAAV6-mediated exon skipping. *Mol Ther* 20: 580–589.
16. Vaillend, C, Perronnet, C, Ros, C, Gruszczynski, C, Goyenvalle, A, Laroche, S *et al.* (2010). Rescue of a dystrophin-like protein by exon skipping *in vivo* restores GABAA-receptor clustering in the hippocampus of the mdx mouse. *Mol Ther* 18: 1683–1688.
17. Baillat, D, Hakimi, MA, Näär, AM, Shilatifard, A, Cooch, N and Shiekhattar, R (2005). Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* 123: 265–276.
18. Patel, SB and Bellini, M (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res* 36: 6482–6493.
19. Grimm, C, Stefanovic, B and Schümperli, D (1993). The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO J* 12: 1229–1238.
20. Brun, C, Suter, D, Pauli, C, Dunant, P, Lochmüller, H, Burgunder, JM *et al.* (2003). U7 snRNAs induce correction of mutated dystrophin pre-mRNA by exon skipping. *Cell Mol Life Sci* 60: 557–566.
21. Salva, MZ, Himeda, CL, Tai, PW, Nishiuchi, E, Gregorevic, P, Allen, JM *et al.* (2007). Design of tissue-specific regulatory cassettes for high-level rAAV-mediated expression in skeletal and cardiac muscle. *Mol Ther* 15: 320–329.
22. Aartsma-Rus, A, Bremmer-Bout, M, Janson, AA, den Dunnen, JT, van Ommen, GJ and van Deutekom, JC (2002). Targeted exon skipping as a potential gene correction therapy for Duchenne muscular dystrophy. *Neuromuscul Disord* 12 Suppl 1: S71–S77.
23. Vincent-Lacaze, N, Snyder, RO, Gluzman, R, Bohl, D, Lagarde, C and Danos, O (1999). Structure of adeno-associated virus vector DNA following transduction of the skeletal muscle. *J Virol* 73: 1949–1955.
24. Bartlett, JS, Sethna, M, Ramamurthy, L, Gowen, SA, Samulski, RJ and Marzluff, WF (1996). Efficient expression of protein coding genes from the murine U1 small nuclear RNA promoters. *Proc Natl Acad Sci USA* 93: 8852–8857.
25. Phillips, SC and Turner, PC (1992). A transcriptional analysis of the gene encoding mouse U7 small nuclear RNA. *Gene* 116: 181–186.
26. Bojak, A, Hammer, D, Wolf, H and Wagner, R (2002). Muscle specific versus ubiquitous expression of Gag based HIV-1 DNA vaccines: a comparative analysis. *Vaccine* 20: 1975–1979.
27. Nakai, H, Thomas, CE, Storm, TA, Fuess, S, Powell, S, Wright, JF *et al.* (2002). A limited number of transducible hepatocytes restricts a wide-range linear vector dose response in recombinant adeno-associated virus-mediated liver transduction. *J Virol* 76: 11343–11349.
28. Azzouz, TN and Schumperli, D (2003). Evolutionary conservation of the U7 small nuclear ribonucleoprotein in Drosophila melanogaster. *RNA* 9: 1532–1541.
29. Suter, D, Tomasini, R, Reber, U, Gorman, L, Kole, R and Schümperli, D (1999). Double-target antisense U7 snRNAs promote efficient skipping of an aberrant exon in three human beta-thalassemic mutations. *Hum Mol Genet* 8: 2415–2423.
30. Seipelt, RL, Zheng, B, Asuru, A and Rymond, BC (1999). U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway. *Nucleic Acids Res* 27: 587–595.
31. Ezzeddine, N, Chen, J, Waltenspiel, B, Burch, B, Albrecht, T, Zhuo, M *et al.* (2011). A subset of Drosophila integrator proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3'-end formation. *Mol Cell Biol* 31: 328–341.
32. VandenDriessche, T, Naldini, L, Collen, D and Chuah, MK (2002). Oncoretroviral and lentiviral vector-mediated gene therapy. *Meth Enzymol* 346: 573–589.
33. Nathwani, AC, Gray, JT, McIntosh, J, Ng, CY, Zhou, J, Spence, Y *et al.* (2007). Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates. *Blood* 109: 1414–1421.
34. Naldini, L, Blömer, U, Gallay, P, Ory, D, Mulligan, R, Gage, FH *et al.* (1996). *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272: 263–267.
35. Allay, JA, Sleep, S, Long, S, Tillman, DM, Clark, R, Carney, G *et al.* (2011). Good manufacturing practice production of self-complementary serotype 8 adeno-associated viral vector for a hemophilia B clinical trial. *Hum Gene Ther* 22: 595–604.
36. Allocca, M, Doria, M, Petrillo, M, Colella, P, Garcia-Hoyos, M, Gibbs, D *et al.* (2008). Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J Clin Invest* 118: 1955–1964.
37. Herson, S, Hentati, F, Rigolet, A, Behin, A, Romero, NB, Leturcq, F *et al.* (2012). A phase I trial of adeno-associated virus serotype 1-γ-Sarcoglycan gene therapy for limb girdle muscular dystrophy type 2C. *Brain* 135(Pt 2): 483–492.

Supplementary Information accompanies this paper on the Molecular Therapy–Nucleic Acids website (http://www.nature.com/mtna)

Supplementary Material

Supplementary Figure S1. MHCK enhancer increases GFP expression in myogenic cells.

(a) Western blot to detect GFP expression obtained from protein extracts of C2C12 cells transduced with self complementary AAV2/5 vectors expressing the PGK-GFP or the MHCK-PGK-GFP cassettes (MOI = 10^4 vg /cell). The non-transduced cells (NT) and the endogenous control β-actin are shown. C2C12 protein extracts were obtained by incubation 30 minutes at 4°C in lysis buffer (1× complete protease inhibitor cocktail (Roche Diagnostics, Basel, Switzerland), 0.5% NP-40, 20 mM Tris pH 7.5, 10% glycerol, 100 mM $(NH_4)_2SO_4$). Equal amounts of protein were separated on a 12% polyacrylamide gel 1 hour at 110 V and transferred to a PVDF membrane (Amersham, UK) 90 minutes at 100 V. The membrane was blocked with PBS 5% milk 0,1% Tween 20 and probed with rabbit polyclonal anti-GFP antibody (sc-8334; Santa Cruz Biotechnology, Santa Cruz, CA - dilution 1:500) and mouse anti-actin antibody (A-5441; Sigma - dilution 1:15000). The secondary antibodies goat anti-mouse and goat anti-rabbit, conjugated with peroxidase (Sigma), were used at a 1:2000 dilution. Signals were detected with the ECL kit (Amersham).

(b) Representation of GFP expression levels relative to β-actin in the transduced cells. The quantification was performed from the western blot presented in (a). The membrane was converted to a numerical picture and band intensities were analysed by densitometry using the GeneTools analysis software (Syngene, UK). GFP levels are expressed as a percentage compared with β-actin levels.

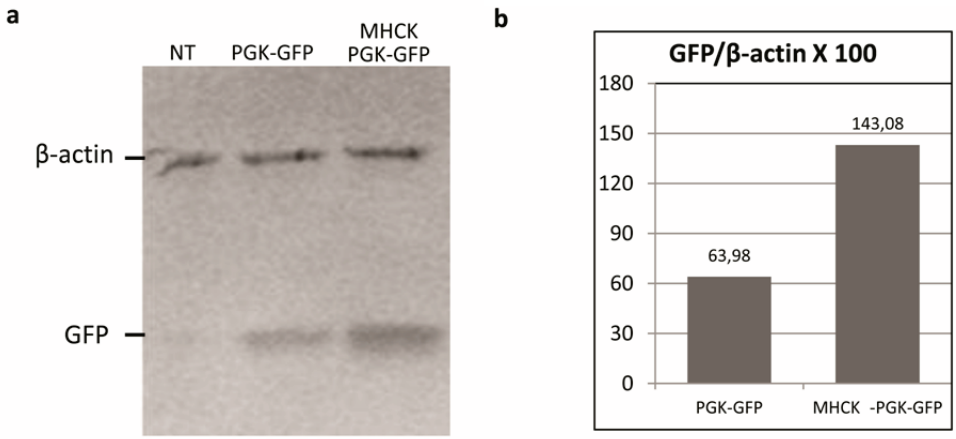

Figure S1

Supplementary Figure S2. Dose-dependent accumulation of readthrough U7 RNAs transcripts generated from AAV vector genome concatemers

Northern blot analysis of U7DTex23 transcripts expressed from C2C12 cells transduced with AAV2/5-U7DTex23 (-MHCK) or AAV2/5-MHCK-U7DTex23 (+MHCK) at a MOI of 10^5 or 10^6 viral genome per cell. The non-transduced cells (NT) and the endogenous U2snRNA transcripts are shown as controls. The same samples as in Fig. 2b were analysed, using a probe complementary to the U7 3' box, that allows detection of unprocessed U7 transcripts. A representation of the primary U7 transcript and the position of the probe are shown on top.

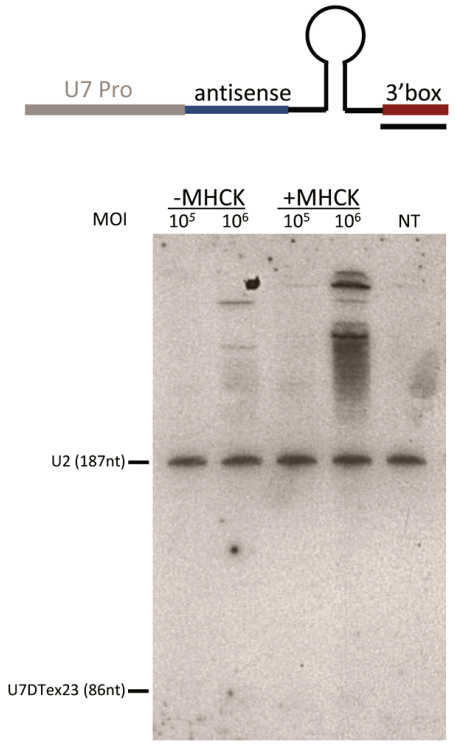

Figure S2.

Supplementary Figure S3. Northern blot analysis of U7DTex23 (a) and U7 (b) transcripts expressed in 293T cells transfected with AAV constructs containing either the SmOPT sequence (AAUUUUGGAG, Fig 1a), or the native U7 Sm sequence (AAUUUGUCUAG).

U7 contains the natural antisense sequence against histone H1 pre-mRNA. The non-transfected samples (NT) are shown as controls. The blots were hybridized with 32P end-labeled probes recognizing the U2 snRNA and either the U7DTex23 chimera (a) or U7 (5'-32P-CTAAAAGAGCTGTAACACTT-3') (b). U2 snRNA shows up as a 187 nt band. Two bands of 86nt and approximately 66 nt react with the U7DTex23 probe. Two bands of 62nt and approximately 42nt react with the U7WT probe. In both cases, the shorter species are detected in presence of the SmOPT sequence only.

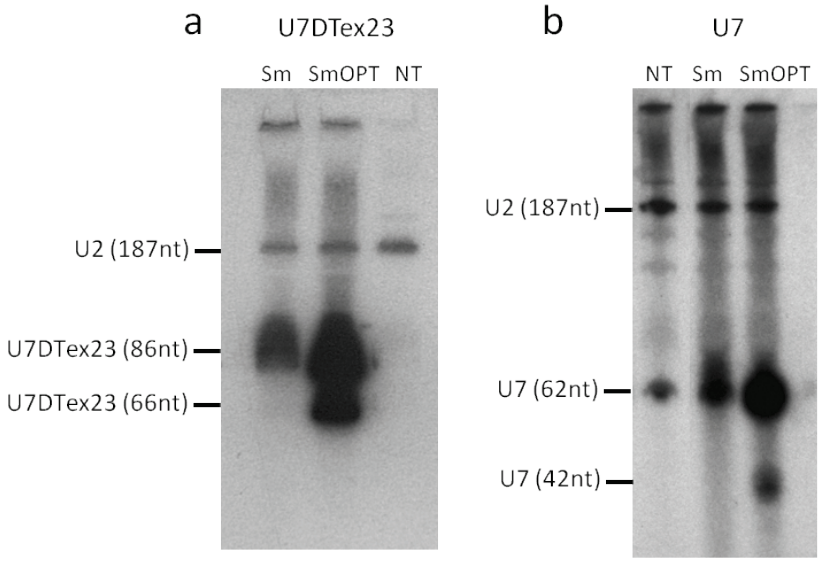

Figure S3.

Supplementary Figure S4. The MHCK enhancer loses its tissue-specificity when coupled with a ubiquitous promoter.

Northern blot analysis of U7DTex23 transcripts expressed in 293T cells transduced with AAV2/5-U7DTex23 (MHCK-) or AAV2/5-MHCK-U7DTex23 (MHCK+) at a MOI of 10^5. The non-transduced cells (NT) and the endogenous U2snRNA transcripts are shown as controls. The expression of the U7DTex23 increases when driven by the MHCK enhancer in non-muscular cells, highlighting the absence of tissue-specificity.

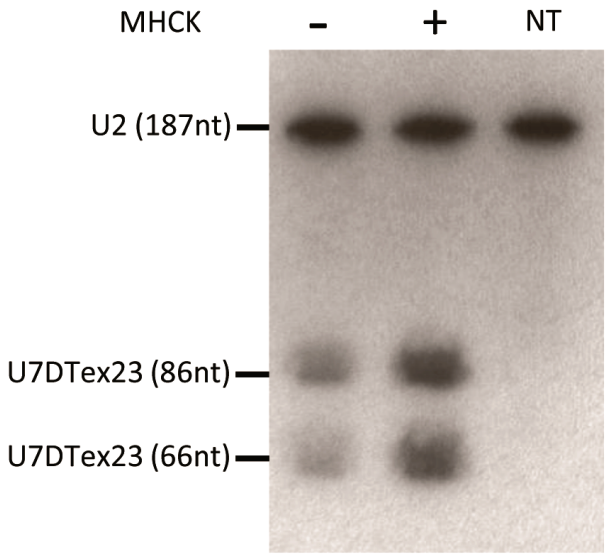

Figure S4.