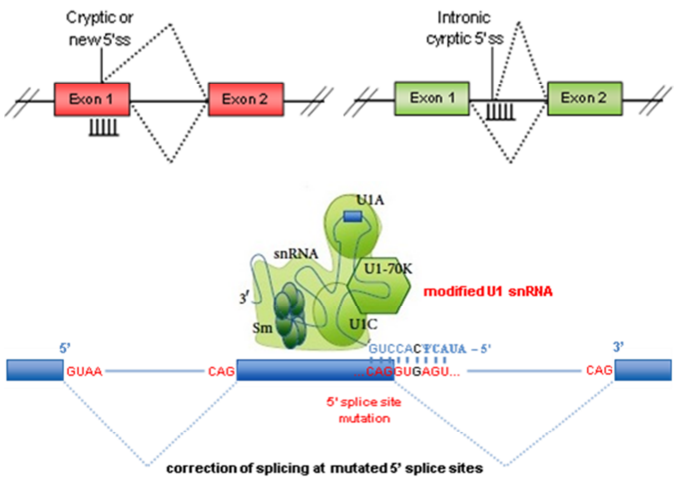


Splicing therapeutics for patients affected by lysosomal storage disorders

Liliana da Silva Matos
Programa Doutoral em Biologia
Departamento de Biologia
2015

Orientadora
Doutora Sandra Catarina da Conceição Alves,
Investigadora Auxiliar,
Instituto Nacional de Saúde Doutor Ricardo Jorge

Coorientadora
Professora Doutora Maria João Prata Martins Ribeiro,
Professora Associada com Agregação,
Faculdade de Ciências da Universidade do Porto

Dissertação apresentada
à Faculdade de Ciências da Universidade do Porto
para cumprimento dos requisitos necessários
à obtenção do grau de Doutor em Biologia

Thesis presented to the
Faculty of Sciences of the University of Porto
to obtain the Doctor degree in Biology

“I was taught that the way of progress is neither swift nor easy”

Marie Curie

Aos meus pais

ACKNOWLEDGEMENTS/AGRADECIMENTOS

O presente trabalho contou com o apoio e colaboração de várias pessoas que de alguma forma contribuíram para a sua realização, e a quem não posso deixar de agradecer.

Ao Conselho Diretivo do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA) e à Direção do Centro de Saúde Pública Doutor Gonçalves Ferreira (CSPGF), assim como à coordenadora do Departamento de Genética Humana, Dra. Glória Isidro e ao responsável pela Unidade de Investigação e Desenvolvimento, Dr. João Lavinha, agradeço por me terem dado a possibilidade de efectuar este trabalho no INSA.

À Fundação para a Ciência e Tecnologia (FCT) agradeço a concessão da bolsa individual de doutoramento (SFRH/BD/64592/2009) sem a qual não teria sido possível a execução dos trabalhos conducentes a esta dissertação.

De uma forma especial gostaria de agradecer:

À Doutora Sandra Alves, minha orientadora, por me ter dado a oportunidade de continuar a trabalhar no seu grupo de investigação e por me proporcionar todas as condições necessárias à realização deste trabalho, pela confiança que depositou nas minhas capacidades e pela autonomia que me deu, tendo permitido efetuar todo o trabalho ao meu ritmo. Agradeço ainda todos os conselhos e sugestões que ajudaram a melhorar a redação desta tese. Obrigada por estar sempre presente e por toda a disponibilidade e amizade.

À Professora Doutora Maria João Prata, minha co-orientadora, pela disponibilidade para comigo e pela atenção e preocupação sempre demonstradas pelo meu trabalho. Agradeço o cuidado que teve e todas as críticas construtivas que fez na correção de cada artigo e, em particular desta tese, sempre na intenção de os valorizar. O meu muito obrigada também pela calma transmitida e pelas palavras de incentivo que tanto me animaram e deram força na fase final de redação desta tese.

Ao Doutor Peter Jordan pela colaboração nos vários estudos aqui apresentados, e à Vânia Gonçalves do seu grupo de investigação pelos conhecimentos práticos transmitidos ainda na fase do meu mestrado, que se mostraram importantes no início da realização deste trabalho. Agradeço ainda pela colaboração no primeiro estudo aqui apresentado.

A Profesora Doctora Belén Pérez, agradezco por darme la oportunidad de trabajar en su laboratorio en el Centro de Biología Molecular Severo Ochoa (CBMSO) – Madrid, por toda la orientación y disponibilidad para discutir el trabajo conmigo, por confiar en mí y dejarme desarrollar libremente mi trabajo práctico. Sólo puedo agradecer todo lo que aprendí porque me hizo crecer tanto personal como profesionalmente. Agradezco también por seguir colaborando con nuestro grupo y por mantener interés en mi trabajo. Una palabra de agradecimiento también a Profesora Doctora Lourdes Desviat por su atención conmigo y por la colaboración en los trabajos aquí presentados. A todos los demás en el grupo, incluso en el laboratorio de diagnóstico en la Universidad Autónoma gracias por la simpatía y por recibirme bien.

To Professors Daniel Grinberg and Lluïsa Vilageliu, and to Isaac Canals from the University of Barcelona as also to Professor Alexey Pshezhetsky and his group from the University of Canada, I would like to thank for the collaboration in a part of this work.

A todos os colegas das várias unidades do Centro de Genética Médica Doutor Jacinto de Magalhães (CGMJM) onde esta caminhada começou e com quem partilhei 5 anos de trabalho, especialmente aos da Unidade de Bioquímica Genética dirigida pela Doutora Lúcia Lacerda, agradeço todo o apoio e colaboração prestados. À Eugénia e Isaura dedico uma palavra especial pela forma como sempre se disponibilizaram para me ajudar, pelas dicas e conhecimentos transmitidos e por todo o apoio e amizade.

A todos os colegas do CSPGF, especialmente aos da Unidade de Rastreio Neonatal que me acompanharam desde o início no CGMJM, agradeço pela boa disposição sempre presente e pelo incentivo, apoio e colaboração disponibilizados. Em particular à Lígia, agora a trabalhar em terras privadas, obrigada pela amizade e pela disponibilidade para ajudar e trocar ideias sobre o meu trabalho sempre que precisei.

Aos colegas informáticos, Paula (CGMJM), Zé Luís e Filipe (CSPGF), agradeço a ajuda técnica prestada sempre que necessário, algumas das vezes tão importante.

Às auxiliares de laboratório, Ana, Dina, D.ª Lúcia, D.ª Fátima, Fernanda e Conceição, a trabalhar connosco em diferentes momentos, agradeço por manterem todos os dias o laboratório e o material necessário sempre preparados para que pudesse realizar o meu trabalho e por estarem sempre disponíveis para ajudar. Especialmente à D.ª Lúcia, Dina e Fernanda, obrigada pela boa disposição e bons

momentos passados, pela vossa amizade e por terem sempre um sorriso. Agradeço ainda à Anabela todo o auxílio prestado ao laboratório sempre que foi necessário.

A todos los lokis del 204 en el CBMSO, Rocio, Ana Jorge, Sandra, Lorena, Pachu, Alfonso, Celia, Patricia Alcaide, Elena, Esther, Raquel y Virginia, así como a Rocio Simón del labo vecino, gracias por acogerme tan bien, por hacerme sentir integrada en el grupo en tan poco tiempo, por compartir conmigo tan buenos y divertidos momentos dentro y fuera del laboratorio y por ser tan majos! A Sven e Borja agradezco por toda la ayuda tecnica prestada. Ha sido un placer conocerlos a todos. Me gustaria todavía de dedicar una palabra especial a algunas personas. À Sandra o meu obrigada pela disponibilidade em me ajudar e esclarecer todas as dúvidas sempre que precisei, por me dar a conhecer um pouco de Madrid e por me fazer sentir mais perto de casa com a sua presença portuguesa. A Ana Jorge, agradezco por guiarme en los primeros días de laboratorio, por su amistad y por corregirme siempre que hablaba “portuñol”. A Rocio, agradezco también por estar siempre disponible para contestar mis dudas, por ayudarme y enseñarme, por todas las conversaciones, apoyo y amistad. A Lorena, mi rubia, muchas gracias por ser siempre tan agradable y amiga, por interesarse en mis clases de portugués que se le da tan genial, por todas las conversaciones tardías y por la compañía mientras trabajábamos casi hasta la cena en el labo. A todos gracias por todo! Aunque corta, nunca voy olvidar mi estancia con vosotros, vos echo de menos!

Às meninas do “muquifo” no CGMJM, Mia, Natália e Liliana e a todos os que por lá passaram desde os tempos de mestrado e com quem convivi diariamente ao longo de vários anos, agradeço a amizade, apoio e todos os momentos de boa disposição e descontração que tantas vezes ajudaram a descomprimir. Guardo boas recordações!

Aos colegas do “Tzero” no CSPGF, obrigada pelo bom convívio diário, pelas gargalhadas e todos os bons momentos de descontração, pelo apoio e amizade. À Olga, agradeço ainda a colaboração num dos estudos aqui apresentados e a revisão de parte desta tese. À Joana e ao Diogo o meu obrigada também pela colaboração no mesmo estudo com tão bons e divertidos momentos no laboratório que tanto ajudaram a amenizar a nossa saga proteica! Ao Pedro um obrigada especial por todas as boas conversas, pelos conselhos tanto pessoais como profissionais que sempre me ajudaram e por todo o incentivo na fase de redação desta tese.

A todos os amigos/as que de alguma forma acompanharam esta longa caminhada, o meu obrigada pelos bons momentos partilhados que ajudaram a suportar melhor este percurso e pelo apoio e amizade.

Às minhas parentes zumbásticas, Ana, Marta, Joana, Vanessa e Catarina, assim como ao meu parente e instrutor Sérgio, obrigada pela vossa boa energia, por tantos bons momentos com gargalhadas e boa disposição, por tantas aulas divertidas que sempre ajudaram a descomprimir depois do trabalho e por todo o vosso apoio, incentivo e amizade!

À Sofia, Nádia, Nanda, Carlos, Luciana e Francisca, cada um com uma amizade particular e especial, o meu obrigada por me acompanharem e apoiarem ao longo destes anos com dedicação, por estarem sempre presentes e me fazerem sorrir nos bons e maus momentos. Um agradecimento especial à Luciana, por tantas horas de boa companhia no laboratório enquanto trabalhavámos até tarde e por estar sempre pronta a ajudar e a trocar ideias quando algo corria mal, assim como à Francisca, por estar sempre disponível para ajudar e por todas as sugestões e “desbloqueios de inglês” que tão importantes foram em algumas fases da escrita desta tese. Meninas, obrigada por na fase final me aturarem e aguentarem tantas vezes, pelos conselhos, pelas palavras de apoio e incentivo que tanto me animaram e deram força para caminhar...

A toda a minha família, em especial às minhas primas Romy e Isabel e ao meu primo Ricardo, obrigada pelo carinho, pelo apoio e incentivo constantes e por me fazerem sorrir sempre presentes. Ao Ricardo agradeço ainda a ajuda na formatação de algumas figuras apresentadas nesta tese.

Aos meus meninos, Gaspar, Lucas e Matilde, obrigada por serem sempre tão “terapêuticos” e pelo miminho, conforto e companhia em tantas horas de escrita.

Aos meus pais, Inocêncio e Elvira, a quem devo tudo o que sou... pelo apoio e carinho incondicionais, por estarem sempre ao meu lado e pela confiança que sempre depositaram em mim e no resultado de todo o meu trabalho e empenho ao longo destes anos, o meu mais sincero obrigada...

SUMMARY

Splicing of pre-mRNA is a crucial regulatory stage in the pathway of gene expression. The majority of human genes that encode proteins undergo alternative splicing, being now known that mutations affecting splicing are more prevalent than previously thought. According to some estimates up to 50-60% of all pathogenic mutations may affect splicing in some way. Thus, the development of therapeutic strategies targeting RNA represents an important opportunity to correct faulty splicing, opening the prospects of treatment for numerous genetic disorders. The vast majority of RNA-based approaches have exploited, *in vitro* and *in vivo*, antisense oligonucleotide sequences to either block the use of natural or new splice sites, inducing the skipping of defective exons, or blocking the newly generated cryptic sites to favour the use of the canonical ones. On the other hand, the use of U1 snRNA complementary to the mutated site has been described as a potentially therapeutic strategy to correct 5' splice donor site (SDS) defects dependent on U1 binding.

Following the rule, many splicing mutations have been identified in cohorts of patients with Lysosomal Storage Disorders (LSDs), a group of approximately 50 inherited metabolic diseases, usually associated with severe and progressive phenotypic manifestations due to specific lysosomal dysfunctions. Some treatment strategies are already available for conventional LSDs, but yet with some limitations. Therefore, in the case of LSDs, therapeutics directed to splicing mutations might represent a crucial option or an important adjunct of other treatments.

In this work attention was focused on the analysis of mechanisms disrupting splicing in a set of mutations involved in the pathogenesis of LSDs, and on the design and development of RNA-based therapies for the correction of the effects of some of those mutations.

The *IDS* gene encodes iduronate-2-sulphatase, an enzyme deficient in the X-linked lysosomal storage disorder - Mucopolysaccharidosis type II (MPS II). For the three nucleotide changes in the *IDS* gene associated with disease; the missense mutation c.257C>T, the nonsense c.241C>T (both in exon 3) and the synonymous c.1122C>T (exon 8), transcript analysis and minigene functional assays revealed that the first two cause the pathogenic activation of a 3' cryptic splice site in exon 3, whereas c.1122C>T leads to the creation of a new 5' splice site in exon 8. The cryptic splice site activation in exon 3 suggests that the use of its 3' constitutive splice site requires the interaction of several auxiliary *trans*-acting factors whose *cis*-binding motifs were changed by the mutations. Given these findings, an *in silico* analysis was performed to search for mutation-induced alterations in auxiliary elements located in

IDS exon 3. The elimination of a binding motif for the SRSF2 and hnRNP E1 and E2 splicing factors and the creation of one for the SRSF1 protein was predicted for the missense mutation, c.257C>T. The overexpression of the SRSF2 and hnRNP E1 proteins along with the mutant minigene, as well as an assay with other minigene presenting the deletion of a 6-base pair sequence which comprises the predicted *cis*-acting motifs for the binding of these proteins showed that they may be involved in the use and repression of the constitutive 3' splice site of exon 3, respectively.

For the synonymous mutation c.1122C>T which creates a new 5' splice site within exon 8 causing the deletion of its last 60 base pairs, we investigated the applicability of antisense therapy to correct the splicing defect. Yet, transfection of three antisense morpholino oligonucleotides (AMO) and one locked nucleic acid (LNA) in patient fibroblasts did not abolish the abnormal transcript, resulting in addition in the production of a new aberrant splicing product. Also, the antisense oligonucleotides transfection in control fibroblasts produced the same aberrant transcript detected in patient cells following treatment, evidencing that the oligonucleotides are masking important *cis*-acting elements for the 5' splice site regulation of exon 8. These results pointed out that the development of antisense therapies remains a challenge in genes under fine splicing regulation, as appears to be *IDS*.

In this work we also explored the use of different modified U1 snRNAs to correct three SDS mutations (c.234+1G>A, c.633+1G>A and c.1542+4dupA) reported in patients with Mucopolysaccharidosis type IIIC (MPS IIIC), a LSD caused by mutations in the *HGSNAT*, which encodes the enzyme acetyl CoA: α-glucosaminide acetyltransferase. The partial correction (almost 50%) of c.234+1G>A (intron 2) was achieved in patients fibroblasts with a modified U1 snRNA matching all the nucleotides of the mutated SDS. In the case of the other mutations, no rescue was observed after the modified U1 snRNAs overexpression.

Finally, we developed antisense oligonucleotide and U1 snRNA-mediated therapeutic strategies for the correction of a SDS mutation in the *CSTB* gene, which encodes cystatin B, a key protein in the protection against the proteinases leaking from lysosomes, whose deficiency is associated with the most common form of progressive myoclonic epilepsy: Unverricht-Lundborg disease (ULD). The synonymous mutation here studied, c.66G>A, located in the last nucleotide of the exon 1, causes the partial inclusion of intron 1 due to the activation of a cryptic splice site inside the intron. A specific LNA oligonucleotide designed to block the 5' cryptic splice site allowed to recover the normal splicing pattern in patient cells, although the therapeutic effect was shown to be dose-dependent. Concerning the U1 approach, the transfection of different adapted U1 vectors failed to restore correct splicing. A careful analysis of the wild-type

and mutant exon 1 SDS suggested that splicing regulation around this exon depends not only on the U1 snRNA binding but also on other splicing factors that could interact with the *CSTB* pre-mRNA compromising the success of the therapy.

Globally, the results obtained in this work allowed to clarify some aspects of the basic molecular mechanisms underlying the pre-mRNA splicing and its dysregulation in the presence of splicing mutations, further contributing to pave the way for innovative RNA-based therapeutic strategies that could ultimately result in patient-tailored correction of splicing mutations.

KEY WORDS: Lysosomal Storage Disorders (LSDs), Splicing regulation, Antisense oligonucleotide therapy, U1 snRNA-mediated therapy

RESUMO

O *splicing* do pré-mRNA é um processo fundamental para a regulação da expressão génica. A maioria dos genes humanos codificadores de proteínas apresenta *splicing* alternativo, sabendo-se hoje que mutações que interferem com o *splicing* conduzindo à produção de proteínas anormais e/ou deficientes são mais frequentes do que inicialmente previsto. De acordo com algumas estimativas, cerca de 50 a 60% de todas as mutações patogénicas descritas afetam de alguma forma o processo. Assim, o desenvolvimento de terapias que têm por alvo o RNA representa uma importante oportunidade de correção direcionada de erros de *splicing,* abrindo perspetivas de tratamento a um grande número de doenças genéticas.

A grande maioria destas abordagens tem explorado *in vitro* ou *in vivo*, o uso de oligonucleótidos *antisense* com o objetivo de bloquear locais de *splicing* constitutivos ou criados *de novo*, induzindo desta forma o *skipping* de exões que apresentam mutações patogénicas ou o bloqueio de locais aberrantes de *splicing* promovendo a utilização dos locais convencionais. Por outro lado, o uso de vetores U1 snRNA com diferentes complementaridades de ligação ao local mutado na região 5' do *splicing*, tem sido considerado uma promissora abordagem terapêutica para a correção de mutações que afetam a ligação do U1 snRNA.

Não fugindo à regra, mutações que alteram o processo de *splicing* também se encontram em doentes diagnosticados com Doenças Lisossomais de Sobrecarga (DLS). As DLS são um grupo de cerca de 50 doenças hereditárias do metabolismo geralmente caracterizadas pela apresentação de um fenótipo grave e progressivo devido a uma disfunção lisossomal específica. Atualmente estão disponíveis algumas estratégias de tratamento, mas dirigidas a um pequeno número de DLS e sempre associadas a sérias limitações. Assim, no caso destas patologias, o desenvolvimento de abordagens terapêuticas baseadas na correção do *splicing* pode constituir uma opção de maior relevância, quer individualmente, quer como adjuvante de tratamento.

Este trabalho teve como objetivos principais: 1) analisar os mecanismos de *splicing* e sua desregulação na presença de mutações envolvidas na patogénese lisossomal, e 2) contribuir para o desenvolvimento de abordagens terapêuticas baseadas na correção dos efeitos causados por algumas dessas mutações.

O gene *IDS* codifica a enzima iduronato-2-sulfatase, cujo défice causa a Mucopolissacaridose tipo II (MPS II), uma DLS com modo de transmissão ligado ao X. Relativamente a três alterações nucleotídicas no gene *IDS* associadas a doença - a mutação *missense* c.257C>T, a *nonsense* c.241C>T (ambas localizadas no exão 3) e a mutação sinónima c.1122C>T (exão 8) - a caracterização dos respetivos transcritos

e a análise funcional através de minigenes permitiram verificar que as mutações c.257C>T e c.241C>T causavam a ativação patogénica de um local de *splicing* críptico localizado na extremidade 3' do exão 3, e que a alteração c.1122C>T estava na origem da criação de um novo local de *splicing* na região 5' do exão 8. A ativação do local críptico no exão 3 sugere que o uso do local de *splicing* constitutivo a 3' requer a interação de vários elementos *cis-acting* e fatores *trans-acting* que devem ter sido alterados pela presença destas mutações. Foi então efetuada uma análise *in silico* para as duas mutações do exão 3, permitindo prever que a eliminação de um motivo de ligação para os fatores de *splicing* SRSF2 e hnRNP E1 e E2, assim como a criação de um motivo de ligação para o fator SRSF1 ocorreriam na presença da mutação c.257C>T. Os resultados tanto da sobrexpressão das proteínas SRSF2 e hnRNP E1 em co-transfeção com um minigene que continha a sequência mutada, como de um ensaio com um minigene onde foi efetuada a deleção de uma sequência de 6 pares de bases que inclui os motivos *cis-acting* previstos para a ligação destas proteínas, indicaram que ambas as proteínas poderão estar envolvidas respetivamente, no uso e repressão do local de *splicing* constitutivo a 3' do exão 3.

No caso da mutação sinónima c.1122C>T, que origina um novo local de *splicing* a 5' no exão 8 causando a deleção dos seus últimos 60 pares de bases, a aplicabilidade da terapia *antisense* foi investigada com o objetivo de corrigir o defeito de *splicing*. Contudo, a transfeção de fibroblastos de um doente com três oligonucleótidos *antisense* do tipo morpholino e de um oligonucleótido do tipo *locked nucleic acid* (LNA) não permitiu a eliminação do transcrito anormal, resultando até na produção de um novo transcrito aberrante. Para além disso, a transfeção dos mesmos oligonucleótidos em fibroblastos de um controlo originou o mesmo transcrito aberrante observado nas células do doente após o tratamento, demonstrando que os oligonucleótidos ocultaram importantes motivos *cis-acting* envolvidos na regulação do local 5' de *splicing* do exão 8. Estes resultados evidenciam não só a complexidade da regulação do *splicing* no gene *IDS*, como a dificuldade em desenvolver com sucesso, em casos deste tipo, terapêuticas dirigidas à correção de anomalias no processo.

Neste trabalho também foi explorado o uso de diferentes vetores U1 snRNA modificados para a correção de três mutações que afetam o local 5' de *splicing* (c.234+1G>A, c.633+1G>A e c.1542+4dupA) presentes em doentes com Mucopolissacaridose tipo IIIC (MPS IIIC), um tipo de DLS causada por mutações no gene *HGSNAT* que codifica a enzima Acetil-CoA: α-glicosamina-N-acetiltransferase. A correção parcial (cerca de 50%) da mutação c.234+1G>A (intrão 2) foi conseguida em fibroblastos de um doente com a utilização de um vetor U1 snRNA modificado que emparelhava na totalidade com o local 5' de *splicing* mutado. Para as restantes

mutações, após a sobrexpressão de diferentes vetores U1 snRNA não se observou a reversão do padrão de *splicing* aberrante.

Por fim, ensaiou-se a correção de uma mutação que afeta um local 5' de *splicing* no gene *CSTB,* codificador da proteína cistatina B que desempenha um papel importante na proteção celular contra a ação de proteases lisossomais. O défice desta proteína encontra-se associado à forma mais comum de epilepsia mioclónica progressiva, a doença de Unverricht-Lundborg. A alteração sinónima alvo deste estudo, c.66G>A localiza-se no último nucleótido do exão 1 do gene *CSTB* e a sua presença leva à inclusão parcial do intrão 1 devido à ativação de um local de *splicing* críptico a 5'. Um oligonucleótido do tipo LNA especificamente desenhado para bloquear o local críptico a 5' permitiu recuperar o normal padrão de *splicing* nas células de um doente, embora o efeito terapêutico fosse dependente da dose de oligonucleótido. Recorreu-se também a outra abordagem com vetores U1 snRNA modificados, que após transfeção não permitiram obter qualquer correção do *splicing* aberrante. A análise detalhada dos locais 5' de *splicing wild-type* e mutado do exão 1 revelou que a regulação do *splicing* envolvendo este exão depende não só da ligação do U1 snRNA, como também de outros fatores de *splicing* que poderão interagir com o pré-mRNA do gene *CSTB* comprometendo assim o sucesso da terapia com estes vetores.

Globalmente, os resultados obtidos com a realização deste trabalho permitiram clarificar alguns aspetos do mecanismo molecular subjacente ao processo de *splicing* e sua desregulação na presença de mutações, assim como avançar no caminho inovador do desenvolvimento de abordagens terapêuticas baseadas na correção do *splicing* que poderão ser aplicadas de forma personalizada em doentes portadores de mutações de *splicing*.

PALAVRAS-CHAVE: Doenças Lisossomais de Sobrecarga (DLS), Regulação do *splicing*, Terapia com oligonucleótidos *antisense*, Terapia mediada por vetores U1 snRNA.

TABLE OF CONTENTS

ABBREVIATIONS

µg	microgram
2'-MOE	2'-O-methoxyethyl
2'-OMe	2'-O-methyl
AAV	adeno-associated virus
AG	adenine/guanine
ALS	amyotrophic lateral sclerosis
AU	adenine/uracil
AO	antisense oligonucleotide
Arg-Rich	arginine-rich
AS	alternative splicing
ATP	adenosine triphosphate
BBB	blood brain barrier
BMT	bone marrow transplantation
BP	branch point
BPS	branch point sequence
cDNA	complementary DNA
CNS	central nervous system
CPP	cell-penetrating peptide
DLS	doença lisossomal de sobrecarga
DMD	Duchenne muscular dystrophy
DNA	deoxyribonucleic acid
EPEI	ethoxylated polyethylenimine
EPM1	progressive myoclonic epilepsy type 1
ERT	enzyme replacement therapy
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
ExSpeU1	exon-specific U1

FTDP-17	fronto-temporal dementia and parkinsonism - 17
GC	guanine/cytosine
GlcNAc	N-acetylglucosamine
GT	guanine/thymine
GU	guanine/uracil
h	hour
hFVII	human coagulation factor VII
HGMD	human gene mutation database
hnRNP	heterogeneous nuclear ribonucleoprotein
HSCT	haematopoietic stem cells transplantation
IMD	inherited metabolic disease
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
KH	K homology
LNA	locked nucleic acid
LSD	lysosomal storage disorder
mer	the length of an oligonucleotide (from Greek *meros* “part”)
ML	mucolipidosis
MPS	mucopolysaccharidosis / mucopolissacaridose
mRNA	messenger ribonucleic acid
NCL	neuronal ceroid lipofuscinose
NMD	nonsense-mediated mRNA decay
ORF	open reading frame
PD	pharmacodynamic
Pip	PMO internalization peptide
PK	pharmacokinetic
PMO	phosphorodiamidate morpholino oligomer
Pol II	polymerase II
poly-A	polyadenylation

PPMO	peptide-linked phosphorodiamidate morpholino oligomer
PPT	polypyrimidine tract
pre-mRNA	precursor mRNA
PS	phosphorothioate
PTC	premature termination codon
RGG	arginine/glycine/glycine
RNA	ribonucleic acid
RNase	ribonuclease
RNP	ribonucleoprotein
RP	retinitis pigmentosa
RRM	RNA recognition motif
RS	arginine/serine rich
RT-PCR	reverse transcriptase-polymerase chain reaction
SDS	splice donor site
SF1	splicing factor 1
SMA	spinal muscular atrophy
SMaRT	spliceosome-mediated RNA *trans*-splicing
SNP	single nucleotide polymorphism
snRNA	small nuclear ribonucleic acid
snRNP	small nuclear ribonucleoprotein
SR	serine/arginine
SRE	splicing regulatory element
SRSF	serine/arginine-rich splicing factor
SRT	substrate reduction therapy
ss	splice site
SSO	splice switching oligonucleotide
tc-DNA	tricyclo-DNA
Tm	melting temperature
U2AF	U2 auxiliary factor

ULD	Unverricht-Lundborg disease
U.S.FDA	United States Food and Drug Administration
UTR	untranslated region
v-PMO	vivo-morpholino
WT	wild-type

CHAPTER 1

INTRODUCTION

1.1 Regulation of gene expression and the splicing process

Since in the late 1950s it was stated that DNA makes RNA and RNA makes proteins (Crick, 1958; Crick, 1970), much research was done to uncover how each step in the sequential transfer of information from nucleic acids to protein works (Gottlieb, 2003). The understanding of gene structure, expression and function has provided fundamental insights into the molecular biology of cells. Over the time it has become clear that cell behaviour is determined not only by the genes, but importantly, also by which of them are expressed at any given time. Expression of eukaryotic genes is a multistep process that includes transcription of the gene, 5' -end capping and 3' -end polyadenylation of the primary transcript, splicing and transport of the fully processed messenger RNA (mRNA) to the cytoplasm, where it can be translated into protein. Regulation of gene expression allows cells to adapt to changes in their environments and is responsible for the distinct activities of the multiple differentiated cell types that make up complex plants and animals (Cooper & Hausman, 2013; Singer & Green, 1997).

One of the most important stages in the chain of processes from DNA to protein is pre-mRNA splicing. Firstly identified in the late 1970s with the discovery of introns (Berget et al., 1977; Chow et al., 1977), pre-mRNA splicing is a highly regulated step-wise process in which intron sequences are removed and exons are joined together to generate a mature protein-coding mRNA transcript (the so-called constitutive splicing). Splicing reactions can be alternative, since exons and introns can be excised or included in variable ways in the final mRNA product. Alternative splicing (AS) is predicted to occur in the vast majority of mammalian genes, being a primary mechanism through which complex organisms can regulate protein expression and generate a diverse proteome from a relatively limited genome (De Conti et al., 2013; House & Lynch, 2008; Kelemen et al., 2013).

1.1.1 Core splicing signals and spliceosome assembly

The splicing reaction is carried out by the spliceosome, a dynamic molecular machine, which consists of five small nuclear ribonucleoprotein complexes (snRNPs) U1, U2, U4, U5 and U6, and up to 300 different structural proteins. Two of the main functions of the spliceosomal snRNPs are to recognise and interact with specific sequence elements (splicing signals) at the exon/intron boundaries such as the 5' and

3' splice sites (ss), branch point sequence (BPS) and polypyrimidine tract (PPT) (Figure 1.1A), and to engage the spliceosomal complex to catalyze an efficient splicing reaction. Spliceosome assembly (Figure 1.1B) begins with the recognition of the 5′ ss by the U1 snRNP and the binding of splicing factor 1 (SF1) to the branch point and of the U2 auxiliary factor heterodimer (U2AF65/35) to the polypyrimidine tract and 3′ terminal AG. This assembly is ATP-independent and results in the formation of the E complex, which is converted into the ATP-dependent, pre-spliceosomal A complex after the replacement of SF1 by the U2 snRNP at the branch point. Further recruitment of the U4/U6•U5 tri-snRNP leads to the formation of the B complex. This is followed by extensive conformational changes and remodelling, including the loss of U1 and U4 snRNPs, ultimately resulting in the formation of the C complex, which is the active spliceosome capable of catalyzing the transesterification chemistry required for exon ligation and lariat release (Chen & Manley, 2009; De Conti et al., 2013; Matera & Wang, 2014). Although the spliceosome catalyzes RNA cleavage and ligation with high fidelity, the inherent flexibility of this enzymatic complex allows it to be highly sensitive to regulation. Such flexibility has played a fundamental role in the evolution of higher organisms by promoting proteome expansion and helping to regulate gene expression (De Conti et al., 2013; House & Lynch, 2008).

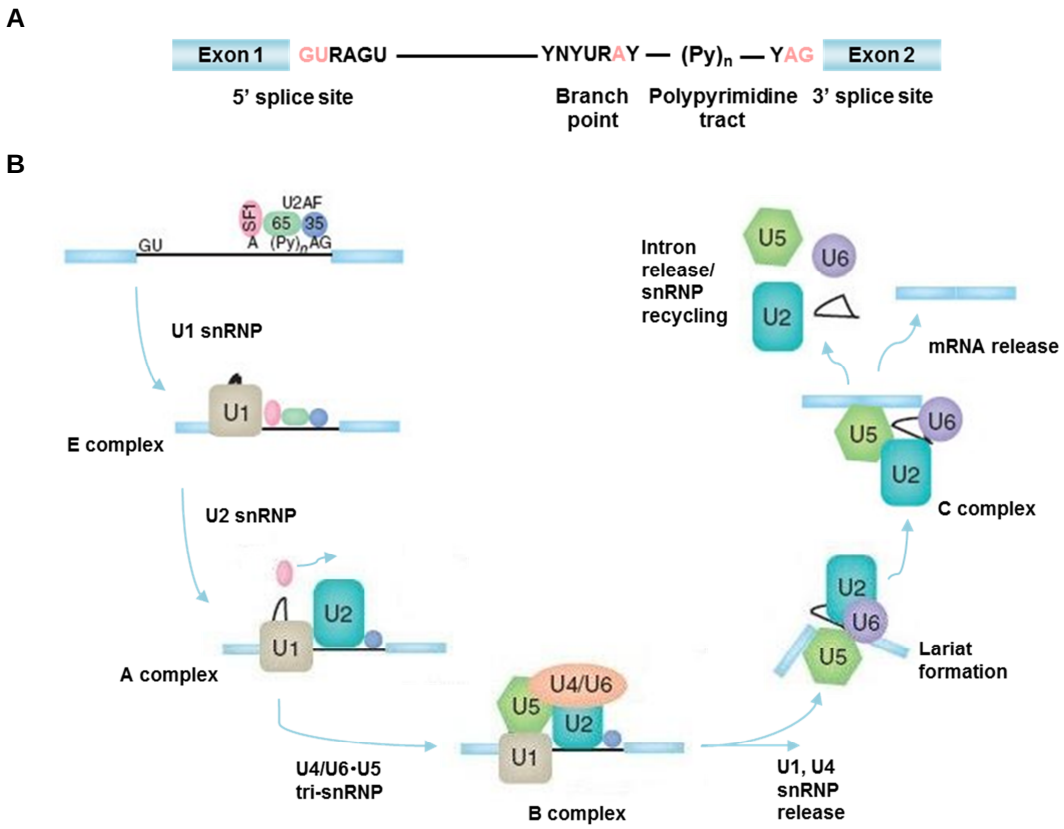


Figure 1.1: A) The architecture of a pre-mRNA and the important *cis*-acting sequence elements that direct the splicing reaction. The consensus sequences at the 5' ss, BPS and 3' ss are as indicated, where N is any nucleotide, R is a

purine and Y is a pyrimidine. The polypyrimidine tract (Py) is a pyrimidine-rich stretch located between the BPS and the 3' ss. **B)** Schematic representation of spliceosome assembly and the splicing of a pre-mRNA. The first step in spliceosome assembly is the formation of early (E) or commitment complex in which the 5' ss is bound by the U1 snRNP and the splicing factors SF1 and U2AF cooperatively recognise the BPS, the Py and the 3' ss. This pairing is stabilized by addition of ATP, which allows binding of the U2 snRNP to the BPS to form the A complex. The U4/U6•U5 tri-snRNP then joins as a single unit to form the B complex. Multiple ATP-dependent rearrangements result in the release of the U1 and U4 snRNPs and formation of the C complex, which catalyzes excision of the intron as a lariat and ligation of the exon sequences. Adapted from Hartmann et al., 2008; Rymond, 2007; Wahl et al., 2009

1.1.2 Exon recognition – a key multifactorial step in splicing regulation

A critical challenge to the spliceosome is to correctly identify exons within the pre-mRNA. Exons make up only one-tenth of the typical pre-mRNA and therefore must be identified within a sea of introns. Despite the splicing mechanism has been determined in great detail, it is not yet fully understood how splice sites are selected. The major problem is the degeneracy of splicing regulatory sequences, such as the 5', 3' ss, branch point and exonic/intronic sequence elements. These can only be described as consensus sequences that are necessary but not sufficient for defining intron-exon junctions (Hertel, 2008; Stamm et al., 2005; Ward & Cooper, 2010). Thus, the accurate mechanism of exon recognition *in vivo* is now believed to be the result of a complex combinatorial control depending on multiple parameters such as splice site strength, the presence or absence of splicing regulators (silencer or enhancer sequences), the exon/intron architecture, RNA secondary structures, transcriptional processivity by RNA polymerase II (pol II), nucleosome positioning and histone modifications at the chromatin level. Depending on local contexts, all these factors will act either antagonistically or synergistically to decide the exon/intron fate of any given RNA sequence (De Conti et al., 2013; Hertel, 2008; Kornblihtt et al., 2013; Wang & Burge, 2008). The factors that present greater relevance for this work will be considered in more detail in the following sections.

1.1.2.1 Splice site strength and gene architecture roles in exon-intron definition

The adjacent nucleotides at each splice site comprise two easily distinguished consensus sequences: the 5' ss junction defined by a single element of 9 bases and the 3' ss loosely defined by three sequence elements (branch point sequence, polypyrimidine tract and the 3' exon/intron junction), which may extend 40 nucleotides into the intron, upstream to the 3' exon/intron junction (Reed, 1996). The initial

recognition of exon/intron junctions is based on the relative strength of the splice sites (i.e., how near to consensus is their sequence) that play a major role in determining whether the basic snRNP factors will be able to interact and bind. Because the sequence specificity of the splice site/snRNP binding is driven by their interaction level and the U2 auxiliary factor binding preference for polypyrimidines, splice sites are classified by their complementarity to U1 snRNA (small nuclear RNA) (5' ss) and the extent of the polypyrimidine tract (3' ss). Greater complementarity with U1 snRNA and longer polypyrimidine tracts translate into higher affinity binding sites for these spliceosomal components and thus more efficient exon recognition (De Conti et al., 2013; Hertel, 2008).

Most human genes contain relatively short exons (typically, 50-250 base pairs in length) separated by much larger introns (frequently hundreds to thousands of base pairs) that on average account for >90% of the primary transcript (Chen & Manley, 2009; Wang & Burge, 2008). This relative length of introns and exons – exon/intron architecture – has also been shown to have an influence on splice site recognition (Berget, 1995). Early experiments aimed to investigate the influence of exon and intron sizes in vertebrate pre-mRNA processing showed that enlarged exons lead to exon skipping, but if the flanking introns are short, the enlarged exon is included (Sterner et al., 1996). Therefore, splice site recognition seems to be more efficient when introns or exons are small. These observations suggested that splice sites are recognised across an optimal nucleotide length, meaning thus that intron length might significantly influence the efficiency of pre-mRNA splicing and alternative splice site choice. On this ground, two models were proposed for the mechanism of exon and intron selection: "intron definition" and "exon definition". The "intron definition" model posits that interactions occur first across the intron between factors recognizing the 5' ss and the downstream 3' ss, whereas the alternative model, "exon definition", sustains that the splice sites flanking the exon are initially recognised and subsequently paired (Berget, 1995). The higher GC content in exons relative to their flanking introns is presumed to be the signal that allows exons to be identified (Keren et al., 2010). Considering that in the human genome the vast majority of the exons are short and introns are long it is likely that most of the human splice sites are recognised across the exon (Sakharkar et al., 2005). In fact, coevolution analyses of the 5' and 3' ss have detected predominant cross-exon interactions in human and mouse, but cross-intron interactions in invertebrates, plants, and fungi, supporting the primacy of exon definition in mammals and intron definition in most other metazoans (Xiao et al., 2007). On a practical level, however, it is important to bear in mind that there is no mechanistic difference in spliceosomal complex assembly over exons or introns and that both exon definition

and intron definition may occur in different parts of the same pre-mRNA (De Conti et al., 2013).

With the increasing knowledge on the process, several bioinformatic algorithmic models have been developed to calculate the splice site consensus values (i.e. to predict the splice site strength). Presently, there are several online tools that analyse the strengths of splice sites and the changes induced by different sequence variations (e.g. mutations). The more widely used are: *GENSCAN*, available at http://genes.mit.edu/GENSCAN.html (Burge & Karlin, 1997), *Splice Site Prediction by Neural Network*, available at http://www.fruitfly.org/seq_tools/splice.html (Reese et al., 1997), *MaxEntScan*, available at http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html (Yeo & Burge, 2004), and *Human Splicing Finder*, available at http://www.umd.be/HSF/ (Desmet et al., 2009).

1.1.2.2 The role of *cis*-acting elements and *trans*-acting factors in exon-intron definition

Besides the core splicing signals, additional sequences are needed in exons and introns to increase the overall fidelity of the splicing reaction. Indeed, it is now widely accepted that the majority of pre-mRNA molecules contains a myriad of auxiliary *cis*-acting regulatory elements that either enhance or inhibit exon-intron recognition. These elements are conventionally classified as exonic splicing enhancers (ESEs) or silencers (ESSs) if from an exonic location they function to promote or inhibit inclusion of the exon they reside in, and as intronic splicing enhancers (ISEs) or silencers (ISSs) if they enhance or inhibit usage of adjacent splice sites or exons from an intronic location (Figure 1.2) (Chen & Manley, 2009; De Conti et al., 2013; Kornblihtt et al., 2013; Wang & Burge, 2008). In general, these splicing regulatory elements (SREs) share several common features: they are small (6-8 nucleotides), individually weak, and present in multiple copies. They are often conserved between species and perhaps between similarly regulated genes, but they contain degenerate sequence motifs, making it difficult to identify them (Hartmann et al., 2008; Jensen et al., 2009; Ladd & Cooper, 2002). Enhancers and silencers are involved in both constitutive and alternative splicing and function by recruiting *trans*-acting splicing factors that activate or supress splice site recognition or spliceosome assembly by various mechanisms (De Conti et al., 2013; Wang & Burge, 2008). The ESE motifs are usually recognised by members of the SR (serine/arginine) protein family that act as positive regulators promoting exon inclusion. These *trans*-acting proteins are characterized by having one or more RNA recognition motifs (RRMs) as well as an RS (arginine/serine-rich) domain

responsible for protein–protein interactions (Busch & Hertel, 2012; De Conti et al., 2013; Jensen et al., 2009; Risso et al., 2012; Shepard & Hertel, 2009). Unlike ESEs, the proteins that mediate the effect of ISEs are less well characterized although several tissue-specific proteins, such as NOVA (Jelen et al., 2007; Ule et al., 2006) and FOX2 (Yeo et al., 2009) have been shown to bind ISEs and to stimulate splicing.

On the other hand, negative regulation of exon recognition by ESS and ISS elements is carried out by the heterogeneous nuclear ribonucleoprotein protein (hnRNP) protein family. Similar to SR proteins, hnRNPs direct their influence on pre-mRNA splicing through site-specific binding mediated by a RRM, with the exception of hnRNPs E/K, which interact with RNA via the K homology (KH) domain. Most hnRNPs also harbor auxiliary Arginine-Glycine-Glycine (RGG) boxes and glycine-patch domains that are often involved in protein–protein interactions (Busch & Hertel, 2012; Chaudhury et al., 2010; De Conti et al., 2013; Jensen et al., 2009). Although much remains to be learned about how these factors function, they can repress splicing either by sterical hindrance or by promoting the formation of inhibitory RNA secondary structures (De Conti et al., 2013).

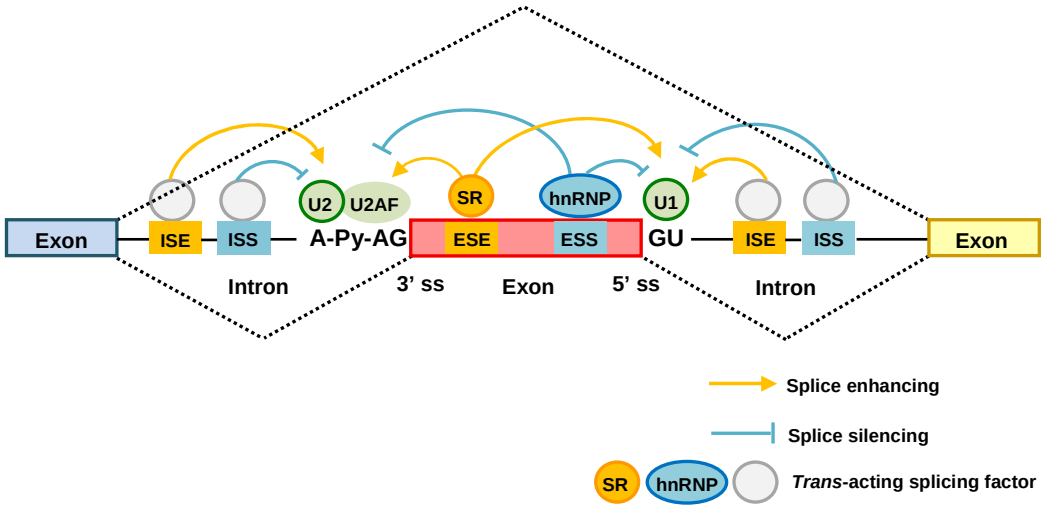


Figure 1.2: Schematic diagram of key pre-mRNA splicing regulatory elements. Splice site choice is regulated not only by the core *cis*-sequence elements (5' and 3' ss) and associated 3' sequences (polypyrimidine tract, Py; and branch point sequence, A) but also by additional enhancer and silencer elements in exons and introns (ESE/ESS: exonic splicing enhancers/silencers; ISE/ISS: intronic splicing enhancers/silencers) and by two main families of *trans*-acting splicing factors, Ser/Arg-rich proteins (SRs) and heterogeneous nuclear ribonucleoproteins (hnRNPs). These regulatory proteins as other *trans*-acting factors bind to enhancers and silencers and interact with spliceosomal components (shown in green) that associate with both the 5' and the 3' ss flanking the alternative exon and can have either activating (yellow arrow) or inhibitory (blue arrow) effects on the recognition and use of that site. Based in Jensen et al., 2009; McManus & Graveley, 2011; Wang & Burge, 2008

Frequently, the binding of SR proteins to enhancer motifs and hnRNPs to silencer elements promote the inclusion or exclusion of an exon, respectively. However, it

should be noted that the same sequence motif can act as an enhancer or as a silencer, depending on its position with respect to the splice sites (Hartmann et al., 2008; Zhou & Fu, 2013). The activities of *cis*-acting elements were shown to be context-specific and there is compelling evidence that SR proteins (e.g. SRSF1, SRSF2, SRSF7) can suppress splicing when bound to sequences located within the intron. There are also examples of members of the hnRNPs (e.g. hnRNP F/H, TIA-1) exhibiting stimulating effects on splicing (Dauksaite & Akusjärvi, 2002; Dembowski et al., 2012; Erkelenz et al., 2013; Schaub et al., 2007). Additionally, the overall affinity, frequency and concentration ratio of the different protein family members can also influence the final decision of whether an exon is constitutively or alternatively included in the mature mRNA transcript (Busch & Hertel, 2012; Chen & Manley, 2009; Jensen et al., 2009). Globally, the sum of SRs and hnRNPs activities to enhance or repress *cis*-acting elements contributes to the overall recognition of an exon or to the overall binding affinity for the spliceosome.

In the post-genomic era, research into the mechanisms of splice site selection is leading toward the establishment of rules that will allow splice patterns to be predicted on the basis of sequence information. Computational methods combined with laboratory experiments have already generated algorithm programs that predict sequence motifs of SREs (Hertel, 2008; Wang & Burge, 2008) as for example *ESEfinder*, available at http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home (Cartegni et al., 2003), *ESRsearch*, available at http://ibis.tau.ac.il/ssat/ESR.htm (Goren et al., 2006) and *RESCUE-ESE*, available at http://genes.mit.edu/burgelab/rescue-ese/ (Fairbrother et al., 2002; Yeo et al., 2004). These programs are useful tools that help to predict specific *cis*-acting elements and the respective *trans*-acting binding proteins in particular genes. However, they still often fail to identify these elements accurately and to make exact predictions of the effect of a given genomic mutation on the splicing process. Therefore, experimental analysis should be always considered to validate “dry-lab” predictions.

The relative contributions of each of these factors, account to control how efficiently splice sites are recognised and flanking introns are removed. Therefore, the splicing process can be viewed as a series of checkpoints that allow the spliceosome to “examine” different choices and to adjust splicing decisions according to local dictates and cellular requirements. Such flexibility and control ultimately provide for the extent of alternative splicing that is now recognised to be pervasive in higher eukaryotes and essential for the functional diversity required in complex organisms (Hertel, 2008; House & Lynch, 2008).

1.1.3 Alternative splicing: from one gene to many proteins

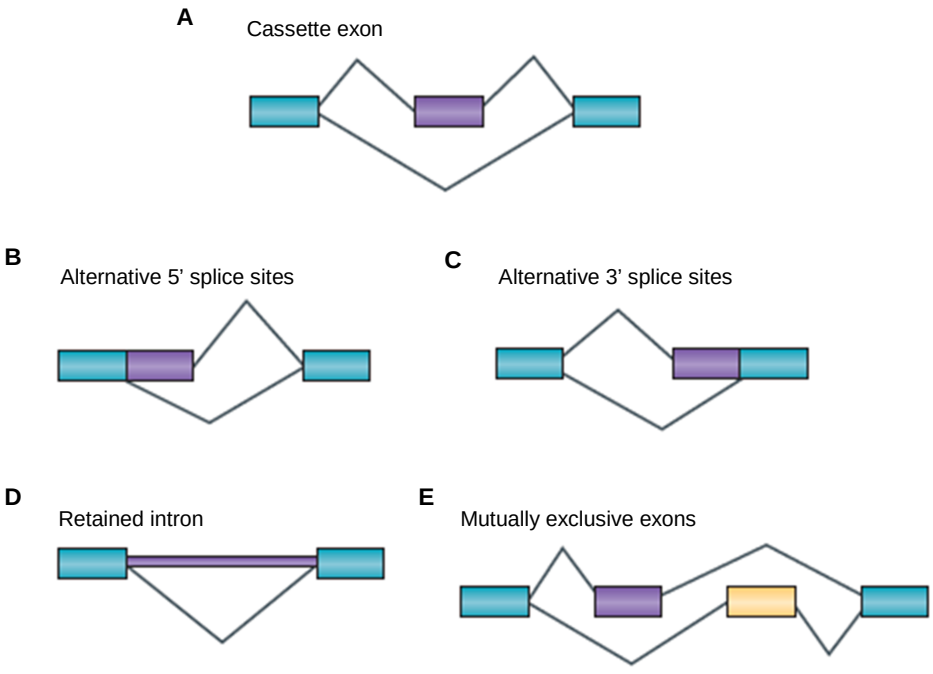
In humans, a myriad of proteins is needed to guarantee cellular functionality. This repertoire of different proteins can be largely explained by the occurrence of AS, a crucial mechanism for gene regulation that allows the production of several structurally and functionally different proteins from a single gene (Chen et al., 2012; Nilsen & Graveley, 2010). With the progressive generation of deep sequencing methodologies, it has been estimated that up to 95% of human multi-exon genes undergo AS (Pan et al., 2008; Wang et al., 2008), leading to an estimated 90,000 protein species in humans, despite there being only roughly 25,000 genes (Consortium, 2004; Kim et al., 2007). Because many AS events occur in translated regions of mRNAs, they can affect the sequence of the encoded proteins. Therefore, changes in the primary structure can impact several aspects of protein biological function and properties such as subcellular localization, stability/expression, enzymatic activity, molecular interactions, binding properties and posttranslational modifications. Consequently, AS seems to be a major factor in coordinating physiologically meaningful changes in protein isoform expression and is a key mechanism to generate the complex proteome of multicellular organisms (Kelemen et al., 2013; Stamm et al., 2005).

The potential to generate variability is enormous. Not only the majority of genes encode pre-mRNAs that are alternatively spliced but also the number of mRNA isoforms encoded by a single gene can vary from two to several thousands. For instance, the human neurexin-3 (*NRXN3*) gene can potentially form 1728 transcripts due to AS at four different sites (Tabuchi & Südhof, 2002) and in Drosophila, the Down syndrome cell adhesion molecule (*Dscam*) gene can potentially generate 38 016 distinct mRNA isoforms, a number far in excess of the total number of genes (~14,500) in the organism (Schmucker et al., 2000). The significance of AS extends beyond the ability to generate different protein isoforms to the ability to modulate the levels of those isoforms. The proportions of different splice forms produced by AS may vary in different cell contexts, such as by cell or tissue type, developmental stage, gender, apoptosis, disease state or in response to external stimuli (Kelemen et al., 2013; Pajares et al., 2007; Stamm et al., 2005).

AS can be regulated at different steps of spliceosome assembly by different splicing factors and by many mechanisms that rely on *cis*-acting elements (Chen & Manley, 2009; House & Lynch, 2008). Correct AS also depends on the stoichiometry and interactions of positive and negative regulatory proteins (SR and hnRNPs). Therefore, differences in the activities or amounts of general splicing factors and/or gene-specific splicing regulators during development or in different tissues can cause

the differential patterns of splicing observed between different cell types (Chen & Manley, 2009; Cáceres & Kornblihtt, 2002; Hertel, 2008; Kalsotra & Cooper, 2011). The production of several isoforms from the same transcription unit by various types of AS (Figure 1.3) is a very common event in mammalian cells, and makes the already formidable task of correctly identifying splice sites even more complex (Cartegni et al., 2002). Currently, there are several different types of AS events, which can be classified into four main subgroups. The first type is exon skipping, in which a type of exon known as a cassette exon is spliced out of the transcript together with its flanking introns (Figure 1.3A). The second and third types are alternative 5' ss and 3' ss selection. These types of AS events occur when two or more splice sites are recognised at one end of an exon (Figure 1.3B and 1.3C). The fourth type is intron retention, whereby an intron is retained in the resultant mature transcript (Figure 1.3D). Other classes of more complex events that give rise to alternative transcript variants include mutually exclusive exons (Figure 1.3E), alternative promoter usage (Figure 1.3F) and alternative polyadenylation sites (Figure 1.3G) (Chen et al., 2012; Gamazon & Stranger, 2014; Keren et al., 2010; Roy et al., 2013). Prevalence of each type of AS has been found to vary between different taxa. Several studies have shown that exon skipping is common in metazoan genomes (Alekseyenko et al., 2007; Kim et al., 2007; Sugnet et al., 2004) whereas intron retention is the most common type of AS among plants (Wang & Brendel, 2006) and fungi (Kim et al., 2008).

Globally, the flexibility of AS contributes not only to expand protein diversity but also to the clockwork functioning of human biology.



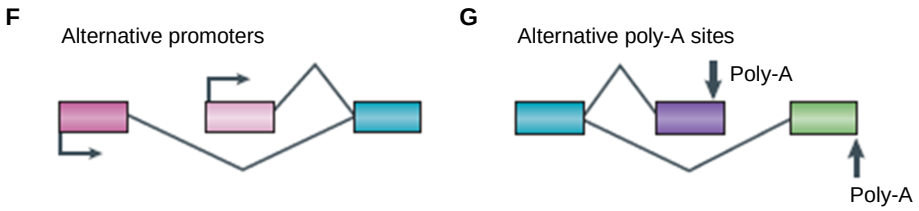


Figure 1.3: Representation of the major forms of alternative splicing. Through alternative use of exons, introns, promoters and polyadenylation sites, alternative splicing acts to greatly increase the diversity of mRNA transcripts. **A)** Cassette exon or exon skipping; **B)** Alternative 5' ss; **C)** Alternative 3' ss; **D)** Intron retention; **E)** Mutually exclusive exons; **F)** Alternative promoters; **G)** Alternative poly-A sites. Adapted from Li et al., 2007

1.1.4 Splicing and disease

Given the widespread functions of splicing, it is not surprising that mutations in either alternatively or constitutively spliced genes can trigger aberrant splicing, which can lead to human disease (Douglas & Wood, 2011; Kelemen et al., 2013).

Splicing defects constitute approximately 9% of all human pathogenic mutations (HGMD®Professional Release 2015.3) according to an estimate based on mutations at highly conserved splice sites. However, a growing understanding of *cis* splicing codes at locations other than splice sites has suggested that more than 50-60% of mutations that cause disease do so by disrupting splicing (Douglas & Wood, 2011; Lewandowska, 2013; López-Bigas et al., 2005).

Alterations in splicing can cause disease directly, modify the severity of the disease phenotype or be linked with disease susceptibility. In each case, the mechanisms causing altered splicing involve disruption of either *cis*-acting elements within the affected gene or *trans*-acting factors that are required for splicing regulation. Disruption of *cis*-acting sequence elements have a direct impact on the expression of only one gene, whereas alterations in *trans*-acting factors such as a component of the splicing machinery or a splicing regulator might affect the expression of multiple genes (Kornblihtt et al., 2013; Wang & Cooper, 2007; Ward & Cooper, 2010). Most of the known splicing mutations occur mainly in the canonical nucleotides GT and AG at the 5' and 3' ss, respectively, but can also disrupt other conserved splicing signals at the exon-intron junctions like the branch point and polypyrimidine tract sequences, and also auxiliary exonic or intronic splicing regulatory elements (Cartegni et al., 2002; Havens et al., 2013; Lewandowska, 2013; Tazi et al., 2005). The presence of such mutations can have different outcomes such as the complete or partial skipping of an exon, intron retention, introduction of a new splice site within an exon or intron or the activation of cryptic splice sites (Havens et al., 2013; Jensen et al., 2009). Mutations occurring deep within introns can also create or activate novel donor and acceptor

splice sites resulting in the inappropriate inclusion of intronic sequences, usually known as pseudoexons (Dhir & Buratti, 2010). In general, the use of unnatural splice sites or intron retention can origin non-functional proteins through the generation of out-of-frame transcripts which can lead to premature termination codons (PTCs) into the mRNA, typically resulting in degradation of the transcripts by nonsense-mediated mRNA decay (NMD) and loss of function of the mutated allele (Faustino & Cooper, 2003; Jensen et al., 2009).

The pathogenicity of exonic mutations is generally assumed to result from the predicted effect on the reading frame and protein function. However, it is likely that the primary pathogenic effect for many exonic mutations is at the level of splicing (Havens et al., 2013; Wang & Cooper, 2007). For example, a single nucleotide change that does not change the encoded amino acid of a protein (a synonymous or silent mutation) might be misclassified as a neutral polymorphism when in fact it causes disruption of a crucial ESE and is a disease-causing mutation. Similarly, base changes that would result in premature stop codons (that is, nonsense mutations) or in amino acid changes (that is, missense mutations) might not only be acting at the level of the encoded protein but also at the pre-mRNA level by affecting the fate of splicing events (Kornblihtt et al., 2013). Therefore, to better understand the impact of every mutation at mRNA level, the study of splicing patterns by cDNA analysis and/or by functional assays using minigenes as a splicing model system is often necessary (Baralle & Baralle, 2005; Cáceres & Kornblihtt, 2002). Minigenes are splice reporter vectors (also known as exon-trapping vectors) which consist of exonic portions of a gene with functional 5' and 3' ss separated by an intronic sequence where a polylinker (i.e. a multiple cloning site into which a fragment of sample DNA can be inserted) is located. Any genomic region of interest, for example, an exon with part of its intronic flanking regions, can be amplified from normal or affected individuals and cloned into the minigene. The resultant constructs are then transiently transfected into an appropriate cell line and the vector splicing patterns are analysed allowing to disclose or confirm the mutations' effect on the splicing process (Baralle et al., 2009; Desviat et al., 2012). In the last decade, minigenes have been widely used as a model system to study splicing mutations in several diseases as Spinal muscular atrophy (*SMN1* and *SMN2* genes) (Cartegni et al., 2006), Familial adenomatous polyposis (*APC* gene) (Gonçalves et al., 2009), Congenital disorder of glycosylation type IA (*PMM2* gene) (Vega et al., 2009) and Duchenne muscular dystrophy (*DMD* gene) (Disset et al., 2006), having often allowed to identify the *cis* and *trans*-acting elements involved in the disease mechanism, thus adding to the understanding of the observed pathogenic splicing alterations. In general, the use of minigene vectors for splicing assays is essential not

only for diagnostic purposes but also for unravelling novel splicing regulatory mechanisms and monitor therapeutic interventions aimed at modulating splicing or at correcting splicing defects (Desviat et al., 2012).

1.2 Lysosomal Storage Disorders: a brief overview and the burden of splicing mutations

Lysosomes are membrane-enclosed compartments within the cytosol of most cell types, involved in the degradation of certain lipids, sugars, proteins, and nucleic acids, as well as foreign substrates, such as substances and pathogens, which enter cells from the extracellular milieu. This function is enabled by acidic hydrolytic enzymes working in concert with small activator molecules within the lysosomal lumen. Also important for lysosomal function are transport proteins and ion pumps in the lysosomal membrane, which maintain a harsh acidic environment in this organelle (Sabatini & Adesnik, 2001). Inherited deficiency of any component of the lysosomal system leads to a group of rare diseases designated Lysosomal Storage Disorders (LSDs). LSDs are a large group of more than 50 different Inherited Metabolic Diseases (IMDs) which, in the great majority of cases, result from the defective function of specific lysosomal enzymes and, in few cases, from defects in lysosomal membrane proteins, non-enzymatic lysosomal proteins or non-lysosomal proteins involved in lysosomal biogenesis and vesicular traffic (Boustany, 2013; Filocamo & Morrone, 2011). The biochemical hallmark of these diseases is the accumulation of undegraded metabolites in the lysosome. This accumulation results in generalized cell and tissue dysfunction which can cause a chain of secondary disruptions to other biochemical and cellular functions, leading to multi-systemic pathology (Filocamo & Morrone, 2011; Parkinson-Lawrence et al., 2010; Platt et al., 2012). Storage may begin during early embryonic development, and the clinical presentation for LSDs can vary from an early and severe phenotype with rapid progressive neurological manifestations to late-onset mild disease (Filocamo & Morrone, 2011; Hopwood, 2012). LSDs can affect all systems of the body, with potential consequences in the skeleton and major organs, as well as the blood-forming and immune systems. Most of the diseases also affect the brain (Cox, 2012).

LSDs can be grouped according to various classifications. While, initially, they were classified according to the nature of the accumulated substrate(s), the more recent trend is to classify them on the basis of the molecular defect. Some LSDs are still referred to by the early class defined by the storage products as is the case of the

group of Mucopolysaccharidoses (MPSs) (accumulation of glycosaminoglycans also called mucopolysaccharides), Sphingolipidoses (storage of sphingolipids), Oligosaccharidoses also known as Glycoproteinoses (mostly characterized by the accumulation of oligosaccharides) and so on (Boustany, 2013; Filocamo & Morrone, 2011; Winchester, 2012). However, some other lysosomal disorders did not fitted easily into this classification scheme and only the elucidation of their molecular basis has permitted a new classification based on the nature of the molecular defect in the lysosomal system, which includes groups of disorders due to: (i) non-enzymatic lysosomal protein defects; (ii) integral membrane protein defects (transporters and structural proteins); (iii) lysosomal enzyme protection defects; (iv) post-translational processing defects of lysosomal enzymes; (v) trafficking defects in lysosomal enzymes; and (vi) polypeptide degradation defects. There is also another group that includes the neuronal ceroid lipofuscinoses (NCLs), which are considered to be lysosomal disorders, even though distinct characteristics exist (Table 1.1) (Filocamo & Morrone, 2011).

Table 1.1: Classification of Lysosomal Storage Disorders

Adapted from Boustany, 2013; Filocamo & Morrone, 2011; Winchester, 2012

Disease	Defective enzyme or protein	Main storage materials	Gene symbol
Sphingolipidoses			
Fabry	α-Galactosidase A	Galactosylated glycolipids	*GLA*
Farber	Acid ceramidase	Ceramide	*ASAH1*
Gaucher (Types I, II and III)	β-Glucosidase	Glucosylceramide	*GBA*
Niemann-Pick (Types A and B)	Sphingomyelinase	Sphingomyelin	*SMPD1*
Sphingolipid-activator deficiency	Sphingolipid activator	Glycolipids	*PSAP*
GM1 gangliosidosis (Types I, II and III)	β-Galactosidase	GM1 ganglioside, Keratan sulfate, oligos, glycolipids	*GLB1*
GM2 gangliosidosis (Tay-Sachs)	β-Hexosaminidase A	GM2 ganglioside, oligos, globoside, glycolipids	*HEXA*
GM2 gangliosidosis (Sandhoff)	β-Hexosaminidase A and B	GM2 ganglioside, oligos	*HEXB*
Krabbe disease (Globoid cell leukodystrophy)	β-Galactosylceramidase	Galactosylceramide	*GALC*
Metachromatic leukodystrophy	Arylsulfatase A	Sulfatides	*ARSA*

Mucopolysaccharidoses (MPS)			
MPS I (Hurler, Scheie, Hurler/Scheie)	α-L-Iduronidase	Dermatan sulfate, heparan sulfate	*IDUA*
MPS II (Hunter)	Iduronate-2-sulfatase	Dermatan sulfate, heparan sulfate	*IDS*
MPS IIIA (Sanfilippo A)	Heparan-N-sulfatase (sulfamidase)	Heparan sulfate	*SGSH*
MPS IIIB (Sanfilippo B)	N-Acetyl-α-glucosaminidase	Heparan sulfate	*NAGLU*
MPS IIIC (Sanfilippo C)	Acetyl-CoA:α-glucosaminide N-acetyltransferase	Heparan sulfate	*HGSNAT*
MPS IIID (Sanfilippo D)	N-Acetylglucosamine-6-sulfatase	Heparan sulfate	*GNS*
MPS IVA (Morquio A)	N-Acetylgalactosamine-6-sulfatase	Keratan sulfate, chondroitin-6-sulfate	*GALNS*
MPS IVB (Morquio B)	β-Galactosidase	Keratan sulfate	*GLB1*
MPS VI (Maroteaux-Lamy)	N-Acetylgalactosamine-4-sulfatase (arylsulfatase B)	Dermatan sulfate	*ARSB*
MPS VII (Sly)	B-Glucuronidase	Dermatan sulfate, heparan sulfate, chondroitin-6-sulfate	*GUSB*
MPS IX (Natowicz)	Hyaluronidase	Hyaluronan	*HYAL1*
Oligosaccharidoses (Glycoproteinoses)			
Aspartylglicosaminuria	Aspartylglucosaminidase (Glycosylasparaginase)	Aspartylglucosamine	*AGA*
Fucosidosis	α-L-Fucosidase	Glycoproteins, glycolipids, Fucoside-rich oligos	*FUCA1*
α-Mannosidosis	α-D-Mannosidase	Mannose-rich oligosaccharides	*MAN2B1*
β-Mannosidosis	β-D-Mannosidase	Man (β1→4) GlnNAc	*MANBA*
Sialidosis (Mucolipidosis I)	Neuraminidase (Sialidase 1)	Oligos, glycopeptides	*NEU1*
Schindler/Kanzaki	α-N-Acetylgalactosaminidase	Sialylated/asialoglyco-peptides, glycolipids	*NAGA*
Glycogenoses			
Pompe (Glycogenosis type II)	α-1,4-glucosidase	Glycogen, oligos	*GAA*
Lipidoses			
Wolman, cholesterol ester storage disease (CESD)	Acid lipase	Cholesterol esters	*LIPA*
Non-enzymatic lysosomal protein defects			
GM2 gangliosidosis (GM2 activator deficiency)	GM2 activator protein	GM2 ganglioside, oligos	*GM2A*

Metachromatic leukodystrophy	Saposin B	Sulfatides	*PSAP*
Krabbe disease, atypical	Saposin A	Galactosylceramide	*PSAP*
Gaucher, atypical	Saposin C	Glucosylceramide	*PSAP*
Integral membrane protein defects			
Transporters			
Cystinosis	Cystinosin (cystine transporter)	Cystine	*CTNS*
Sialic acid storage disease; infantile form (ISSD) and adult form (Salla)	Sialin (sialic acid transporter)	Sialic and uronic acids	*SLC17A5*
Niemann-Pick type C1	Niemann-Pick type C1 protein (proton-driven transporter)	Cholesterol and sphingolipids	*NPC1*
Niemann-Pick type C2	Niemann-Pick type C2 protein (proton-driven transporter)	Cholesterol and sphingolipids	*NPC2*
Structural proteins			
Danon disease	Lysosome-associated membrane protein 2 (LAMP-2)	Cytoplasmic debris and glycogen	*LAMP2*
Mucolipidosis IV	Mucolipin	Lipids	*MCOLN1*
Lysosomal enzyme protection defects			
Galactosialidosis	Protective protein cathepsin A (PPCA)	Sialyloligosaccharides and glycopeptides	*CTSA*
Post-translational processing defects			
Multiple sulfatase deficiency	Formylglycine generating enzyme (SUMF1)	Sulfatides, GAGs, glycolipids	*SUMF1*
Trafficking defects in lysosomal enzymes			
Mucolipidosis II α/β (ML II or I-cell) Mucolipidosis III α/β (ML IIIA or pseudo-Hurler polydystrophy)	N-acetylglucosamine-1-phosphotransferase – α/β subunit	Oligos, GAGs, lipids	*GNPTAB*
Mucolipidosis III γ (ML III variant)	N-acetylglucosamine-1-phosphotransferase – γ subunit	Oligos, GAGs, lipids	*GNPTG*
Polypeptide degradation defects			
Pycnodysostosis	Cathepsin K	Collagen and other bone proteins	*CTSK*
Neuronal ceroid lipofuscinoses (NCLs)			
NCL 1	Palmitoyl protein thioesterase 1 (PPT1)	Lipofuscin, Saposins A and D	*PPT1*
NCL 2	Tripeptidyl peptidase 1 (TPP1)	Lipofuscin, Subunit c of ATP synthase	*TPP1*
NCL 3	CLN3, lysosomal and/or Golgi transmembrane protein	Lipofuscin, Subunit c of ATP synthase	*CLN3*

NCL 4	DnaJ homologue subfamily C member 5	Lipofuscin, Subunit c of ATP synthase	*DNAJC5*
NCL 5	CLN5, soluble lysosomal protein	Lipofuscin, Subunit c of ATP synthase	*CLN5*
NCL 6	CLN6, transmembrane protein in endoplasmic reticulum	Lipofuscin, Subunit c of ATP synthase	*CLN6*
NCL 7	CLN7/MFSD8 (major facilitator superfamily domain-containing protein 8), transporter	Lipofuscin, Subunit c of ATP synthase	*MFSD8*
NCL 8	CLN8, transmembrane protein in endoplasmic reticulum	Lipofuscin, Subunit c of ATP synthase	*CLN8*
NCL 10	Cathepsin D	Lipofuscin, Saposins A and D	*CTSD*
NCL 11	Progranulin	*	*GRN*
NCL 12	P type ATPase	*	*ATP13A2*
NCL 13	Cathepsin F	*	*CTSF*
NCL 14	Potassium channel tetramerization domain-containing protein 7	*	*KCTD7*

*The underlying genes of these four (NCL 11 – 14) variants have already been identified; however, data on their associated storage materials is still scarce.

The frequency of LSDs varies across populations and geographical regions. Although individually rare, the prevalence of LSDs is significant when the group is considered as a whole, varying from one case in every 4000 to 8000 births across different studies (Fuller et al., 2006; Meikle et al., 1999; Pinto et al., 2004; Poorthuis et al., 1999; Poupetová et al., 2010). Comparatively to other IMDs, the class of LSDs is relatively frequent in the Portuguese population, presenting a prevalence of 1:4000 live births (Pinto et al., 2004).

Most LSDs are autosomal recessive diseases, with the exception of Hunter syndrome (MPS II) and Fabry disease which are X-linked recessive; Danon disease, which shows X-linked dominant inheritance and the Parry disease (a variant type of adult NCL4), which is autosomal dominant (Boustany, 2013; Hopwood, 2012). Although being monogenic diseases, for most LSDs, numerous mutations have been described for the same gene in different patients. These mutations include missense, nonsense and splice site mutations, partial deletions and insertions (Futerman & van Meer, 2004). In general, mutations resulting in null or very low residual enzyme activity cause the most severe early onset forms of the diseases. In contrast, higher residual enzyme activity delays disease onset. However, no obvious genotype–phenotype correlation has been found for most LSDs and prediction of the clinical course of the disease cannot usually be made only on the basis of mutational analysis (Futerman & van Meer, 2004; Hopwood, 2012). Different epigenetic factors, genetic variability in the

synthesis of the stored substrate, and the overall functional efficiency and control of an individual patient's endosome-lysosome-network are other factors that may influence clinical outcome (Hopwood, 2012).

As mentioned before, splicing mutations are only one of the underlying causes of LSDs, even though having been identified in many patients. Specifically for a group of well-known lysosomal diseases (Fabry, Gaucher, Pompe, Niemann-Pick type C, Gangliosidosis GM1 and GM2, MPSs and NCLs) splicing mutations account for ~9% of all described mutations (HGMD®Professional Release 2015.3). This figure may likely be an underestimate given that only genomic DNA analysis is routinely performed in the diagnostic setting, uncovering the real impact that many silent, missense and nonsense mutations may have on the splicing process.

Great strides have been made in the characterization of biochemical and molecular basis of most LSDs, which naturally generated excitement and anticipation regarding direct approaches to therapy. However, despite the accumulated knowledge about disease mechanisms and much biopharmacological investment, for most LSDs, no curative therapy is available, although a number of specific treatment strategies do exist for some, aiming to improve quality of life (Boustany, 2013; Cox, 2012). Essentially, the primary aim of treatment is the preservation or restoration of organ function, which can be achieved by preventing or reducing tissue substrate build up. Clinically proven approaches include the restoration of intracellular activity through cross-correction, by transplanting healthy donor cells (bone marrow or umbilical cord blood as a source of haematopoietic stem cells) or infusion of the relevant exogenous enzyme (enzyme replacement therapy – ERT), and by substrate synthesis inhibition (substrate reduction therapy – SRT). Haematopoietic stem cells transplantation (HSCT) have the potential to supply complementing enzyme activity over the long-term and has led to good results in some cases (MPS I and VI; Metachromatic leukodystrophy; Krabbe disease and α-Mannosidosis), but only if the procedure is performed at an early phase of disease progression. Also, it is limited by the availability of compatible donors and by procedural complications with significant morbidity and mortality (Cox, 2012; Lachmann, 2010; Pastores, 2010). ERT on the other hand, uses recombinant versions of the deficient enzymes to overcome the cellular defect. This approach is only available for a restrict number of LSDs (Gaucher type I, Fabry, Pompe and MPS I, II, IVA and VI), even though it will probably be extended to others for which clinical trials are already in course (α-Mannosidosis, Wolman disease and Niemann-Pick type B). ERT has generally shown to be effective in treating systemic symptoms and slowing down the disease progression in later onset or milder variants of LSDs. Unfortunately, ERT cannot treat the neurological manifestations present in patients affected by the

more severe types of LSDs, given the inability of recombinant enzymes to cross the blood brain barrier (BBB) and reach the brain (Boustany, 2013; Cox, 2012; Cox, 2015). Concerning the SRT strategy, whose principle is to reduce the biosynthesis of new storage material, it is currently available for Gaucher disease and Niemann-Pick type C, and being evaluated for some other lysosomal disorders (MPS, Fabry, late onset GM2 Gangliosidosis and Sandhoff disease) (Cox, 2012; Cox, 2015; Pastores, 2010; Platt & Jeyakumar, 2008).The ineffectiveness of ERT and, in some cases of bone marrow transplantation (BMT) and SRT, to overcome neurological symptoms is one of the main reasons prompting research on additional therapeutic approaches to address the needs of the more severe types. Other strategies presently under study include pharmacological chaperones, gene therapy and stop-codon read-through technology (Boustany, 2013; Cox, 2015; Pastores, 2010; Tomanin et al., 2012). Still, if early diagnosis of these disorders before the onset of irreversible pathologies is already crucial to obtain better outcomes for current therapies, it certainly will continue to be for the innovate medicines that hopefully will appear in the next years.

1.3 RNA-based therapies

A strong conviction is being formed that as diagnostic approaches become individualized through whole genome sequencing as well as transcriptome and epigenetic profiling, the fraction of the detected disease-causing mutations and single nucleotide polymorphisms (SNPs) predisposing to disease or modulating disease severity that affect splicing will increase (Cooper et al., 2009; Singh & Cooper, 2012). The continuous discovery of new disease-causing mutations in RNAs is yielding a wealth of new therapeutic targets, and the growing understanding of RNA biology and chemistry is also providing new tools that can be exploited in the development of RNA-based therapeutics (Cooper et al., 2009; Douglas & Wood, 2011; Singh & Cooper, 2012). Concerning approaches to manipulate splicing for therapy, significant advances have been achieved by now. Splicing can be manipulated with a number of tools including antisense oligonucleotides (AOs), overexpression or silencing of specific splicing factors, modified U1 snRNAs, spliceosome-mediated RNA *trans*-splicing (SMaRT) and using small molecule compounds (e.g. pentamidine, sodium butyrate, kinetin and tetracycline derivatives), all of them applied to increase specific alternatively spliced isoforms or to correct aberrant gene expression resulting from gene mutations that alter splicing (Arechavala-Gomeza et al., 2014; Douglas & Wood, 2011; Havens et al., 2013; Pérez et al., 2012; Rigo et al., 2014; Wally et al., 2012). Since the AOs and

the modified U1 snRNA therapeutic approaches were the two options selected for this work, they will be the focus of the next sections.

1.3.1 Antisense oligonucleotides technology

Antisense oligonucleotide-based therapy covers a range of methods for modifying gene expression, which have the potential to revolutionize the development of therapeutics and biomedical practice (Magen & Hornstein, 2014). The modification of gene expression, using a synthetic single stranded DNA, resulting in inhibition of mRNA translation was demonstrated for the first time by Paterson and colleagues in 1977 in a cell-free system (Paterson et al., 1977). Almost a year later, Zamecnik and Stephenson showed that in chicken fibroblast tissue culture containing Rous Sarcoma virus, the addition of a synthetic 13-mer oligonucleotide complementary to the 3' end of the virus, could inhibit its replication and transformation of fibroblasts into sarcoma cells (Zamecnik & Stephenson, 1978). Since then, remarkable progress has been made in oligonucleotide drug development and currently antisense technology is a powerful tool that can be used for target validation and to correct or alter RNA expression for therapeutic benefit (Bennett & Swayze, 2010; Evers et al., 2015; Havens et al., 2013; Rigo et al., 2014).

AOs are short, synthetic, and modified nucleic acids that are able to bind to mRNA or pre-mRNA with high specificity via base-pairing and to interfere in its function through a variety of post-binding events (Bennett & Swayze, 2010; Rigo et al., 2014). Depending on the chemistry and target site, AOs are able to perform their function through different mechanisms, being important to distinguish between those that knockdown gene expression by degrading the target RNA via enzymes such as RNase H, or Argonaute 2 (RNA interference), and those that act in splicing modulation without promoting RNA degradation as the splice switching oligonucleotides (SSOs). SSOs are single stranded AOs typically 15-25 nucleotides long which direct pre-mRNA splicing to a new pathway by binding sequence elements and sterically blocking access to the transcript by the spliceosome and other splicing factors (Bauman et al., 2009; Bennett & Swayze, 2010; Kole et al., 2012; Rigo et al., 2014).

1.3.2 Antisense oligonucleotides chemistry

To be suitable for therapeutic application, AOs ideally require a number of intrinsic properties. To start off with, the AO should bind in a sequence specific manner to the target RNA transcript; the higher the specificity, the less the chance of unwanted

off-target effects. Secondly, the AO should be of a chemistry that facilitates cellular uptake and activity in the appropriate intracellular compartment (e.g. nucleus or cytoplasm). Thirdly, because of the plethora of nucleases present *in vivo*, a well-designed AO should be resistant to nuclease degradation in order to allow it to reach its desired target intact and to maximize its potential duration of action once there. Fourthly, as with any drug, the ideal AO should have favorable pharmacokinetics (PK) and pharmacodynamics (PD). Linked to this is, of course, the prerequisite that the AO should not be a toxic compound. Finally, the design of the AO must allow its effective delivery to the target tissues, whether that be a localized area such as a specific organ or a body-wide systemic delivery (Bennett & Swayze, 2010; Douglas & Wood, 2013; Geary et al., 2015; Mansoor & Melendez, 2008; Rigo et al., 2014). To try to cope with these desired AO properties, several chemical modifications have been made to the backbone and sugar component of AOs allowing the development of different oligonucleotide chemistries with features that make them potential therapeutic tools (Bennett & Swayze, 2010; Mansoor & Melendez, 2008; Rigo et al., 2014).

1.3.2.1 'First generation oligonucleotides': the PS backbone modification

The precise positioning of chemical modifications within an AO fundamentally determines its mechanism of action. One of the first successful modifications, which remains widely used in antisense therapeutics, is the phosphorothioate (PS) backbone (Figure 1.4). Termed the first generation of oligonucleotide modifications, the PS backbone is accomplished by replacement of one of the non-bridging oxygen atoms of the natural phosphodiester linkage in the backbone with a sulfur atom. This simple chemical modification improves resistance to nucleolytic degradation, elicits RNase H-mediated cleavage of the target mRNA, carries negative charges that facilitate their cell delivery *in vitro* an *in vivo*, and increases affinity for plasma proteins leading to reduced renal clearance and an increased circulation time of the AO (Bennett & Swayze, 2010; Evers et al., 2015; Kurreck, 2003; Mansoor & Melendez, 2008; Rigo et al., 2014). However, the binding of PS AOs not restrictively to serum proteins can also be a disadvantage due to the production of undesirable side effects such as nonspecific activation of the immune system (Evers et al., 2015; Rigo et al., 2014). The first AO drug approved for clinical use, the intravitreous Fomivirsen® (Vitravene) was a first generation PS-modified oligonucleotide used for repression of cytomegalovirus mRNA translation. It gained United States Food and Drug Administration (U.S. FDA) approval for intraocular treatment of Cytomegalovirus retinitis in immunosuppressed patients in

1998 (Roehr, 1998) and was discontinued later due to commercial considerations (Magen & Hornstein, 2014).

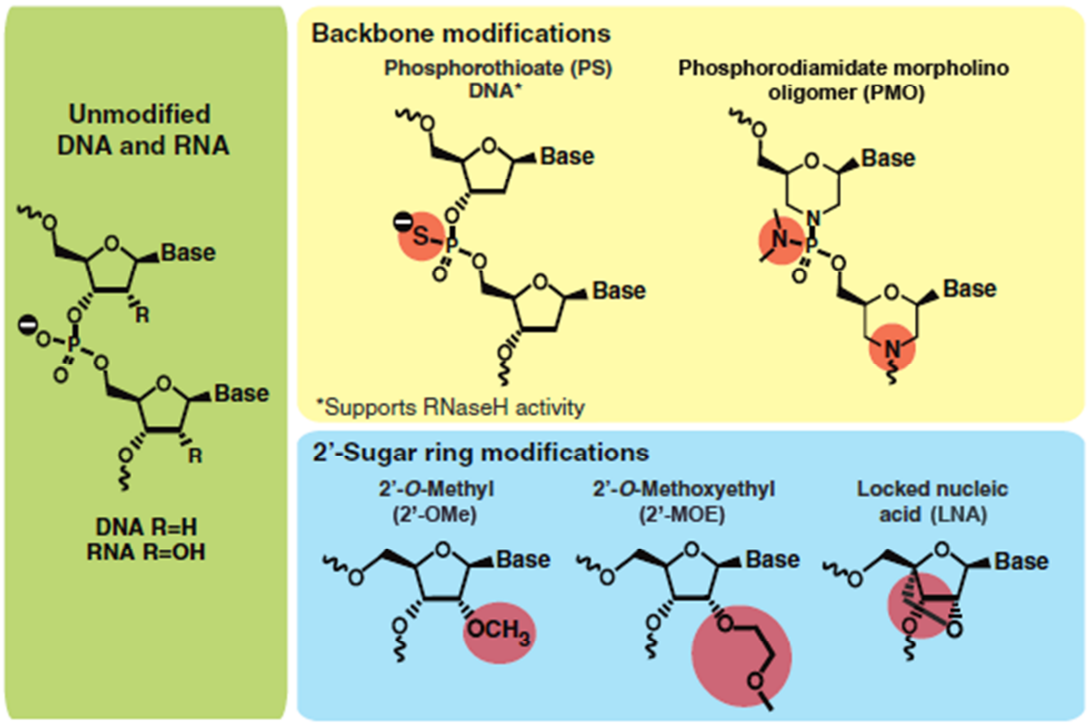


Figure 1.4: Antisense oligonucleotides chemical structures. Schematic representation of unmodified DNA/RNA base pair (left). Different backbone modifications that can be applied (top row) and different 2'-sugar modifications that can be used (bottom row) to increase nuclease resistance and RNA binding affinity of the antisense oligonucleotide. Deviations from the original unmodified DNA/RNA are highlighted by circles. Adapted from DeVos & Miller, 2013

1.3.2.2 'Second generation oligonucleotides': the sugar group modifications

Another class of oligonucleotide modifications includes those made at the 2' position of the sugar ring, which in combination with the PS backbone has highly improved oligonucleotide safety and pharmacologic properties. The class represents the second generation of oligonucleotides, whose most successful members are the 2'-O-methyl (2'-OMe) and 2'-O-methoxyethyl (2'-MOE) modifications (Figure 1.4). They have improved the PS backbone oligonucleotides, with enhancements in nuclease resistance and RNA-binding affinity, better tissue uptake, longer *in vivo* half-life and reduction of nonspecific protein binding, which in turn, contributed to reduce the AO toxicity profile (Bennett & Swayze, 2010; Chan et al., 2006; Evers et al., 2015; Kurreck, 2003; Mansoor & Melendez, 2008).

These desirable properties are, however, counterbalanced by the fact that 2'-O-alkyl fully modified oligonucleotides cannot induce RNase H cleavage of a target RNA

(RNase H requires a free 2′-oxygen), which restricts the use of 2' modifications for antisense downregulation purposes (Bennett & Swayze, 2010; Evers et al., 2015; Kurreck, 2003). This limitation has been minimized with the development of "gapmer" structures where 2' sugar-modified residues are present on either side of a central "gap" region comprising 8-10 PS-modified nucleotides. The external sugar modified residues thus increase affinity and nuclease resistance, while the internal "gap" region allows RNase H-mediated cleavage of the target RNA (Bennett & Swayze, 2010; Kole et al., 2012; Monia et al., 1993; Rigo et al., 2014).

Recently, a systemically administered AO of the 2'-MOE gapmer type - Kynamro®, which targets apolipoprotein B (*APOB*) transcripts, was approved by the U.S. FDA for the treatment of Familial hypercholesterolemia (Crooke & Geary, 2013). Also, the phase I clinical trials were recently concluded for another 2'-MOE gapmer oligonucleotide (ISIS-333611) targeting the exon 1 of *SOD1* gene mutated in Amyotrophic lateral sclerosis (ALS) disease (Miller et al., 2013).

1.3.2.3 'Second and third generation oligonucleotides': designing AOs for splicing modulation

For splicing modulation, oligonucleotides were developed that do not induce RNase H-mediated cleavage of RNA but instead act by blocking a specific site in the targeted RNA without inducing its degradation. The most commonly used splice-switching oligonucleotide chemistries include the second generation PS-modified 2'-OMe and 2'-MOE AOs and also the phosphorodiamidate morpholino oligomer (PMO) and locked nucleic acid (LNA) designs which are part of a third generation class of AOs (Figure 1.4) (Aartsma-Rus & van Ommen, 2007; Bauman et al., 2009; DeVos & Miller, 2013; Rigo et al., 2014). This third class was developed mainly by chemical modifications of the furanose ring of the nucleotide, along with modifications of phosphate linkages aiming to further enhance nuclease stability, target affinity, pharmacokinetic and toxicology profiles of the AOs (Chan et al., 2006; Evers et al., 2015; Mansoor & Melendez, 2008).

The PMOs represent a successful modification in which the ribose sugar is replaced by a six-membered morpholine ring and the negatively charged phosphodiester backbone is replaced with a neutral phosphorodiamidate backbone (Figure 1.4) (Summerton & Weller, 1993; Summerton, 2007). Morpholinos are usually constituted by a 25 nucleotide chain which presents high target specificity and affinity, good solubility in aqueous solution, high resistance to nucleases and do not activate RNase H, being used primarily in splicing modulation and translational interference

through steric hindrance of the ribosomal machinery (Bennett & Swayze, 2010; Summerton, 1999; Summerton, 2007).

Unlike PS oligonucleotides, PMOs have no net electrical charge and, therefore, do not tend to interact with nucleic acid-binding proteins. On the one hand, this reduces the chance of side effects, but on the other hand, poor protein binding causes the drug to be rapidly cleared from the body and high doses of the oligonucleotide are required to elicit a pharmacological response (Mansoor & Melendez, 2008; Rigo et al., 2014). Also, their electrostatically neutral backbones at physiological pH make difficult the cellular uptake. For this reason their delivery into tissue cell culture can be improved by non-toxic endocytosis-assisted delivery reagents like ethoxylated polyethylenimine (EPEI) (Morcos, 2001) and Endo-Porter® (Summerton, 2005) as also by the scrape delivery method (Partridge et al., 1996). As expected, these antisense molecules also do not readily cross cell membranes *in vivo* without delivery techniques. To overcome this problem, another chemical modification was made to allow the oligonucleotides to more easily penetrate cell membranes and to improve their PK properties (Evers et al., 2015; Lee & Yokota, 2013; Moulton & Jiang, 2009). The first effective chemically-mediated method for systemic delivery of morpholino antisense was based on covalently linking the oligonucleotides to arginine-rich (arg-rich) cell-penetrating peptides (CPPs), resulting in the peptide-linked phosphorodiamidate morpholino oligomers, or PPMO (Figure 1.5A) (Moulton et al., 2003). The presence of positively charged guanidinium moieties as part of the arginine chains in this conjugated oligonucleotides is responsible for its increased cellular uptake (Moulton & Moulton, 2004; Wender et al., 2000).

Another interesting conjugated PMO, which exhibits superior performance in transporting morpholino oligomers across biological barriers is the vivo-morpholino (vPMO) (Figure 1.5B). vPMOs are morpholino oligonucleotides covalently linked to a dendrimer scaffold which carries eight guanidinium head groups that enable delivery into cells (Li & Morcos, 2008; Morcos et al., 2008). This molecular transporter showed efficient delivery of coupled morpholino oligomers to a wide spectrum of tissues without evidence of toxicity after its intravenous injection in transgenic mice (Morcos et al., 2008) and in a dystrophic *mdx* mouse model (Aoki et al., 2012). However, other recent studies targeting the *SMN2* (Zhou et al., 2013) and *PAH* (Gallego-Villar et al., 2014) genes have reported severe toxicity after the systemic administration of v-PMOs in mice, even though an effective splicing modulation has also been observed.

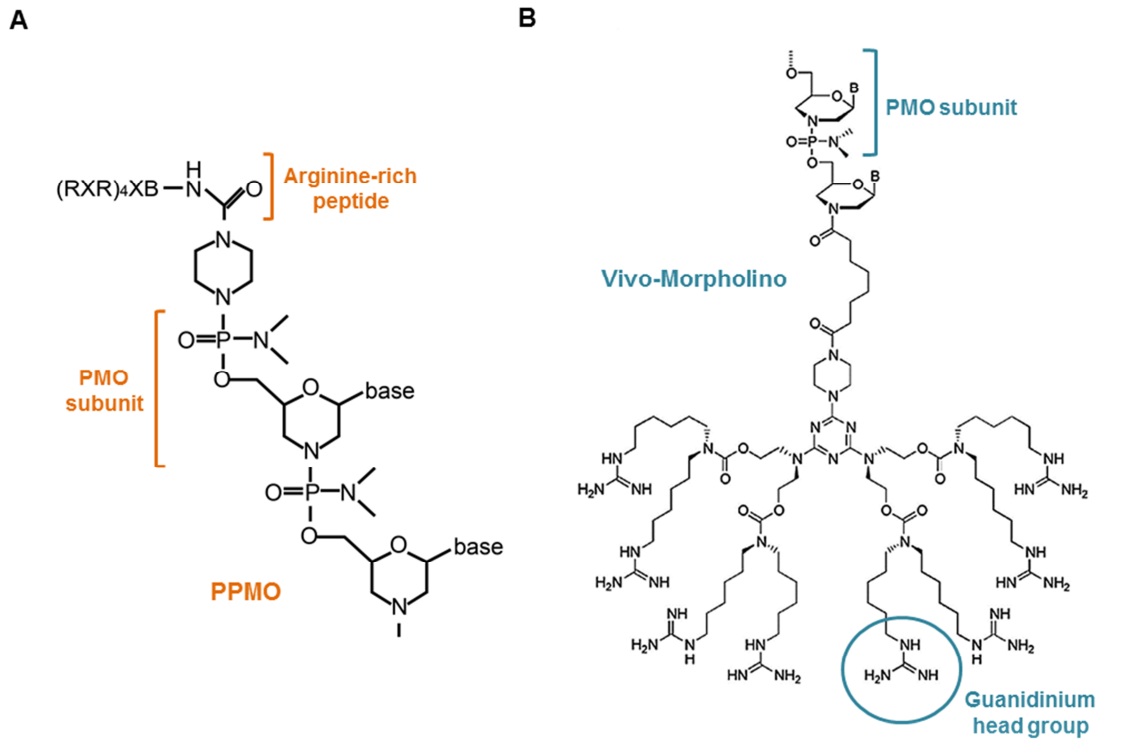


Figure 1.5: Chemical structure of a peptide-linked phosphorodiamidate morpholino (PPMO) and a vivo-morpholino (vPMO). **A)** Representation of the PMO subunits linked to the arginine-rich peptide in the PPMO structure. R – arginine; X – 6-aminohexanoic acid, B – β-alanine; Base – A, C, G or T. **B)** Representation of a v-PMO structure. Shown are two 3′ end PMO subunits (top portion) linked to the molecular transporter with eight guanidinium head groups (bottom portion). B – base. Adapted from Goyenvalle & Davies, 2011; Morcos et al., 2008

LNA oligonucleotides constitute a class of chemically modified RNA nucleotides that has been also widely used seeking therapeutic applications. LNAs contain a methylene bridge connecting the 2'-*O* with the 4'-*C* position in the furanose ring, which enables it to form a strictly *N*-type conformation that offers to the nucleic acid an unprecedented binding affinity and excellent specificity toward complementary RNA or DNA (Figure 1.4) (Obika et al., 1997; Singh et al., 1998; Vester & Wengel, 2004). LNAs show structural resemblance with RNA, high resistance to nucleolytic degradation with consequent high bio-stability, good aqueous solubility, great thermal stability of the duplexes and are easy to synthesize (Kaur et al., 2007; Veedu & Wengel, 2009; Vester & Wengel, 2004). Furthermore, the charged phosphate backbone of LNAs allows their ready delivery into cells using standard cationic transfection agents (Kaur et al., 2007).

The high RNA binding affinity and potency of LNAs was demonstrated *in vivo* without evidence of toxicity after systemic administration in mice (Gupta et al., 2010; Roberts et al., 2006) or when applied directly in rat brain (Wahlestedt et al., 2000). However, profound levels of hepatotoxicity were reported in mice after intraperitoneal delivery (Swayze et al., 2007). As interestingly revealed in a recent work of Burdick and

colleagues (2014), the hepatotoxicity observed after LNA administration in mice seems to be strongly associated with the presence of specific trinucleotide sequence motifs in the oligonucleotide. AOs containing these sequence motifs tend to exhibit higher binding to mouse liver proteins, which results in the activation of transcription factors that trigger specific stress pathways leading to hepatotoxicity. This finding suggests that *in silico* approaches can be used to evaluate structure–toxicity relationships of LNA-modified AOs aimed at decreasing the likelihood of hepatotoxicity in preclinical testing (Burdick et al., 2014).

1.3.3 Factors influencing antisense oligonucleotides activity

All antisense mechanisms have in common the binding of the oligonucleotide to the targeted RNA. What happens after the oligonucleotide binding is dictated not only by the oligonucleotide chemistry but also by additional factors that may influence its activity. Among them are the oligonucleotide length, the melting temperature (Tm), the guanine-cytosine content, the presence of secondary or tertiary structures formed by RNA that can compromise the accessibility of the AO to the target, or sequence motifs for the binding of auxiliary factors like RNA-binding proteins, as well as the local in the cell (nucleus or cytoplasm) where the AO binds to the RNA. To design an AO all these aspects should be taken into account and every option should firstly be screened for activity in cell culture. This step is essential to find the most effective oligonucleotide(s) for a particular antisense mechanism, which can subsequently be tested *in vivo* (Aartsma-Rus, 2012; Aartsma-Rus et al., 2009; Rigo et al., 2014).

1.3.4 The use of antisense oligonucleotides for splicing modulation

The ultimate goal in the development of many AO-based therapies is to block the production of a toxic form of a protein or to restore i) the production of a protein that is not translated, or ii) the activity of a nonfunctional protein. For this, SSOs are designed to base-pair to pre-mRNA, in order to influence splicing through the alteration of the protein-coding mRNA (Havens et al., 2013). The first demonstration of the ability of a SSO to correct a splicing mutation appeared two decades ago for the *β-globin* gene, where SSOs were successfully used to mask the activated intronic cryptic splice sites leading to abnormal splicing and causing β-thalassemia (Dominski & Kole, 1993). Since then, SSOs have been fruitfully employed in numerous applications targeting many disease-relevant genes including a number coding for IMDs (Arechavala-

Gomeza et al., 2014; Bauman & Kole, 2011; Lee & Yokota, 2013; Pérez et al., 2010; Pérez et al., 2012; Siva et al., 2014; Veltrop & Aartsma-Rus, 2014).

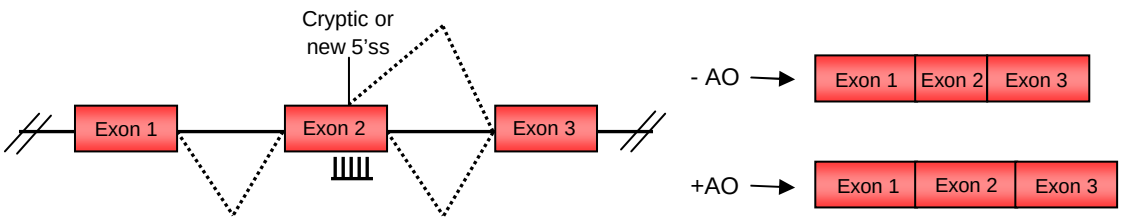
Splicing modulation can be achieved through different approaches depending on the nature of the mutation to correct. For instance, SSOs can be used to block cryptic, new, or pseudo splice sites in the pre-mRNA (Figure 1.6A), as was already done to efficiently neutralize aberrant splice sites in cells of patients with Ataxia-telangiectasia (Du et al., 2007; Nakamura et al., 2012), Pyridoxine-dependent epilepsy (Pérez et al., 2013), Hypomyelinating leukodystrophy (Regis et al., 2013), Congenital disorders of glycosylation (Vega et al., 2009; Yuste-Checa et al., 2015) and Niemann-Pick C disease (Rodríguez-Pascau et al., 2009) or in animal models of Hutchinson-Gilford progeria syndrome (Osorio et al., 2011) and Usher syndrome (Lentz et al., 2013). SSOs can also be used to promote the inclusion of an exon by targeting splicing regulatory sequences (Figure 1.6B). A case of success is illustrated by Spinal muscular atrophy (SMA), the leading genetic cause of mortality in infants under the age of 2 years, in which SSOs were used to block an ISS promoting exon inclusion (Hua et al., 2010; Hua et al., 2008). For this disease, SSOs antisense therapy is now in phase II clinical trials (ISIS-SMNRx; 2'-MOE; ClinicalTrials.gov), showing potential to become a treatment for SMA patients (Rigo et al., 2012; Zanetta et al., 2014). Another approach is based on the use of SSOs to remove exons either to eliminate a nonsense mutation or to restore the reading frame around a genomic deletion (Figure 1.6C). It was the strategy applied in Duchenne muscular dystrophy, an early onset, severe neuromuscular disorder caused in the majority of patients by deletions or duplications of one or more exons that result in disruption of the open reading frame (ORF) and deficit of its encoded protein, dystrophin (Muntoni et al., 2003). In a model of "exon skipping", SSOs were used to "mask" specific exons to the splicing machinery, restoring the ORF with consequent production of internally deleted but partially functional proteins for DMD patients. The developments in the use of SSOs as a potential therapeutic agent in DMD have been fast since the early proof of concept experiments (Alter et al., 2006; Lu et al., 2003; van Deutekom et al., 2001; Yokota et al., 2009) to the phase II (Eteplirsen (AVI-4658); PMO; targeting exon 51) and III (Drisapersen; 2'-OMe; targeting exons 51, 44, 45 and 53) clinical trials that are currently underway revealing encouraging results (Aartsma-Rus, 2014; Cirak et al., 2011; Goemans et al., 2011; Mendell et al., 2013). However, AO-mediated exon-skipping for DMD still faces major hurdles such as extremely low efficacy in the cardiac muscle, no crossing of the BBB, poor cellular uptake and relatively rapid clearance from circulation. Therefore, to overcome these limitations new modifications and chemistries are being explored in pre-clinical tests, including for dystrophin exon

skipping. One such example is the recent study of Betts and colleagues (2012), where the development of novel Pip (PMO internalization peptide) transduction peptides, termed Pip6-PMOs, led in a DMD mouse model to efficient exon skipping and dystrophin restoration in multiple muscle groups (Betts et al., 2012). In another study, Goyenvalle and co-workers (2015) used a new class of conformationally constrained DNA analogues: the tricyclo-DNAs (tc-DNAs), which displayed unprecedented uptake after systemic administration and induced efficient exon skipping in a DMD mouse, promoting a high degree of dystrophin rescue in several tissues including the brain (Goyenvalle et al., 2015). This makes tc-DNA AO chemistry particularly attractive as a potential future therapy for patients with DMD or with other diseases that are eligible for exon-skipping approaches requiring whole-body treatment. The exon-skipping process of splicing modulation was also applied with success in cells from patients suffering from dysferlinopathies (Aartsma-Rus et al., 2010; Wein et al., 2010). SSOs can also be applied to manipulate alternative splicing from one splice variant to another (Figure 1.6D) as was demonstrated in Fronto-temporal dementia and Parkinsonism linked to chromosome 17 (FTDP-17) (Kalbfuss et al., 2001; Peacey et al., 2012), in some cancers (Bauman et al., 2010; Wan et al., 2009; Zammarchi et al., 2011), and in inflammatory disorders (Graziewicz et al., 2008; Vickers et al., 2006).

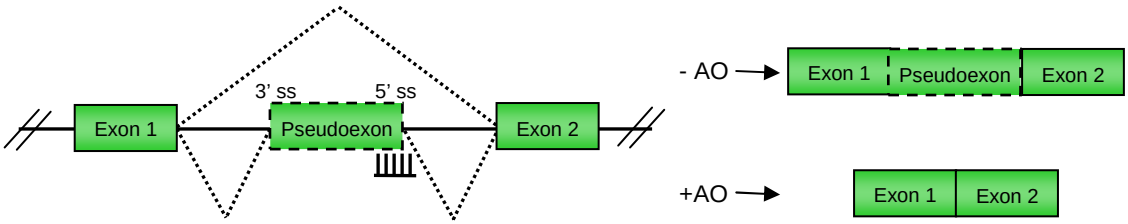
In addition to splicing modulation, AOs have been used to disrupt interactions between a protein splicing factor and a toxic repeat-containing RNA in myotonic dystrophy (Leger et al., 2013; Wheeler et al., 2009), to inhibit translation in Huntington, Machado Joseph (Evers et al., 2011; Gagnon et al., 2010; Hu et al., 2009) and Alzheimer (Erickson et al., 2012; Kumar et al., 2000) diseases as well as to inhibit the viability of bacteria (Geller et al., 2005) and virus (Warren et al., 2010). AOs of the PMO class have been extensively used in studies of gene function and development in a variety of organisms like frog, zebrafish, chicken, mouse and sea urchin (Eisen & Smith, 2008; Ferguson et al., 2013) as in the generation of disease models, mainly in mouse and zebrafish (Cline et al., 2012; Gallego-Villar et al., 2014; Mahmood et al., 2013; Moro et al., 2010).

A) Cryptic and Pseudo splice sites

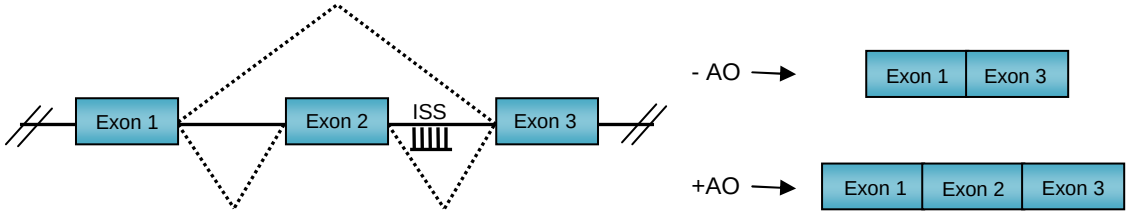
Cryptic splice site



Pseudoexon

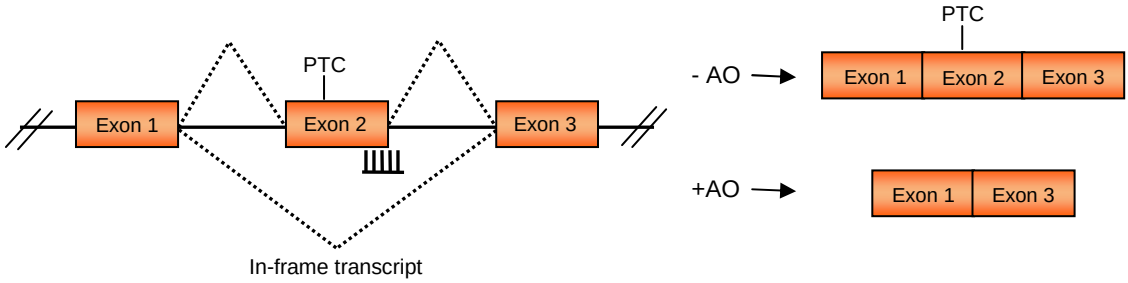


B) Exon inclusion

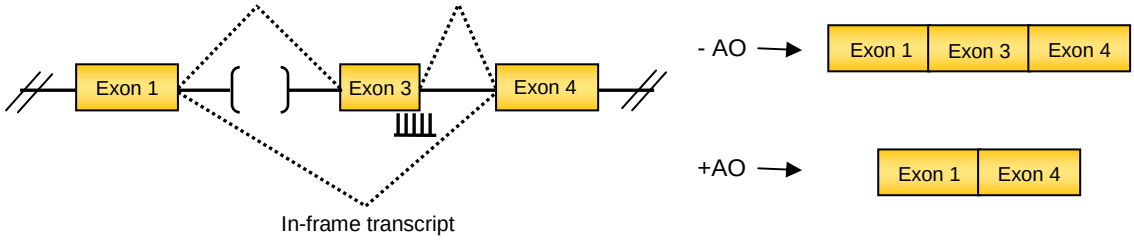


C) Exon exclusion

Exclusion of exons with PTC mutations

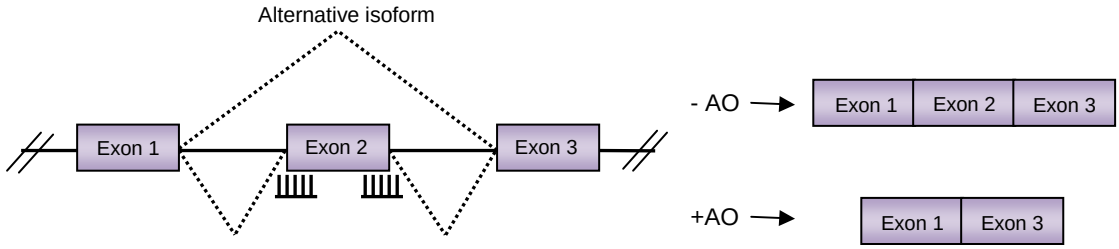


Restoration of ORF due to a deletion



D) Alternative splicing

Forcing the exclusion of an alternative exon



Forcing the selection of an alternative splice site

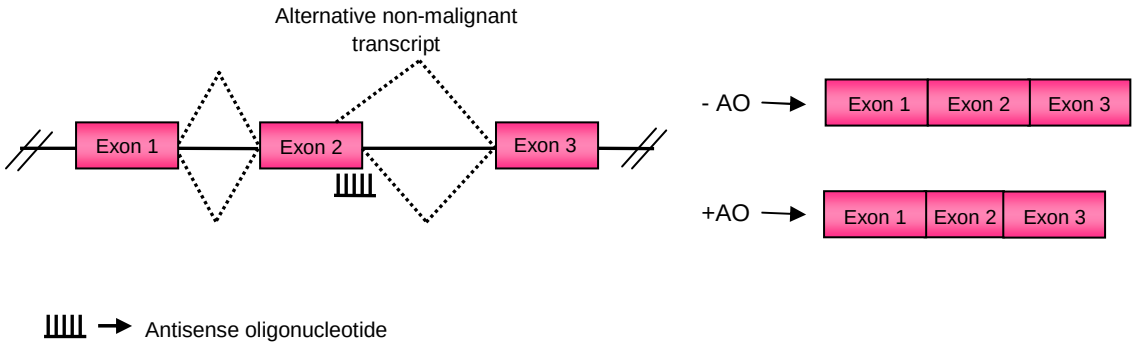


Figure 1.6: Methods of AO-mediated modification of mutation-generated aberrant splicing. AOs have been used to revert aberrant splicing of activated cryptic splice sites, new splice sites or pseudoexons **(A)**, to promote the inclusion of an exon by targeting splicing regulatory sequences **(B)**, to exclude an exonic sequence to either eliminate a nonsense mutation or restore the reading frame caused by a frameshift deletion **(C)**, or to force the exclusion of an alternative exon or the selection of an alternative splice site **(D)**. Based on Hammond & Wood, 2011; Pérez et al., 2010; Siva et al., 2014

1.3.5 U1 snRNA basic principles

Since its discovery in the early days of splicing research, U1 snRNA has been recognised as a crucial player in the first stages of the splicing process (Lerner et al., 1980; Mount et al., 1983; Rogers & Wall, 1980). U1 snRNA is a 164 nucleotides long molecule with a well-defined structure consisting of four stem–loops, which primarily exerts its function in the form of an ribonucleoprotein (RNP) complex (termed U1 snRNP) containing seven Sm proteins and three U1-specific proteins U1A, U1C and U1-70k (West, 2012) (Figure 1.7A). It is now well-established that U1 snRNP initiates spliceosome assembly by binding to the 5' splice donor site (SDS) through base pairing

between the single stranded terminal sequence of the U1 snRNA molecule and the moderately conserved stretch of nucleotides at the 5' SDS (CAG/GURAGU; R-purine) marking the exon-intron boundary (Buratti & Baralle, 2010). However, not all base pairs at different 5' ss positions are equally important, and their contribution to splicing roughly correlates with their conservation (Figure 1.7B). In the 9 nucleotides consensus sequence (which sometimes is expanded to include 11 base pairs), the most conserved 5' ss positions lie at the first two intronic nucleotides (+1 and +2), which determine the 5' ss subtype. The GU subtype, with Watson-Crick complementarity with A7 and C8 in U1, accounts for 99% of 5' ss. The minor subtypes have a mismatch to U1 at either +1 or +2 and include the GC (0.9%) and the very rare AU 5' ss recognised by the spliceosome. The next most conserved 5' ss positions (>75% in humans) are -1G (the last exonic nucleotide) and +5G, which form strong G-C base pairs with U1, with three hydrogen bonds. Consensus nucleotides -2A, +3A, +4A, and +6U are also conserved but have a lesser although important contribution to 5' ss strength because their base pairing to U1 involves only the formation of two hydrogen bonds. The 5' ss positions +7 and +8 do not exhibit substantial conservation in humans, yet several lines of evidence indicate that these positions can base-pair to U1 and contribute to splicing (Roca et al., 2013). Once the donor site does not always conform to the consensus sequence, but can instead have a degenerate pattern feature, it is understandable that many other additional elements such as splicing silencer and enhancer motifs, the presence of alternative splice sites, secondary structures and regulatory proteins can influence the splice site selection (Figure 1.7A) (Raponi & Baralle, 2008; Roca et al., 2013).

U1 snRNA is classically known for its role in pre-mRNA splicing events. However, the finding that U1 snRNA levels far exceed other spliceosomal associated snRNA levels led to the notion that it may have additional roles in the cell apart from splicing regulation (Guiro & O'Reilly, 2015; West, 2012). Indeed, emerging evidence suggests that U1 snRNA plays a key role in transcription initiation and in the protection of pre-mRNAs from degradation, as also has a regulatory function in the 3'-end formation, protecting pre-mRNA transcripts against premature polyadenylation and contributing to the regulation of alternative polyadenylation (Guiro & O'Reilly, 2015; Spraggon & Cartegni, 2013; Valadkhan & Gunawardane, 2013; West, 2012). Furthermore, the binding of U1 snRNA to 5' ss like sequences present in the 3' untranslated region (UTR) of some mRNA sequences (eg. viral mRNA), can lead to the suppression of the polyadenylation process, resulting in mRNA degradation (Buratti & Baralle, 2010; Valadkhan & Gunawardane, 2013; West, 2012).

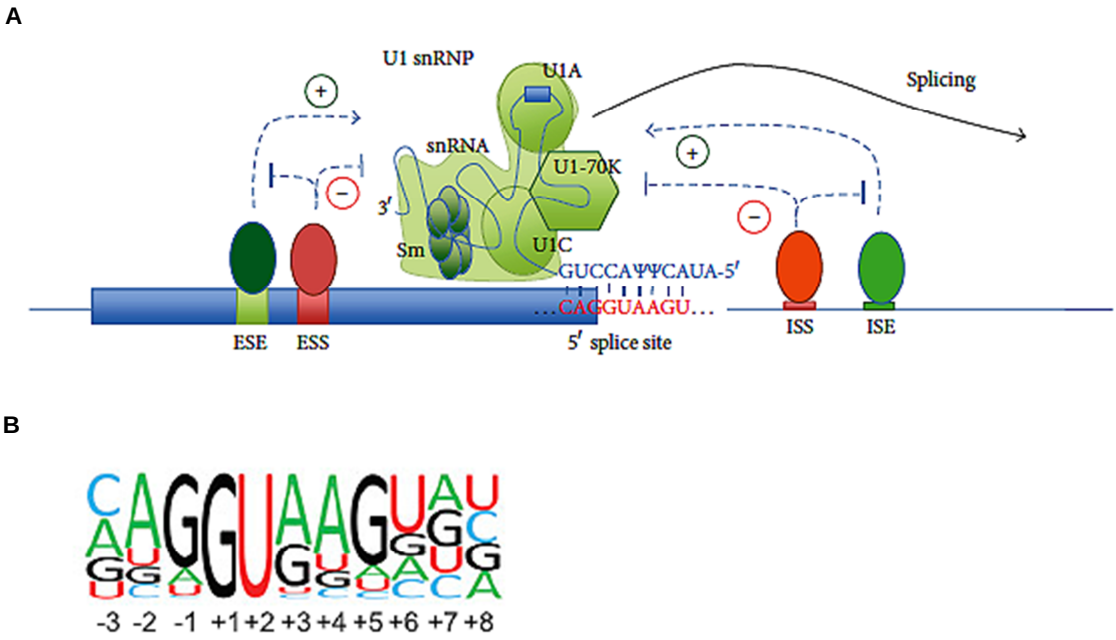


Figure 1.7: A) Role of U1 snRNP in splicing. The 5' end of U1 snRNA base pairs to the 5' ss, to define the functional SDS. The process is positively and negatively modulated by splicing factors binding to exonic and intronic splicing enhancer and silencers motifs (ESE, ISE, ESS, and ISS, respectively). **B)** The 5' ss motif. The height of each nucleotide corresponds to its conservation at the corresponding position (-3 to -1 are exonic positions and +1 to +8 corresponds to intronic positions). Adapted from Roca et al., 2013; Spraggon & Cartegni, 2013

1.3.6 U1 snRNA-mediated therapy

Splicing mutations at the 5' SDS, which are frequent among defects that cause human disease, compromise U1 snRNA binding and can prevent spliceosome assembly and subsequent splicing, which results in exon skipping, intron retention or activation of cryptic splice sites (Buratti et al., 2007; Roca et al., 2013). The most deleterious mutations at a 5' ss are those affecting the nearly invariant GU dinucleotide at the positions +1 and +2. For the remaining nine positions the effects on splicing and consequently the possibility of causing disease are less clear. In fact, nucleotide substitutions affecting the less conserved positions can cause splicing defects in some but not all 5' ss, suggesting that the remaining 5' ss positions and/or the overall context dictate the extent to which splicing is disrupted (Roca et al., 2013; Roca et al., 2008).

As donor splice site mutations disrupt the complementarity of the donor site with the endogenous U1 snRNA, restoring the complementarity through engineered modification of the U1 snRNA represents a valuable approach. For this purpose, the normal 5' end sequence of the U1 snRNA cloned in a vector is substituted with a sequence complementary to the target mutated region, and the modified U1 snRNA gene is tested in appropriate splicing assays in an attempt to correct the splicing defect (Figure 1.8) (Havens et al., 2013; Pinotti et al., 2011).

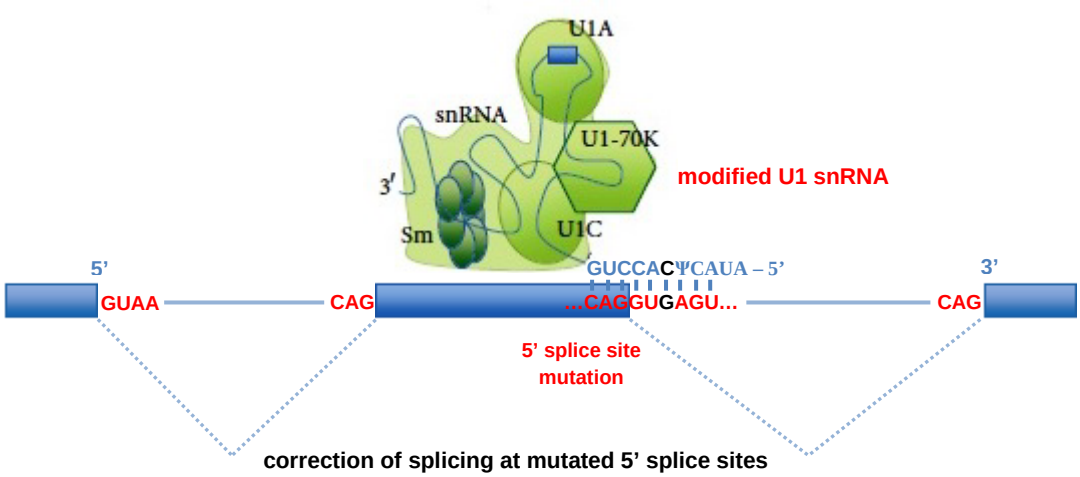


Figure 1.8: Modified U1 snRNA compensation of 5' ss mutations. Exogenous U1 snRNA with a compensatory mutation (black) allows for base-pairing with the 5' ss and the restoration of exon recognition and inclusion. Based in Havens et al., 2013

Once the U1 snRNA approach acts on the pre-mRNA level, it has the advantage of maintaining the expression regulation of the targeted gene in the normal chromosomal context. Also, given that the U1 snRNA gene used for splicing rescue includes promoter and regulatory sequences, it has the capability of guaranteeing long term correction of the genetic defect (Pinotti et al., 2011).

In common with other rescue strategies based on targeting RNA by complementarity (e.g. AOs), modified U1 snRNAs have to deal with potential off-target effects that might affect splicing of other genes. This could be dangerous for modified U1 snRNAs that have only one base change from the natural U1 snRNA, and thus might activate normally silent cryptic donor splice sites and induce aberrant splicing in other genes (Pinotti et al., 2011). The consequences of such unwanted side reactions are hard to predict and depend on the gene function of the spliced transcript. However, screening and mapping of the binding site sequence against the human genome to exclude sequence homologies should widely reduce nonspecific events even though their complete elimination cannot be guaranteed. Therefore, experimental analysis should be performed whenever possible to test the effect of the U1 treatment on nontarget transcripts. In a mutation-adapted U1 snRNA approach to correct a 5' splicing defect in Retinitis pigmentosa (RP) this type of test was performed to search for the presence of unwanted side effects. The splicing pattern of five RP-associated genes containing potential binding sites to a fully adapted U1 was analysed after U1 treatment and no missplicing events were found in the nontarget transcripts (Glaus et al., 2011).

Additionally, to reduce the possible interaction of modified U1 snRNAs with non-target 5' SDSs, an interesting model of an adapted U1 snRNA called Exon-Specific U1 (ExSpeU1) was developed. The ExSpeU1s have engineered 5' tails that direct their loading onto non-conserved intronic regions downstream of the SDS of a specific exon, and are expected to improve specificity and reduce potential off-target effects (Fernandez Alanis et al., 2012; Pinotti et al., 2011). This novel strategy has been applied successfully in different studies allowing an efficient rescue of exon skipping due to different types of splicing mutations associated with defective exon definition, constituting therefore a possible approach for correction of 5' splicing defects with less probable unwanted side effects (Dal Mas et al., 2015a; Dal Mas et al., 2015b; Fernandez Alanis et al., 2012).

In the U1-mediated approach, the U1 snRNA gene must be incorporated into an expression vector and delivered to cells. In tissue culture experiments, the plasmid vector is usually delivered using cationic lipids as transfection agents (Sánchez-Alcudia et al., 2011), but the U1 snRNA gene can also be inserted in lentivirus and adeno-associated virus (AAV) vectors to be delivered effectively into *in vitro* cell culture or *in vivo* animal models (Balestra et al., 2014; Schmid et al., 2011). However, the potential for antiviral immunity is still a major limitation of using viral vectors for therapy. Possible alternatives to viruses are exosomes, liposomes and nanoparticle delivery (Hammond & Wood, 2011).

1.3.6.1 The use of U1 snRNA-mediated therapy to correct 5' splice site defects

Since the original observation that engineered U1 snRNA mutations can suppress 5' ss mutations (Zhuang & Weiner, 1986), modified U1 snRNAs have been assayed as a possible therapy for numerous diseases caused by 5' ss mutations affecting different positions of the donor site. The great majority of the approaches was successfully performed using different modified-U1's co-transfected along with minigenes or directly delivered into patients' cells in diseases like Cystic fibrosis, Haemophilia B, SMA (Fernandez Alanis et al., 2012), Retinitis pigmentosa (Glaus et al., 2011; Tanner et al., 2009), Fanconi anemia (Hartmann et al., 2010; Mattioli et al., 2014), Propionic acidemia (Sánchez-Alcudia et al., 2011), Bardet-Biedl syndrome (Schmid et al., 2011; Schmid et al., 2013) and Netherton syndrome (Dal Mas et al., 2015a). Modified U1's also allowed to rescue splicing at mutated 5' ss in mouse models of DMD (Denti et al., 2006), SMA (Dal Mas et al., 2015b) and Human coagulation factor VII (hFVII) deficiency (Balestra et al., 2014). Most of the U1

therapeutic reports did not comprise toxicity studies. However, in the study of Dal Mas and colleagues for SMA (Dal Mas et al., 2015b), the U1 (ExSpeU1) therapeutic effects were also tested in different cellular models and no toxicity was observed. Yet, in the recent *in vivo* study for hFVII (Balestra et al., 2014), the authors reported that the expression of a modified U1 snRNA was associated with hepatotoxicity in a dose-dependent manner, which probably arose due to the binding of the engineered U1 to similar consensus 5' ss in other genes.

We must recognise that so far the extent of rescue of expression levels obtained with strategies targeting the RNA level has been only moderate. Moreover, the *in vivo* application of AOs and modified U1 snRNA depends on a safe and efficient delivery of oligonucleotides and expression cassettes. However, in the case of LSDs it can be argued that even a very low increase in enzymatic activity of lysosomal enzymes, would result in a significant amelioration of the patients' clinical phenotype. Therefore, for these disorders, RNA-based therapeutic studies aimed at demonstrating their clinical translatability have been greatly encouraged.

1.4. Challenges for delivery and the future of antisense therapy

Despite the significant progresses achieved in the field of RNA-based therapies, their clinical application to rare genetic disorders such as IMDs remains limited due to a number of important constrains that still need to be overcome. Improvements in different fields of antisense technology will be fundamental to achieve a safe and effective systemic delivery of antisense agents to target tissues in different organs, ensuring prolonged therapeutic effects and absence of drug-induced toxicities. The establishment of proof of concept efficacy in IMD animal models of disease will be also mandatory to make possible the translation of antisense-mediated splicing modulation therapies to the clinical realm (Du & Gatti, 2009; Pérez et al., 2010; Pérez et al., 2014).

As in several other IMDs, most LSDs include a clinical variant characterized by primary central nervous system (CNS) involvement. At present, control of the CNS manifestations remains a major challenge because of the inability to deliver therapeutic agents of large molecular weight across the intact BBB (Pastores, 2010). One way to solve this problem that was already used in animal models, is by the local injection in the desired brain region or by injections in the cerebrospinal fluid if broad distribution in the brain is deemed more important (Zalachoras et al., 2011). However, given the

invasive character of this methodology and since the treatment of neurodegenerative disorders with antisense drugs may require life-long repeated dosing, it would be highly preferable to be able to administer such agents systemically. Strategies to improve brain uptake after peripheral delivery are being assayed, including increasing the permeability of the BBB, complexing or conjugating to cationic nanoparticles or CPPs, encapsulation in liposomes conjugated to monoclonal antibodies, or use of exosomes, constituting a promise for the delivery of antisense molecules to CNS tissues (Douglas & Wood, 2013; Pérez et al., 2014). Nevertheless, more studies are needed to optimize these strategies or find new others that in a near future could allow not only the transport of different antisense drugs across the BBB as also their body-wide distribution through noninvasive systemic delivery protocols. Clearly, systemic delivery would be the choice for diseases involving multiorgan pathology as is the case of the majority of LSDs and several other IMDs.

Another hurdle that is delaying the recognition of antisense drugs as effective therapeutics is related to the number of affected patients necessary to conduct robust clinical trials. If in several pre-mRNA processing diseases there is a sufficient number of affected patients (Rigo, 2014), in rare genetic disorders like IMDs, the mutations found in clinical practice, including those affecting splicing, are largely “private” mutations present in only one or few patients, which in some cases may compromise the development of clinical trials (Douglas & Wood, 2011; Pérez et al., 2014). Additionally, if every new antisense designed sequence is classed as a novel therapeutic agent, it will be unfeasible to subject each one individually to all the rigorous drug development tests and trials used in current pharmaceutical practice. These are two of the major challenges facing personalized medicine that must be soon resolved to derive the full benefit promised by AO-based therapies (Douglas & Wood, 2011). To achieve that, concerted efforts must be fostered and the drug regulatory processes must also be adapted to encompass these new medicines. A possibility that is being discussed is to treat all AOs as a unique drug class, rather than considering the specific sequences for each case, which would help to eliminate an important obstacle in the current pipeline of the development of antisense drugs (Douglas & Wood, 2013; Lee & Yokota, 2013; Pérez et al., 2010).

CHAPTER 2

OBJECTIVES

With the recognition of the importance of splicing defects in human disease has come the realization that constitutive and regulated splicing reactions are potential therapeutic targets. This is stimulating the development of new therapeutic approaches to modify or eliminate an mRNA bearing a disease-causing mutation. Therefore, in many rare genetic disorders for which there is to date no effective treatment, the emergence of RNA-based approaches constitutes a powerful alternative or adjunct therapeutic strategy.

Also, an understanding of the mechanisms through which splicing mutations are associated to disease together with the increase of the knowledge about the mechanisms that underlie pre-mRNA splicing regulation is being crucial to disclose new therapeutic targets and approaches.

In this context, the present work had, as major goals, the genetic analysis and splicing mechanisms comprehension of a panel of gene mutations involved in the pathogenesis of LSDs as well as the design and application of splicing therapeutic approaches.

More specifically the following objectives were set:

1. To refine the knowledge on the genetic mechanisms disrupted by a number of LSDs splicing mutations previously selected for study.

2. To identify the *cis* and *trans*-acting factors altered in the presence of the splicing mutations under study, applying both *in silico* bioinformatic analyses and *in vitro* functional studies, thus contributing to a better understanding of the role of specific sequences and proteins in the regulation of the splicing process in general.

3. To design and develop splicing therapeutic approaches (antisense-U1 snRNAs and antisense oligonucleotides) for the selected LSDs splicing mutations and to provide the *in vitro* proof of concept that exploring such strategies would result in splicing correction.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 Article 1

Liliana Matos, Vânia Gonçalves, Eugénia Pinto, Francisco Laranjeira, Maria João Prata, Peter Jordan, Lourdes R. Desviat, Belén Pérez, Sandra Alves. **Functional analysis of splicing mutations in the *IDS* gene and the use of antisense oligonucleotides to exploit an alternative therapy for MPS II**.

Biochimica et Biophysica Acta – Molecular Basis of Disease, 2015; 1852(12):2712-21

3.2 Article 2

Liliana Matos, Vânia Gonçalves, Eugénia Pinto, Francisco Laranjeira, Maria João Prata, Peter Jordan, Lourdes R. Desviat, Belén Pérez, Sandra Alves. **Data in support of a functional analysis of splicing mutations in the *IDS* gene and the use of antisense oligonucleotides to exploit an alternative therapy for MPS II**.

Data in Brief, 2015; 5:810-817

Synopsis

Background and motivation for the study

For the last years our group has been working on the genetic characterization of several LSDs. One of the various studies performed focused on a group of patients with MPS II, a LSD caused by mutations in the *IDS* gene, which led to the identification of a considerable number of mutations that impair the normal splicing process of the gene. A wide range of splicing alterations was observed in the presence of the different mutations. Some occurred at the conventional 5' and 3' ss, having a rather quite predictable outcome, but others had an exonic location (like missense and nonsense mutations) and were further proven to be less typical splicing mutations since they affected the normal splicing pattern due to the creation or disruption of *cis*-acting sequence elements. These observations in addition to the fact that for the *IDS* gene several alternative transcripts are annotated in Vega Genome Browser (http://vega.sanger.ac.uk/) pointed to a complex regulation of the splicing of some regions of this gene that probably required the interaction of several auxiliary SREs. The *IDS* gene seemed therefore to be a good model to deepen the analysis of disrupted splicing mechanisms and its disease-related phenotypes. Indeed, a substantial part of the knowledge on the molecular mechanisms that regulate splicing has been provided by functional studies of specific mutations, as they represent an important field on observation to clarify which elements are disturbed by their presence, ultimately contributing to a better knowledge of the splicing process in general.

Under these grounds, we decided to analyse two mutations in *IDS* exon 3. A missense mutation (c.257C>T; p.P86L) for which the impact on splicing was already known (Alves et al., 2006) and a nonsense mutation (c.241C>T, p.Q81X) that we also suspected to be implicated in the deregulation of the *IDS* exon 3 splicing. The main questions we wanted to address were why and how was the splicing affected by these mutations.

We additionally studied another *IDS* gene mutation, a synonymous mutation located in exon 8 (c.1122C>T) whose pathogenicity was previously associated with a splicing alteration. The c.1122C>T leads to the creation of a new 5' ss inside exon 8 with a higher score than the normal 5' ss. Regarding this mutation, our main interest was to develop a splicing therapeutic strategy using AOs to neutralize the use of the new splice site and restore the normal splicing of exon 8.

Study design and methods overview

- Evaluation of the splicing mechanism of each disease-causing mutation through cDNA analysis and minigene reporter vector functional studies.
- *In silico* identification of the *cis*-acting elements and *trans*-acting splicing factor proteins that were altered in the presence of the mutations in exon 3.
- Elaboration of depletion (RNA interference) and overexpression assays of specific predicted SR and hnRNP proteins to analyse if their binding to *cis*-acting motifs correlated with the ability to affect splicing of the *IDS* exon 3.
- Design and construction, by site-directed mutagenesis, of an *IDS* minigene construct appropriate to assess if the specific *cis*-acting motifs found *in silico* were related to the splicing pattern alteration on exon 3.
- Exploration of AOs therapy (AMO and LNA oligonucleotide) to correct the splicing defect caused by the synonymous mutation on exon 8 of *IDS* gene in patient fibroblasts.

Major results

We have achieved a better comprehension about the molecular mechanisms regulating *IDS* gene splicing. Additionally, a hypothetical model for the regulation of the constitutive splicing mechanism around the 3′ ss of *IDS* exon 3 was proposed.

The application of the antisense therapeutic strategy to correct the splicing defect caused by the exonic mutation c.1122C>T was not successful. Besides the need of additional experiments, the results so far obtained indicate that the application of this type of therapies may be particularly difficult for mutations located in gene regions under complex splicing regulation.

Biochimica et Biophysica Acta 1852 (2015) 2712–2721

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbadis

Functional analysis of splicing mutations in the *IDS* gene and the use of antisense oligonucleotides to exploit an alternative therapy for MPS II

Liliana Matos [a,b], Vânia Gonçalves [c], Eugénia Pinto [d], Francisco Laranjeira [d], Maria João Prata [b,e], Peter Jordan [c], Lourdes R. Desviat [f,g,h], Belén Pérez [f,g,h,1], Sandra Alves [a,*,1]

[a] Research and Development Unit, Department of Human Genetics, INSA, Porto, Portugal
[b] Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal
[c] Research and Development Unit, Department of Human Genetics, INSA, Lisbon, Portugal
[d] Biochemical Genetics Unit, Center for Medical Genetics Jacinto Magalhães, Porto Hospital Center, Porto, Portugal
[e] i3S - Instituto de Investigação e Inovação em Saúde/IPATIMUP - Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal
[f] Centro de Diagnóstico de Enfermedades Moleculares, Centro de Biología Molecular Severo Ochoa, UAM-CSIC, Universidad Autónoma de Madrid, Madrid, Spain
[g] CIBER de Enfermedades Raras (CIBERER), Madrid, Spain
[h] IDIPaz, Madrid, Spain

ARTICLE INFO

Article history:
Received 10 April 2015
Received in revised form 16 September 2015
Accepted 21 September 2015
Available online 25 September 2015

Keywords:
Lysosomal storage disorders
IDS gene
Splicing regulation
Antisense therapy

ABSTRACT

Mucopolysaccharidosis II is a lysosomal storage disorder caused by mutations in the *IDS* gene, including exonic alterations associated with aberrant splicing. In the present work, cell-based splicing assays were performed to study the effects of two splicing mutations in exon 3 of *IDS*, i.e., c.241C>T and c.257C>T, whose presence activates a cryptic splice site in exon 3 and one in exon 8, i.e., c.1122C>T that despite being a synonymous mutation is responsible for the creation of a new splice site in exon 8 leading to a transcript shorter than usual. Mutant minigene analysis and overexpression assays revealed that SRSF2 and hnRNP E1 might be involved in the use and repression of the constitutive 3′ splice site of exon 3 respectively. For the c.1122C>T the use of antisense therapy to correct the splicing defect was explored, but transfection of patient fibroblasts with antisense morpholino oligonucleotides (n = 3) and a locked nucleic acid failed to abolish the abnormal transcript; indeed, it resulted in the appearance of yet another aberrant splicing product. Interestingly, the oligonucleotides transfection in control fibroblasts led to the appearance of the aberrant transcript observed in patients' cells after treatment, which shows that the oligonucleotides are masking an important *cis*-acting element for 5′ splice site regulation of exon 8. These results highlight the importance of functional studies for understanding the pathogenic consequences of mis-splicing and highlight the difficulty in developing antisense therapies involving gene regions under complex splicing regulation.

1. Introduction

Mucopolysaccharidosis type II (MPS II, Hunter syndrome, OMIM# 309900) is an X-linked, recessive, lysosomal storage disorder (LSD) caused by mutations in the *IDS* gene (OMIM* 300823). *IDS* codes for the enzyme iduronate-2-sulfatase (IDS, EC 3.1.6.13), which is required for the degradation of the glycosaminoglycans dermatan and heparan sulfate. The progressive abnormal storage of these compounds within the lysosomes leads to the eventual malfunction of different organs and systems [1]. Although two clinical forms, mild and severe, are generally recognized in MPS II, they are just the ends of a wide spectrum of clinical severity, likely related to the diversity of mutations affecting the *IDS* gene [1].

IDS is located on chromosome Xq28 and spans nine exons [2]. About 20 kb telomeric to the functional gene lies a functional pseudogene showing homology with exons 2 and 3 and introns 2, 3 and 7 [3-5]. More than 500 MPS II causal mutations have been reported in *IDS*, of which around 9% are single nucleotide substitutions that affect splice site signals (www.hgmd.org; [6]). However, this figure is probably an underestimate since the direct analysis of mRNA is not routinely performed in the diagnostic setting. Thus, the impact of missense, nonsense and synonymous mutations involved in the downregulation of splicing

Abbreviations: AMO, antisense morpholino oligonucleotide; AO, antisense oligonucleotide; C, healthy control; ESE, exonic splicing enhancer; ESS, exonic splicing silencer; hnRNP, heterogeneous nuclear ribonucleoprotein; ISE, intronic splicing enhancer; ISS, intronic splicing silencer; LNA, locked nucleic acid; LSD, lysosomal storage disorder; MPS II, mucopolysaccharidosis type II; NC, negative control; P, patient; siRNA, small interfering RNA; SR, arginine–serine; V, vector sequence; WT, wild-type.

* Corresponding author.
E-mail addresses: liliana.matos@insa.min-saude.pt (L. Matos), vania.goncalves@insa.min-saude.pt (V. Gonçalves), eugenia.pinto@chp.min-saude.pt (E. Pinto), francisco.laranjeira@chp.min-saude.pt (F. Laranjeira), mprata@ipatimup.pt (M.J. Prata), peter.jordan@insa.min-saude.pt (P. Jordan), lruiz@cbm.csic.es (L.R. Desviat), bperez@cbm.csic.es (B. Pérez), sandra.alves@insa.min-saude.pt (S. Alves).
[1] Co-last authors.

http://dx.doi.org/10.1016/j.bbadis.2015.09.011
0925-4439/

mechanisms may remain unknown; this is especially plausible for genes with a fine balance between constitutive and cryptic splice sites such as *IDS*, which requires the interaction of several auxiliary *cis*-acting elements for proper splicing to occur [7-9].

Splicing mutations can interfere with accurate pre-mRNA splicing if they affect nucleotides at (or in the vicinity of) acceptor (AG) or donor (GT) splice sites (detectable at the genomic DNA level). Correct consensus sequence recognition also relies on so-called exonic and intronic splicing enhancers (ESE and ISE) or silencers (ESS and ISS). These are usually bound by arginine–serine (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNP) respectively. Regardless of their type (nonsense, missense or synonymous), single nucleotide mutations can interfere with all these internal splicing regulatory elements, causing partial or total exon skipping, intron retention, the creation of ectopic splice sites, and the activation of cryptic sites [7,10,11]. None of these effects can be deduced from the analysis of genomic DNA.

The understanding of basic splicing mechanisms is important if therapeutic strategies that might correct splicing defects are to be designed. Several studies have explored the potential of targeting aberrant pre-mRNA resulting from splicing mutations. These include the use of small molecule compounds that modulate alternative splicing, the overexpression or silencing of splice factors, *trans*-splicing approaches (spliceosome-mediated RNA *trans*-splicing, SMaRT), use of adapted U1 snRNAs complementary to the mutated site and antisense oligonucleotides used to correct mutation-induced aberrant splicing (Splice-switching oligonucleotides) or to redirect splicing to knock-down gene expression [12,13]. One of the most promising genetic therapies of this kind is pre-mRNA targeting by antisense oligonucleotides (AOs). These have been used to mask abnormal or cryptic splice sites and regulatory regions away from the mutated site aiming to restore pre-mRNA processing. AOs hybridize to a selected site in the pre-mRNA, and thus sterically hinder the recognition of a specific region by the spliceosomal machinery. The usage of the natural splice site(s) is thus promoted, recovering the production of normally spliced transcripts [14,15].

The modification of splicing by AOs has been shown effective in β-thalassemia (*HBB* gene) [16], ataxia-telangiectasia (*ATM* gene) [17,18], pyridoxine-dependent epilepsy (*ALDH7A1* gene) [19], hypomyelinating leukodystrophy (*PLP1* gene) [20], Niemann–Pick disease (*NPC1* gene) [21] and organic acidemias (*PCCA*, *PCCB* and *MUT* genes) [22,23]. AOs have also been used to modify alternative splicing in spinal muscular atrophy (*SMN2* gene) [24], and to target the pre-mRNA of genes involved in cancer such as *Bcl-x* [25] and *C-myc* [26,27]. Phenotypic improvements have been achieved using AOs *in vitro* and in animal models of Duchenne muscular dystrophy for which antisense therapy is now in the early phases of clinical trials [14,28].

In a previous study [29] involving Portuguese patients with MPS II, we detected several splicing mutations, some of which gave rise to alternative transcripts. Some involved exon 3 of *IDS*; the pre-mRNA region of this exon is particularly vulnerable to defects in splicing regulation. In the present work, functional analysis based on reporter minigenes was performed for two exon 3 nucleotide changes, c.257C>T [29-31] and c.241C>T [32], which confirmed these to be involved in exon 3 splicing dysregulation. Further, mutant minigene analysis and overexpression assays revealed that the SRSF2 (formerly SC35) and hnRNP E1 proteins might be involved in the use and repression of the constitutive 3′ splice site of exon 3 respectively.

Although enzyme replacement therapy is already available for MPS II, it has some serious limitations, justifying, therefore, the development of novel therapies for this disease. Hence, for MPS II patients' affected by splicing mutations, splicing therapeutics is a potential alternative or an adjunct therapeutic strategy. With this aim, tests were also performed to see whether AOs could be used to circumvent the effect of c.1122C>T [29,33], an exonic nucleotide change that creates a new 5′ splice site inside exon 8. The transfection of three antisense morpholino oligonucleotides (AMOs) and one locked nucleic acid (LNA) in patient fibroblasts did not abolish the abnormal transcript.

2. Material and methods

2.1. Patients and healthy individuals

Fibroblasts were obtained (during routine genetic analysis and under informed consent) from three male Portuguese patients - all carriers of one known mutation (Table 1) - with a biochemical diagnosis of MPS II, and from anonymous healthy donors. The confidentiality of personal data was respected at all times.

2.2. Cell culture and cycloheximide treatment of patients' fibroblasts

Fibroblasts, used as a source of mRNA and genomic DNA, were cultured following standard procedures in DMEM medium (Gibco Invitrogen, Carlsbad, USA) with 10% FBS and 1% kanamycin at 37 °C in a 5% CO_2 atmosphere.

To perform the nonsense-mediated mRNA decay (NMD) assays, fibroblast cells from Patient 1 (c.241C>T) and 2 (c.257C>T) were cultured in the presence of three different concentrations of cycloheximide (0.75, 1 and 2 mg/ml) for 6 h. Total RNA was then isolated and reverse-transcribed as described above. RT-PCR was performed using specific primers for exons 2 and 3 (Table I in [34]).

2.3. Molecular investigation

For molecular genetic analysis, genomic DNA was extracted automatically using a Bio Robot EZ1® system and the EZ1 DNA Tissue Extraction Kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. Specific primers were used to amplify DNA fragments containing the nine exons of the *IDS* gene and their intronic flanking regions [29]. Total mRNA was isolated using the High Pure RNA Isolation Kit (Roche Applied Science, Indianapolis, USA) and reverse-transcribed using the "Ready-To-Go You-Prime First-Strand Beads" Kit (GE Healthcare, Buckinghamshire, UK), following the manufacturer's protocol. For RT-PCR, the specific primers shown in Table I [34] were designed using the Ensembl database (www.ensembl.org; Transcript: IDS-001 ENST00000340855). cDNA amplifications were performed in a total volume of 25 μl using the Hot-Start PCR Mastermix Immomix Red Kit (Bioline, London, UK), following the manufacturer's instructions, employing 0.2 μM of each primer and an annealing temperature of 60 °C (35 cycles). All PCR products were purified directly with Illustra ExoStar 1-Step™ (GE Healthcare, Buckinghamshire, UK), or after gel extraction using the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, USA). Sequencing reactions were then performed in an ABI Prism 3130xl genetic analyzer (Applied Biosystems, Foster City, USA). Sequencing profiles were compared with the *IDS* reference sequence ENSG00000010404 (www.ensembl.org) using the ClustalW2 bioinformatic tool (www.ebi.ac.uk/Tools/msa/clustalw2/).

2.4. Mutation nomenclature

The sequence used as reference for numbering the residues was the Ensembl sequence ENSG00000010404, and the mutation nomenclature was done in accordance with the recommendations of the Human Genome Variation Society (www.hgvs.org/mutnomen/).

2.5. Bioinformatic analysis

Prediction of the splicing scores for the natural, cryptic donor and acceptor splice sites was obtained using the *MaxEntScan* software: Maximum Entropy Model with ideal *MaxEnt* splice site scores of 11.81 for 5′ splice site and 13.59 for 3′ splice site (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) [35]. Modeling of the

Table 1
Genotype and *in silico* analysis of genomic changes in the studied patients with mucopolysaccharidosis type II (MPS II).

Patient	cDNA change[a]	Protein change[a]	Mutation type	*In silico* effect of gDNA change[b]	Reference
1	c.241C>T	p.Q81X (CAA-TAA)	Nonsense/splicing	Reduction in splicing score for 3′ splice site (8.78 to 8.13)	Brusius-Facchin et al. [32] and present study
2	c.257C>T	p.P86L (CCG-CTG)	Missense/splicing	Creates an SRSF1 binding site and eliminates the hnRNP E1, hnRNP E2 and SRSF2 binding sites	Bunge et al. [30], Flomen et al. [31] and Alves et al. [29]
3	c.1122C>T	p.G374G (GGC-GGT)	Synonymous/ splicing	Increase in splicing score for 5′ splice site (7.46 to 9.22)	Popowska et al. [33] and Alves et al. [29]

[a] The mutation nomenclature is that recommended by HGVS. cDNA and protein numbering is based on the reference sequence in Ensembl (Ensembl sequence ENSG00000010404), taking nucleotide +1 as the A of the ATG translation initiation.
[b] *In silico* effect analyzed by *Splicing Rainbow* and *ESEfinder* 3.0 software.

alterations in exonic splicing enhancer or silencer sequences was performed using *ESEfinder* 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home) [36,37] and *Splicing Rainbow* (http://www.ebi.ac.uk/asd) [38] software.

2.6. Minigene construction

For the *in vitro* splicing analysis of the variants c.241C>T and c.257C>T in *IDS* exon 3, the respective regions of patient and healthy control genomic DNA were amplified. Since an *IDS* pseudogene [3-5] exists that shows high homology with the *IDS* gene in the region corresponding to exons 2, 3 and introns 2 and 3, primers for these amplifications were designed to anneal at intron 1 and 3 in such a way that the amplified fragment matched the corresponding gene and not the pseudogene. A BamHI (Forward primer – Intron 1 BamHI F) and XhoI (Reverse primer - Intron 3 XhoI R) restriction enzyme cleavage site tail was added to each oligonucleotide (Table I in [34]). The PCR conditions were as described above. The fresh products were then subcloned into the TOPO vector (Invitrogen, Carlsbad, USA) following the manufacturer's protocol. The inserts were cut with BamHI and XhoI, and purified after gel excision. Using the Rapid Ligation Kit (Roche Applied Science, Indianapolis, USA) they were also cloned into pcDNA3.1-myc, a modified plasmid vector (Invitrogen, Carlsbad, USA) that contains the sequence for the Myc epitope (to be used as a sequence tag) upstream of the multiple cloning site (Fig. 1A in [34]).

To validate the in silico predicted changes in the splicing factors SRSF2 (formerly SC35), hnRNP E1 and hnRNP E2 (also known as α-CP1 or PCBP1 and α-CP2 or PCBP2) in the presence of the c.257C>T mutation, a mutant minigene was designed containing a small deletion (GCCCCGAGCC) in the mutated region which comprises the predicted *cis*-acting motifs for the binding of both mentioned *trans*-acting factors. The wild-type (WT) exon 3 minigene was used as template and the mutated minigene (ex3 del ESE/ESS motifs) was ordered from Nzytech (Lisbon, Portugal).

To functionally investigate the splicing defects caused by c.1122C>T in exon 8, wild-type and mutant minigenes were constructed in vector pSPL3 (Exon Trapping System, Life Technologies, Gibco, NY, USA), a specific vector for splicing analysis in which the multiple cloning site

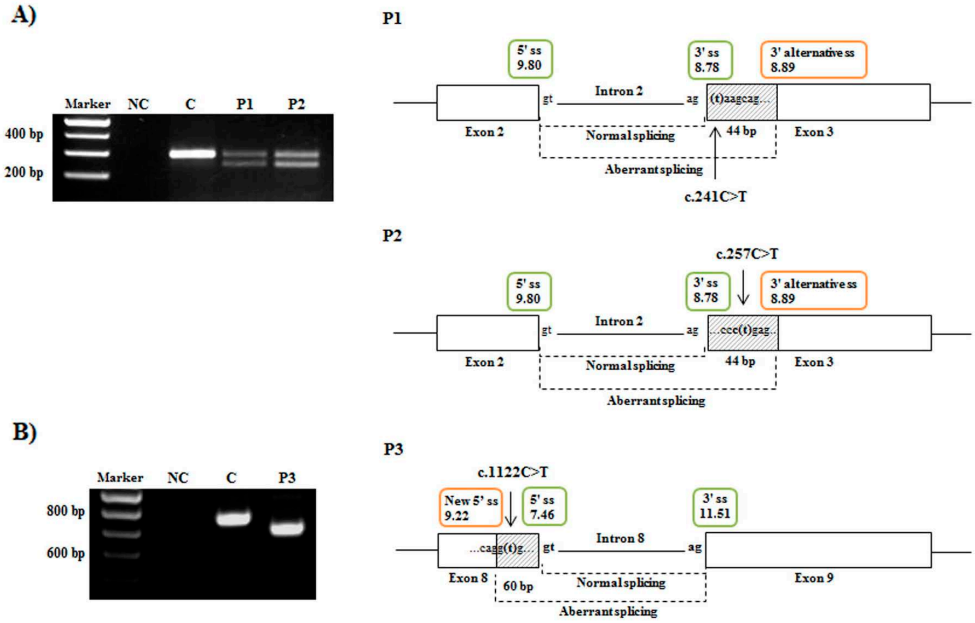


Fig. 1. Agarose gel showing the *IDS* transcripts observed in the three MPS II patients together with schematic views of each transcript's constitution. Sequencing analysis identified the structure of each amplified fragment. The diagrams show the position of the three splicing mutations and its effects on splicing. The splicing scores calculated using *MaxEntScan* software are shown in colored boxes above the corresponding 5′ and 3′ splice sites. A) Negative control (NC); healthy control (C) and Patients 1 (P1) and 2 (P2). Agarose gel electrophoresis after RT-PCR showed one band for the control fibroblasts and two for those of Patients 1 and 2. In both patients the upper band revealed a sequence with exons 2 and 3, with the nucleotide changes c.241C>T (P1) and c.257C>T (P2) in exon 3. The lower band reveals the deletion of the first 44 bp of exon 3 in both patients. B) Negative and healthy control, and Patient 3 (P3). RT-PCR amplification of patient mRNA between exons 7 and 9 revealed an aberrant transcript lacking the last 60 bp of exon 8.

is flanked by functional splice donor and acceptor sites. For this, gene fragments comprising exon 8 and its intronic flanking regions were amplified from the DNA of both control and patient fibroblasts using the primers shown in Table I [34], employing the PCR conditions described above. After subcloning into the TOPO vector, each insert was excised with EcoRI, purified, and cloned into the pSPL3 final vector (Fig. 1C in [34]). All clones were verified by DNA sequencing.

2.7. In vitro splicing assays

The *in vitro* splicing assays were performed in COS-7 and Hep3B cell lines to investigate which one best reproduces the endogenous splicing pattern observed in patients. For that, cells from both lines (4×10^5) were grown in 6-well plates and transfected with WT or mutant minigenes (2 μg) using 4 μl of Lipofectamine 2000 (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions (performed in triplicate). Following the same conditions the mutant minigene (ex3 del ESE/ESS motifs) was transfected in Hep3B cells. At 24 h post-transfection, total RNA was extracted from the cells and used as a template for cDNA synthesis. RT-PCR was then performed as described above, using vector specific pcDNA3.1-myc (K-myc-Bam and pfGHr1) or pSPL3 (SD6 and SA2) primers (Table I in [34]). The amplified products were separated by agarose gel electrophoresis. Following band excision and purification they were subjected to direct sequencing using an ABI Prism 3130xl genetic analyzer (Applied Biosystems, Foster City, USA).

2.8. Overexpression and depletion of splicing factors

To confirm the predicted changes in the splicing factors SRSF2, hnRNP E1 and hnRNP E2, overexpression and depletion studies were performed using plasmids coding for them (kindly provided by Dr. B. Andersen and Dr. S. Liebhaber) and siRNA molecules respectively. Some 4×10^5 Hep3B cells grown in 6-well plates were co-transfected with plasmids encoding the SRSF2, hnRNP E1 or E2 *trans*-acting factors (using 1.5–3.5 μg of each minigene), plus 2 μg of WT or mutant minigenes. All transfections were performed using Lipofectamine 2000 reagent. The amount of plasmid DNA was adjusted as necessary using empty vector (i.e., without any insert). At 48 h post-cotransfection, the cells were harvested and the transcript pattern analyzed by RT-PCR as described above.

Depletion studies were also performed in the Hep3B cell line. Cells at 30–40% confluence were transfected (using Lipofectamine 2000) with siRNA (200 μM) targeting the mRNA of SRSF1 (formerly ASF/SF2) and luciferase (control). These cells were then transfected with the WT or mutant minigene (c.257C>T) 24 h later. RT-PCR analysis was performed 48 h later. All siRNA oligos came from MWG Biotech (Ebersberg, Germany); the sequences are described in Table I [34].

The overexpression of SRSF2 and hnRNP E1 and hnRNP E2 was confirmed in whole cell lysates by quantitative Real-Time PCR (qRT-PCR). Briefly, 1.5 μg of total RNA was reverse-transcribed using NOT *I-d*$(T)_{18}$ bifunctional primer with first-strand mix beads from the Ready-To-Go You-Prime First-Strand Beads Kit (GE Healthcare, Buckinghamshire, UK) according to the manufacturer's protocol. Relative levels of gene expression were analyzed using TaqMan Universal PCR Master Mix 1× (Applied Biosystems) and the TaqMan Gene Expression Kit (which includes primers and probes for each specific splicing factor gene –Table I in [34]), following the recommended PCR conditions.

The relative mRNA levels of target genes were calculated using standard curves (values ranging from 0.05 ng to 50 ng of RNA converted into cDNA). A standard curve was constructed for each target gene relating Ct values (the fractional cycle number at which the amount of amplified target reached a fixed threshold) to log RNA quantities. The normalization of expression was given by the ratio between the RNA concentrations of each target gene and the endogenous gene *PGK1*. The relative amount of RNA was determined via the ratio of the normalized expressions of the target and control samples.

Depletion of SRSF1 was confirmed in whole cell lysates through a qRT-PCR assay as described above and also by Western blotting. For protein analysis, samples were boiled for 10 min and resolved in a 12% SDS-PAGE mini-gel. After electrophoresis, proteins were transferred to a polyvinylidene difluoride membrane using a Mini Trans-Blot cell (Bio-Rad). Immunodetection was carried out using the primary antibodies mouse anti-SF2/ASF clone 96 from Zymed (San Francisco, CA) and anti-α-tubulin from Sigma-Aldrich (Switzerland).

2.9. Antisense oligonucleotide treatment and analysis

AMOs were designed to target the new 5′ splice site created in the presence of the mutation c.1122C>T in exon 8. AMOs were synthesized and purified by Gene Tools (Philomath, OR). A standard 25-mer morpholino control oligonucleotide (also provided by Gene Tools) was used as a negative control (Table I in [34]). Endo-Porter Reagent (Gene Tools) was used as the peptide delivery system, following the manufacturer's recommendations. For the AMO treatment, between 3 and 4×10^5 fibroblasts from control and patient were grown in 6-well plates and after overnight culture different concentrations of each AMO were added with 9 μl/ml of Endo-Porter. The scrape delivery method was also used to help cells incorporate AMOs [39].

To target and block the new 5′ splice site in exon 8, an LNA was designed and synthesized by Exiqon (Vedbaek, Denmark) (Table I in [34]). For LNA exposure, fibroblasts from control and patient were plated under the same conditions as described for AMO treatment. After overnight incubation, different concentrations of LNA were added to cells using Lipofectamine LTX or 2000 (Invitrogen, Carlsbad, USA) as delivery reagents.

After 24 h and 48 h treatment with the different AMOs and LNA, cells were harvested, total mRNA isolated, and RT-PCR performed as described above using the forward and reverse primers for exon 7 and 9 respectively (Table I in [34]).

3. Results

3.1. Genetic analysis of patient-derived fibroblasts

Table 1 shows the genetic defects carried by the patients. Patient 1 was hemizygous for a nucleotide change in exon 3, c.241C>T (p.Q81X) already erroneously reported as p.Q80R by Brusius-Facchin et al. [32]. A detailed analysis of the codon in the HGMD database (Professional 2014.3 Release) showed the mutation reported by these authors was, in fact, Q81X. The impact of this change at the RNA level and its effects on *IDS* pre-mRNA splicing is here addressed for the first time.

Transcriptional analysis of *IDS* cDNA from Patient 1 revealed two transcripts, one corresponding to a normal splicing product with the C>T change in the first nucleotide of exon 3 (predicted to introduce an early stop codon at position 81 [p.Q81X]), and the other corresponding to an aberrant transcript lacking the first 44 bp of exon 3 (predicted to code for a truncated protein with only 83 amino acids) (Fig. 1A). This additional transcript suggests that this point mutation interferes with splicing. Further, since both transcripts bear a premature STOP codon, they are likely targets for nonsense-mediated mRNA decay (NMD), with no protein being produced. When patients' cells were treated with cycloheximide, an inhibitor of the NMD mechanism, a clear increase in the intensity of the band which corresponds to the transcript bearing the nonsense mutation Q81X (Fig. 2) was observed, meaning that this transcript is targeted by NMD. Regarding the abnormal splicing product (with the absence of 44 bp), despite being a potential target for NMD (once leads to a truncated protein, p.Q81EfsX2), no significant changes were observed in the intensity of the corresponding band. The classical NMD pathway is thought to be trigged during the pioneer round of translation when the first ribosome translates the processed mRNA and stalls at a PTC that is located more than 50–55 nucleotides

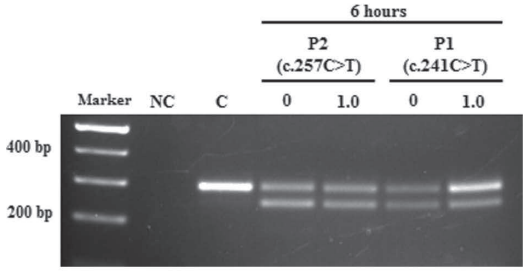


Fig. 2. Cycloheximide treatment. RT-PCR amplification of RNA extracted from mucopolysaccharidosis II patients with the nonsense c.241C>T (P1) and missense c.257C>T (P2) mutations untreated (0 mg/ml) or treated with cycloheximide (1.0 mg/ml) for 6 h. The same results were obtained with 0.75 and 2.0 mg/ml (data not shown). The amplification of a wild-type sample (C) is also presented. NC — negative control; C — control; P — patient; CHX — cycloheximide.

upstream of the last EJC (Exon Junction Complex) [40]. However, there is growing evidence that mutation, codon, gene, cell and tissue-specific differences in NMD efficiency can alter the underlying disease pathology. Moreover, inter-individual variation in NMD efficiency among patients carrying identical mutations have also been reported [41,42]. The differences in the NMD mechanism activation here observed between two transcripts having PTCs could be explained by this recent data.

cDNA analysis was performed for fibroblasts from Patients 2 and 3, who carried mutations already described and previously analyzed in our group [29]. The transcript pattern for Patient 2, who carried the nucleotide change c.257C>T in exon 3, was similar to that observed for Patient 1 (Fig. 1A). Cycloheximide experiments were also performed in this case and, as described for Patient 1, no significant changes in the intensity of the band corresponding to the transcript lacking the 44 bp were observed (Fig. 2). The nucleotide change c.1122C>T carried by Patient 3 (Fig. 1B) was expected to be synonymous in terms of amino acid coding [p.G374G (GGC-GGT)], but the cDNA analysis revealed the presence of a transcript lacking the last 60 bp of exon 8. As previously described [29,30], the normal transcript was also present but in lower quantities. The presence of this transcript even in lower quantities is responsible for the mild phenotype of this patient. In our previous study [29] we have determined patients' enzymatic activities and for this patient we have observed a relatively higher IDS enzymatic activity, in comparison with other MPS II cases.

3.2. Bioinformatic analysis

Splicing scores for the splice site junctions were found compatible with the splicing changes observed in cDNA. For the mutation c.241C>T, *in silico* analysis predicted a reduction in the 3′ splice site score value from 8.78 to 8.13. Under these circumstances the constitutive splice site continued to be used, but the regulation of splicing in this region was perturbed and a downstream cryptic splice site (8.89) was also used, producing a second transcript without the first 44 bp of exon 3 (Fig. 1A). In contrast to what we have observed before [29] for two other *IDS* mutations (c.241 − 2A>G and c.241 − 1G>A) which affect the same splice site, in the case of the c.241C>T change the presence of a transcript with the total skipping of the exon 3 was not detected. In fact, the two other mutations affect the 3′ invariable splice site sequence, being responsible for a drastic reduction of the splice site score. That is not the case of c.214C>T change, once this mutation alters the first base of the exon 3, causing only a slightly decrease in the 3′ splice site score (from 8.78 to 8.13) which might explain the aberrant splicing pattern differences observed.

For the c.257C>T mutation, *in silico* analysis returned a higher splicing score for the 3′ cryptic splice site in exon 3 (8.89) than for its constitutive counterpart (8.78) (Fig. 1A).

The change in c.1122C>T was shown to be associated with the activation of a new 5′ splice site in exon 8, which returned a higher splicing score (9.22) than the original splice site (7.46) (Fig. 1B).

Since the *MaxEntScan*-estimated splicing scores did not completely account for the perturbation in constitutive splicing, further bioinformatic tools were used to search for differences in exonic splicing regulatory elements between the WT and mutated sequences of exon 3. For the nucleotide change c.241C>T, neither the *Splicing Rainbow* [38] nor the *ESEfinder* 3.0 [36,37] software's predicted alterations in any exonic splicing regulatory element. In the presence of the c.257C>T nucleotide change, analysis with *Splicing Rainbow* software predicted the elimination of a binding motif for the hnRNP E1 and hnRNP E2 silencing proteins, and the creation of an SRSF1 (formerly ASF/SF2) recognition motif (Fig. 2A in [34]). Analysis with *ESEfinder* 3.0 software indicated the recognition motif for SRSF1 would be slightly altered, and that a binding site for SRSF2 (formerly SC35) would no longer be recognized in the presence of the mutation (Fig. 2B in [34]).

3.3. Functional analysis of IDS gene splicing mutations

To further study the effect of the three exonic changes on splicing patterns, minigene assays were performed using constructs generated with WT or mutant exons (plus their respective intronic flanking regions) in the pcDNA3.1-myc or pSPL3 vectors (Fig. 1A and C in [34]). Fig. 1B and 1D [34] shows the results of RT-PCR splicing analysis after the transfection of COS-7 and Hep3B cells: the amplified transcripts derived from the WT and mutant minigenes are shown in 1B) for the nucleotide changes in exon 3, and in 1D) for the mutation in exon 8.

The transcript patterns for the three *IDS* mutations were similar in both the Hep3B and COS-7 cells, although slight qualitative differences in expression were detected in the minigene assays.

3.4. Splicing regulatory elements analysis

To validate the predictions for the *trans*-acting factor proteins involved in the 3′ splice site choice of *IDS* exon 3, they were either overexpressed or silenced in Hep3B cells transfected with either WT or mutant minigenes constructed for c.257C>T. cDNA analysis after overexpression of the SRSF2 protein in Hep3B cells with the mutant c.257C>T minigene revealed the rescue of the normal transcript and the disappearance of the aberrant form lacking the first 44 nucleotides of exon 3 (Fig. 3B). This strongly suggests that the choice of the 3′ constitutive splice site of *IDS* exon 3 may be dependent on an ESE site recognized by SRSF2. This recognition is compromised by the mutation c.257C>T.

The c.257C>T mutation was further predicted to eliminate a binding motif for the splicing silencers hnRNP E1 and hnRNP E2. The overexpression of hnRNP E1 with the mutant (c.257C>T) minigene resulted in a substantially increased production of the aberrant transcript, showing that the cryptic splice site activated downstream was preferentially chosen (Fig. 3C). This suggests the involvement of hnRNP E1 in the repression of the 3′ constitutive splice site of exon 3. When hnRNP E2 was overexpressed with the mutant (c.257C>T) minigene, no alteration in the mutant splicing pattern was seen (Fig. 3C). Individual co-transfection of SRSF2, hnRNP E1 or hnRNP E2 plus the WT minigene did not modify the WT splicing pattern (Fig. 3A).

Finally, to determine whether the c.257C>T change interfered with the recognition of the splicing enhancer SRSF1, as bioinformatically predicted, endogenous protein in the Hep3B cells was depleted through RNA interference. The cells were then transfected with WT or mutant minigenes. No changes were detected in the transcriptional profiles of either the WT or mutant sequences (Fig. 3D).

The degree of overexpression or suppression achieved for each splicing factor was analyzed by qRT-PCR using specific probes, and by Western blotting in the case of SRSF1 depletion (see Material and Methods). The relative quantification of each RNA, via the ratio of the

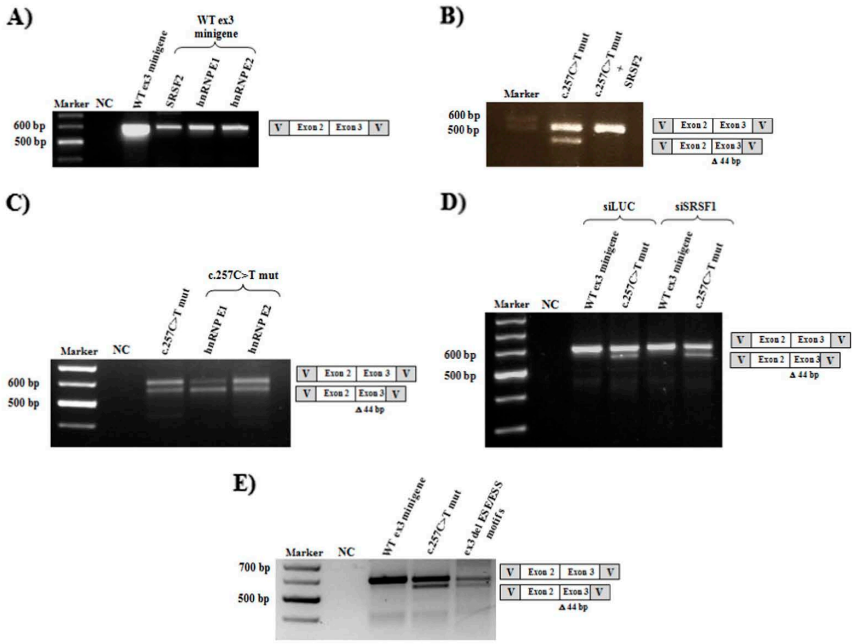


Fig. 3. Analysis of the involvement of specific SR and hnRNP *trans-acting* factors in the regulation of the 3′ splice site of exon 3. The results of RT-PCR analysis using vector-specific primers are shown along with a diagram of the transcripts obtained in Hep3B cells after 24 h or 48 h of incubation. A) Co-transfection of the wild-type exon 3 minigene (2 μg) with each *trans-acting* factor (3.5 μg). B) Overexpression of SRSF2 (formerly SC35) (3.5 μg) after transfection with the c.257C>T mutant minigene (2 μg). C) RT-PCR analysis of processed transcripts after co-transfection with hnRNP E1 or hnRNP E2 (3.5 μg) plus the mutant minigene c.257C>T (2 μg). D) Splicing pattern of the minigene reporter transcripts following depletion of endogenous SRSF1 (formerly ASF/SF2) and luciferase (control). For the silencing of both proteins, 200 μM of each siRNA were used. Note that only the SRSF2 and hnRNP E1 *trans*-acting proteins had an effect (respectively) on the constitutive and alternative splicing of *IDS* exon 3. The same results were obtained when between 1.5 μg and 3 μg of the *trans*-acting factors were overexpressed. E) Transcript pattern obtained after expression of 2 μg of the exon 3 mutant minigene (ex3 del ESE/ESS motifs) with the deletion of 6 bp (GCCCCGAGCC) of predicted exonic motifs with overlapping sequences for the binding of SRSF2 and hnRNP E1 and hnRNP E2 proteins. NC — negative control; V — vector sequence; siRNA — small interfering RNA.

normalized expressions of the target and control samples, confirmed the efficacy of SRSF2, hnRNP E1 and hnRNP E2 overexpression, as well as SRSF1 depletion (Table II in [34]). SRSF1 depletion was also confirmed by immunodetection (Fig. 3 in [34]).

The expression in Hep3B cells of the mutant minigene (ex3 del ESE/ESS motifs) containing the deletion of 6 bp (GCCCCGAGCC) of the predicted *cis*-acting motifs for the binding of SRSF2 and hnRNP E1 and E2 showed the aberrant splicing pattern observed in the presence of the c.257C>T mutation (Fig. 3E). This result, as the ones obtained with the overexpression of SRSF2 and hnRNP E1 proteins, points to the involvement of these two *trans*-acting factors in the regulation of the 3′ constitutive splice site of *IDS* exon 3.

3.5. Antisense oligonucleotide studies in control and patient fibroblasts with the mutation c.1122C>T in IDS gene: a therapeutic approach

Given that the mutation c.1122C>T in exon 8 has a silent effect on the amino acid sequence, the possibility of redirecting transcript processing using modified antisense oligonucleotides of two different types was examined. Three AMOs and one LNA complementary to the region of the newly created 5′ splice site (Fig. 4A) were designed in an attempt to block the access of the splicing machinery to the mutant pre-mRNA, and consequently circumvent the formation of the aberrantly spliced transcript. cDNA analysis was performed 24 h and 48 h after transfection of the patient and control fibroblasts. In patients' fibroblasts, under none of conditions tested (30 or 50 μM of each AMO, and 1, 5, 25, 50 or 100 nM of LNA) was the abnormally spliced mRNA ever abolished. The RT-PCR results (Fig. 4B and C) also showed a transcript skipping exon 8 in the presence of the different AOs, alongside the abnormal transcript showing partial skipping of exon 8. In control fibroblasts transfected with 50 μM of each AMO, the RT-PCR results revealed three different transcripts: one of normal molecular weight, one lacking the last 60 bp of exon 8 and one with the total skipping of exon 8. On the other hand, the transfection of 50 and 100 nM of LNA lead to the appearance of one aberrant transcript presenting the total skipping of exon 8 (Fig. 4D).

The AMO treatment proved to be sequence-specific since an AMO with a scrambled sequence had no effect on the exon 8 splicing pattern in patient fibroblasts (data not shown).

4. Discussion

A number of diseases have been linked to mutations that disrupt the splicing process, which leads to abnormal and/or deficient protein production [9,43]. Splicing defects represent more than 9% of all currently published mutations [6], although this figure might well be underestimated since our knowledge on the regulatory elements that influence splicing in many genes is far from complete. Genomic DNA and cDNA analyses (including functional studies addressing the effect of specific mutations on splicing) are needed if we are to develop therapeutic strategies for correcting molecular defects. In the present work, one of the mutations causing aberrant splicing was a synonymous substitution. This led us to explore the possible use of antisense oligonucleotide therapy for redirecting abnormal pre-mRNA splicing.

Under physiological conditions, the *IDS* gene is associated with at least 8 transcripts produced by alternative splicing (Vega Genome Browser; VEGA: http://vega.sanger.ac.uk/), four of them are protein coding, two are thought to undergo NMD and the last two are processed

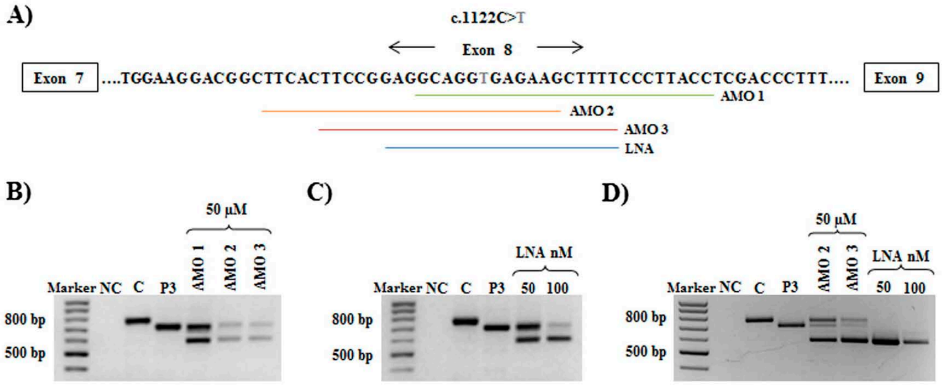


Fig. 4. Antisense oligonucleotide treatment of control and patient fibroblasts. A) Illustration of the *IDS* exon 8 region where the nucleotide change c.1122C>T is located (T change marked in bold gray). The underlined sequences represent each blocking AMO or LNA. B) RT-PCR analysis of mRNA from untreated control cells (C) and from Patient 3 (P3) fibroblasts (0 µM), as well as from P3 cells treated for 24 h with 50 µM of AMO 1, 2 or 3 targeting the new 5′ splice site. C) Same analysis as in B but performed for P3 cells treated for 24 h with 50 and 100 nM of the specific LNA. D) cDNA analysis of control fibroblasts after treatment with 50 µM of AMOs or 50 and 100 nM of LNA. Samples of untreated control (C) and P3 fibroblasts are also presented. The same results were obtained after 48 h of treatment with the three AMOs and LNA (data not shown). NC — negative control.

transcripts that do not have an ORF. This number of transcripts suggests a complex underlying regulatory mechanism that probably requires the interaction of multiple splicing factors. This, of course, also renders this system sensitive to mutations that disrupt splicing. Regulatory sequence elements act *in vivo* as binding sites for specific SR and hnRNP *trans*-acting factors that, respectively, promote and inhibit splice site recognition and spliceosome assembly around a given exon. hnRNPs can disturb the protein-protein interactions of SR proteins, and thus compete with them in splice site recognition [11,44].

For the three disease-causing mutations analyzed in the present work, the *in vitro* splicing analysis of specific minigenes closely reproduced the pattern seen in patient fibroblasts. Therefore, the *cis* and *trans*-acting elements involved in exon 3 splicing regulation were examined in further studies.

The c.257C>T mutation seemed to abolish the exonic splicing enhancer (ESE) for the SRSF2 (formerly SC35) protein, which is predicted to be necessary for the constitutive transcription of exon 3 in the WT context. The overexpression of this SR protein along with the mutant minigene compensated for the aberrant splicing by favoring the selection of the constitutive 3′ splice site by the splicing machinery. In agreement, the *MaxEntScan* splicing score for the 3′ constitutive splice site (8.78) was lower than for the cryptic site situated downstream (8.89). All these findings are consistent with the documented model for ESE function, in which ESE-bound SR proteins are thought to activate weak 3′ splice sites by increasing the recruitment of $U2AF^{65/35}$ and improving its binding stabilization via protein-protein interactions between the SR protein and the $U2AF^{35}$ RS domains [45,46].

The c.257C>T mutation was further predicted to eliminate a binding site for the splicing silencers hnRNP E1 and hnRNP E2; this site overlaps with the interrupted sequence motif for SRSF2 recognition (Fig. 2A in [34]). Overexpression of the hnRNP E1 factor with the mutant minigene resulted in the increased expression of the aberrant transcript, showing that this *trans*-acting factor, which has been reported as a splicing regulator [47-49], may indeed inhibit the recognition of the 3′ constitutive splice site, thus favoring abnormal alternative splicing. Moreover, the transfection of Hep3B cells with a mutant minigene (ex3 del ESE/ESS motifs) presenting the deletion of a 6 bp sequence (CCCGAG) which comprises the predicted *cis*-acting motifs for the binding of SRSF2 and hnRNP E1 lead also to the appearance of the aberrant splicing product with the absence of the 44 bp. This observation further supports the hypothesis that this region contains important *cis*-acting elements for the regulation of the 3′ splice site of exon 3.

The *in vitro* expression or silencing of splicing factors can modulate splicing patterns, revealing that precise combinations and relative ratios of factors dictate regulatory decisions [10,50]. Overall, the present results suggest that the binding of hnRNP E1 to exon 3 can modulate the interaction of SRSF2 with an overlapping site (Fig. 5). The mutant minigene analysis together with the overexpression data for both proteins indicate that SRSF2 binding to the ESE seems to be important for splicing of the upstream intron, and that it might depend on changes in the concentration ratio between the SR protein and hnRNP E1. Further, and notably, the overexpression of hnRNP E1 via the WT minigene showed no alteration of the normal splicing pattern; thus, an excess of hnRNP E1 may be insufficient to antagonize the binding of SRSF2 and disturb constitutive splicing. This suggests that differences in the concentrations of these RNA-binding proteins may be required for exon 3 to be silenced by hnRNP E1. A similar occlusion model involving the hnRNP A1 and SRSF2 protein splicing factors has been described in which the function of hnRNP in *tat* exon 2 splicing was to inhibit the binding of SRSF2 to an overlapping site [51,52]. The ability of hnRNPs to interfere with the binding of SR proteins has also been described in other pre-mRNAs [53-55].

hnRNP E1 and hnRNP E2 are highly homologous proteins, with 89% amino acid similarity [56]. Nonetheless, the overexpression of hnRNP E2 with the WT and mutant constructs did not alter the splicing pattern. This shows that, despite being very alike, they may recognize different sequence motifs. Distinct functions for these proteins have been reported by other authors [49,57].

As a whole, these findings contribute towards our better understanding of the molecular splicing mechanism of *IDS*. The results, like others reported for disease-related mutations [58-61], are of importance for clarifying the basic mechanisms underlying *IDS*-associated splicing, but improve our knowledge of the splicing process in general. This is a prerequisite for the design of new gene-specific therapeutic strategies.

This study also examined whether the aberrant splicing effect produced by the synonymous mutation c.1122C>T in *IDS* exon 8 could be neutralized via the use of specific antisense oligonucleotides. In patient cells, assays were performed with different AMOs and an LNA designed to bind to different sites around the new donor splice site. However, the abnormal transcript was still produced, and a new transcript skipping exon 8 was induced. This total absence of exon 8 led to the in-frame loss of 58 amino acids instead of the 20 lost through the absence of the last 60 nucleotides of exon 8. In addition, the skipping of

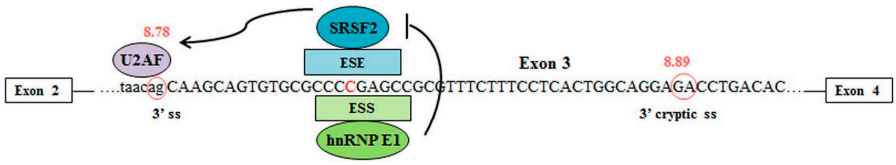


Fig. 5. Hypothetical model for the regulation of the constitutive splicing mechanism around the 3′ splice site of *IDS* exon 3. Exon 3 contains an exonic splicing silencer (ESS) motif recognized by hnRNP E1 that overlaps with an exonic splicing enhancer (ESE) element bound by SRSF2 (formerly SC35). Both the ESE and ESS motifs are eliminated by the c.257C>T mutation. SRSF2 overexpression can rescue the use of the 3′ splice site while hnRNP E1 overexpression increases the use of the downstream cryptic splice site.

exon 8 (recorded at the cDNA level) has been reported to be disease-causing in one patient [62]. Thus, the induction of this transcript would be of no therapeutic use.

Although the tested AOs fulfilled most of the criteria for efficacy, none of them directed exon 8 splicing towards correction. This shows that even though an AO strategy may be appropriately designed, there is no guarantee of success. Negative results have been obtained for 50% (17 of 34) of the AOs designed to inhibit *ICAM-1* expression [63], and 33% (52 of 156) of a series of exon-internal AOs designed to induce exon skipping in the dystrophin gene [64] failed to have the desired effect. Problems of target accessibility caused by the secondary or tertiary mRNA structure, and/or by proteins bound to it, might account for some failures of oligonucleotides [65]. Further, short LNA oligonucleotides (16-mer or even 12-mer) are known to be more specific and effective. The size of the LNA used (20-mer) in the present work, might, therefore explain (at least partly) its lack of effectiveness. It was not possible to use a shorter oligonucleotide to specifically block the region of the new "gt" donor splice site due to the lack of gene specificity of <20-mer sequences; this would certainly derail any attempt at therapy.

It might be thought that the different oligonucleotides obstruct the access of specific SR proteins to their respective ESE binding motifs, thus promoting the skipping of exon 8 [66]. Such a scenario is also predicted by the *in silico* analysis results (summarized in Table 2), which indicate exon 8 to contain a region where AO binding occurs, in which distinct binding motifs are likely present at different positions. This prediction is supported by the results of the AMOs and LNA transfection in control fibroblasts wherein the splicing pattern is changed towards to aberrant forms (including exon 8 skipping), pointing to an important role of the predicted exonic binding motifs in the 5′ splice site regulation of exon 8. In addition, since two of the alternative *IDS* transcripts lack exon 8 (www.ensembl.org), regulation around this exon may be rather complex, compromising the efficacy of the AOs assayed. Studies are required to explore the use of oligonucleotides designed to bind to other regions of exon 8; an oligonucleotide distant from the mutation might have a positive effect on the retention of exon 8, helping more effective solutions for exon 8 recovery to be designed. Until then we should be careful in consider definitively that an antisense therapeutic strategy is not applicable to the splicing rescue of this mutation. Nevertheless, and as is highlighted by the present study, one such approach may be particularly challenging for this region of the *IDS* gene.

Since antisense therapy is mutation-targeted, and the target defects usually involve 'private' mutations (i.e., they affect only one person and his or her family members), it can be understood as personalized medicine. Efforts are currently underway to allow AOs to be regarded as a single drug class; this could dramatically accelerate progress in the development of antisense drug therapies for RNA mis-splicing diseases [67,68].

To our knowledge this is the first attempt to modulate an MPS II-causing splicing mutation using antisense oligonucleotides. Although little success was enjoyed, the results raise new questions regarding the use of antisense therapy in diseases involving finely regulated gene systems.

Table 2
In silico prediction of the effect of exonic splicing enhancer (ESE) binding motifs on the *IDS* exon 8 mutant sequence (c.1122C>T) using *ESR*search (http://ibis.tau.ac.il/ssat/ESR.htm) [69], *RESCUE-ESE* (http://genes.mit.edu/burgelab/rescue-ese/) [70] and *ESEfinder* 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home) [36,37] software. Each column shows the ESE binding motifs present in the mutant exon 8 region where antisense oligonucleotides (AOs) binding occurs. Number of base pairs of exon 8 = 174 bp. The AO binding region lies between nucleotides 97 and 134. Numbers in brackets refer to the first nucleotide position of the predicted binding motif.

IDS exon 8 sequence	*ESR*search	*RESCUE-ESE*	*ESEfinder* 3.0
SR protein binding motifs	TTCCGG (103) TGAGAAGCTT (116) TGAGAAGC (116) GAGAAG (117) TACCTC (132)	TGAGAA (116) GAGAAG (117) AGAAGC (118)	GGCAGGT (110)

Transparency document

The Transparency document associated with this article can be found in the online version.

Acknowledgments

LM was supported by a grant (SFRH/BD/64592/2009) from the *Fundação para a Ciência e Tecnologia IP* (FCT)/POPH/FSE, Portugal.

Research involving human participants

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Disclosure of potential conflicts of interest

The authors have no conflicts of interest to disclose.

References

[1] E.F. Neufeld, J. Muenzer, Part 12: 78 The Mucopolysaccharidoses, in: C.R. Scriver, A.L. Beaudet, W.S. Sly, D. Valle (Eds.), The Metabolic and Molecular Bases of Inherited Disease, 8th e, vol. II, McGraw-Hill, New York 2001, pp. 2465–2495.
[2] P.J. Wilson, C.P. Morris, D. Anson, T. Occhiodoro, J. Bielicki, P.R. Clements, J.J. Hopwood, Hunter syndrome: isolation of an iduronate-2-sulfatase cDNA clone and analysis of patient DNA, Proc. Natl. Acad. Sci. 87 (1990) 8531–8535.
[3] M.L. Bondeson, H. Malmgren, N. Dahl, B.M. Carlberg, U. Petterson, Presence of an IDS-related locus (IDS2) in Xq28 complicates the mutational analysis of Hunter syndrome, Eur. J. Hum. Genet. 3 (1995) 219–227.
[4] M. Rathmann, S. Bunge, C. Steglich, E. Schwinger, A. Gal, Evidence for an iduronate-sulfatase pseudogene near the functional Hunter syndrome gene in Xq27.3-q28, Hum. Genet. 95 (1995) 34–38.

[5] K.M. Timms, F. Lu, Y. Shen, C.A. Pierson, D.M. Muzny, Y. Gu, D.L. Nelson, R.A. Gibbs, 130 kb of DNA sequence reveals two new genes and a regional duplication distal to the human iduronate-2-sulfate sulfatase locus, Genome Res. 5 (1995) 71–78.
[6] P.D. Stenson, M. Mort, E.V. Ball, K. Shaw, A.D. Phillips, D.N. Cooper, The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine, Hum. Genet. 133 (2014) 1–9.
[7] L. Cartegni, S.L. Chew, A.R. Krainer, Listening to silence and understanding nonsense: exonic mutations that affect splicing, Nat. Rev. Genet. 3 (2002) 285–298.
[8] S. Lualdi, M.G. Pittis, S. Regis, R. Parini, A.E. Allegri, F. Furlan, B. Bembi, M. Filocamo, Multiple cryptic splice sites can be activated by IDS point mutations generating misspliced transcripts, J. Mol. Med. (Berl.) 84 (2006) 692–700.
[9] A.J. Ward, T.A. Cooper, The pathobiology of splicing, J. Pathol. 220 (2010) 152–163.
[10] A.J. Matlin, F. Clark, C.W.J. Smith, Understanding alternative splicing: towards a cellular code, Nat. Rev. Mol. Cell Biol. 6 (2005) 386–398.
[11] Z. Wang, C.B. Burge, Splicing regulation: from a parts list of regulatory elements to an integrated splicing code, RNA 14 (2008) 802–813.
[12] M.A. Havens, M.D. Dominik, M.L. Hastings, Targeting RNA splicing for disease therapy, Wiley Interdiscip. Rev. RNA 4 (3) (2013) 247–266.
[13] B. Pérez, M. Ugarte, L.R. Desviat, V.A. Erdmann, RNA-based therapies for inherited metabolic diseases, in: J. Barciszewski (Ed.), RNA Technologies. From Nucleic Acids Sequences to Molecular Medicine, Springer-Verlag Berlin Heidelberg 2012, pp. 357–370.
[14] R. Kole, A.R. Krainer, S. Altman, RNA therapeutics: beyond RNA interference and antisense oligonucleotides, Nat. Rev. Drug Discov. 11 (2012) 125–140.
[15] B. Pérez, L. Rodríguez-Pascau, L. Vilageliu, D. Grinberg, M. Ugarte, L.R. Desviat, Present and future of antisense therapy for splicing modulation in inherited metabolic disease, J. Inherit. Metab. Dis. 33 (2010) 397–403.
[16] S. Svasti, T. Suwanmanee, S. Fucharoen, H.M. Moulton, M.H. Nelson, N. Maeda, O. Smithies, R. Kole, RNA repair restores hemoglobin expression in IVS2-654 thalassemic mice, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 1205–1210.
[17] L. Du, J.M. Pollard, R.A. Gatti, Correction of prototypic ATM splicing mutations and aberrant ATM function with antisense morpholino oligonucleotides, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 6007–6012.
[18] L. Du, R. Kayali, C. Bertoni, F. Fike, H. Hu, P.L. Iversen, R.A. Gatti, Arginine-rich cell-penetrating peptide dramatically enhances AMO-mediated ATM aberrant splicing correction and enables delivery to brain and cerebellum, Hum. Mol. Genet. 20 (2011) 3151–3160.
[19] B. Pérez, G.-.S. LG, A. Verdú, B. Merinero, P. Yuste-Checa, P. Ruíz-Sala, R. Calvo, A. Jalan, L.L. Marín, O. Campos, M.A. Ruíz, M. San Miguel, M. Vázquez, M. Castro, I. Ferrer, R. Navarrete, L. Ruiz-Desviat, P. Lapunzina, M. Ugarte, C. Pérez-Cerdá, Clinical, biochemical, and molecular studies in pyridoxine-dependent epilepsy. Antisense therapy as possible new therapeutic option, Epilepsia 54 (2013) 239–248.
[20] S. Regis, F. Corsolini, S. Grossi, B. Tappino, D.N. Cooper, M. Filocamo, Restoration of the normal splicing pattern of the *PLP1* gene by means of an antisense oligonucleotide directed against an exonic mutation, PLoS One 8 (2013), e73633.
[21] L. Rodríguez-Pascau, M.J. Coll, L. Vilageliu, D. Grinberg, Antisense oligonucleotide treatment for a pseudoexon-generating mutation in the NPC1 gene causing Niemann–Pick type C disease, Hum. Mutat. 30 (2010) E993–E1001.
[22] B. Pérez, A. Rincón, A. Jorge-Finnigan, E. Richard, B. Merinero, M. Ugarte, L.R. Desviat, Pseudoexon exclusion by antisense therapy in methylmalonic aciduria (MMAuria), Hum. Mutat. 30 (2009) 1676–1682.
[23] A. Rincón, C. Aguado, L.R. Desviat, R. Sánchez-Alcudia, M. Ugarte, B. Pérez, Propionic and methylmalonic acidemia: antisense therapeutics for intronic variations causing aberrantly spliced messenger RNA, Am. J. Hum. Genet. 81 (2007) 1262–1270.
[24] Y. Hua, T.A. Vickers, H.L. Okunola, C.F. Bennett, A.R. Krainer, Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice, Am. J. Hum. Genet. 82 (2008) 834–848.
[25] D.R. Mercatante, C.D. Bortner, J.A. Cidlowski, R. Kole, Modification of alternative splicing of Bcl-x pre-mRNA in prostate and breast cancer cells, J. Biol. Chem. 276 (2001) 16411–16417.
[26] R.V. Giles, D.G. Spiller, R.E. Clark, D.M. Tidd, Antisense morpholino oligonucleotide analog induces missplicing of C-myc mRNA, Antisense Nucleic Acid Drug Dev. 9 (1999) 213–220.
[27] H.S. Sekhon, C.A. London, M. Sekhon, P.L. Iversen, G.R. Devi, c-MYC antisense phosphorodiamidate morpholino oligomer inhibits lung metastasis in a murine tumor model, Lung Cancer 60 (2008) 347–354.
[28] T. Koo, M.J. Wood, Clinical trials using antisense oligonucleotides in Duchenne muscular dystrophy, Hum. Gene Ther. 24 (2013) 479–488.
[29] S. Alves, M. Mangas, M.J. Prata, G. Ribeiro, L. Lopes, H. Ribeiro, J. Pinto-Basto, M.R. Lima, L. Lacerda, Molecular characterization of Portuguese patients with mucopolysaccharidosis type II shows evidence that the IDS gene is prone to splicing mutations, J. Inherit. Metab. Dis. 29 (2006) 743–754.
[30] S. Bunge, C. Steglich, C. Zuther, M. Beck, C.P. Morris, E. Schwinger, A. Schinzel, J.J. Hopwood, A. Gal, Iduronate-2-sulfatase gene mutations in 16 patients with mucopolysaccharidosis type II (Hunter syndrome), Hum. Mol. Genet. 2 (1993) 1871–1875.
[31] R.H. Flomen, P.M. Green, D.R. Bentley, F. Giannelli, E.P. Green, Detection of point mutations and a gross deletion in six hunter syndrome patients, Genomics 13 (1992) 543–550.
[32] A.C. Brusius-Facchin, S. Iv, C. Zimmer, M.G. Ribeiro, A.X. Acosta, D. Horovitz, I.L. Monlleó, M.I. Fontes, A. Fett-Conte, R.P. Sobrinho, A.R. Duarte, R. Boy, P. Mabe, M. Ascurra, M. de Michelena, K.L. Tylee, G.T.N. Besley, M.C.V. Garreton, R. Giugliani, S. Leistner-Segal, Mucopolysaccharidosis type II: identification of 30 novel mutations among Latin American patients, Mol. Genet. Metab. 111 (2014) 133–138.
[33] E. Popowska, M. Rathmann, A. Tylki-Szymanska, S. Bunge, C. Steglich, E. Schwinger, A. Gal, Mutations of the iduronate-2-sulfatase gene in 12 Polish patients with mucopolysaccharidosis type II (Hunter syndrome), Hum. Mutat. 5 (1995) 97–100.
[34] L. Matos, V. Gonçalves, E. Pinto, F. Laranjeira, M.J. Prata, P. Jordan, L.R. Desviat, B. Pérez, S. Alves, Data in support of a functional analysis of splicing mutations in the *IDS* gene and the use of antisense oligonucleotides to exploit an alternative therapy for MPS II, Data Brief (2015) Submitted.
[35] G. Yeo, C.B. Burge, Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals, J. Comput. Biol. 11 (2004) 377–394.
[36] L. Cartegni, J. Wang, Z. Zhu, M.Q. Zhang, A.R. Krainer, ESEfinder: a web resource to identify exonic splicing enhancers, Nucleic Acids Res. 31 (2003) 3568–3571.
[37] P.J. Smith, C. Zhang, J. Wang, S.L. Chew, M.Q. Zhang, A.R. Krainer, An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers, Hum. Mol. Genet. 15 (2006) 2490–2508.
[38] S. Stamm, J.J. Riethoven, V. Le Texier, C. Gopalakrishnan, V. Kumanduri, Y. Tang, N.L. Barbosa-Morais, T.A. Thanaraj, ASD: a bioinformatics resource on alternative splicing, Nucleic Acids Res. 34 (2006) D46–D55.
[39] M. Partridge, A. Vincent, P. Matthews, J. Puma, D. Stein, J. Summerton, A simple method for delivering morpholino antisense oligos into the cytoplasm of cells, Antisense Nucleic Acid Drug Dev. 6 (1996) 169–175.
[40] J.N. Miller, D.A. Pearce, Nonsense-mediated decay in genetic disease: friend or foe? Mutat. Res. Rev. Mutat. Res. 762 (2014) 52–64.
[41] L.S. Nguyen, M.F. Wilkinson, J. Gecz, Nonsense-mediated mRNA decay: inter-individual variability and human disease, Neurosci. Biobehav. 46 (2014) 175–186.
[42] C. Seoighe, C. Gehring, Heritability in the efficiency of nonsense-mediated mRNA decay in humans, PLoS ONE 5 (2010), e11657.
[43] M.A. Lewandowska, The missing puzzle piece: splicing mutations, Int. J. Clin. Exp. Pathol. 6 (2013) 2675–2682.
[44] A. Busch, K.J. Hertel, Evolution of SR protein and hnRNP splicing regulatory factors, Wiley Interdiscip. Rev. RNA 3 (2012) 1–12.
[45] B.R. Graveley, K.J. Hertel, T. Maniatis, The role of U2AF^{35} and U2AF^{65} in enhancer-dependent splicing, RNA 7 (2001) 806–818.
[46] P. Zuo, T. Maniatis, The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing, Genes Dev. 10 (1996) 1356–1368.
[47] S.A. Akker, S. Misra, S. Aslam, E.L. Morgan, P.J. Smith, B. Khoo, S.L. Chew, Pre-spliceosomal binding of U1 small nuclear ribonucleoprotein (RNP) and heterogeneous nuclear E1 is associated with suppression of a growth hormone receptor pseudoexon, Mol. Endocrinol. 21 (2007) 2529–2540.
[48] Q. Meng, S.K. Rayala, A.E. Gururaj, A.H. Talukder, B.W. O'Malley, R. Kumar, Signaling-dependent and coordinated regulation of transcription, splicing, and translation resides in a single coregulator, PCBP1, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 5866–5871.
[49] K. Woolaway, K. Asai, A. Emili, A. Cochrane, hnRNP E1 and E2 have distinct roles in modulating HIV-I gene expression, Retrovirology 4 (2007) 28.
[50] J.F. Cáceres, A.R. Kornblihtt, Alternative splicing: multiple control mechanisms and involvement in human disease, Trends Genet. 18 (2002) 186–193.
[51] R. Martinez-Contreras, P. Cloutier, L. Shkreta, J.-F. Fisette, T. Revil, B. Chabot, Chapter 8: hnRNP proteins and splicing control, in: B.J. Blencowe, B.R. Graveley (Eds.), Alternative Splicing in the Postgenomic era, Advances in experimental medicine and biology series, vol. 623, Landes Bioscience, New York 2007, pp. 123–147.
[52] A.M. Zahler, C.K. Damgaard, J. Kjems, M. Caputi, SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 *tat* exon 2 splicing, J. Biol. Chem. 279 (2004) 10077–10084.
[53] J.B. Crawford, J.G. Patton, Activation of α-tropomyosin exon 2 is regulated by the SR protein 9G8 and heterogeneous nuclear ribonucleoproteins H and F, Mol. Cell. Biol. 26 (2006) 8791–8802.
[54] A.J. Pollard, A.R. Krainer, S.C. Robson, N.G. Europe-Finner, Alternative splicing of the adenylyl cyclase stimulatory G-protein Gα$_s$ is regulated by SF2/ASF and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and involves the use of an unusual TG 3′-splice site, J. Biol. Chem. 277 (2002) 15241–15251.
[55] J. Zhu, A. Mayeda, A.R. Krainer, Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins, Mol. Cell 8 (2001) 1351–1361.
[56] A. Chaudhury, P. Chander, P.H. Howe, Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: focus on hnRNP E1's multifunctional regulatory roles, RNA 16 (2010) 1449–1462.
[57] Y. Zhu, Y. Sun, X.O. Mao, K.L. Jin, D.A. Greenberg, Expression of poly(C)-binding proteins is differentially regulated by hypoxia and ischemia in cortical neurons, Neuroscience 110 (2002) 191–198.
[58] L. Cartegni, M.L. Hastings, J.A. Calarco, E. de Stanchina, A.R. Krainer, Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2, Am. J. Hum. Genet. 78 (2006) 63–77.
[59] V. Gonçalves, P. Theisen, O. Antunes, A. Medeira, J.S. Ramos, P. Jordan, G. Isidro, A missense mutation in the APC tumor suppressor gene disrupts an ASF/SF2 splicing enhancer motif and causes pathogenic skipping of exon 14, Mutat. Res. 662 (2009) 33–36.
[60] A.J. Richards, A. McNinch, B. Whitakker Treacy, K. Oakhill, A. Poulson, M.P. Snead, Splicing analysis of unclassified variants in COL2A1 and COL11A1 identifies deep intronic pathogenic mutations, Eur. J. Hum. Genet. 20 (2012) 552–558.
[61] G. Ulzi, V.A. Sansone, F. Magri, S. Corti, N. Bresolin, G.P. Comi, S. Lucchiari, In vitro analysis of splice site mutations in the CLCN1 gene using the minigene assay, Mol. Biol. Rep. 41 (2014) 2865–2874.

[62] V. Ricci, S. Regis, M. Di Duca, M. Filocamo, An Alu-mediated rearrangement as cause of exon skipping in Hunter disease, Hum. Genet. 112 (2003) 419–425.
[63] V. Patzel, U. Steidl, R. Kronenwett, R. Haas, G. Sczakiel, A theoretical approach to select effective antisense oligodeoxyribonucleotides at high statistical probability, Nucleic Acids Res. 27 (1999) 4328–4334.
[64] A. Aartsma-Rus, L. van Vliet, M. Hirschi, A.A. Janson, H. Heemskerk, C.L. de Winter, S. de Kimpe, J.C. van Deutekom, P.A. 't Hoen, G.J. van Ommen, Guidelines for antisense oligonucleotide design and insight into splice-modulating mechanisms, Mol. Ther. 17 (2009) 548–553.
[65] A. Dias, C.A. Stein, Antisense oligonucleotides: basic concepts and mechanisms, Mol. Cancer Ther. 1 (2002) 347–355.
[66] J.C. Long, J.F. Caceres, The SR protein family of splicing factors: master regulators of gene expression, Biochem. J. 417 (2009) 15–27.
[67] S.M. Hammond, M.J. Wood, Genetic therapies for RNA mis-splicing diseases, Trends Genet. 27 (2011) 196–205.
[68] J.A.J. Lee, T. Yokota, Antisense therapy in neurology, J. Perspect. Med. 3 (2013) 144–176.
[69] A. Goren, O. Ram, M. Amit, H. Keren, G. Lev-Maor, I. Vig, T. Pupko, G. Ast, Comparative analysis identifies exonic splicing regulatory sequences —the complex definition of enhancers and silencers, Mol. Cell 22 (2006) 769–781.
[70] W.G. Fairbrother, R.F. Yeh, P.A. Sharp, C.B. Burge, Predictive identification of exonic splicing enhancers in human genes, Science 297 (2002) 1007–1013.

Data in Brief 5 (2015) 810–817

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data article

Data in support of a functional analysis of splicing mutations in the *IDS* gene and the use of antisense oligonucleotides to exploit an alternative therapy for MPS II

Liliana Matos [a,b], Vânia Gonçalves [c], Eugénia Pinto [d], Francisco Laranjeira [d], Maria João Prata [b,e], Peter Jordan [c], Lourdes R. Desviat [f,g,h], Belén Pérez [f,g,h,1], Sandra Alves [a,*,1]

[a] *Research and Development Unit, Department of Human Genetics, INSA, Porto, Portugal*
[b] *Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal*
[c] *Research and Development Unit, Department of Human Genetics, INSA, Lisbon, Portugal*
[d] *Biochemical Genetics Unit, Center for Medical Genetics Jacinto Magalhães, Porto Hospital Center, Porto, Portugal*
[e] *i3S - Instituto de Investigação em Saúde/IPATIMUP - Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal*
[f] *Centro de Diagnóstico de Enfermedades Moleculares, Centro de Biología Molecular Severo Ochoa, UAM-CSIC, Universidad Autónoma de Madrid, Madrid, Spain*
[g] *CIBER de Enfermedades Raras (CIBERER), Madrid, Spain*
[h] *IDIPaz, Madrid, Spain*

ARTICLE INFO

Article history:
Received 23 September 2015
Accepted 12 October 2015
Available online 28 October 2015

ABSTRACT

This data article contains insights into the methodology used for the analysis of three exonic mutations altering the splicing of the *IDS* gene: c.241C > T, c.257C > T and c.1122C > T.

We have performed splicing assays for the wild-type and mutant minigenes corresponding to these substitutions. In addition, bioinformatic predictions of splicing regulatory sequence elements as well as RNA interference and overexpression experiments were conducted.

The interpretation of these data and further extensive experiments into the analysis of these three mutations and also into the

DOI of original article: http://dx.doi.org/10.1016/j.bbadis.2015.09.011
* Corresponding author.
E-mail addresses: liliana.matos@insa.min-saude.pt (L. Matos), vania.goncalves@insa.min-saude.pt (V. Gonçalves), eugenia.pinto@chp.min-saude.pt (E. Pinto), francisco.laranjeira@chp.min-saude.pt (F. Laranjeira), mprata@ipatimup.pt (M.J. Prata), peter.jordan@insa.min-saude.pt (P. Jordan), lruiz@cbm.csic.es (L.R. Desviat), bperez@cbm.csic.es (B. Pérez), sandra.alves@insa.min-saude.pt (S. Alves).
[1] Co-last authors.

http://dx.doi.org/10.1016/j.dib.2015.10.011

methodology applied to correct one of them can be found in "Functional analysis of splicing mutations in the *IDS* gene and the use of antisense oligonucleotides to exploit an alternative therapy for MPS II" Matos et al. (2015) [1].

Specifications table

Subject area	*Biology*
More specific subject area	*Molecular genetics*
Type of data	*Text, figures, tables and graphs*
How data was acquired	*Agarose gel images*, quantitative real-time PCR, *western blotting, in silico analyses (Splicing rainbow and ESEfinder 3.0 software)*
Data format	*Raw, analyzed*
Experimental factors	*Cell lines (Hep3B and COS-7) treated with different minigenes*
Experimental features	*Wild-type and mutant minigenes were transfected in COS-7 and Hep3B cell lines and the splicing patterns analysed by RT-PCR. SRSF1 (formerly ASF/SF2) siRNAs and plasmids coding for SRSF2 (formerly SC35), hnRNP E1 and hnRNP E2 splicing factors were transfected in Hep3B cells and the experiments analysed by quantitative real-time PCR and/or Western blotting. In silico predictions were done using Splicing rainbow and ESEfinder 3.0 software.*
Data source location	*INSA, Porto, Portugal*
Data accessibility	*All data provided within the article*

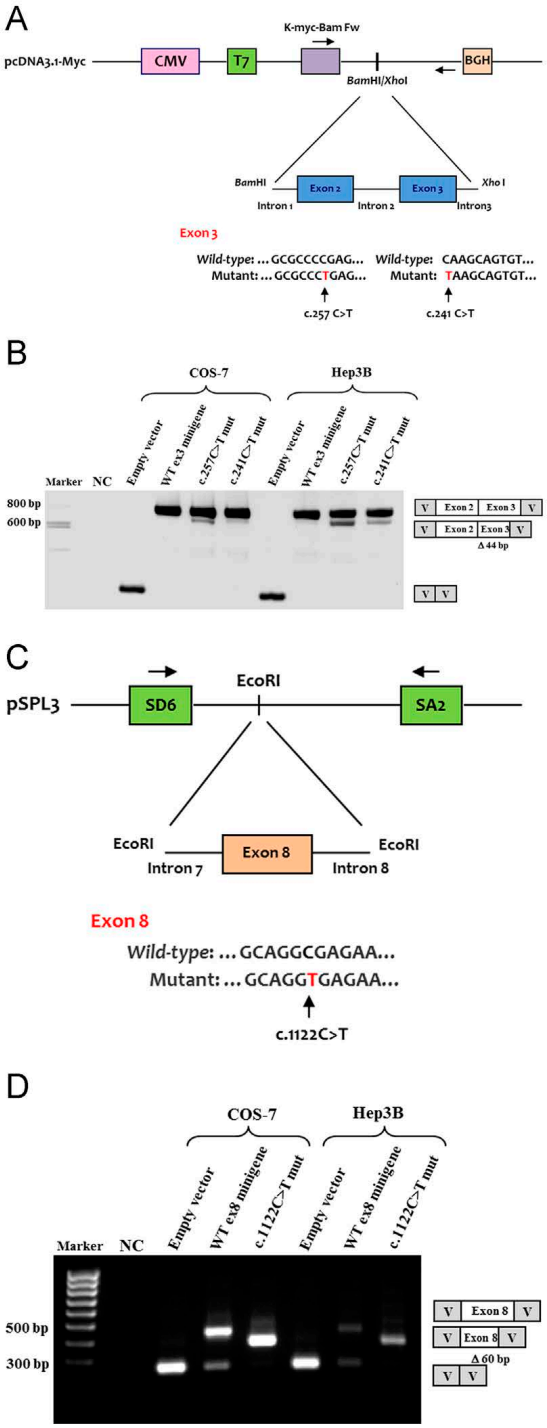
Value of the data

- Data show the methodology for the analysis of the effect of three exonic mutations on splicing.
- Experimental and *in silico* data are presented.
- The data may be valuable for future studies addressing the impact of mutations in splicing.

1. Data, materials and methods

The *IDS gene* encodes the lysosomal hydrolase iduronate-2-sulphatase, the enzyme that is deficient in the X-Linked Lysosomal Storage Disease; Mucopolysaccharidosis type II [2].

Here, we performed cell-based functional splicing assays to deeper analyze the effects of two splicing mutations located in exon 3 of *IDS*, c.241C > T and c.257C > T and one in exon 8, c.1122C > T that were also studied in Matos et al. [1]. The pathogenic effects of these mutations are shown in Fig. 1. Also, all the data relative to oligonucleotides sequences used in the work are depicted in Table 1. Furthermore, to identify the putative SR proteins involved in the splicing regulation we have undertaken bioinformatic predictions of splicing regulatory elements (SREs) in the *IDS* exon 3 (where


A
pcDNA3.1-Myc
CMV
T7
K-myc-Bam Fw
BGH
BamHI/XhoI
BamHI
Exon 2
Exon 3
Xho I
Intron 1
Intron 2
Intron3
Exon 3
Wild-type: ... GCGCCCCGAG...
Mutant: ... GCGCCCTGAG...
c.257 C>T
Wild-type: CAAGCAGTGT...
Mutant: TAAGCAGTGT...
c.241 C>T
B
COS-7
Hep3B
Empty vector
WT ex3 minigene
c.257C>T mut
c.241C>T mut
Marker
NC
800 bp
600 bp
V Exon 2 Exon 3 V
V Exon 2 Exon 3 V
Δ 44 bp
V V
C
pSPL3
SD6
EcoRI
SA2
EcoRI
Intron 7
Exon 8
EcoRI
Intron 8
Exon 8
Wild-type: ... GCAGGCGAGAA...
Mutant: ... GCAGGTGAGAA...
c.1122C>T
D
COS-7
Hep3B
Empty vector
WT ex8 minigene
c.1122C>T mut
Marker
NC
500 bp
300 bp
V Exon 8 V
V Exon 8 V
Δ 60 bp
V V

the mutations c.241C > T and c.257C > T are located) using *ESEfinder* 3.0 and Splicing Rainbow software (Fig. 2 and 3). Finally, we have conducted RNAi and overexpression experiments that were quantified by Real time PCR and Western blot (Tables 1, 2 and Fig. 4).

1.1. Oligonucleotides sequences

See Table 1.

1.2. Minigenes construction and in vitro functional splicing analysis

For the *in vitro* splicing analysis of the variants c.257C > T and c.241C > T in *IDS* exon 3, the respective regions of patient and healthy control genomic DNA were amplified and cloned into pcDNA3.1-myc, a modified plasmid vector (Invitrogen, Carlsbad, USA) (Table 1 and Fig. 1A).

Also, to functionally investigate the splicing defects caused by c.1122C > T in exon 8, wild-type (WT) and mutant minigenes were constructed in vector pSPL3 (Exon Trapping System, Life Technologies, Gibco, NY, USA) (Table 1 and Fig. 1C). To perform the functional splicing assays, Hep3B and COS-7 cell lines (4×10^5) were grown in 6-well plates and transfected with WT or mutant minigenes (2 μg) using 4 μl of Lipofectamine 2000 (Invitrogen, Carlsbad, USA). At 24 h post-transfection, total RNA was extracted from the cells and used as a template for cDNA synthesis. RT-PCR was then performed using vector specific primers (Table 1) and the amplified products were separated by agarose gel electrophoresis (Fig. 1B and D).

1.3. Bioinformatic analysis of SREs in IDS exon 3;

In silico predictions for alterations in exonic splicing enhancer or silencer sequences (splicing regulatory elements – SREs) in the presence or absence of c.257C > T and c.241C > T mutations in *IDS* exon 3 were performed using *ESEfinder* 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home) [3,4], and *Splicing Rainbow* (http://www.ebi.ac.uk/asd) [5] software (Figs. 2 and 3).

1.4. Quantitative Real time PCR and Western blot assays to confirm overexpression and depletion of splicing factors

To confirm the predicted changes in the splicing factors SRSF2 (formerly SC35), hnRNP E1 and hnRNP E2, overexpression studies were performed using plasmids coding for them which were co-transfected in Hep3B cells with WT or mutant c.257C > T minigenes. All transfections were performed using Lipofectamine 2000 reagent. At 48 h post-cotransfection, the cells were harvested and the transcript pattern analyzed by RT-PCR. Depletion studies were also performed to verify the predicted changes for the SRSF1 (formerly ASF/SF2) splicing factor. Hep3B cells were firstly transfected with siRNAs targeting the mRNA of SRSF1 and luciferase (control) and 24 h later transfected with the WT

Fig. 1. Splicing assays for the wild-type (WT) and mutant minigenes corresponding to the *IDS* nucleotide changes c.257C > T and c.241C > T in exon 3, and c.1122C > T in exon 8. (A, C) Diagrams of the reporter minigenes used in the functional splicing experiments. Normal and mutated genomic *IDS* sequences were cloned into the pcDNA3.1-myc or pSPL3 vectors to generate the indicated minigenes. Exons are shown by boxes and introns by straight lines. For all exonic alterations, the WT and mutant regions are shown and the specific changes marked by an arrow. (B, D) Wild-type and mutant minigenes were transfected into COS-7 and Hep3B cells and the splicing pattern analyzed by RT-PCR using the indicated vector-specific primers (arrows in diagrams A and C). Minigene expression of the splicing mutations in exon 3 (c.257C > T and c.241C > T) revealed two transcripts, a predominant one with exon 2 and a mutated exon 3, and a minor transcript of smaller size in which the first 44 nucleotides of exon 3 were missing. The WT minigene produced a single transcript of normal size (B). For the synonymous c.1122C > T change in exon 8, the mutant minigene showed a single transcript lacking the last 60bp of exon 8. The WT construct produced two bands, one corresponding to the transcript with exon 8 inserted, the other to a transcript resulting from an expected splicing event between the vector splice sites (D). A diagram of the bands characterized by sequence analysis is also provided. NC – negative control; V – vector sequence.

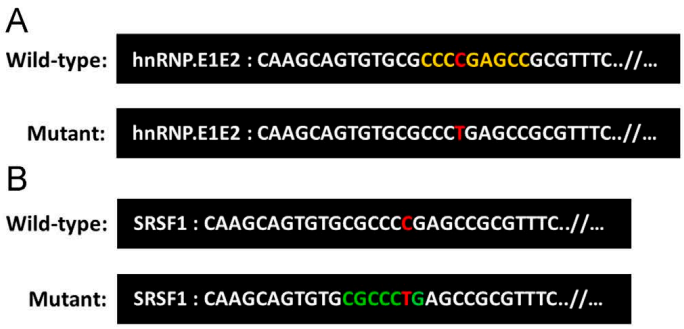


Fig. 2. Bioinformatic predictions of splicing regulatory sequence elements in IDS exon 3 using the Splicing Rainbow software. For the mutation c.241C > T, no alterations were predicted. In the presence of the c.257C > T mutation, no hnRNP E1 or hnRNP E2 silencer motif is recognized (A) and a SRSF1 (formerly ASF/SF2) motif is created (B). The mutated nucleotide is marked in red.

or mutant c.257C > T minigenes. RT-PCR analysis was performed 48 h later. The sequences of all siRNAs are described in Table 1.

To confirm the overexpression of SRSF2, hnRNP E1 and hnRNP E2 as also the depletion of SRSF1 in whole cell lysates, quantitative Real-Time PCR (qRT-PCR) assays were performed. Relative levels of gene expression were analyzed using Taqman Universal PCR Master Mix 1x (Applied Biosystems) and the Taqman Gene Expression Kit (which includes primers and probes for each specific splicing factor gene – Table 1). The relative mRNA levels of target genes were calculated using standard curves (values ranging from 0.05 ng to 50 ng of RNA converted into cDNA). A standard curve was constructed for each target gene relating Ct values to log RNA quantities. The normalization of expression was given by the ratio between the RNA concentrations of each target gene and the endogenous gene

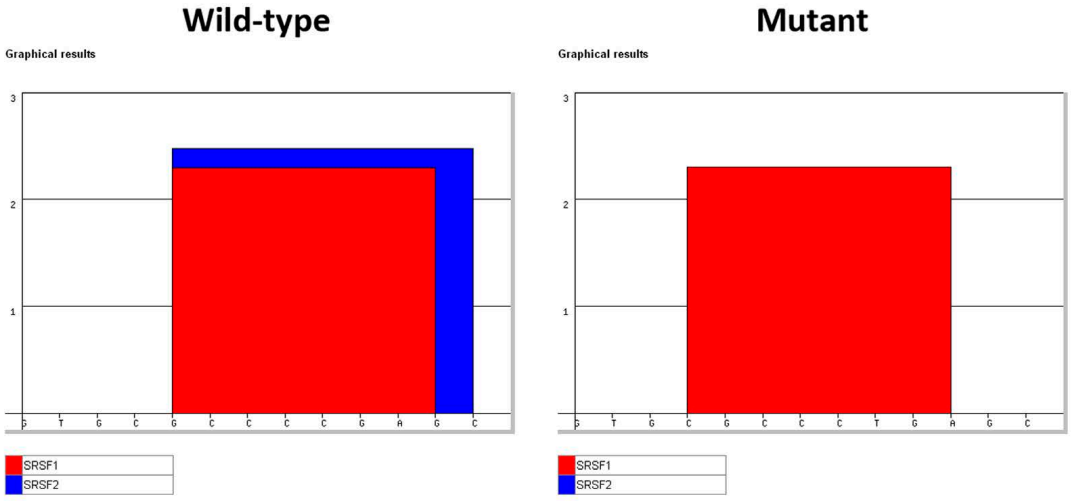


Fig. 3. Bioinformatic predictions of splicing regulatory sequence elements in *IDS* exon 3 using the *ESEfinder* 3.0 software. For the mutation c.241C > T, no alterations were predicted. In the presence of the c.257C > T mutation, a putative binding motif for SRSF1 (formerly ASF/SF2) is slightly altered (red square) and a binding motif for SRSF2 (formerly SC35) is eliminated (blue square).

Table 1
Description of the different sequences used in the work: primers for PCR amplifications[a]; small interfering RNA's for silencing of specific genes; probes used for real-time quantitative PCR; antisense oligonucleotide sequences.

Oligonucleotide	Sequence (5′ to 3′)
Cloning fragments	
Intron 1 BamHI F	ATATATGGATCCTCCAGCCTTGGGCCTCTT
Intron 3 XhoI R	ATATATCTCGAGGAATGCTGGATTCAGACA
Intron 7 IDS F	TATCTCGAGGAACCGCCACAGAGTCCT
Intron 8 IDS R	TATGGATCCGCACCTGTTCCTTTGTCC
RT-PCR fragments	
Exon 2 IDS F	TCATCATCGTGGATGACCTG
Exon 3 IDS R	AAAGACTTTTCCCACCGACA
Exon 7 IDS F	GGAAAATCCGCCAGAGCTAC
Exon 9 IDS R	GATCTCCACCTTGGGAATCA
Plasmid vector primers	
K-myc-Bam F	TACCGCCACCATGGAGCAGAA
pfGHr1 R	TTTATTAGGAAAGGACAGTGGG
SD6 F	TCTGAGTCACCTGGACAACC
SA2 R	ATCTCAGTGGTATTTGTGAGC
Small interfering RNA's	
siSRSF1(formerly ASF/SF2)-a	AGAAGAUAUGACCUAUGCA
siSRSF1 (formerly ASF/SF2)-b	GCAGGUGAUGUAUGUUAUG
siLUC (control)	CGUACGCGGAAUACUUCGA
TaqMan gene expression assays	
SRSF2 (formerly SC35)	Hs 00427515_g1 (Applied Biosystems)
SRSF1 (formerly ASF/SF2)	Hs 00199471_m1 (Applied Biosystems)
hnRNP E1 (PCBP1)	Hs 00362410_s1 (Applied Biosystems)
hnRNP E2 (PCBP2)	Hs 01590472_mH (Applied Biosystems)
Antisense Oligonucleotides	
AMO 1 IDS	GTAAGGGAAAAGCTTCTCACCTGCC
AMO 2 IDS	TTCTCACCTGCCTCCGGAAGTGAAG
AMO 3 IDS	AAAGCTTCTCACCTGCCTCCGGAAG
AMO Standard	CCTCTTACCTCAGTTACAATTTATA
LNA IDS	AAAGCTTCTCACCTGCCTCC

[a] Primers were designed according to the sequence described in the ENSEMBL database (www.ensembl.org; ENSG00000010404). F – Forward; R – Reverse.

PGK1. The relative amount of RNA was determined *via* the ratio of the normalized expressions of the target and control samples (Table 2).

The depletion of the SRSF1 splicing factor was also confirmed through Western blotting analysis. The immunodetection was carried out using the primary antibodies mouse anti-SF2/ASF clone 96 from Zymed (San Francisco, CA) and anti-α-tubulin from Sigma-Aldrich (Switzerland) (Fig. 4).

Acknowledgments

LM was supported by a Grant (SFRH/BD/64592/2009) from the Fundação para a Ciência e Tecnologia IP (FCT)/POPH/FSE, Portugal.

Table 2
Quantification and normalization ratio for PGK1, SRSF2 (formerly SC35), hnRNP E1, hnRNP E2 and SRSF1 (formerly ASF/SF2) expression levels obtained through quantitative real-time PCR. Calculations of relative amounts of each target and endogenous reference RNA were determined using the appropriate standard curve.

Sample	Amount of *PGK1* RNA (ng)	Amount of *trans*-acting factor RNA (ng)	Normalized amount of *trans*-acting factor RNA (ng)	Ratio
SRSF2 overexpression				
WT ex3 minigene + Empty vector	1.2	1.4	1.2	1
c.257C > T mut + Empty vector	1.3	1.8	1.3	1
WT ex3 minigene + SRSF2	1.5	14.6	9.7	8.1
c.257C > T mut + SRSF2	1.2	17.9	14.6	11.2
hnRNP E1 overexpression				
WT ex3 minigene + Empty vector	0.45	0.2	0.45	1
c.257C > T mut + Empty vector	0.9	0.95	1.06	1
WT ex3 minigene + hnRNP E1	0.68	6.34	9.32	20.71
c.257C > T mut + hnRNP E1	0.46	8.73	18.98	17.91
hnRNP E2 overexpression				
WT ex3 minigene + Empty vector	0.76	0.22	0.29	1
c.257C > T mut + Empty vector	1.03	0.78	0.76	1
WT ex3 minigene + hnRNP E2	0.96	2.04	2.13	7.34
c.257C > T mut + hnRNP E2	0.76	3.89	5.12	6.74
SRSF1 silencing				
siLUC + WT ex3 minigene	0.58	0.55	0.95	1
siLUC + c.257C > T mut	0.99	0.77	0.78	1
siSRSF1 + WT ex3 minigene	1.16	0.15	0.13	0.14
siSRSF1 + c.257C > T mut	0.56	0.08	0.14	0.18

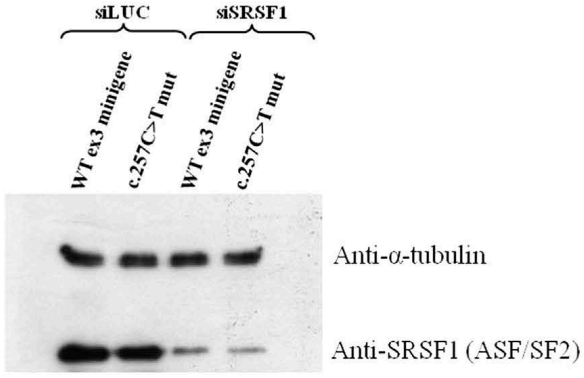


Fig. 4. Western blot analysis of the degree of depletion achieved after transfection of siRNAs for luciferase (control) and SRSF1 (formerly ASF/SF2) plus the WT exon 3 or c.257C > T mutation minigenes. The immunoblot shows the suppression of the endogenous SRSF1 when silencing was performed using 200 μM of the siRNA pool. α-tubulin was used as a loading control.

References

[1] L. Matos, V. Gonçalves, E. Pinto, F. Laranjeira, M.J. Prata, P. Jordan, L.R. Desviat, B. Pérez, S. Alves, Functional analysis of splicing mutations in the *IDS* gene and the use of antisense oligonucleotides to exploit an alternative therapy for MPS II, Biochim. Biophys. Acta 1852 (2015) 2712–2721.
[2] E.F. Neufeld, J. Muenzer, Part 12: 78. The mucopolysaccharidoses, in: 8th ed.,in: C.R. Scriver, A.L. Beaudet, W.S. Sly, D. Valle (Eds.), The Metabolic and Molecular Bases of Inherited Disease, Vol. II, McGraw-Hill, New York, 2001, pp. 2465–2495.

[3] L. Cartegni, J. Wang, Z. Zhu, M.Q. Zhang, A.R. Krainer, ESEfinder: a web resource to identify exonic splicing enhancers, Nucleic Acids Res. 31 (2003) 3568–3571.
[4] P.J. Smith, C. Zhang, J. Wang, S.L. Chew, M.Q. Zhang, A.R. Krainer, An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers, Hum. Mol. Genet. 15 (2006) 2490–2508.
[5] S. Stamm, J.J. Riethoven, V. Le Texier, C. Gopalakrishnan, V. Kumanduri, Y. Tang, N.L. Barbosa-Morais, T.A. Thanaraj, ASD: a bioinformatics resource on alternative splicing, Nucleic Acids Res. 34 (2006) D46–D55.

3.3 Article 3

Liliana Matos, Isaac Canals, Larbi Dridi, Yoo Choi, Maria João Prata, Peter Jordan, Lourdes R. Desviat, Belén Pérez, Alexey V. Pshezhetsky, Daniel Grinberg, Sandra Alves, Lluïsa Vilageliu. **Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations**.

Orphanet Journal of Rare Diseases, 2014; 9:180

Synopsis

Background and motivation for the study

This study involved the development of innovative therapeutic strategies for Sanfilippo C splicing mutations. Sanfilippo C or MPS IIIC is a very rare LSD caused by mutations in the *HGSNAT,* a gene that was cloned in 2006 (Hrebícek et al., 2006). Since then, many molecular defects associated to MPS IIIC have been reported worldwide. The Portuguese MPS IIIC patients were molecularly characterised by our group and two splicing mutations were reported: c.372-2A>G (Coutinho et al., 2008) and c.234+1G>A (unpublished data). In this work two distinct therapeutic approaches were developed for different splicing mutations in the *HGSNAT* gene, one potentially able to correct effects at the RNA level (using modified U1 snRNAs) and the other one at the protein level (using chaperones). Therefore, the target mutations were previously evaluated in terms of those that could be corrected at RNA level and those whose effect could be ameliorated at protein level. The first group included three nucleotide changes affecting distinct 5' SDSs (c.234+1G>A, c.633+1G>A and c.1542+4dupA) and the second group only had one mutation affecting a 3' acceptor splice site (c.372-2A>G).

The study was performed in collaboration with two research groups also working in MPS IIIC, the group of Professor Daniel Grinberg from the University of Barcelona, Spain and the group of Professor Alexey Pshezhetsky from the University of Montreal, Canada. Our group together with the group of Professor Daniel Grinberg was dedicated to the application of modified U1 snRNAs therapeutic approaches for the correction of the mutations c.234+1G>A, c.633+1G>A and c.1542+4dupA. The group of Professor Alexey Pshezhetsky was involved in the use of a pharmacological chaperone (glucosamine) to correct the aberrant folding of a mutant HGSNAT protein lacking 4 amino acids that was encoded from an abnormal splicing transcript originated in the presence of the mutation c.372-2A>G.

Study design and methods overview

- Construction of minigene reporter vectors for each donor site mutation under study and their transfection in COS-7 cells (Portuguese and Spanish groups).
- Construction through site-directed mutagenesis of different U1 snRNA vectors adapted to each mutated SDS under study (Portuguese group).

- Development of U1 snRNA therapeutic approaches for the correction of the different 5' SDS mutations in *HGSNAT*, both through the transfection of the different adapted U1 vectors along with the mutation-disease minigenes in COS-7 cells as well as directly in control and patients' fibroblasts (Portuguese and Spanish groups).
- Assessment of the effect of a pharmacological chaperone in the correction of the aberrant folding of a mutant protein with the loss of four amino acids (Canadian group).

Major results

A partial correction of the splicing pattern (almost 50%) was achieved for the c.234+1G>A mutation with the application of a modified U1 snRNA that completely matched the mutant SDS in fibroblasts of patients carrying this mutation, which is one of the most prevalent MPS IIIC causing mutations in Spanish and Portuguese patients. For the mutations c.633+1G>A and c.1542+4dupA, no rescue was observed after the modified U1 snRNAs overexpression in patients' cells.

Moreover, the glucosamine treatment resulted in an increase in the enzymatic activity, indicating a partial recovery of the correct folding. Nevertheless, as this part of the study goes beyond the scope of this thesis, the obtained results will not be mentioned or discussed in the other sections of this thesis.

Matos et al. Orphanet Journal of Rare Diseases (2014) 9:180
DOI 10.1186/s13023-014-0180-y

ORPHANET JOURNAL OF RARE DISEASES

RESEARCH Open Access

Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations

Liliana Matos[1,2†], Isaac Canals[3,4,5†], Larbi Dridi[6], Yoo Choi[6], Maria João Prata[2,7], Peter Jordan[8], Lourdes R Desviat[4,9], Belén Pérez[4,9], Alexey V Pshezhetsky[6,10], Daniel Grinberg[3,4,5*], Sandra Alves[1†] and Lluïsa Vilageliu[3,4,5†]

Abstract

Background: Mutations affecting RNA splicing represent more than 20% of the mutant alleles in Sanfilippo syndrome type C, a rare lysosomal storage disorder that causes severe neurodegeneration. Many of these mutations are localized in the conserved donor or acceptor splice sites, while few are found in the nearby nucleotides.

Methods: In this study we tested several therapeutic approaches specifically designed for different splicing mutations depending on how the mutations affect mRNA processing. For three mutations that affect the donor site (c.234 + 1G > A, c.633 + 1G > A and c.1542 + 4dupA), different modified U1 snRNAs recognizing the mutated donor sites, have been developed in an attempt to rescue the normal splicing process. For another mutation that affects an acceptor splice site (c.372-2A > G) and gives rise to a protein lacking four amino acids, a competitive inhibitor of the HGSNAT protein, glucosamine, was tested as a pharmacological chaperone to correct the aberrant folding and to restore the normal trafficking of the protein to the lysosome.

Results: Partial correction of c.234 + 1G > A mutation was achieved with a modified U1 snRNA that completely matches the splice donor site suggesting that these molecules may have a therapeutic potential for some splicing mutations. Furthermore, the importance of the splice site sequence context is highlighted as a key factor in the success of this type of therapy. Additionally, glucosamine treatment resulted in an increase in the enzymatic activity, indicating a partial recovery of the correct folding.

Conclusions: We have assayed two therapeutic strategies for different splicing mutations with promising results for the future applications.

Keywords: Splicing mutations, Modified U1 snRNAs, Glucosamine, Sanfilippo C syndrome, Lysosomal storage disorder

Background

Sanfilippo syndrome or Mucopolysaccharidosis III (MPS III), is a group of autosomal recessive lysosomal storage disorders caused by mutations in genes encoding enzymes responsible for heparan sulfate degradation [1]. There are four types of the disease, depending on the gene affected. They all present similar clinical symptoms, including severe central nervous system degeneration accompanied by mild somatic manifestations [2]. Sanfilippo syndrome type C (MPS IIIC) is caused by mutations in the *HGSNAT* gene. This gene codes for the acetyl-CoA: α-glucosaminide N-acetyltransferase (EC 2.3.1.78), a protein localized in the lysosomal membrane which catalyses the acetylation of the terminal glucosamine residues of heparan sulphate prior to their hydrolysis by α-*N*-acetyl glucosaminidase [3]. The *HGSNAT* gene, identified by two independent groups in 2006 [4,5], is located at the chromosome 8 (8p11.1) and contains 18 exons. The cDNA encodes a polypeptide of either 635 or 663 amino acids, since there is a controversy concerning the real initiation codon [6,7]. To date, 64 mutations have been reported, 16 of which (25%) involve splicing alterations: 13 are described as splicing mutations and three as small

* Correspondence: dgrinberg@ub.edu
†Equal contributors
[3]Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain
[4]CIBER de Enfermedades Raras (CIBERER), Madrid, Spain
Full list of author information is available at the end of the article

deletions and duplications that affect the splicing process (HGMD®Professional Spring 2014.1 Release).

Splicing is an essential step for the expression of most of human genes, in which the 5' and 3' splice sites (ss), the branch point sequence and the polypyrimidine tract (both the last two within 50 nucleotides upstream of the 3' ss) play a fundamental role. These sites present sequence variability throughout the human genome. The splicing process is conducted by the spliceosome, which is formed of five small nuclear RNAs (snRNAs) and more than 200 different proteins (reviewed in [8]). The U1 snRNA, which presents a partially complementary sequence to the 5' ss, is essential for the recognition of the 5' ss consensus motif (CAG/GTRAGT, exon/intron, R = purine). The U1 snRNP is composed of a 164-nt long U1 snRNA and several protein factors. The 5' region of the U1 snRNA is involved in the recognition of the 5' ss, with the C8 nucleotide being the one that binds the first nucleotide (G) of the intron (reviewed in [9]). Application of modified U1 snRNAs to improve recognition of mutated 5' ss represents a new strategy for recovering the normal splicing process. They have been assayed as a therapeutic approach for different diseases and splicing mutations affecting different positions of the 5' ss with variation in the efficacy of the treatment [10-19]. Recently, assays correcting an exogenous injected construct have been performed using modified U1 snRNAs in mice as a treatment for severe human factor VII deficiency [20].

In some cases, alternative splicing caused by specific mutations can give rise to misfolded proteins, which may be prone to rapid intracellular degradation (reviewed in [21]). Molecular chaperones are proteins that act on the correct folding of other polypeptides in cells. Pharmacological and chemical chaperones are small compounds that can be used, in a similar way, to avoid the misfolding of mutant proteins. They are principally potent enzyme inhibitors which interact specifically with their active sites to restore the correct folding and to increase stability [22]. In the case of lysosomal storage disorders, once in the lysosome, the enzyme substrate replaces the chaperone, thereby completing restoration of enzyme activity [23]. Aminosugars and iminosugars are the most common pharmacological chaperones used in enzyme enhancement therapy (EET) for lysosomal disorders. EET have been reported for several of these diseases including Fabry disease, G_{M1}-gangliosidosis, Morquio B disease, Pompe disease, Gaucher disease, Krabbe disease, Niemann-Pick A/B and C diseases; as well as for many other types of disorders such as retinitis pigmentosa, cystic fibrosis, Parkinson's disease, Alzheimer's disease and cancer (reviewed in Ref. [21]). In the case of Sanfilippo syndrome type C, glucosamine, a competitive HGSNAT inhibitor, has been shown to increase residual enzyme activity in cultured skin fibroblasts from patients affected with a number of missense mutations [24].

In this work we focus on *HGSNAT* mutations that affect the splicing process. Three previously described splicing mutations [5,25,26] were the object of the study with different modified U1 snRNAs to improve recognition of the donor ss and enhance the correct splicing process. The studies were performed using cells transfected with minigene constructs bearing the specific mutation as well as cultured patients' skin fibroblasts. Mutations include the most frequent change in Moroccan patients (c.234 + 1G > A), a mutation found in Spanish patients (c.633 + 1G > A) and a mutation found in a French patient (c.1542 + 4dupA). Furthermore, EET approach, was tested for a splicing mutation c.372-2A > G. This mutation is the most prevalent in Spanish and Portuguese patients [25,27] and affects the acceptor site at the end of the 4th intron of the *HGSNAT* gene [27], thereby altering the splicing process. The use of a downstream alternative cryptic ss generates an mRNA with an in-frame deletion that codes for a protein with the loss of four amino acids (p. [L125_R128del]). Here, we show the effect of this mutation, which reduces enzyme activity, and the recovery of a part of that activity through treatment with glucosamine as a chaperone.

Our results show that, depending on the context of the mutated donor site, modified U1 snRNAs can be a promising therapeutic tool. The use of glucosamine as a chaperone improved enzyme activity, suggesting a therapeutic effect of this compound for Sanfilippo C patients.

Methods

Mutation analysis of the *HGSNAT* gene

This study included five MPS IIIC patients: three previously described, two Spanish and one Moroccan [25]; and two recently diagnosed, one French and one Portuguese, carrying mutations already reported by us [25,26] (Table 1). Studies were approved by the authors' Institutional Ethics Committee and conducted under the Declaration of Helsinki. Patients were encoded to protect their confidentiality. Genetic analysis was performed using control and patients' fibroblast cell lines as the source of RNA and genomic DNA whenever necessary. Total RNA was

Table 1 Genotype and origin of MPS IIIC affected patients

Patient	Origin	Allele 1*	Allele 2*	Reference
SFCP	Portugal	c.234 + 1G > A	c.234 + 1G > A	This study
SFC3	Morocco	c.234 + 1G > A	c.234 + 1G > A	[25]
SFC6	Spain	c.633 + 1G > A	c.1334T > C	[25]
SFC7	Spain	c.372-2A > G	c.372-2A > G	[25]
SFC13	France	c.1542 + 4dupA	c.1150C > T	This study

*Mutation nomenclature is based on cDNA sequence (NM_152419.2), with nucleotide +1 corresponding to the A of the ATG translation initiation codon.

extracted using High Pure RNA Isolation Kit (Roche, Basel, Switzerland) and converted into cDNA using High-Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, CA) following the manufacturers' instructions. The RT-PCR amplifications were performed using the primers described in Additional file 1: Table S1 and information regarding specific conditions for each cDNA amplification and sequence analysis is available upon request.

Minigene cloning and U1 snRNA expression constructs

For the *in vitro* splicing approaches, wild type (WT) and mutant minigenes were constructed for each mutation under study. A gene fragment including exon 2 and the flanking intronic regions was amplified from the DNA of the c.234 + 1G > A patient's fibroblast cells using the primers described in Additional file 1: Table S1, and cloned into the TOPO vector (Life Technologies). The insert was excised with EcoRI, purified using Wizard® SV Gel and PCR clean-up system (Promega, Madison, WI), and subsequently cloned into the pSPL3 vector (Exon Trapping System, Life Technologies; kindly provided by Dr. B. Andresen) using the Rapid Ligation Kit (Roche Applied Science, Mannheim, Germany). Theclone containing the desired mutant insert in the correct orientation was identified by restriction enzyme analysis and DNA sequencing. The exon 6 (c.633 + 1G > A) and exon 15 (c.1542 + 4dupA), together with their intronic flanking sequences, were also cloned in the pSPL3 vector. These mutated minigenes were ordered from GenScript (Piscataway, NJ). In all cases, the WT minigenes were obtained by site-directed mutagenesis using the QuikChange II XL site-directed Kit (Agilent Technologies, Santa Clara, CA) following the manufacturer's instructions and nucleotide changes were confirmed by sequencing analysis. The primers used are listed in Additional file 1: Table S1.

To express WT U1 snRNA (U1-WT), we used the vector pG3U1, which includes the sequence coding for human U1 ([28]; kindly provided by Dr. F. Pagani). The different U1 vectors adapted to the donor ss of *HGSNAT* exon 2, exon 6 and exon 15 were obtained by site-directed mutagenesis using the QuikChange II XL site-directed Kit (Agilent Technologies). For each construct (U1 suppressor 1 to 9) the PCR reaction was performed with the specific primers shown in Additional file 1: Table S1. The desired mutations were confirmed by sequence analysis.

Cell culture and U1 transfection experiments

To perform the splicing assays, COS-7 cells and fibroblasts were grown in Dulbecco's modified Eagle's medium (Sigma-Aldrich, St. Louis, MO) supplemented with 10% fetal bovine serum (Life Technologies) and 1% PenStrep (Life Technologies) at 37°C with 5% CO_2. For co-transfection experiments, COS-7 cells at 90% of confluence were transfected in 6 or 12-well plates with either 1 or 2 μg of each WT or mutant minigene and 1 to 4 μg of each different mutation adapted U1 snRNA vector, using Lipofectamine 2000 (Life Technologies) according to the manufacturer's protocol. When required, the amount of DNA was adjusted with the pSPL3 empty vector. For splicing analysis of the endogenous *HGSNAT* transcripts, healthy control and patients' fibroblasts at 90% of confluence were transfected in 6-well plates with 1, 2.5 or 3.5 μg of each modified U1 snRNA vector using either Lipofectamine 2000 or LTX (Life Technologies) according to the manufacturer's instructions. To estimate transfection efficiency, healthy control and patients' cells were transfected with a control plasmid encoding either GFP or RFP and fluorescent cells were monitored by microscopy.

RT-PCR transcript analysis after transfection of modified U1 snRNAs

Cells were harvested 24 or 48 h after transfection. Total RNA was extracted using High Pure RNA Isolation Kit (Roche Applied Science) and converted into cDNA using High-Capacity cDNA Reverse Transcription Kit (Life Technologies). The RT-PCR splicing analysis for minigene transfections was performed using the typical pSPL3 primers SD6 and SA2 (Additional file 1: Table S1). For the endogenous experiments, RT-PCRs were performed using the following primers: HGSNAT-Exon 2F/HGSNAT-Exon 3R for mutation c.234 + 1G > A, HGSNAT-Exon 5F/HGSNAT-Exon 13R for c.633 + 1G >A/c.1334T > C and HGSNAT-Exon 12F/HGSNAT-Exon 16R for the c.1542 + 4dupA/c.1150C > T mutation. In the case of the c.234 + 1G > A mutation, the forward primer was designed to anneal in the middle of exon 2 to amplify only the transcripts in which the correct splicing process was recovered. For the other mutations, the last nucleotide of one of the primers corresponded to the point mutation of the other allele (but with the WT nucleotide) to favour that only the cDNA from the splice-mutation allele was amplified. The RT-PCR products were sequenced to confirm their identity. All the primers used are listed in Additional file 1: Table S1 and information regarding the amplification conditions is available upon request.

Expression of recombinant human mutant and wild type *HGSNAT* in COS-7 cells

COS-7 cells cultured in Eagle's minimal essential medium supplemented with 10% fetal bovine serum to 70% confluence were transfected with the pcTAP-HGSNAT plasmid and mutant pcDNA-HGSNAT-L125_R128del plasmid as previously described [24]. 48 h after transfection the cells were harvested and analysed for N-acetyltransferase enzymatic activity or by Western blot.

Glucosamine-mediated refolding of mutant HGSNAT

Twenty-four h after transfection with pcTAP-HGSNAT or mutant pcDNA-HGSNAT-L125_R128del plasmids, COS-7 cells were grown for 72 h in Eagle's minimal essential medium supplemented with 10% fetal bovine serum containing 10 mM glucosamine in 6-well plates. Then the cells were kept for another 24 h in the normal culture medium without glucosamine. The cells were harvested, lysed by freeze-thaw in 750 μl of water and assayed for HGSNAT activity. Three independent experiments (each with 2 cell plates) were performed on 3 different occasions.

Confluent primary cultures of skin fibroblasts of the MPS IIIC patient homozygous for the c.372-2A > G mutation and healthy control fibroblasts (n = 5) were grown in 6-well plates for 72 h in Dulbecco's minimal essential medium supplemented with 10% fetal bovine serum containing 10 mM glucosamine. Then the cells were kept for another 24 h in the medium without glucosamine, harvested, lysed in 500 μl of water and assayed for HGSNAT activity. Three independent experiments (each with 2 cell plates) were performed on 3 different occasions.

Enzyme assay

HGSNAT N-acetyltransferase enzymatic activity was measured using the fluorogenic substrate 4-methylumbelliferyl β-D-glucosaminide (4MU-β-GlcN; Moscerdam, Oegstgeest, Netherlands) as previously described [24]. The protein concentration was measured according to the method of Bradford using a commercially available reagent (BioRad, Hercules, CA).

Statistical analysis

Statistical analysis of the data has been performed by two-tailed unpaired t-test using the Prism GraphPad software.

Western blotting

Cell lysates from 3 independent transfections (20 μg of total protein each) were analysed by Western blot as previously described [24] using rabbit polyclonal antibodies raised against human HGSNAT Q52-N156 peptide (Sigma-Aldrich HPA029578, dilution 1:5000, incubation overnight at 4°C). Detection was performed with an anti-rabbit IgG antibodies-HRP conjugate (ref. 7074S, Cell Signalling, Beverly, MA), and the enhanced chemiluminescence reagent (ref. 32106, Thermo Scientific, Waltham, MA).

Results

Mutation analysis in MPS IIIC patients' fibroblasts

This study involved four splicing mutations found in five different patients: SFCP and SFC3 (both homozygous for c.234 + 1G > A), SFC6 (compound heterozygous for c.633 + 1G > A and the missense mutation c.1334T > C; p.L445P), SFC13 (compound heterozygous for c.1542 + 4dupA and the nonsense mutation c.1150C > T; p.R384*) and SFC7 (homozygous for c.372-2A > G) (see Table 1). For patients SFC3 (previously reported [25]) and SFCP (novel patient), the RT-PCR analysis of cDNA with primers in exon 1 and 3 revealed a fragment skipping exon 2 (252 bp) due to the presence of the c.234 + 1G > A mutation in homozygosity (Figure 1A). For patient SFC6 (previously described [25]) the cDNA analysis with primers in exons 5 and 7 showed two transcripts: one of a normal length of 172 bp arising from the missense mutation allele (c.1334T > C; p.L445P); and a second one resulting from the skipping of exon 6 (102 bp), due to the splicing alteration c.633 + 1G > A in the other allele (Figure 1B). For a novel patient, SFC13, bearing the intron 15 mutation c.1542 + 4dupA in compound heterozygosity with the c.1150C > T nonsense change in exon 11, the RT-PCR analysis with primers in exons 11 and 16 revealed two aberrant transcripts. The first one showed the skipping of exon 15 (395 bp), and in the second, exon 15 was extended for the first 5 nucleotides of intron 15 (478 bp) due to the presence of a cryptic ss in the beginning of this intron (Figure 1C). For the fifth patient, SFC7 (previously reported [25]), who is homozygous for the c.372-2A > G mutation, the transcript analysis using primers in exons 3 and 5 revealed the presence of two transcripts: one of lower size (79 bp) in which exon 4 was skipped; and the other of higher size (189 bp) corresponding to a fragment with the deletion of the first 12 nucleotides of exon 4 due to the use of a cryptic 3' ss localized downstream of the constitutive one (Figure 1D), which gives rise to the p.[L125_R128del] mutant protein.

Development of splicing therapy approaches for *HGSNAT* gene mutations: minigene assays with modified U1 snRNA vectors

To reproduce the splicing defects in a cellular model that could be used to test U1 snRNA overexpression as a therapeutic strategy we constructed several mutant minigenes in the pSPL3 vector and expressed them in COS-7 cells. Each of the minigenes bore either one of the three mutations that affect the donor ss or the WT allele. The post-transfection cDNA analysis and sequencing revealed the skipping of exon 2 and exon 6 in the cases of the c.234 + 1G > A and c.633 + 1G > A mutant minigenes, respectively (Figure 2C,F). For the c.1542 + 4dupA minigene, the cDNA showed a band that corresponded to the skipping of exon 15 and another band that corresponded to the inclusion of this exon plus the first five nucleotides of intron 15 (Figure 2I). Meanwhile, the constructs with the WT sequences (relative to each mutation) revealed a single transcript for each of the three cases with the inclusion of exon 2, exon 6 or exon 15, respectively. These results

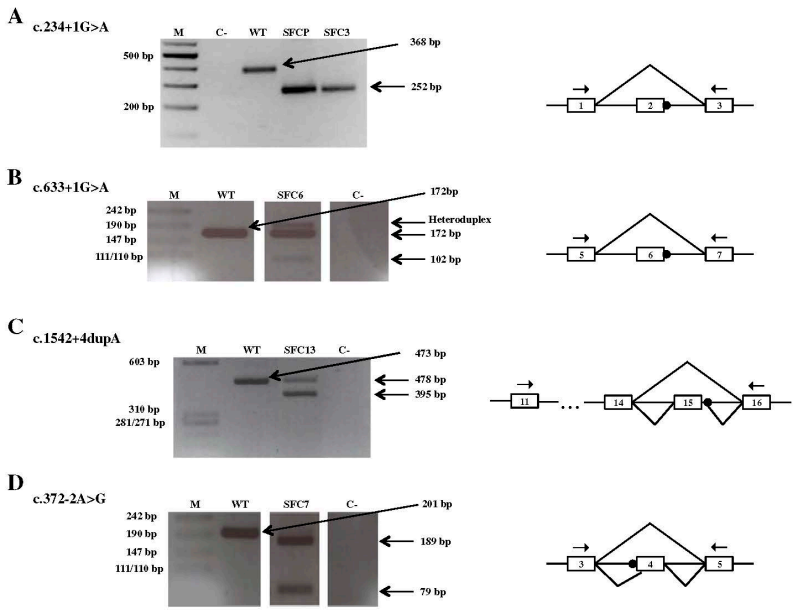


Figure 1 Analysis of alternative *HGSNAT* transcripts from the mRNA of five MPS IIIC patients by RT-PCR. RT-PCR amplification of mutations c.234 + 1G > A **(A)**, c.633 + 1G > A **(B)**, c.1542 + 4dupA **(C)**, and c.372-2A > G **(D)**, and of the WT allele, using the primers indicated in the schemes by arrows. The different splicing patterns are depicted in the schemes (right), in which the mutations are indicated by black circles. M: marker, C-: negative control, SFCP, –3, – 6, –13, –7: patients. Boxes in the schemes correspond to exons.

show that the minigene-derived splicing patterns closely resemble the patterns observed in control and patients' cDNAs obtained from fibroblasts. Thus, the minigenes are reliable tools to test and optimize the overexpression of modified U1 snRNAs in our attempts to correct the splicing defect.

To evaluate the efficiency of different U1s for correcting missplicing of the *HGSNAT* exons 2, 6 and 15, several constructs, with different degrees of complementarity to each mutated donor ss, were generated (Figure 2A,D,G). For each specific case, the WT and mutant splicing reporter constructs were co-transfected with the different U1 snRNAs into COS-7 cells for 24 and 48 h. RT-PCR analysis was then performed and showed that U1-WT and the different U1 snRNA modifications do not interfere with the splicing pattern of any WT minigene (Figure 2B, E,H).

In the case of the mutant minigenes, different splicing patterns were observed after overexpression of the different U1 snRNAs. In the case of the mutant c.234 + 1G > A minigene, the cDNA analysis revealed no change in splicing after co-expression with the U1-WT and U1-sup1 (+ 1T) constructs. With the remaining modified U1s [U1-sup2 (–1G + 1T); U1-sup3 (–1G + 4A); U1-sup4 (–1G + 1T + 4A)], a band of the size of 376 bp, expected for the normal splicing, was observed. However, sequence analysis showed that the fragment included exon 2 and the first four base pairs of intron 2 (ATAT), due to the use of an alternative donor site with the canonical "gt" in positions +5 and +6. A very faint band corresponding to exon 2 skipping was still observed when U1-sup2 and U1-sup3 were overexpressed (Figure 2C).

For the mutant c.633 + 1G > A minigene, the splicing pattern was not altered after overexpression of U1-WT and U1-sup5 (+ 1T), while an apparently normal band was detected with the U1 matching all the nucleotides of the mutated donor ss [U1-sup6 (- 1A - 2G - 3A + 1T + 7A + 8T)]. Sequencing of this band demonstrated aberrant splicing which includes, apart from exon 6, the first four nucleotides (ATAA) of intron 6. Again, this was due to the presence of an alternative donor site with the canonical "gt" in positions +5 and +6. A band of low molecular weight, which corresponded to the skipping of exon 6 was also detected upon U1-sup6 treatment (Figure 2F).

In the c.1542 + 4dupA mutant minigene, the co-transfection of the totally complementary U1 [U1-sup9 (- 1A - 2A + 5T + 6C + 7A)] induced the appearance of a single band of the predicted normal size. Sequence analysis revealed an abnormal fragment including exon 15 and the first five base pairs (GTAAA) of intron 15. As in the two previous cases, this was due to the presence of a "gt" in positions +6 and +7 (the duplication of an A in

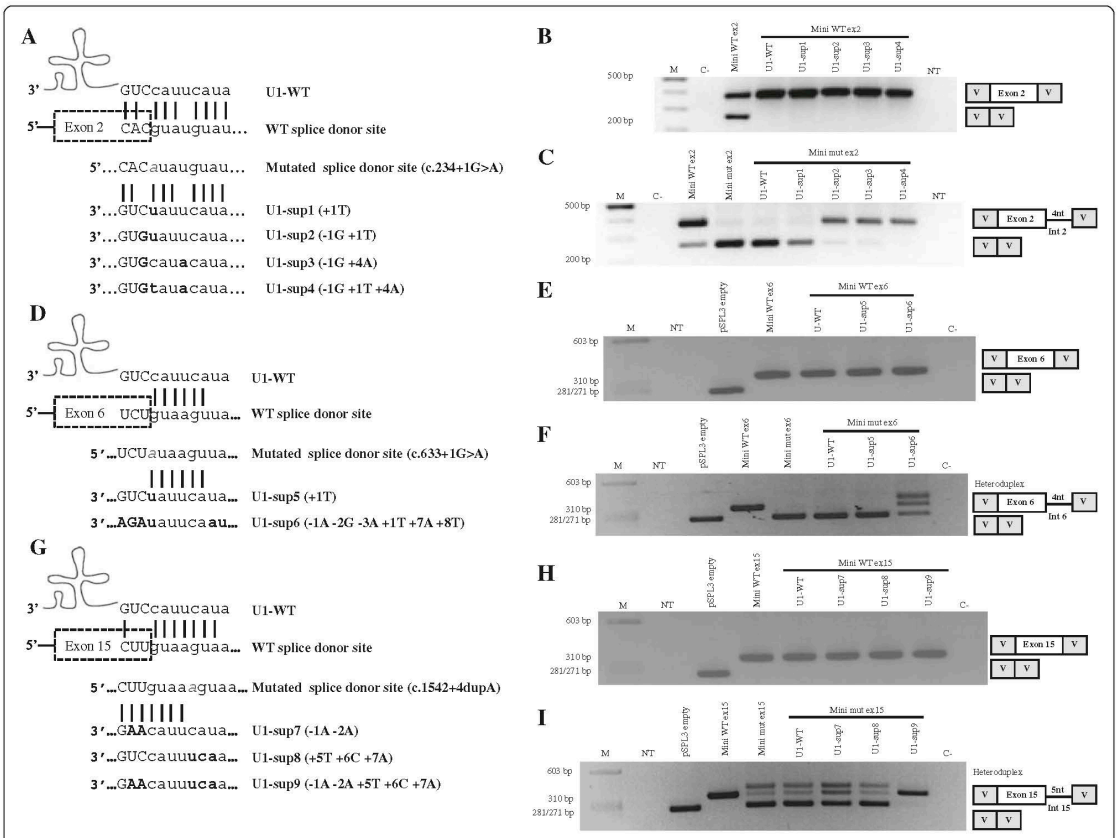


Figure 2 Minigene therapeutic approaches with U1 adaptations to correct the pathogenic effect of different donor ss mutations on the *HGSNAT* gene. (A), **(D)** and **(G)** Schematic illustrations of base pairing between the wild type U1 (U1-WT) and the 5′ ss of wild type and mutant exons 2, 6 and 15 of the *HGSNAT* gene. The position of each mutation in the 5′ ss is marked in grey and it is in italics. The different U1 snRNAs used for each mutated 5′ ss of *HGSNAT* (designated as U1-sup, for suppressor) are also shown. The U1 sequence changes are illustrated in bold. **(B)**, **(C)**, **(E)**, **(F)**, **(H)** and **(I)** RT-PCR analysis of COS-7 cells in which splicing reporter minigenes for each mutation were co-transfected with the different modified U1 constructs as appropriate. Neither the U1-WT nor the various adapted U1s showed an effect on the splicing of exon 2, 6 or 15 wild type minigenes **(B, E and H)**. In the case of the mutant minigenes, some of the U1 co-transfections [U1-sup2, 3 and 4 **(C)**; U1-sup6 **(F)**; U1-sup7, 8 and 9 **(I)**] changed the splicing pattern, not to the correct one, but mainly towards an alternative pattern using a "gt" 5′ ss immediately downstream of the constitutive one. Importantly, the wild type minigene for exon 2 (Mini WT ex2) shows an extra lower band due to the amount of empty vector added to adjust the total quantity of DNA. Sequencing results for all RT-PCR products are illustrated by schematic drawings. M: molecular weight marker; NT: non-treated cells; C-: negative control.

this mutant allele moved the "gt" from the original +5 and +6 positions). This band was also observed in the untreated mutant minigene. The overexpression of U1-sup7 (−1A-2A) produced an increase in intensity of the band that corresponded to the use of the alternative site; while no effect of U1-sup8 (+ 5T + 6C + 7A) or U1-WT was detected (Figure 2I).

Development of splicing therapy approaches for *HGSNAT* gene mutations: treatment of patients' cells with modified U1 snRNA vectors

Despite the data of minigene assays showing that the three ss defects were not corrected by the expression of the different modified U1 vectors, we explored the feasibility of this approach to correct the endogenously mis-spliced *HGSNAT* transcripts through transfection of the same U1 variants in patients' derived cell lines. The spliced transcripts obtained after transfection of each fibroblast cell line were analysed by RT-PCR (Figure 3A-F).

Interestingly, for patients SFCP and SFC3, homozygous for the c.234 + 1G > A mutation, treatment with different quantities of fully adapted U1 [U1-sup4 (−1G + 1T + 4A)] resulted in a partial recovery from the splicing defect. Sequence analysis revealed two different sequences: one with the normal splicing; and the other which included the first four base pairs of intron 2 (Figure 3B and Additional file

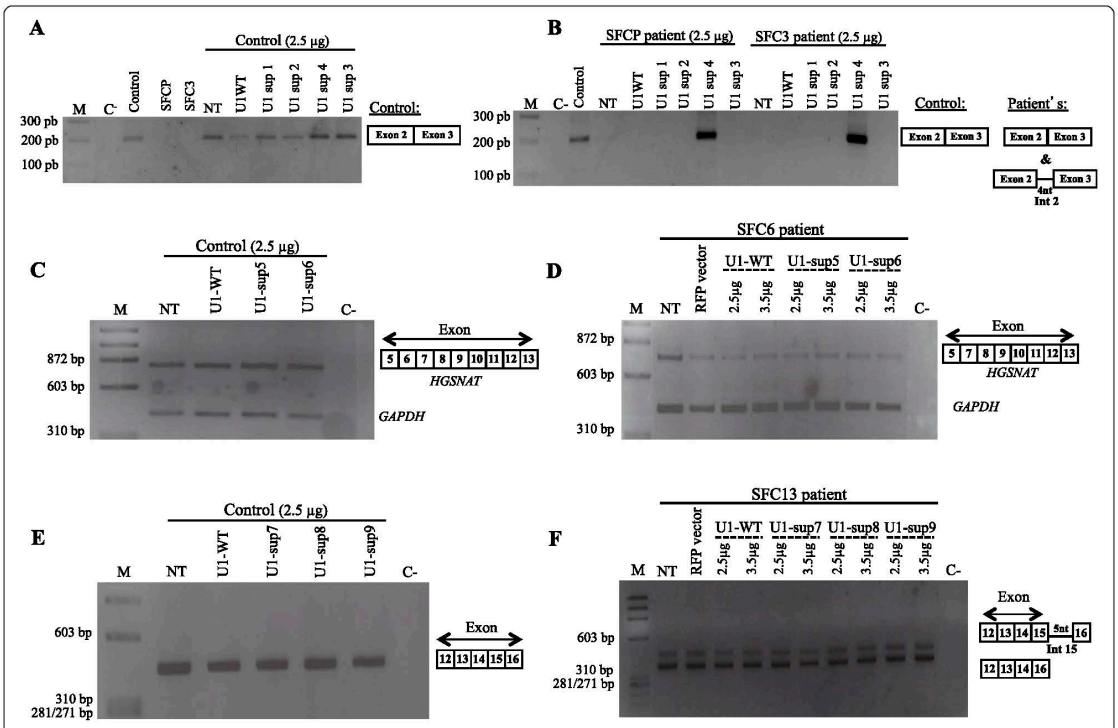


Figure 3 Analysis by RT-PCR of the endogenous splicing pattern of control and patients SFCP, SFC3, SFC6 and SFC13 derived fibroblasts after transfection with different U1 isoforms. (A), (C) and **(E)** The constitutive splicing of exons 2, 6 and 15 of the *HGSNAT* gene was not altered in control fibroblasts after overexpression of U1-WT or any of the modified U1 constructs. **(B)** In the patients SFCP and SFC3, bearing the homozygous mutation c.234 + 1G > A, only the fully adapted U1 (2.5 µg of U1-sup4) resulted in partial correction of exon 2 skipping. The same result was obtained with 1 or 3.5 µg of the U1-sup4 construct (data not shown). **(D)** and **(F)** For patients SFC6 and SFC13, bearing genotypes c.633 + 1G >A/c.1334T > C and c.1542 + 4dupA/c.1150C > T, respectively, the transfection of 2.5 or 3.5 µg of U1-WT or any generated U1 suppressor vector did not produced any change in the endogenous aberrant splicing pattern. Sequencing results for all RT-PCR products are illustrated by schematic drawings. M: molecular weight marker; NT: non-treated cells; C-: negative control; RFP: red fluorescent protein.

2: Figure S1), as detected in the minigene approaches in COS-7 cells (Figure 2C). To estimate the levels of correct splicing, the obtained PCR product was cloned in a pUC19 vector and amounts of around 50% were detected for the correct sequence (10 out of 22 clones with the correct spliced fragment). Due to this partial correction obtained, an experiment to measure the enzymatic activity in patient's cells after U1-sup4 transfection was carried out. However, no improvement in enzyme activity was observed. The values obtained for both patients, without treatment and using 2.5 µg U1sup4, were negligible and far below the normal range of controls (1.95-13.4 nmol/17 h/mg). The expression of the remaining U1 modifications had no detectable effect on exon 2 splicing.

In the case of the patients SFC6 (c.633 + 1G > A/c.1334T > C) and SFC13 (c.1542 + 4dupA/c.1150C > T), the RT-PCR analysis showed no effect of any modified U1 snRNA on the endogenous splicing process (Figure 3D,F). Similarly, WT *HGSNAT* gene splicing was not affected by any of the U1 isoforms (Figure 3A,C,E).

To rule out the possibility of the negative results being due to poor fibroblast transfection efficiency, fluorescence expression GFP/RFP vectors were used; efficient acquisition of the vectors was observed (data not shown). The uptake of the U1 snRNA was confirmed by PCR (Additional file 3: Figure S2). Moreover, the electroporation technique was also tried but no improvement in transfection efficiency was achieved (data not shown).

Analysis of the protein encoded by the mutant *HGSNAT* cDNA with 12 nucleotides deleted

In order to test whether mutant cDNA with a 12-nucleotide in-frame deletion (encoding p.[L125_R128del]) produces a stable protein and, if so, whether it is enzymatically active, we transfected cultured COS-7 cells with a corresponding pcDNA-HGSNAT-L125_R128del plasmid

and measured N-acetyltransferase activity in the cell homogenate. Cells transfected with an empty plasmid or those transfected with the pcTAP-HGSNAT plasmid described previously [6] encoding WT human HGSNAT were used as positive and negative controls, respectively. Our data (Figure 4A) showed a small but significant increase of N-acetyltransferase activity in the cells transfected with the mutant *HGSNAT* compared to that in the cells transfected with the empty plasmid. The activity in the cells transfected with WT pcTAP-HGSNAT increased about 10-fold (Figure 4A).

We further analysed cell homogenates by Western blot and detected a cross-reacting band in the cells transfected with both WT and mutant plasmids (Figure 4B). The amount of 160 kDa dimer of the HGSNAT precursor was comparable in both homogenates, but the amount of 44 kDa and 25 kDa enzymatically cleaved mature HGSNAT chains was reduced in the cells transfected with the pcDNA-HGSNAT-L125_R128del plasmid. Since the enzymatic cleavage of the HGSNAT precursor into the 2-chain form occurs in the lysosome [6], this result is consistent with misfolding of the majority of the HGSNAT-L125_R128del mutant and its retention in the endoplasmic reticulum (ER), as previously demonstrated for most HGSNAT mutants with amino acid substitutions [24].

Partial recovery of the enzyme activity of p.[L125_R128del] by glucosamine as a pharmacological chaperone

Our previous data demonstrated that the treatment of cultured cells of MPS IIIC patients bearing missense mutations with the competitive HGSNAT inhibitor (Ki = 0.28 mM) glucosamine, significantly increased the level of the residual N-acetyltransferase activity [24]. In order to determine whether glucosamine can also restore the folding and activity of the p.[L125_R128del] mutant protein, we treated COS-7 cells transfected with the pcDNA-HGSNAT-L125_R128del plasmid and cultured primary fibroblasts of the patient homozygous for the c.372-2A > G mutation with 10 mM glucosamine. Our data (Figure 4C) show that glucosamine significantly increases residual N-acetyltransferase activity in both COS-7 cells transfected with the pcDNA-HGSNAT-L125_R128del plasmid and in the patient's fibroblasts. This suggests that the active conformation of the mutant HGSNAT can be stabilized by glucosamine, resulting in part of the mutant enzyme pool being properly processed and targeted at the lysosomes.

Discussion

In this study we describe two therapeutic approaches specific for several splicing mutations in the *HGSNAT*

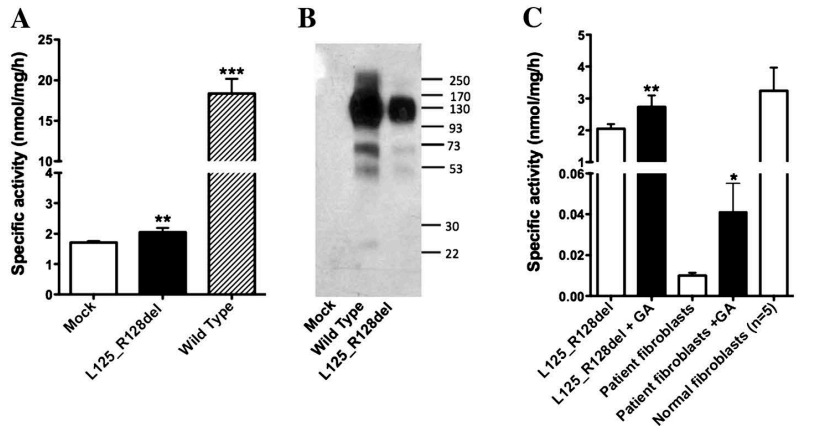


Figure 4 Enzymatic activity of mutant HGSNAT-L125_R128del protein can be partially restored by the pharmacological chaperone, glucosamine. (A) N-acetyltransferase activity of COS-7 cells transfected with the pcDNA-HGSNAT-L125_R128del plasmid is significantly increased compared to that of cells transfected with the empty pcDNA plasmid (mock). The data show means (±S.D.) of individual measurements. Three transfections (each in duplicate) were performed on separate occasions.** and ***: statistically different from mock-transfected cells ($p < 0.01$ and $p < 0.001$, respectively) according to unpaired t-test. **(B)** COS-7 cells transfected with the pcDNA-HGSNAT-L125_R128del plasmid produce 160 kDa dimmers and 78 kDa monomers of HGSNAT precursor protein but show drastically reduced amounts of 44 kDa and 25 kDa mature HGSNAT chains produced by intra-lysosomal enzymatic cleavage. Panel shows representative data of 3 independent transfections. **(C)** N-acetyltransferase activity of COS-7 cells transfected with the pcDNA-HGSNAT-L125_R128del plasmid and of cultured primary fibroblasts of the patient homozygous for the c.372-2A > G mutation is significantly increased after treating the cells in culture with 10 mM glucosamine for 72 h (+GA). The data show means (±S.D.) of individual measurements. Three independent experiments measurements were performed each of them with 2 cell plates. * and **: statistically different from untreated cells ($p < 0.05$ and $p < 0.01$, respectively) by unpaired t-test.

gene that lead to defects in mRNA processing in five Sanfilippo C patients, each carrying at least one splicing mutation. To our knowledge, this is the first attempt to treat Sanfilippo syndrome type C splicing mutations with modified U1 snRNAs. The chaperone treatment, the second therapeutic strategy examined in the present work, was previously tested *in vitro* for several missense Sanfilippo C syndrome mutations with promising results [24] and it is applied here for the treatment of a mutant protein lacking four amino acids.

For patients carrying mutations in the donor ss (SFCP, SFC3, SFC6 and SFC13), modified U1 snRNAs have been tested as a therapeutic tool to recover the normal splicing process. This approach has previously been tested for different mutations in several disorders showing different efficiencies at rescuing the normal transcripts [10-19]. A total of 9 different U1s have been developed, as well as the U1-WT. None of them affected the normal splicing process in WT minigenes or healthy control fibroblasts when overexpressed.

Few assays to correct a +1 5' ss mutation have been reported. In some of those, the efficiency of modified U1 snRNAs for splicing mutations in canonical positions +1 and +2 has been shown to be inexistent [10,13]. However, Hartmann *et al.* [19] showed a partial correction of a +1 mutation in a case in which some degree of normal splicing was conserved in the mutant allele and the alternative donor site with a "gt" dinucleotide in positions +5 and +6 presented a low score according to different predictors. In this report, we present three cases with splicing mutations at position +1, two homozygous patients with the c.234 + 1G > A mutation (SFCP and SFC3) and one heterozygous patient with the c.633 + 1G > A mutation (SFC6). When we analysed the treatment efficiency of different modified U1 snRNAs in minigene constructions carrying these mutations, we were not able to detect restoration of the normal splicing process. Instead, we detected alternative splicing patterns due to the use of highly conserved "gt" nucleotides at positions +5 and +6 as a donor site, giving rise to a transcript that includes 4 intronic nucleotides. This result was obtained using three of the four modified U1 snRNAs for the c.234 + 1G > A mutation and the total complementary U1 snRNA for the c.633 + 1G > A mutation. Due to the presence of these "gt" dinucleotides at position +5 and +6 in many introns, it is important to sequence putative rescued transcripts to check whether this alternative site was used.

Despite these results, and taking into account that minigenes only included partial intronic sequences which could lack some splicing regulatory sites and that they were assayed in non-human cells, modified U1 snRNAs were tested on patients' fibroblasts. Partial rescue (almost 50%) of the normal splicing for the c.234 + 1G > A mutation was observed in patients SFCP and SFC3 overexpressing the U1-sup4. In the case of the c.633 + 1G > A mutation (patient SFC6), no rescue was observed after the modified U1 snRNAs overexpression. Analysis of the ss sequences using the Human Splicing Finder predictor [29], indicated that the alternative site of patient SFC6 had a high score (68.17/100), while that of patients SFCP and SFC3 had a null score. This could explain the difference in the rescue results between these two "+1" mutations: the mutant allele of patient SFC6 would efficiently use the alternative site and, thus, would not be rescued.

The molecular reason why the modified U1 promotes, in general, a new splicing process using the +5 + 6 "gt" donor site when its sequence perfectly matches the mutated +1 + 2 site remains unclear. One explanation could be the involvement of U6 snRNA complementarity, in addition to that of U1, in the choice of alternative ss in proximity to the normal one, which has been previously described [30,31].

The rescue of the splicing mutation shown here is one of the few positive results for a "+1" mutation using a modified U1 snRNA and the first one performed on an allele that did not produce any of the correctly spliced mRNAs when untreated. It is important to note that a few natural U2-type 5' ss that present "au" instead of the normal "gu" at the first two nucleotides of the intron have been described [32,33]. These introns present "ac" instead of "ag" at the last two nucleotides. Thus, one could think that mutated "au" sites could promote an alternative splicing using an "ac" acceptor site. However, for this +1 mutation the "au-ac" alternative splicing type was not observed. On the contrary, we found that the "au" 5' ss correctly paired with the normal "ag" and not with a possible cryptic "ac" 3'ss, restoring the normal splicing of intron 2.

Further studies are needed to improve the efficacy and to test the toxicity and side effects of U1 overexpression in order to confirm the feasibility of the use of this modified U1 snRNAs in therapy. Overexpression of U1 snRNA vectors introduced with adeno-associated virus into mice liver has been shown to be toxic at high viral doses but safe at low doses, suggesting the viability of this treatment when low amounts of U1 snRNA viral particles are injected [20].

In the case of the c.1542 + 4dupA mutation (patient SFC13), it is important to note that the canonical "gt" site is not affected by the mutation. However, in this case there is also another "gt" dinucleotide in positions +6 and +7 (one nucleotide downstream due to the duplication). The minigene analysis showed that the splicing that uses the alternative donor site occurs, even without treatment, together with that causing exon 15 skipping. This alternative splicing is slightly enhanced with U1-sup7 and greatly enhanced with U1-sup9. Many different ss predictors were used in order to estimate the score for each donor site in

this case (Additional file 4: Table S2). Clearly, in the absence of mutation, the scores of both sites are similar; but when the mutation is present, the normal site presents a lower score while the alternative site increases its score. This is consistent with the fact that in the presence of the mutation the alternative "gt" site is the only one used (see Figure 2I). As it was previously discussed, it remains unclear why the modified U1 snRNAs, designed to perfectly match the mutated site, enhance the splicing using the alternative site. When the modified U1 snRNAs were tested on patient's fibroblasts, no rescue was detected: either for the normal mRNA or for the one including the five intronic nucleotides. The latter was detected without any treatment, indicating that it takes place in normal conditions in the patient's cells. In any case, the alternative transcript would not restore the enzyme activity since it does not keep the frame, so it could not have any therapeutic use. Many different mutations in the +3 to +6 positions have been partially or totally rescued using U1 snRNAs [10-13,15,17,18], while one +5 mutation was not corrected [15]. Our current results point to the importance of the presence of an additional "gt" dinucleotide in the region which, depending on the sequence context, may be used as an alternative donor site. In previous reports where mutations have been partially or totally corrected and the "gt" dinucleotide was present at positions +5 and +6 [15,19], the alternative site presented a low score in accordance with different predictors, in contrast to the cases described here.

Finally, we attempted to rescue the c.372-2G > A mutation (present in SFC7), using molecular chaperone. This mutation gives rise to two different splicing processes in fibroblasts: one causing exon 4 skipping; the other using an alternative acceptor site, 12 nucleotides downstream of the normal site and producing a transcript with an in-frame deletion. We showed that this alternative splicing produces a protein lacking 4 amino acids p. [L125_R128del] that has some residual activity but the majority of the protein does not reach the lysosome, remaining in the ER due to its misfolding.

In order to restore the correct protein folding, cells transfected with the plasmid carrying the mutant protein and the patient's fibroblasts were treated with glucosamine chaperone, which resulted in a significant increase in the residual HGSNAT activity. This indicates the feasibility of this therapeutic approach for patients carrying this splicing mutation. Different chaperones assays for lysosomal disorders have been performed before and some have been tested in humans, showing their safety and potential as a therapeutic tool (reviewed in Ref. [21]).

Conclusions

In conclusion, the results of two therapeutic approaches for different splicing mutations varied depending on the nature of the mutation. For the treatment of donor ss mutations, U1 snRNAs could represent a feasible option. This would depend on the presence of alternative donor sites close to the normal site that could interfere with the correction process and, thus, with the success of the therapy. This is important since many introns present a "gt" in positions +5 and +6. In the present study, we have shown that it is possible to partially recover the normal splicing process for +1 mutations, which was reported only once before. Additionally, a chaperone treatment using glucosamine for a mutant protein with a loss of 4 amino acids, caused by an acceptor ss mutation, has been shown to result in a significant increase in the enzyme activity. These promising results encourage further research into the therapeutic use of U1 snRNAs and chaperones to treat Sanfilippo syndrome type C patients.

Availability of supporting data

Supporting data are included in the article and as additional files.

Additional files

Additional file 1: Table S1. Oligonucleotide sequences for PCR amplifications. List of the oligonucleotides used in this work.

Additional file 2: Figure S1. Electropherograms of the bands obtained in RT-PCR after transfection of the U1-sup4 in control and c.234 + 1G > A patient's fibroblasts. (A) Wild type (control) sequence showing the identity of the normal fragment with exon 2 and 3. (B) Normal and aberrant sequence of the rescue band obtained with U1-sup4 transfection in patient's fibroblasts.

Additional file 3: Figure S2. PCR amplification to test the uptake of a modified U1 snRNA vector by patient fibroblasts. Agarose gel electrophoresis shows bands corresponding to the U1 vector after fibroblast transfection (by 2.5 μg and 3.5 μg of vector as indicated) and lower molecular weight bands which correspond to the *HGSNAT* exon 4 amplification, as a control. M: molecular weight marker; NT: non-treated cells; C-: negative control.

Additional file 4: Table S2. Intron 15 donor ss scores using different predictors. Comparison of the ss scores for the normal site and the alternative site either in the presence or in the absence of the mutation.

Abbreviations

Bp: Base pair; EET: Enzyme enhancement therapy; MPS III: Mucopolysaccharidoses type III; MPS IIIC: Mucopolyssacharidoses type IIIC or Sanfilippo syndrome type C; snRNAs: Small nuclear ribonucleic acids; ss: Splice site; WT: Wild type.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LM and IC were involved in the conception and design of the study, performed the experimental work of the U1 snRNA part, the analysis and interpretation of the data, and participated in the drafting and revising of the manuscript. LD and YC were involved in the conception, design and experimental work, as well as in the interpretation of the data and the drafting of the manuscript regarding to the glucosamine part. SA, DG, LV and AVP supervised all the research contributing critically to the design of the work and data interpretation as well as in the revision of the manuscript. MJP, PJ, LRD and BP are collaborators with experience in the field and supervised the research. All authors read and approved the final manuscript.

Acknowledgements
The authors would like to thank Dr. Lúcia Lacerda from Centro de Genética Médica Dr. Jacinto Magalhães - Centro Hospitalar do Porto, Portugal for providing the SFCP patient fibroblasts sample and Helena Ribeiro from the same institution for the technical support. We thank also Dr. Mónica Sousa and Dr. Elsa Logarinho research groups from Instituto de Biologia Molecular e Celular (IBMC), Porto for the collaboration in the electroporation studies. We would like equally to thank Xavier Roca from the School of Biological Sciences, Nanyang Technological University, Singapore, for advice on the U1s constructs and to the Institut de Bioquímica Clínica, Barcelona, for their collaboration. The authors are also grateful for the support of the *Centro de Investigación Biomédica en Red de Enfermedades Raras* (CIBERER), which is an initiative of the ISCIII. This study was partially funded by a grant from the Spanish Ministry of Science and Innovation (SAF2011-25431) and from the Catalan Government (2009SGR971). We are also grateful for the permanent support, including financial aid, from 'patient-support' associations, such as Jonah's Just Begun-Foundation to Cure Sanfilippo Inc. (USA), Association Sanfilippo Sud (France), Fundación Stop Sanfilippo (Spain), Asociación MPS España (Spain). LM was supported by a grant (SFRH/BD/64592/2009) from the Fundação para a Ciência e Tecnologia IP (FCT) /POPH/FSE, Portugal, IC by a grant from the University of Barcelona (APIF), Spain and AVP by an operating grant MOP111068 from Canadian Institutes of Health Research.

Author details
[1]Department of Human Genetics, Research and Development Unit, INSA, Porto, Portugal. [2]Department of Biology, Faculty of Sciences, Porto, Portugal. [3]Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain. [4]CIBER de Enfermedades Raras (CIBERER), Madrid, Spain. [5]Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain. [6]Department of Medical Genetics, Sainte-Justine University Hospital Centre, University of Montreal, Montreal, Canada. [7]IPATIMUP, Porto, Portugal. [8]Department of Human Genetics, Research and Development Unit, INSA, Lisbon, Portugal. [9]Centro de Diagnóstico de Enfermedades Moleculares, Centro de Biología Molecular Severo Ochoa, UAM-CSIC, Universidad Autónoma de Madrid, Madrid, Spain. [10]Department of Anatomy and Cell Biology, Faculty of Medicine, McGill University, Montreal, Canada.

Received: 13 August 2014 Accepted: 4 November 2014
Published online: 10 December 2014

References

1. Neufeld EF, Muenzer J: **The Mucopolysaccharidoses.** In *The metabolic and molecular bases of inherited disease*, Volume 3. 8th edition. Edited by Scriver CR, Beaudet AL, Sly WS, Valle D. New York: McGraw-Hill; 2001:3421–3452.
2. Valstar MJ, Ruijter GJG, van Diggelen OP, Poorthuis BJ, Wijburg FA: **Sanfilippo syndrome: a mini-review.** *J Inherit Metab Dis* 2008, **31**:240–252.
3. Klein U, Kresse H, von Figura K: **Sanfilippo syndrome type C: deficiency of acetyl-CoA:alpha-glucosaminide N-acetyltransferase in skin fibroblasts.** *Proc Natl Acad Sci U S A* 1978, **75**:5185–5189.
4. Fan X, Zhang H, Zhang S, Bagshaw RD, Tropak MB, Callahan JW, Mahuran DJ: **Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C).** *Am J Hum Genet* 2006, **79**:738–744.
5. Hřebíček M, Mrázová L, Seyrantepe V: **Mutations in TMEM76 Cause Mucopolysaccharidosis IIIC (Sanfilippo C Syndrome).** *Am J Hum Genet* 2006, **79**:807–819.
6. Durand S, Feldhammer M, Bonneil E, Thibault P, Pshezhetsky AV: **Analysis of the biogenesis of heparan sulfate acetyl-CoA:alpha-glucosaminide N-acetyltransferase provides insights into the mechanism underlying its complete deficiency in mucopolysaccharidosis IIIC.** *J Biol Chem* 2010, **285**:31233–31242.
7. Fan X, Tkachyova I, Sinha A, Rigat B, Mahuran D: **Characterization of the biosynthesis, processing and kinetic mechanism of action of the enzyme deficient in mucopolysaccharidosis IIIC.** *PLoS One* 2011, **6**:e24951.
8. Wahl MC, Will CL, Lührmann R: **The spliceosome: design principles of a dynamic RNP machine.** *Cell* 2009, **136**:701–718.
9. Roca X, Akerman M, Gaus H, Berdeja A, Bennett CF, Krainer AR: **Widespread recognition of 5′ splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides.** *Genes Dev* 2012, **26**:1098–1109.
10. Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, Rogalska ME, Bernardi F, Pagani F: **An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects.** *Hum Mol Genet* 2012, **21**:2389–2398.
11. Pinotti M, Rizzotto L, Balestra D, Lewandowska MA, Cavallari N, Marchetti G, Bernardi F, Pagani F: **U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency.** *Blood* 2008, 111:2681–2684.
12. Sánchez-Alcudia R, Pérez B, Pérez-Cerdá C, Ugarte M, Desviat LR: **Overexpression of adapted U1snRNA in patients' cells to correct a 5′ splice site mutation in propionic acidemia.** *Mol Genet Metab* 2011, **102**:134–138.
13. Schmid F, Hiller T, Korner G, Glaus E, Berger W, Neidhardt J: **A gene therapeutic approach to correct splice defects with modified U1 and U6 snRNPs.** *Hum Gene Ther* 2013, **24**:97–104.
14. Tanner G, Glaus E, Barthelmes D, Ader M, Fleischhauer J, Pagani F, Berger W, Neidhardt J: **Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA.** *Hum Mutat* 2009, **30**:255–263.
15. Susani L, Pangrazio A, Sobacchi C, Taranta A, Mortier G, Savarirayan R, Villa A, Orchard P, Vezzoni P, Albertini A, Frattini A, Pagani F: **TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA.** *Hum Mutat* 2004, **24**:225–235.
16. Mattioli C, Pianigiani G, De Rocco D, Bianco AMR, Cappelli E, Savoia A, Pagani F: **Unusual splice site mutations disrupt FANCA exon 8 definition.** *Biochim Biophys Acta* 1842, **2014**:1052–1058.
17. Glaus E, Schmid F, Da Costa R, Berger W, Neidhardt J: **Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells.** *Mol Ther* 2011, **19**:936–941.
18. Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, Ffrench-Constant C, Baralle FE: **Identification of a mutation that perturbs NF1 gene splicing using genomic DNA samples and a minigene assay.** *J Med Genet* 2003, **40**:220–222.
19. Hartmann L, Neveling K, Borkens S, Schneider H, Freund M, Grassman E, Theiss S, Wawer A, Burdach S, Auerbach AD, Schindler D, Hanenberg H, Schaal H: **Correct mRNA processing at a mutant TT splice donor in FANCC ameliorates the clinical phenotype in patients and is enhanced by delivery of suppressor U1 snRNAs.** *Am J Hum Genet* 2010, **87**:480–493.
20. Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, Bernardi F, Arruda VR, Pinotti M: **An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice.** *J Thromb Haemost* 2014, **12**:177–185.
21. Suzuki Y: **Emerging novel concept of chaperone therapies for protein misfolding diseases.** *Proc Japan Acad Ser B* 2014, **90**:145–162.
22. Bernier V, Lagacé M, Bichet DG, Bouvier M: **Pharmacological chaperones: potential treatment for conformational diseases.** *Trends Endocrinol Metab* 2004, **15**:222–228.
23. Sawkar AR, Cheng W-C, Beutler E, Wong C-H, Balch WE, Kelly JW: **Chemical chaperones increase the cellular activity of N370S beta -glucosidase: a therapeutic strategy for Gaucher disease.** *Proc Natl Acad Sci U S A* 2002, **99**:15428–15433.
24. Feldhammer M, Durand S, Pshezhetsky AV: **Protein misfolding as an underlying molecular defect in mucopolysaccharidosis III type C.** *PLoS One* 2009, **4**:e7434.
25. Canals I, Elalaoui SC, Pineda M, Delgadillo V, Szlago M, Jaouad IC, Sefiani A, Chabás A, Coll MJ, Grinberg D, Vilageliu L: **Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel HGSNAT mutations and characterization of the mutant alleles.** *Clin Genet* 2011, **80**:367–374.
26. Feldhammer M, Durand S, Mrázová L, Boucher R-M, Laframboise R, Steinfeld R, Wraith JE, Michelakakis H, van Diggelen OP, Hrebícek M, Kmoch S, Pshezhetsky AV: **Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: alpha-glucosaminide N-acetyltransferase (HGSNAT) gene.** *Hum Mutat* 2009, **30**:918–925.
27. Coutinho MF, Lacerda L, Prata MJ, Ribeiro H, Lopes L, Ferreira C, Alves S: **Molecular characterization of Portuguese patients with mucopolysaccharidosis IIIC: two novel mutations in the HGSNAT gene.** *Clin Genet* 2008, **74**:194–195.
28. Lund E, Dahlberg JE: **True genes for human U1 small nuclear RNA. Copy number, polymorphism, and methylation.** *J Biol Chem* 1984, **259**:2013–2021.
29. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Béroud C: **Human splicing finder: an online bioinformatics tool to predict splicing signals.** *Nucleic Acids Res* 2009, **37**:e67.
30. Brackenridge S, Wilkie AO, Screaton GR: **Efficient use of a 'dead-end' GA 5′splice site in the human fibroblast growth factor receptor genes.** *EMBO J* 2003, **22**:1620–31.

31. Hwang DY, Cohen JB: U1 snRNA promotes the selection of nearby 5' splice sites by U6 snRNA in mammalian cells. *Genes Dev* 1996, 10:338–50.
32. Wu Q, Krainer AR: AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes. *Mol Cell Biol* 1999, 19(5):3225–36.
33. Kubota T, Roca X, Kimura T, Kokunai Y, Nishino I, Sakoda S, Krainer AR, Takahashi MP: A mutation in a rare type of intron in a sodium-channel gene results in aberrant splicing and causes myotonia. *Hum Mutat* 2011, 32(7):773–82.

Additional file 1: Table S1

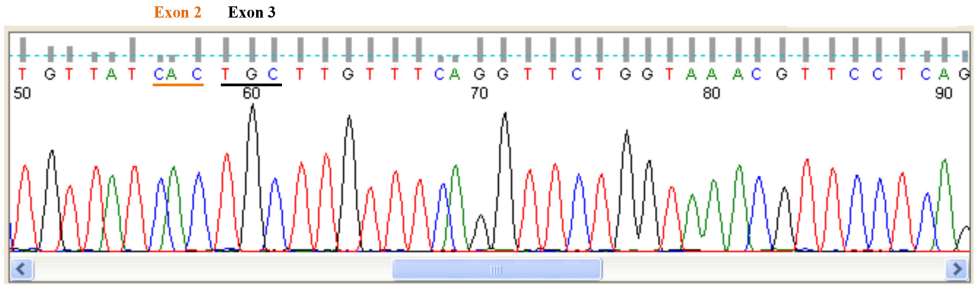
Oligonucleotide	Sequence (5' to 3')
RT-PCR fragments*	
HGSNAT-Exon 1F	GCAGCGGGCAGGCAAGGGCGG
HGSNAT-Exon 2F	ACATGCAGAGCTGAAGATGGA
HGSNAT-Exon 3R	GATAGATCCGTGCTGGGTG
HGSNAT-Exon 4F	AGCTGAACGACACCTTGGAA
HGSNAT-Exon 4R	CAATGATGACAGCAAGACCAA
HGSNAT-Exon 6F	TGCATTCCTTATTGGTCTTGC
HGSNAT-Exon 6R	ATCACCATCGAGAGGGTCTG
c.1542+4dupA F	ACCTGTGCCTGAACATTGTG
c.1542+4dupA R	CAGAAACCTTCGTCAGAGCA
HGSNAT-Exon 5F	CCTTTCTGAGGCTCTTGTTG
HGSNAT-Exon 12F	AGGAGCTGCCTTTCTCTTC
HGSNAT-Exon 13R	TGATCGTCTCCCAGCAGCA
HGSNAT-Exon 16R	ACCTTCGTCAGAGCAACAG
DNA fragments*	
HGSNAT Exon 4 F	TTATTCTGCCTCCATGATATTAGC
HGSNAT Exon 4 R	CTACAGAAAGCGTCATGGACTGC
Cloning fragments*	
Intron 1 F	GCAAAAGGAGACCTGTGTGTG
Intron 3 R	TCATCCCTGAGAACTGGCTTT
Plasmid Vectors	
SD6 F	TCTGAGTCACCTGGACAACC
SA2 R	ATCTCAGTGGTATTTGTGAGC
pGEMQ	ATCGAAATTAATACGACTCA
U1-QR	CTGGGAAAACCACCTTCGT
Site-directed mutagenesis	
pSPL3 WT minigenes	
Exon 2 F	GGAAATCTGAATGCTGTTATCACATATGTATCAGTTCACACTCAG
Exon 2 R	CTGAGTGTGAACTGATACATATGTGATAACAGCATTCAGATTTCC
Exon 6 F	CTGATCGCCTCATCAATTCTGTAAGTTATGAGATGCATAGTG
Exon 6 R	CACTATGCATCTCATAACTTACAGAATTGATGAGGCGATCAG
Exon 15 F	GGTGTTGTATTCTTGTAAGTAAGCAGCATTCCTCGC
Exon 15 R	GCGAGGAATGCTGCTTACTTACAAGAATACAACACC
U1snRNA vectors	
U1 suppressor 1 F	GATCTCATACTTATCTGGCAGGGGAGATAC
U1 suppressor 1 R	GTATCTCCCCTGCCAGATAAGTATGAGATC
U1 suppressor 2 F	GATCTCATACTTATGTGGCAGGGGAGATAC
U1 suppressor 2 R	GTATCTCCCCTGCCACATAAGTATGAGATC
U1 suppressor 3 F	CAAGATCTCATACATACGTGGCAGGGGAGATAC
U1 suppressor 3 R	GTATCTCCCCTGCCACGTATGTATGAGATCTTG
U1 suppressor 4 F	CAAGATCTCATACATATGTGGCAGGGGAG
U1 suppressor 4 R	CTCCCCTGCCACATATGTATGAGATCTTG
U1 suppressor 5 F	GATCTCATACTTATCTGGCAGGGGAGATAC
U1 suppressor 5 R	GTATCTCCCCTGCCAGATAAGTATGAGATC
U1 suppressor 6 F	GAGGCCCAAGATCTCTAACTTATAGAGCAGGGGAGATACCATG
U1 suppressor 6 R	CATGGTATCTCCCCTGCTCTATAAGTTAGAGATCTTGGGCCTC
U1 suppressor 7 F	AGGCCCAAGATCTCATACTTACAAGGCAGGGGAGATACC
U1 suppressor 7 R	GGTATCTCCCCTGCCTTGTAAGTATGAGATCTTGGGCCT
U1 suppressor 8 F	GCAGAGGCCCAAGATCTCAACTTTACCTGGCAGGGGAGATA
U1 suppressor 8 R	TATCTCCCCTGCCAGGTAAAGTTGAGATCTTGGGCCTCTGC
U1 suppressor 9 F	GTGTCGGGGCAGAGGCCCAAGATCTCAACTTTACAAGGCAGGGGAGATAC
U1 suppressor 9 R	GTATCTCCCCTGCCTTGTAAAGTTGAGATCTTGGGCCTCTGCCCCGACAC

* Primers were designed according to the sequence described in the ENSEMBL database (www.ensembl.org; ENSG00000165102). **F** – Forward; **R** – Reverse

Additional file 2: Figure S1

A)

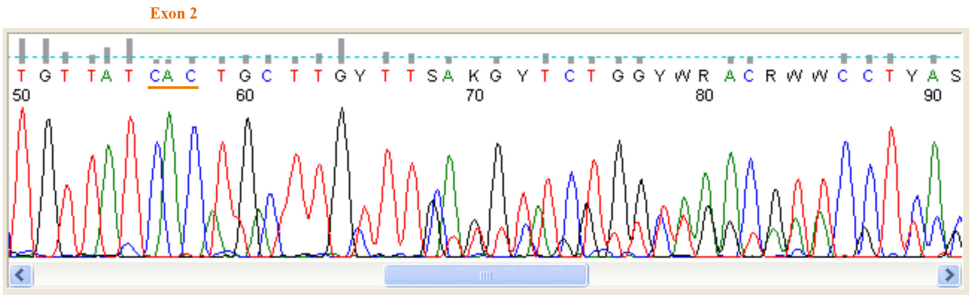
Wild-type sequence



Normal exon 2 – exon 3 sequence …TATCACTGCTTGTTTCAGGTTCTGGTAAACGTTCCTCAG…

B)

c.234+1G>A sequence after U1 sup 4 treatment

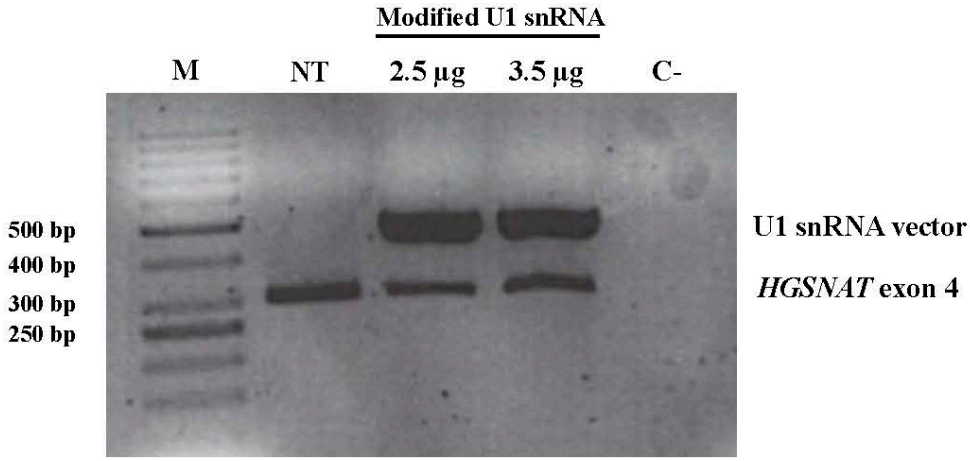


Two sequences:

Normal exon 2 – exon 3 sequence …TATCACTGCTTGTTTCAGGTTCTGGTAAACGTTCCTCAG…

Aberrant exon 2 – exon 3 sequence …TATCACatatTGCTTGTTTCAGGTTCTGGTAAACGTTCCTCAG…

Additional file 3: Figure S2



Additional file 4: Table S2

Predictor	Normal	Alternative	Normal mutated	Alternative mutated
	CTT/gtaagt	taa/gtaagc	CTT/gtaaag	aaa/gtaagc
Splice site score calculation (Max = 12.6)[1]	6.9	6.2	2	7.6
Analyzer splice tool (Max = 100)[2]	78.94	78.17	61.92	81.99
Splice site prediction by neuronal network (Max = 1) [29]	0.95	0	0	0.97
MaxEntScan (Max = 12) [30]	8	5.66	-3	7.31
Human splicing finder (Max = 100) [31]	84.38	85.27	70.28	86.29

1. http://rulai.cshl.edu/new_alt_exon_db2/HTML/score.html
2. http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm

3.4 Article 4

Liliana Matos, Ana Joana Duarte, Diogo Ribeiro, Peter Jordan, Maria João Prata, João Chaves, Lourdes R. Desviat, Belén Pérez, Olga Amaral, Sandra Alves. **Correction of a splicing mutation affecting an Unverricht-Lundborg disease patient by antisense therapy.**

Submitted

Synopsis

Background and motivation for the study

Our therapeutic target in this study was the c.66G>A mutation present in the *CSTB* gene, which encodes the protein cystatin B. Cystatin B is an inhibitor of several papain-family cysteine proteases, cathepsins, which have a key role as lysosomal enzymes (Merwick et al., 2012). The main function of cathepsins is the non-selective degradation of intracellular proteins, but they also participate in antigen processing and apoptosis (Lehesjoki & Gardiner, 2012).

Mutations in the *CSTB* gene are responsible for the Unverricht-Lundborg disease (ULD), an autosomal recessive neurodegenerative disorder also known as progressive myoclonic epilepsy type 1 (EPM1) (Joensuu et al., 2008).

Some of the symptoms presented by ULD patients, such as myoclonus, epilepsy and progressive neurologic deterioration are also observed in some LSDs like the NCL's, Sialidosis type I and Neuronopathic Gaucher disease, which explains the classification as lysosome-related progressive myoclonic epilepsies (Ramachandran et al., 2009). This means that, whilst screening for LSDs in patients with that sort of symptoms, we may come across cases of different aetiology. The case that caught our attention was an ULD patient homozygous for a new synonymous mutation (c.66G>A; p.Q22Q) located at the last nucleotide of *CSTB* exon 1. It affects the 5' SDS of exon 1, as was evidenced by cDNA analysis that allowed the identification of two splice variants, one of normal size with the G>A change and other with partial inclusion of intron 1 due to the activation of a cryptic splice site inside the intronic sequence. Since this mutation is a synonymous change that impairs normal splicing, it appeared to be a good candidate for correction through RNA therapeutic approaches, having been therefore considered for splicing correction in the scope of this study. Two different RNA-based therapeutic strategies were tested: mutation-adapted U1 constructs with increased complementarity to the mutated SDS of exon 1 and AOs (LNA oligonucleotide) specifically designed to block the activated cryptic splice site in *CSTB* intron 1.

Study design and methods overview

- Design and construction through site-directed mutagenesis of the different U1 snRNA vectors adapted to the mutated SDS of exon 1.

- Test the efficacy on splicing pattern correction of the different mutation-adapted U1 snRNA vectors through their transfection in patient fibroblasts.
- Test the efficacy on splicing pattern correction of the LNA oligonucleotide designed to block the activated cryptic splice site in *CSTB* intron 1 through its transfection in patient fibroblasts.

Major results

A successful correction of the aberrant splicing pattern using a LNA oligonucleotide that blocked the use of the cryptic intronic splice site was achieved.

Using adapted U1 vectors with increased complementarity to the mutated splice site, no correction of the aberrant splicing pattern was observed.

Correction of a splicing mutation affecting an Unverricht-Lundborg disease patient by antisense therapy

Liliana Matos[1,2], Ana Joana Duarte[1], Diogo Ribeiro[1], Peter Jordan[3], Maria João Prata[2,4], João Chaves[5], Lourdes R. Desviat[6,7], Belén Pérez[6,7], Olga Amaral[1], Sandra Alves[1*]

[1] Research and Development Unit, Department of Human Genetics, INSA, Porto, Portugal

[2] Department of Biology, Faculty of Sciences, Porto, Portugal

[3] Research and Development Unit, Department of Human Genetics, INSA, Lisbon, Portugal

[4] IPATIMUP, Porto, Portugal

[5] Neurology service, Santo António Hospital, CHP, Porto, Portugal

[6] CIBER de Enfermedades Raras (CIBERER), Spain

[7] Centro de Diagnóstico de Enfermedades Moleculares, Centro de Biología Molecular Severo Ochoa, UAM-CSIC, Universidad Autónoma de Madrid, Madrid, Spain

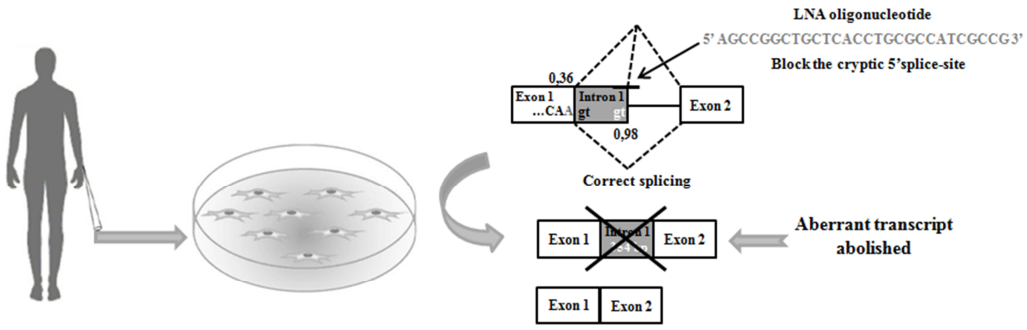
Corresponding author:
Sandra Alves
Affiliation: Research and Development Unit, Department of Human Genetics, INSA – Porto
Address: Rua Alexandre Herculano 321, 4000-055, Porto, Portugal
Phone: +351 223401100; Fax: +351 223401109; Email: sandra.alves@insa.min-saude.pt

Abstract

Unverricht-Lundborg disease (ULD) is a common form of progressive myoclonic epilepsy caused by mutations in the cystatin B gene (*CSTB*) that encodes an inhibitor of several lysosomal cathepsins. Presently, only pharmacological treatment and psychosocial support is available for ULD patients. Although the dodecamer expansion in the promoter region of the *CSTB* gene is the most common cause associated to ULD, a patient without the expansion mutation but homozygous for a unique splice defect was identified in Portugal. This case seemed to provide a good model for testing the feasibility of splicing correction. Therefore, to overcome the effect of the *CSTB* splicing mutation c.66G>A (exon 1), we investigated two strategies for the correction of the defect in patient's cells. One attempt relied on the use of U1 snRNA mediated correction and involved the generation of three constructs with increasing complementarity to the splice donor site of *CSTB* exon 1. In the other strategy used, a specific locked nucleic acid (LNA) antisense oligonucleotide was designed to block a cryptic 5' splice site in intron 1. This latter LNA oligonucleotide approach allowed the restoration of the normal splicing pattern and the recovery was sequence and dose-specific. The failure of the U1 snRNA therapeutic approach may rely on the scarce knowledge on how the correct splicing in the exon 1 of *CSTB* occurs, evidence has suggested that it depends not only on the U1 complementarity, but probably also on enhancer splicing factors.

This work constitutes a proof of concept for correcting a *CSTB* RNA defect involved in ULD with a mutation-specific antisense approach. It contributes to the growing evidence on the feasibility of this type of potential therapy. The insights here obtained make mutation-based correction a candidate for personalized treatment of patients with this type of mutation, encouraging similar investigations in other genetic diseases.

Graphical Abstract



Efficient correction of the aberrant splicing pattern in the fibroblasts of a patient with Unverricht-Lundborg disease using a locked nucleic acid oligonucleotide.

Introduction

Unverricht-Lundborg disease (ULD), also known as progressive myoclonic epilepsy type 1 (EPM1), is an autosomal recessive neurodegenerative disorder characterized by young-onset stimulus sensitive myoclonus and tonic-clonic seizures.[1] Although occurring worldwide, it shows the highest prevalence in Finland, North Africa and the Western Mediterranean region.[2] Presently, symptomatic pharmacologic and rehabilitative management are the mainstay of patient care.[2,3]

ULD is caused by mutations in the *CSTB* gene, localized on chromosome 21q22.3[4], which encodes cystatin B, a protease inhibitor that *in vitro* has shown the ability to inhibit several lysosomal cysteine proteases, known as cathepsins, by tight reversible binding.[5]

Up to now, thirteen *CSTB* mutations have been described as causal of EPM1 (HGMD®Professional 2015.3 Release), the most common of which is an unstable expansion of a dodecamer repeat in the promoter region[6] that down-regulates *CSTB* mRNA levels.[7] The remaining mutations are missense, nonsense, frameshift and splice site (ss) mutations that lead to abnormal RNA processing (HGMD®Professional 2015.3 Release).

The increasing knowledge of RNA biology is stimulating the development of new approaches relying on RNA-based strategies to modify or eliminate mRNA bearing disease-causing mutations. On one hand, the use of U1 snRNA complementary to the mutated splice-site has become an attractive strategy to correct 5'ss defects in U1 binding, as shown for the *PCCA*[8], *RPGR*[9] and *BBS1*[10] genes. On the other hand, oligonucleotide-based approaches are also emerging as efficient alternatives that can be employed in the future as new treatments for a number of diseases. Currently, several strategies are already under investigation for many conditions, including Duchenne muscular dystrophy[11], spinal muscular atrophy[12], β-Thalassaemia[13] and pyridoxine dependent epilepsy[14], as well as for Alzheimer[15], Parkinson[16], Huntington[17] and other neurodegenerative diseases.[18]

Recently, our group described a Portuguese ULD patient who is homozygous for a new synonymous mutation (c.66G>A; p.Q22Q; r.[66g>a,65_66ins66+354pb]) which leads to missplicing of *CSTB* pre-mRNA.[19] Two transcripts were found in the patient-derived fibroblasts, a normal transcript with the synonymous G>A change at the last nucleotide of exon 1 and a mutant one including 354 bp of intron 1 due to the activation of a cryptic 5' splice site, which predicts an abnormal peptide with a premature truncation. Nevertheless, in terms of protein, no abnormal peptide was previously described as being detected. Although the presence of the normal and aberrant

transcripts is consistent, only the decreased amount of normal protein was established, and if an abnormal peptide does exists its level may be below resolution limits. In this work, we have explored both antisense oligonucleotide and U1 snRNA mediated strategies in an attempt to correct the splice defect associated with the c.66G>A splice donor site mutation in patient's cells. We report that the use of a specific locked nucleic acid (LNA) antisense oligonucleotide, designed to block the activated cryptic splice site in intron 1, succeeded to restore the normal splicing pattern since only the normal transcript was then produced.

Although most ULD patients have the dodecamer expansion, and would not be candidates for this approach, the experiment here described demonstrates the feasibility of LNA correction in the particular case under study and may provide evidence for its potential use in other cases with similar mutations. As far as we know this is the first report of a patient-tailored genetic lesion correction in cells from an individual with ULD.

Material and methods

Biological material

A Portuguese patient with a clinical diagnosis of ULD was previously reported to be homozygous for the c.66G>A change in the *CSTB* gene, in a study where the mutation effects at mRNA and protein levels were also analyzed.[19] Skin fibroblasts were obtained from the patient and one healthy control with appropriate informed consent. For all the samples confidentiality of personal data were protected. Both fibroblast cells were cultured and expanded following standard procedures in DMEM medium (Gibco Invitrogen, Carlsbad, USA) with 10% FBS and 1% kanamycin at 37°C and 5% CO_2. Average cell viability and proliferation was assessed at each passage, and at each experimental treatment, by the trypan blue vital dye exclusion test.

Cycloheximide treatment of control and patient fibroblasts

To perform the nonsense-mediated mRNA decay (NMD) assays, both fibroblast cells were cultured in the presence of two different concentrations of cycloheximide (100 and 600 µg/ml) for 8 hours. Total RNA was then isolated using the High Pure RNA Isolation kit (Roche Applied Science, Indianapolis, USA) and reverse-transcribed using the "Ready-To-Go You-prime first-strand beads" kit (GE Healthcare, Buckinghamshire, UK) following the manufacturer's protocols. Specific primers for exon 1 (Fw: 5'-

GCCGAGACCCAGCACATC-3') and exon 2 (Rv: 5'-TGACACGGCCTTAAACACAG-3') were used to amplify cDNA fragments encompassing the wild-type (WT) and the mutation region where the *CSTB* gene splicing was altered as described in Pinto *et al.*.[19] After RT-PCR amplification and electrophoretic separation, the obtained bands were cut and purified using the Wizard® SV Gel and PCR clean-up system (Promega, Madison, USA) and then sequenced.

U1 snRNA constructs treatment and analysis

The pG3U1 vector, which includes the sequence coding for human U1[20] (kindly provided by Dr. F. Pagani) was used to express the wild-type U1 snRNA (U1-WT). The different mutant U1 constructs (U1-CSTB-mut; U1-CSTB-wt and U1-CSTB-tc) adapted to the mutant splice donor site (SDS) of *CSTB* exon 1 were generated by site direct mutagenesis using the QuikChange® Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. PCR reactions were performed with specific primers for each construct. The U1-CSTB-mut construct was obtained using primers U1-mut-F (5'-GATCTCATACTTACTTGGCAGG GGAGATAC-3') and U1-mut-R (5'-GTATCTCCCCTGCCAAGTAAGTATGAGATC-3'), the U1-CSTB-wt was generated with the primers U1-wt-F (5'-GATCTCCCACCCACCT GGCAGGGGAGATAC-3') and U1-wt-R (5'-GTATCTCCCCTGCCAGGTGGGTGGGA GATC-3') and for the U1-CSTB-tc construct we used the primers U1-tc-F (5'-GAGGCCCAAGATCTCCCACCCACTTGGCAGGGGAG-3') and U1-tc-R (5'-CTCCCC TGCCAAGTGGGTGGGAGATCTTGGGCCTC-3'). The presence of the mutation under investigation (Figure 2A) was confirmed by sequence analysis.

To perform the U1 snRNA treatment, both healthy control and patient fibroblast cells at 90% of confluence were transfected in 6-well plates with 2.5, 3.5 or 4.5 µg of each modified U1 snRNA using either Lipofectamine 2000 or LTX (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions. The electroporation technique with the 4D-Nucleofector kit (Lonza, Basel, Switzerland) was also employed. In parallel, to estimate transfection efficiency, healthy control and patient cells were transfected with a control plasmid encoding GFP and fluorescent cells were monitored by microscopy. After 24 h or 48 h of transfection cells were harvested and total RNA was isolated using the High Pure RNA Isolation kit (Roche Applied Science, Indianapolis, USA). Conversion into cDNA and RT-PCR analysis was done as mentioned above.

Antisense oligonucleotide treatment and analysis

The antisense locked nucleic acid (LNA) oligonucleotide targeted to the donor cryptic splice site in the *CSTB* intron 1 was designed, synthesized and purified by Exiqon (Vedbaek, Denmark). The sequence of the LNA used is: 5' AGCCGGCTGCTCACCTGCGCCATCGCCG 3', as shown in Figure 3A. For the LNA treatment, between 2.5 and $3.5x10^5$ fibroblast cells were grown in 6-well plates and after 16 h treated with different concentrations of LNA (5, 25, 50 and 100 nM) using Lipofectamine LTX (Invitrogen, Carlsbad, USA) as delivery reagent. Cells were harvested 24 h later, total RNA was isolated and cDNA synthesized as described above. The RT-PCR analysis of the patient derived transcripts was performed using the primers described above.

CSTB Protein Immunodetection

For Western Blot analysis, between 1.4 and $1.6x10^6$ patient fibroblast cells grown in 75 cm^2 flasks were transfected with 100 nM of LNA using Lipofectamine LTX (Invitrogen, Carlsbad, USA) and harvested after 24 h. Total protein extracts, obtained from patient fibroblasts with and without LNA treatment, as well as from a healthy control sample without treatment, were homogenized in a lysis buffer solution (supplemented with a protease inhibitor cocktail) and ruptured by three cycles of 5 minute freeze thawing and then briefly centrifuged. Protein concentration was determined measuring the absorbance at 280 nm in a NanoDrop® ND-1000 Spectrophotometer. Equal amounts of total protein crude extracts were loaded on a 4-12% NuPAGE® Novex® Bis-Tris precast gel (Invitrogen, Carlsbad, USA). After electrophoresis, proteins were transferred to a nitrocellulose membrane in a semi-dry transfer system for 2 h at 150 mA. Ponceau staining was used to monitor equal loading of protein. Immunodetection was carried out using anti-Stefin B antibody (mouse monoclonal IgG1 reacting with human Stefin B: ab54566 from Abcam – Cambridge Science Park, UK) followed by an incubation with goat anti-mouse IgG-HRP as the secondary antibody (sc-2005, Santa Cruz Biotechnology, CA). Total amount of protein loaded was controlled by incubation with monoclonal anti-α-tubulin antibody (T6199 – Sigma-Aldrich, St.Gallen, Switzerland). The protein signal was developed using the Enhanced Chemiluminescence System (GE Healthcare, Buckinghamshire, UK).

Results:

The *CSTB* c.66G>A mutation generates one aberrant transcript in ULD patient

The *CSTB* transcription profile was obtained by RT-PCR and showed an aberrant transcript present in patient but not in control fibroblasts. To determine whether the expression level of this or other aberrant transcripts were altered by NMD, healthy control and patient fibroblasts were treated with two concentrations of cycloheximide (100 or 600 µg/ml) for a period of eight hours. As shown in Figure 1, no extra band or increased expression was observed in addition to the previously described transcripts. Direct sequencing of the bands confirmed the sequence pattern originally described in Pinto *et al.*.[19] These experiments indicated that the *CSTB* c.66G>A mutation generates one aberrant transcript that is not degraded by NMD.

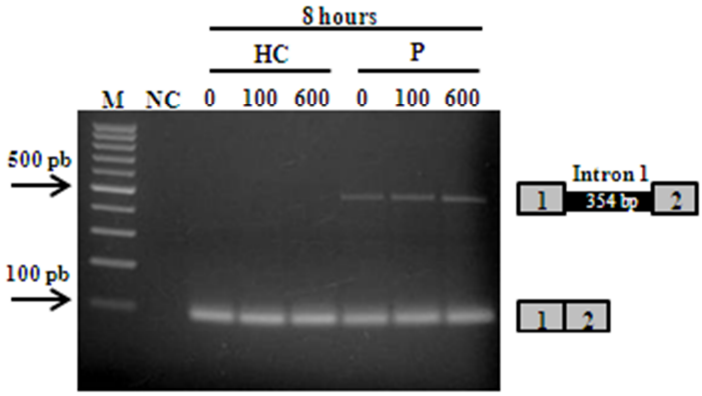


Figure 1: RT-PCR amplification of RNA extracted from control and patient fibroblasts untreated and treated with two different concentrations of cycloheximide (100 and 600 µg/ml) for 8 hours. M – molecular marker; NC – negative control; HC – healthy control; P – patient.

U1 snRNA therapeutic approach in patient cells

Since the recognition of the splice donor site occurs essentially through U1 snRNA complementary base pairing to the 5' splice site[21], any mutation in the SDS can disturb the U1 binding and consequently induce aberrant splicing. Here, we were dealing with exon 1 of the *CSTB* gene, wherein its 5´ splice site is complementary to the 5´end of the U1 snRNA in 7 out of the 9 consensus positions at the exon-intron border (Figure 2A). Mutation c.66G>A introduces a mismatch in the consensus G position at -1 of the exon, thus lowering the complementarity to U1 snRNA. Given these features, we sought to restore normal *CSTB* splicing through the generation of

mutant U1 expression constructs with increased complementarity to the mutated SDS of exon 1. Three different mutation-specific U1 adaptations were generated (Figure 2A): U1-CSTB-mut, which compensates only for the c.66G>A mutation; U1-CSTB-wt, which matches all 11 nucleotides of the wild-type *CSTB* SDS of exon 1; and U1-CSTB-tc engineered to match all 11 nucleotides of the mutated SDS. To control for the effect of the adaptations, the U1-WT expression construct was also used (Figure 2A).

After 24 h or 48 h of transfection with the different U1 constructs, RT-PCR analysis showed that in the case of the ULD patient cells, no changes occurred in the endogenous *CSTB* splicing pattern regardless of the types and concentrations of U1 isoforms tested. Also, in the control cell line none of the U1 variants concentrations tested induced alterations in normal *CSTB* splicing and only correctly spliced transcripts were detected (Figure 2B). The identity of the amplified bands was confirmed by sequencing analysis.

To discard the hypothesis of poor fibroblasts transfection efficiency as a possible cause for the negative results achieved, a fluorescence GFP vector was transfected in parallel and an efficient acquisition of the vector was observed. Furthermore, the electroporation technique was also tried but no improvement in transfection efficiency was achieved and the same results were obtained (data not shown).

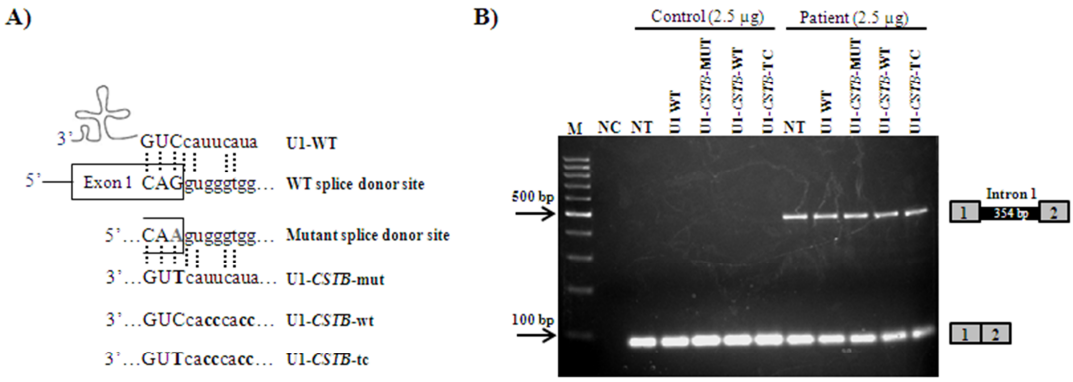


Figure 2: U1 snRNA-mediated Gene Therapeutic Approach in ULD patient-derived fibroblasts. **A)** Molecular base-pairing between the 5' end of the different U1 constructs and the wild-type and mutant splice donor sites of *CSTB* exon 1. The position of the mutation c.66G>A is highlighted in dark grey. Additionally to the construct expressing the U1-WT, the three different U1 variants generated: U1-*CSTB*-mut (-1T), U1-*CSTB*-wt (+3C;+4C;+7C;+8C) and U1-*CTSB*-tc (-1T;+3C;+4C;+7C;+8C) are represented. The U1 snRNA sequence changes are illustrated in bold. **B)** RT-PCR analysis of control and ULD patient cells untreated and after treatment with 2.5 µg of the different U1 constructs. Neither the U1-WT construct nor the three U1 variants tested showed any effect on *CSTB* exon 1 splicing. The same result was observed when overexpression of 3.5 or 4.5 µg of each adapted U1 was performed (data not shown). The healthy control normal splicing pattern was not altered after transfection of the various U1 adaptations. M – molecular marker; NC – negative control; NT – no treatment; mut – mutant; wt – wild-type; tc – totally complementary.

Correction of *CSTB* mutation-induced missplicing using a LNA oligonucleotide in patient-derived fibroblasts

As an alternative approach to correct *CSTB* exon 1 splicing, a specific 28-mer locked nucleic acid (LNA) oligonucleotide was used to block the activated cryptic 5' splice site within intron 1 (Figure 3A). Different quantities of the LNA were transfected in patient fibroblasts and then the endogenously spliced *CSTB* transcripts from 24 h treated cells were submitted to RT-PCR. The cDNA pattern showed that treatment with increasing amounts of LNA abolished the aberrantly spliced transcript with the insertion of 354 bp of intron 1 in a dose-dependent manner, with total correction being achieved at 100 nM of LNA (Figure 3B).

In order to analyze whether the observed splicing correction led to alteration of cystatin B protein levels, the protein was detected by Western blot in untreated control fibroblasts and in patient fibroblasts treated with 100 nM of the specific LNA. The Western blot protein analysis (Figure 3C) revealed the presence of a band with 11 kDa, which corresponds to the normal CSTB protein in control (untreated with LNA) and patient fibroblasts (treated and untreated with LNA). The treatment seems therefore not to interfere with CSTB protein expression. The abnormal truncated protein was not detected in patient fibroblasts, suggesting low amounts of protein below the detection limit of the antibody.

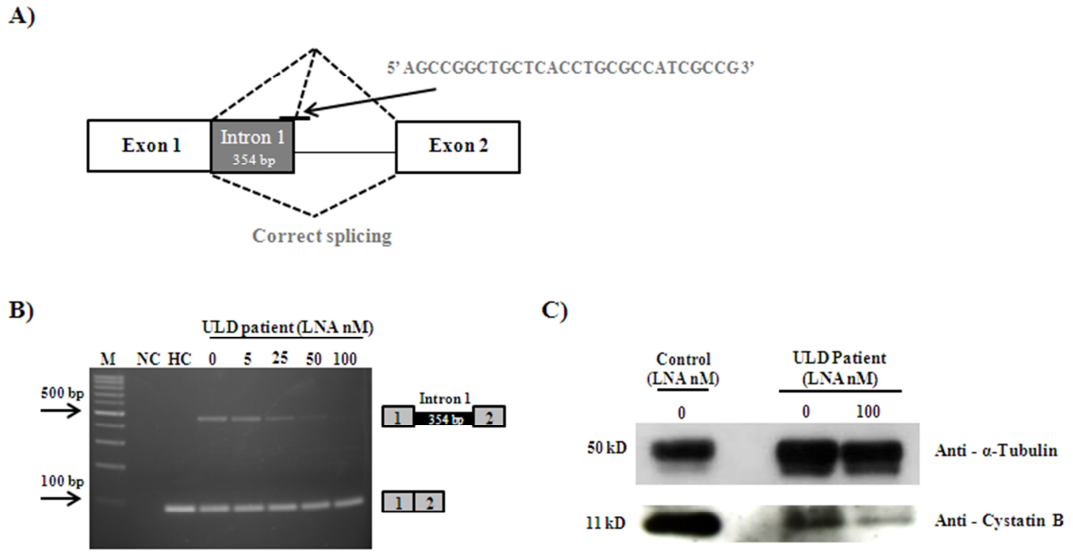


Figure 3: Antisense locked nucleic acid (LNA) therapeutic approach in ULD patient-derived fibroblasts. **A)** Schematic representation of the splicing downregulation observed in the presence of the c.66G>A *CSTB* mutation. The sequence of the LNA complementary to the cryptic donor site activated by the mutation and used in order to block the recognition of the intronic alternative 5' splice site in fibroblasts from the patient is shown. **B)** Transcriptional profile obtained for healthy control (HC) and patient fibroblasts untreated (0 nM) and treated with quantities between 5 to 100 nM of LNA

oligonucleotide. The RT-PCR analysis showed the disappearance of the aberrantly spliced transcript (451 bp) when cells were treated with 100 nM of the LNA oligonucleotide. Correctly spliced mRNA was obtained 24 h after transfection in a dose-dependent manner. **C)** CSTB protein expression in control and patient fibroblasts untreated (0 nM) and treated with 100 nM of the LNA. The Alfa-tubulin protein was used as loading control. M – molecular marker; NC – negative control; HC – healthy control.

Discussion

RNA missplicing diseases account for up to 15% of all inherited diseases, including neurological, myogenic and metabolic disorders.[22]

Mutations affecting splicing have been often neglected, either because they appear to be silent synonymous changes with no effect on amino acid sequence, or due to their misleadingly innocent intronic location. However, with the increasing knowledge arising from fine cDNA analysis and genomic sequencing of patients suffering from a wide range of diseases, the number of known exonic or intronic mutations that affect splicing has increased to 50-60% of all annotated disease-causing mutations.[23] It is therefore understandable that during the last decade, genetic therapy directed toward correction of RNA missplicing has progressed from theoretical work in cultured cells to promising clinical trials.[22]

In this work, in order to mitigate the splice defect present in a ULD Portuguese patient, we have developed and tested both antisense oligonucleotide and U1 snRNA mediated therapeutic strategies. Endogenous U1 snRNA is a major component of the spliceosome which targets 5' splice site introns within poly (A) pre-mRNAs in a sequence-specific manner recruiting the splicing machinery. The modification of U1 snRNA to increase complementarity affinities to mutated splice donor sites is being explored to correct 5' splice site defects dependent on U1 binding. In parallel, modified Antisense Oligonucleotides (AOs) that hybridize by complementarity to a selected site in the pre-mRNA have been also used to redirect splicing allowing restoration of the gene function.[24,25]

The ULD patient here addressed was homozygous for c.66G>A, a mutation that does not alter the coding amino acid sequence (p.Q22Q), but instead activates a cryptic splice site downstream in *CSTB* intron 1, probably because its presence weakens the recognition of the normal splice donor site by the U1. An aberrant transcript with the inclusion of 354 bp of intron 1 is produced and was found to escape nonsense-mediated mRNA decay since, as shown in figure 1, upon treatment with cycloheximide, the intensity of the expressed transcript did not increased. Since the transcript is not degraded, it is predictable that it will give rise to a truncated protein with only 36 amino acids, 13 of them not present in the normal CSTB protein that is

made up of 98 residues. At least in the patient fibroblasts, the normal transcript is also generated and normal CSTB protein is expressed although at lower levels than in control cells (Figure 3C). This raises the possibility that the cause of the disease in the patient can be the combined effect of the reduced levels of the normal CSTB protein and the production of a truncated protein that possibly exerts a toxic or dominant negative role, as it was evidenced for other *CSTB* mutant proteins.[26,27] Recent data indicated that *EPM1* mutants that affect protein sequence are prone to aggregate in cells, without being clear whether the proteins misfolding and aggregation are responsible for augmenting progression of the disease and neurodegenerative changes or whether it is the lack of the protein's function, or a combination of both.[26] It was also found that in ULD patients who are compound heterozygous bearing the dodecamer expansion in the promoter region together with a point or indel mutation, the clinical features are more severe than in patients homozygous for the dodecamer expansion, a genotype simply associated with low CSTB protein levels.[28,29] This finding further reinforces that mutant EPM1 proteins can indeed modulate disease phenotype by a toxic gain of function,[26,27] which, as a plausible hypothesis, may happen with the aberrant protein associated with the mutation c.66G>A. In any case, our reasoning was that the elimination of the aberrant transcript and consequently of the abnormal protein would represent a key step toward a personalized therapeutic correction for this patient. Furthermore, since transcript splicing patterns in other cell types may differ from the pattern observed in fibroblasts, the possibility of a different ratio of normal versus truncated transcript still remains as a plausible explanation for the disease and a target amenable by this type of approach.

The first attempt here report relied on the U1 strategy and failed to correct the aberrant transcript. Since the rescue of splicing mutations in the last base of the exon has been achieved using U1 construct isoforms[10,30], the lack of success in the case of c.66G>A suggests that the mutation might affect splicing through mechanisms that are not only dependent of the U1 complementarity. This interpretation gains support from previous studies[31,32] where mutations affecting the 5' splice site of *TCIRG1* and *BBS1* genes were demonstrated to have different rescue efficiencies after treatment with U1 mutant vectors, providing evidence that mutations close to the GT donor sites are mechanistically different. Symptomatically, according to the *in silico* predictions before performed[19] in the WT context of *CSTB* exon 1, the constitutive 5' splice site presented a lower score value (0.78) than the downstream cryptic splice site (0.98) only used in the presence of the G>A change. This observation suggests that the constitutive splice donor site recognition depends not only on the U1 snRNA binding, but also on other splicing factors that could interact with the *CSTB* pre-mRNA privileging the 5' canonical

use. In fact, the analysis of the exon 1 sequence by the *ESEfinder* and *ESRsearch* software's[33,34] predicted the presence of ESE binding motifs (CGACCAG or GACCAG) in the WT terminal region of exon 1 that were no longer recognized when the G>A change is present, also pointing to an involvement of specific SR proteins in the recognition of the WT exon 1 donor site. Therefore, the failure of the U1 snRNA therapeutic approach in neutralizing the effect of c.66G>A may be explained by the complexity of the *CSTB* exon 1 donor site splicing regulation, which might depend not only on the U1 complementarity, but also on the presence of enhancer splicing factors that were disturbed by the mutation.

As an alternative strategy to suppress the splicing defect, we used a specific LNA oligonucleotide to block the activated cryptic splice site in intron 1, which effectively abolished the aberrant splicing process of *CSTB* pre-mRNA in patient-derived cells. The normal splicing pattern leading to a single transcript with the synonymous change G>A was successfully rescued and so the therapeutic effect was expected to be achieved. Furthermore, the recovery was sequence and dose-specific (Figure 3B). In this way, this work adds evidence on the feasibility of antisense therapy, joining the many studies that are paving the way for translating the technology into the clinical practice.[35-39]

In conclusion, this work points to the promising *in vitro* correction of a CSTB protein defect causing ULD with a mutation-specific antisense therapy. Further studies are needed to assess the therapeutic potential of AOs, including how they could be delivered to specific organs and tissues without limitations. Even so, the insights obtained from this study make mutation-based therapy a clear candidate for personalized treatment of ULD patients encouraging similar investigations in other genetic diseases.

List of Abbreviations

AOs – Antisense Oligonucleotides; bp: base pair; EPM1 – Progressive myoclonic epilepsy type 1; GFP – green fluorescent protein; LNA – locked nucleic acid; NMD – nonsense-mediated mRNA decay; SDS – splice donor site; snRNAs: Small nuclear ribonucleic acids; ss – splice-site; ULD – Unverricht-Lundborg disease; WT: Wild-type

Conflict of Interest

The authors confirm that this article content has no conflict of interest.

Acknowledgements

The authors would like to thank Dr. Mónica Sousa and Dr. Elsa Logarinho research groups from Instituto de Biologia Molecular e Celular (IBMC), Porto for the collaboration in the electroporation studies. LM was supported by a grant (SFRH/BD/64592/2009) from the Fundação para a Ciência e Tecnologia IP (FCT) / POPH/FSE, Portugal.

References:

1. Merwick A, O'Brien M, Delanty N. Complex single gene disorders and epilepsy. Epilepsia 2012; 53 Suppl 4:81-91.

2. Kälviäinen R, Khyuppenen J, Koskenkorva P, et al. Clinical picture of EPM1-Unverricht-Lundborg disease. Epilepsia 2008; 49(4):549-556.

3. Genton P. Unverricht-Lundborg disease (EPM1). Epilepsia 2010; 51 Suppl 1:37-39.

4. Lehesjoki AE, Koskiniemi M, Sistonen P, et al. Localization of a gene for progressive myoclonus epilepsy to chromosome 21q22. Proc Natl Acad Sci USA 1991; 88(9):3696-3699.

5. Turk V, Bode W. The cystatins: protein inhibitors of cysteine proteinases. FEBS Lett 1991; 285(2):213-219.

6. Lalioti MD, Scott HS, Buresi C, et al. Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. Nature 1997; 386(6627):847-851.

7. Joensuu T, Lehesjoki AE, Kopra O. Molecular background of EPM1-Unverricht-Lundborg disease. Epilepsia 2008; 49(4):557-563.

8. Sánchez-Alcudia R, Pérez B, Pérez-Cerda C, et al. Overexpression of adapted U1snRNA in patients' cells to correct a 5' splice site mutation in propionic acidemia. Mol Genet Metab 2011; 102(2):134-138.

9. Glaus E, Schmid F, Da Costa R, et al. Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells. Mol Ther 2011; 19(5):936-941.

10. Schmid F, Glaus E, Barthelmes D, et al. U1 snRNA-mediated gene therapeutic correction of splice defects caused by an exceptionally mild BBS mutation. Hum Mutat 2011; 32(7):815-824.

11. Koo T, Wood MJ. Clinical trials using antisense oligonucleotides in duchenne muscular dystrophy. Hum Gene Ther 2013; 24(5):479-488.

12. Rigo F, Hua Y, Krainer AR, et al. Antisense-based therapy for the treatment of spinal muscular atrophy. J Cell Biol 2012; 199(1):21-25.

13. Roberts J, Palma E, Sazani P, et al. Efficient and Persistent Splice Switching by Systemically Delivered LNA Oligonucleotides in Mice. Mol Ther 2006; 14(4):471-475.

14. Pérez B, Gutiérrez-Solana LG, Verdú A, et al. Clinical, biochemical, and molecular studies in pyridoxine-dependent epilepsy. Antisense therapy as possible new therapeutic option. Epilepsia 2013; 54(2):239-248.

15. Chauhan NB, Siegel GJ. Antisense inhibition at the beta-secretase-site of beta-amyloid precursor protein reduces cerebral amyloid and acetyl cholinesterase activity in Tg2576. Neuroscience 2007; 146(1):143-151.

16. Kalbfuss B, Mabon SA, Misteli T. Correction of alternative splicing of tau in frontotemporal dementia and parkinsonism linked to chromosome 17. J Biol Chem 2001; 276(46):42986-42993.

17. Kay C, Skotte NH, Southwell AL, et al. Personalized gene silencing therapeutics for Huntington disease. Clin Genet 2014; 86(1):29-36.

18. Magen I, Hornstein E. Oligonucleotide-based therapy for neurodegenerative diseases. Brain Res Epub 2014 Apr 12.

19. Pinto E, Freitas J, Duarte AJ, et al. Unverricht-Lundborg disease: homozygosity for a new splicing mutation in the cystatin B gene. Epilepsy Res 2012; 99(1-2):187-190.

20. Lund E, Dahlberg JE. True genes for human U1 small nuclear RNA. Copy number, polymorphism, and methylation. J Biol Chem 1984; 259:2013-2021.

21. Zhuang Y, Weiner AM. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell 1986; 46:827-835.

22. Hammond SM, Wood MJ. Genetic therapies for RNA mis-splicing diseases. Trends Genet 2011; 27(5):196-205.

23. Douglas AG, Wood MJ. RNA splicing: disease and therapy. Brief Funct Genomics 2011; 10(3):151-164.

24. Havens MA, Duelli DM, Hastings ML. Targeting RNA splicing for disease therapy. Wiley Interdiscip Rev 2013; 4(3):247–266.

25. DeVos SL, Miller TM. Antisense oligonucleotides: treating neurodegeneration at the level of RNA. Neurotherapeutics 2013; 10(3); 486-497.

26. Polajnar M, Ceru S, Kopitar-Jerala N, et al. Human stefin B normal and patho-physiological role: molecular and cellular aspects of amyloid-type aggregation of certain EPM1 mutants. Front Mol Neurosci 2012; 5:88.

27. Polajnar M, Zavašnik-Bergant T, Kopitar-Jerala N, et al. Gain in toxic function of stefin B EPM1 mutants aggregates: Correlation between cell death, aggregate number/size and oxidative stress. Biochim Biophys Acta Epub 2014 Jun 5.

28. Koskenkorva P, Hyppönen J, Aikiä M, et al. Severe phenotype in Unverricht-Lundborg disease (EPM1) patients compound heterozygous for the dodecamer repeat

expansion and the c.202C>T mutation in the CSTB gene. Neurodegener Dis 2011; 8(6):515-522.

29. Canafoglia L, Gennaro E, Capovilla G, et al. Electroclinical presentation and genotype-phenotype relationships in patients with Unverricht-Lundborg disease carrying compound heterozygous CSTB point and indel mutations. Epilepsia 2012; 53(12):2120-2127.

30. Tanner G, Glaus E, Barthelmes D, et al. Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. Hum Mutat 2009; 30(2):255-263.

31. Susani L, Pangrazio A, Sobacchi C, et al. TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA. Hum Mutat 2004; 24(3):225-235.

32. Schmid F, Hiller T, Korner G, et al. A gene therapeutic approach to correct splice defects with modified U1 and U6 snRNPs. Hum Gene Ther 2013; 24(1):97-104.

33. Cartegni L, Wang J, Zhu Z, et al. ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acids Res 2003; 31(13):3568-3571.

34. Goren A, Ram O, Amit M, et al. Comparative analysis identifies exonic splicing regulatory sequences - The complex definition of enhancers and silencers. Mol Cell 2006; 22:769-781.

35. Brasil S, Viecelli HM, Meili D, et al. Pseudoexon exclusion by antisense therapy in 6-pyruvoyl-tetrahydropterin synthase deficiency. Hum Mutat 2011; 32(9):1019-1027.

36. Osorio FG, Navarro CL, Cadiñanos J, et al. Splicing-directed therapy in a new mouse model of human accelerated aging. Sci Transl Med 2011; 3(106):106ra107.

37. Miller TM, Pestronk A, David W, et al. An antisense oligonucleotide against SOD1 delivered intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study. Lancet Neurol 2013; 12(5):435-442.

38. Regis S, Corsolini F, Grossi S, et al. Restoration of the normal splicing pattern of the *PLP1* Gene by means of an antisense oligonucleotide directed against an exonic mutation. PLoS One 2013; 8:e73633.

39. Bestas B, Moreno PM, Blomberg KE, et al. Splice-correcting oligonucleotides restore BTK function in X-linked agammaglobulinemia model. J Clin Invest Epub 2014 Aug 8.

3.5 Ongoing work

Development of U1 snRNA-mediated therapeutic strategies to correct 5' splice site defects in Mucopolysaccharidosis I and Mucolipidosis III alpha/beta

In the context of this work, we have begun to explore the use of antisense U1 snRNA-mediated therapeutic strategies to correct two 5' ss mutations reported in patients with MPS I and Mucolipidosis type III alpha/beta (ML III alpha/beta). The selected mutations were c.1650+5G>A (*IDUA* gene) and c.3335+6T>G (*GNPTAB* gene), which affect the 5' ss positions +5 and +6, respectively, and were considered good candidates for the application of this type of therapeutic approach.

MPS I is an autosomal recessive disease caused by the deficiency of the enzyme α-L-iduronidase that can result in a wide range of phenotypic involvement, usually classified in three major clinical entities: Hurler (OMIM# 607014), Hurler-Scheie (OMIM# 607015), and Scheie (OMIM# 607016) syndromes. All of them are caused by mutations in the *IDUA* gene (Neufeld & Muenzer, 2001). Currently, HSCT and ERT using laronidase (Aldurazyme®) are the two available therapeutic options for MPS I that, however, still present some limitations (Coutinho et al., 2012; Cox, 2012).

ML IIII alpha/beta (OMIM# 252600) is an autosomal recessive disease caused by mutations in the *GNPTAB* gene. This gene encodes the α-and β-subunits of the N-acetylglucosamine (GlcNAc)-1-phosphotransferase, an enzyme that plays a key role in the formation of mannose-6-phosphate residues on lysosomal enzymes required for efficient transport to lysosomes (Kornfeld & Sly, 2001). Mutations in the *GNPTAB* gene can also cause a more severe form of mucolipidosis, the ML II alpha/beta (OMIM# 252500), whilst mutations in the *GNPTG* gene, which encodes another subunit of the same enzyme (γ subunit), are associated with the milder ML III gamma (OMIM# 252605) (Raas-Rothschild et al., 2000; Tiede et al., 2005). Both types of mucolipidoses are rare LSDs which share similar clinical features, including skeletal abnormalities. Since to date, no causal therapies are available for the treatment of patients, the management of the disease is essentially symptomatic (Raas-Rothschild et al., 2012).

According to the information contained in the Human Gene Mutation Database (HGMD® Professional 2015.3), splicing mutations are frequent genetic defects in MPS I and ML II and III alpha/beta amounting to ~ 16 % and 10 %, respectively. From these, ~ 64 % in MPS I and 47 % in ML II and III alpha/beta correspond to 5' SDS mutations,

thus constituting a vast target repertory for mutation specific U1 snRNA-mediated therapeutic approaches.

The mutations addressed in this work, c.1650+5G>A in the *IDUA* gene and c.3335+6T>G in the *GNPTAB* gene, were found in a MPS I and ML III alpha/beta patient's respectively. Both mutations were already described as well as was deciphered their impact in splicing (Kudo et al., 2006; Venturi et al., 2002).

Fibroblast samples from the two patients were obtained from the biobank of Institute G. Gaslini, in the scope of a collaborative work with Doctor Mirella Filocamo (Genova – Italy). The MPS I patient is a compound heterozygous for a nonsense mutation (c.1205G>A; p.W402X) and the 5' ss mutation c.1650+5G>A in intron 11 of *IDUA* responsible for the skipping of exon 11. The ML III alpha/beta patient is also a compound heterozygous for a missense mutation (c.2864C>T; p.A955V) and the 5´ ss mutation c.3335+6T>G in the intron 17 of *GNPTAB* whose presence leads to the skipping of the exon 17 (Table 3.1; Figure 3.1 A and B).

Table 3.1: Genotypes of the MPS I and ML III alpha/beta affected patients.

Patient	Allele 1	Allele 2	Location	Reference
P1 (MPS I)	c.1650+5G>A	c.1205G>A (p.W402X)	Intron 11/Exon 9	(Venturi et al., 2002) / (Scott et al., 1992)
P2 (ML III alpha/beta)	c.3335+6T>G	c.2864C>T (p.A955V)	Intron 17/Exon 14	(Kudo et al., 2006) / (Zarghooni & Dittakavi, 2009)

P1 – patient 1; P2 – patient 2

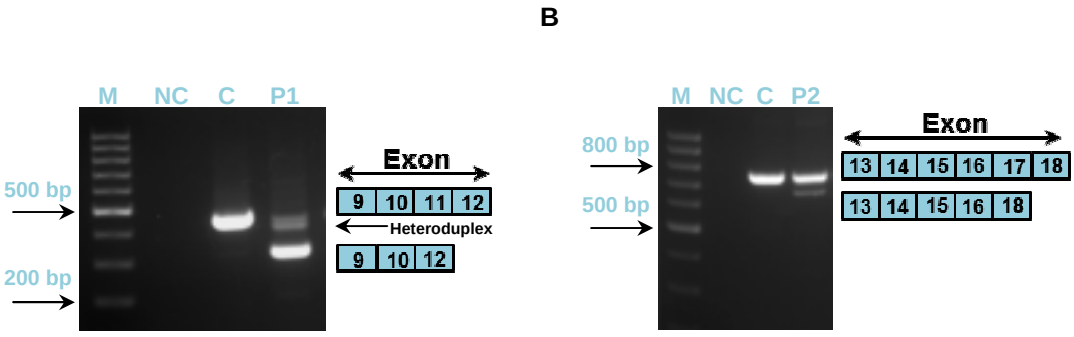
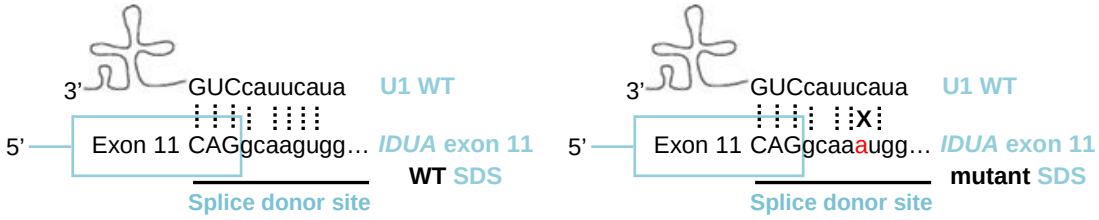


Figure 3.1: Agarose gel showing the transcripts observed in control (C) and in patients (P1 and P2) together with schematic views of each transcript's constitution. **A)** Patient P1 (MPS I) carries a 5' ss mutation located on intron 11 of the *IDUA* gene (c.1650+5G>A) and a nonsense change on exon 9 (c.1205G>A; p.W402X). *IDUA* transcripts were amplified between exons 9 and 12 and the transcriptional pattern showed three bands: one corresponding to the skipping of exon 11 transcribed from the allele bearing the donor site change c.1650+5G>A (lower band), a normal band arising from the allele with the nonsense mutation (p.W402X) (upper band) and an heteroduplex extra product (intermediate band – black arrow). The healthy control showed a single amplified band of normal molecular weight. **B)** Patient P2 (ML III alpha/beta) carries a donor site mutation located on intron 17 of the *GNPTAB* gene (c.3335+6T>G) and a missense mutation (pA955V) on exon 14. For the amplification of the *GNPTAB* transcripts primers for exons 13

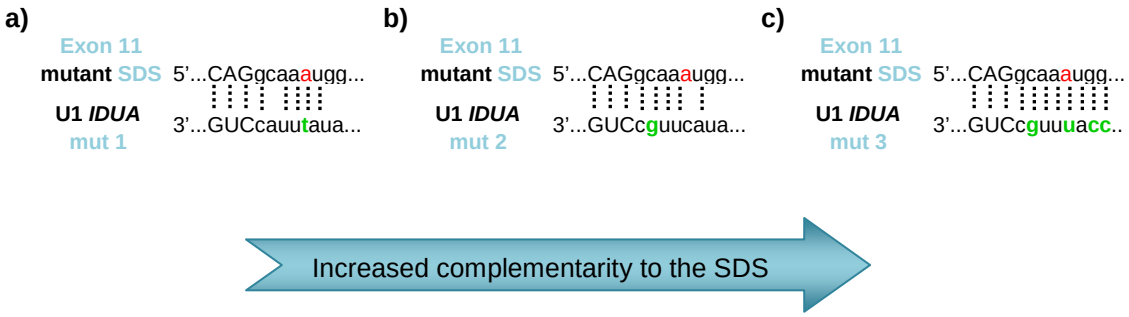
and 18 were used and the RT-PCR pattern showed a smaller band with the skipping of exon 17 resulting from the 5' ss change c.3335+6T>G, and a larger one with normal molecular weight from the allele bearing the missense change. The healthy control showed a single amplified band of normal molecular weight. Sequencing results of the obtained products are illustrated by schematic drawings. M – molecular marker; NC – negative control; C – control; P1 and P2 – patient 1 and 2 respectively.

To proceed with the development of the antisense-snRNA therapeutic strategies for both mutations, an analysis was performed to identify the specific base pairing between the U1 snRNA cDNA sequence and the wild-type (WT) and mutant 5' ss of *IDUA* exon 11/intron 11 and *GNPTAB* exon 17/intron 17. (Figure 3.2 A1 and B1). Subsequently, to construct the different U1 snRNA mutation-adapted vectors, to bind specifically to both mutant SDSs, the pG3U1 vector (Lund & Dahlberg, 1984) was modified through mutagenesis. For the *IDUA* mutation three different U1 variants were constructed (U1 *IDUA* mut 1 – complementary to +5; U1 *IDUA* mut 2 – complementary to +2; U1 *IDUA* mut 3 – totally complementary to the SDS) (Figure 3.2 A2) and for *GNPTAB* mutation two U1 mutant vectors were constructed (U1 *GNPTAB* mut 1 – complementary to +6; U1 *GNPTAB* mut 2 – totally complementary to the SDS) (Figure 3.2 B2).

A1



A2



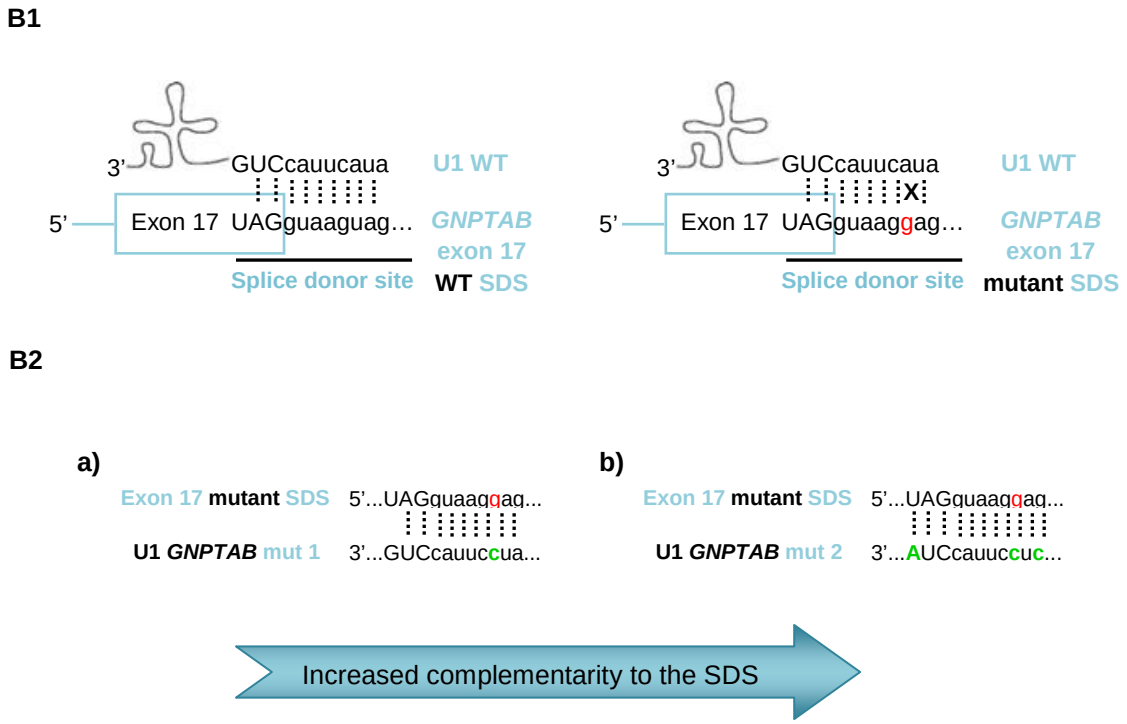
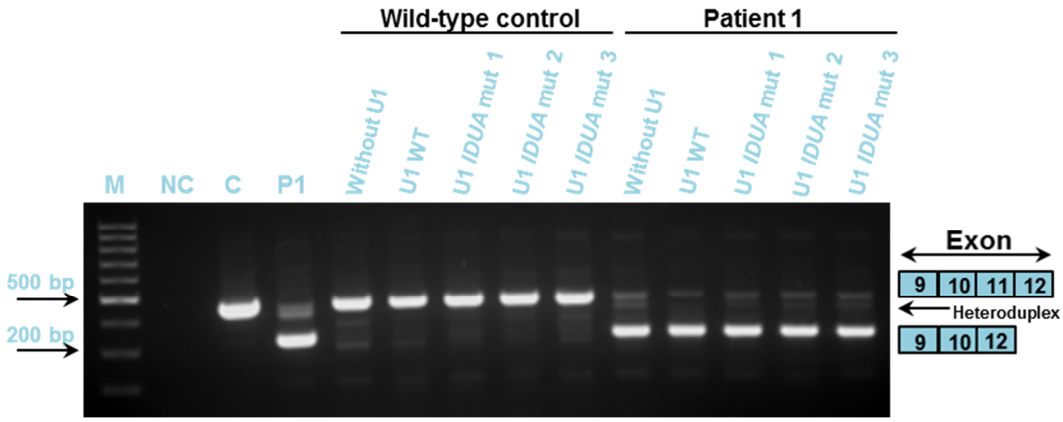


Figure 3.2: A1 and B1) Schematic representation of base-pair interactions between the U1 snRNA and the WT and mutant SDSs of exon 11 and exon 17 of *IDUA* and *GNPTAB* genes, respectively. **A2 and B2)** Illustration of the strategy followed to increase the complementarity of U1 snRNA with the mutated SDS of *IDUA* and *GNPTAB* genes. U1 complementarity was increased stepwise, and to try to compensate for the *IDUA* mutation at position +5, three different U1-adaptations were designed [U1 *IDUA* mut 1 (+5T) – **A2a**; U1 *IDUA* mut 2 (+2G) – **A2b** and U1 *IDUA* mut 3 (+2G; +5T; +7C; +8C) – **A2c**], whereas for the *GNPTAB* mutation at position +6, two different U1-adaptations were engineered [U1 *GNPTAB* mut 1 (+6C) – **B2a** and U1 *GNPTAB* mut 2 (-3A; +6C; +8C) – **B2b**]. Upper case letters show exonic nucleotides, whereas the lower case letters denote intronic nucleotides. Base-pairing is indicated by vertical lines and its loss by an X. The mutant nucleotide is highlighted in red and the changed nucleotides in the U1 sequence are illustrated in green.

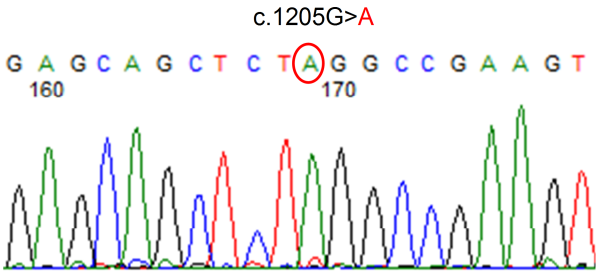
Finally, to achieve therapeutic functional rescue of the mutations under study, different amounts (between 2.5 µg and 4.5 µg) of the U1 WT vector (as control) and of the different mutation-adapted U1 vectors were transfected using Lipofectamine LTX (Invitrogen) into control, MPS I and ML III alpha/beta patients' fibroblasts for 24 h or 48 h. After transfection, the RT-PCR analysis of control samples for both cases revealed that the WT splicing pattern was not affected by any of the U1 isoforms (Figure 3.3 A1 and B1). For patient samples, the skipping of exon 11 (Figure 3.3 A1) and of exon 17 (Figure 3.3 B1) generated by the mutated allele in the SDS positions +5 and +6 respectively, was still observed. Additionally, in both cases the gel band corresponding to the transcript with the normal length was excised, purified and sequenced, in order to check if in the nucleotide positions where the nonsense (*IDUA*) and missense (*GNPTAB*) mutations occur, the WT bases were also present. Their presence would

point to a partial correction of the alleles carrying the SDS mutations. However, no such peaks were detected (Figure 3.3 A2 and B2) allowing to conclude that no correction of exon 11 or exon 17 skipping was achieved after treatment with the different U1-modified vectors.

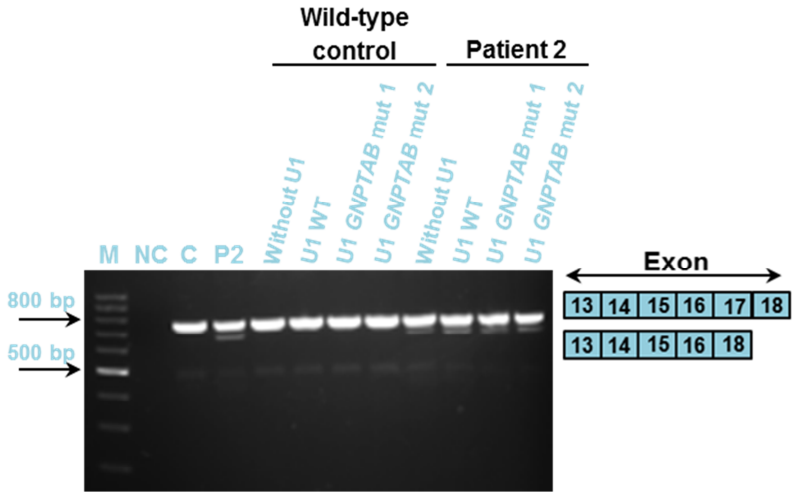
A1



A2



B1



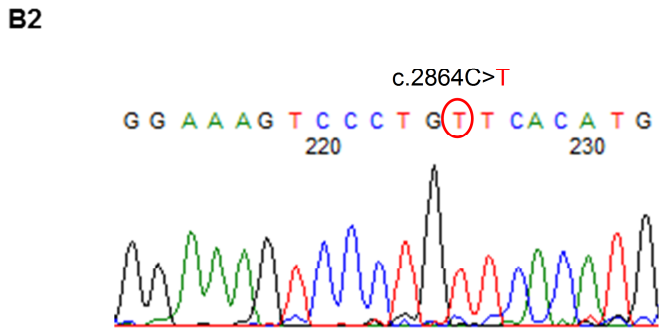


Figure 3.3: A1 and B1) Splicing pattern of WT control fibroblasts and MPS I and ML III alpha/beta patients' cells untransfected and after transfection of the different U1 vectors. The RT-PCR amplification of endogenous transcripts was performed with primers for exons 9 and 12 of the *IDUA* gene [WT control (C) and MPS I patient (P1)] and with primers for exons 13 and 18 of the *GNPTAB* gene [WT control (C) and ML III alpha/beta patient (P2)]. The constitutive splicing of exon 11 of the *IDUA* gene **(A1)** and of exon 17 of *GNPTAB* gene **(B1)** was not altered in control fibroblasts after overexpression of U1 WT or any of the modified U1 constructs. For patients P1 **(A1)** and P2 **(B1)** the transfection of 2.5 µg of U1 WT or any generated U1 adapted vector for 24 h did not produced any change in the endogenous aberrant splicing pattern. The same results were obtained with 3.5 or 4.5 µg of the different U1 modified constructs as well as in the 48 h treatment (data not shown). Band identity is illustrated by schematic drawings on the right of the gels. **A2 and B2)** Electropherograms of the high molecular weight band obtained in RT-PCR after transfection of the U1 *IDUA* mut 3 and of the U1 *GNPTAB* mut 2 in the MPS I patient bearing the genotype c.1650+5G>A/c.1205G>A and in the ML III alpha/beta patient bearing the genotype c.3335+6T>G/c.2864C>T respectively. The region of the nonsense (c.1205G>A) **(A2)** and missense (c.2864C>T) **(B2)** mutations is specifically represented. In these regions no heterozygous peak was observed, showing that not even the partial correction of the alleles carrying the SDS mutations has occurred.

It is well established that the 5' SDS does not always conform to the consensus sequence (CAG/GURAGU; R-purine), but can instead have a degenerate pattern feature (Raponi & Baralle, 2008; Valadkhan & Gunawardane, 2013). Therefore, not all positions of the SDS are equally important to enable the recognition by U1 and to ensure correct splicing. Furthermore, various base-pair combinations within the SDS show increased binding to U1, indicating mutual relationships between specific nucleotides of the SDS (Schmid et al., 2013). These findings highlight the importance of the 5' ss sequence context as a key factor to influence not only the donor site recognition by U1 but also the success of a U1 snRNA-mediated therapy.

Some U1 snRNA therapeutic interventions, targeting SDS mutations at position +5, have been performed showing partial (Schmid et al., 2013) or complete correction of aberrant splicing (Pinotti et al., 2009; Pinotti et al., 2008). However, in a study comprising a +5 mutation in the *TCIRG1* gene (Susani et al., 2004), the application of a mutant U1 snRNA engineered to bind specifically at the mutant donor site did not correct aberrant splicing, but in turn induced an increase in intron retention. These results suggest that 5' ss mutations at this position could affect splicing with different mechanisms and that this may compromise the success of the U1-based therapy.

To the best of our knowledge, there is no successful report of U1 snRNA therapeutic approaches for mutations at position +6. However, it is important to mention that in a study of Schmid and coworkers (Schmid et al., 2013) an attempt was carried out to investigate the effect on splicing of a +6 mutation and its possible correction through U1-based therapy. To evaluate the effect of different sequence alterations on splicing of the exon 5 SDS of *BBS1* gene, minigenes mutated at nine positions (between -3 and +6) were generated. Specifically, for the +6 mutation no alteration in the normal splicing pattern was observed after the expression of the mutated minigene. Despite this result, the authors performed the co-transfection of the +6 mutant minigene along with two different modified U1's, though, as expected, no effect on splicing was observed.

In the current work the transfection of different U1 adaptations in patients' fibroblasts with the c.1650+5G>A (*IDUA* gene) and c.3335+6T>G (*GNPTAB* gene) mutations did not rescued the induced aberrant splicing patterns. It is of note that the U1's transfection was associated with a high level of cell death, which could be due to the presence of the modified U1's, but we cannot rule out that it may also have arisen due to the cells contact with the chemical reagent Lipofectamine used for transfection. This observation led us to hypothesize that low transfection efficiency may have occurred compromising an efficient acquisition of the different U1's by the cells.

In order to understand if the absence of aberrant splicing correction was indeed related with low transfection efficiency, we intend to perform further investigations. At first, we plan to co-transfect the different U1 modifications with WT and +5 and +6 mutant minigenes (already constructed) into an established continuous cell line (e.g. COS-7). Since these cells generally allow good transfection efficacy, its use could help to disclose easily possible splicing pattern changes due to the U1's adaptations. In a further approach, we pretend to implement in our laboratory the viral transduction technique in order to test the mutation therapeutic rescue through the AAV or lentiviral treatment of patients' fibroblasts with the different U1 variants. Viral vectors are considered more efficient and less toxic than other delivery systems, minimizing cell death and side effects that usually come along with transfection due to the use of chemical reagents. The viral transduction of U1 constructs in patients' fibroblasts was applied successfully in some diseases allowing the total or partial recovery of misspliced transcripts (Glaus et al., 2011; Hartmann et al., 2010; Schmid et al., 2011).

Additionally, it would be interesting to test the effects of the U6 snRNA in a similar way to the one tested for the U1 snRNA. The U6 has been described as essential for an accurate splicing process. After U1 dissociation, the correct recognition of the exon at 5' ss is assured through the interaction of the U6 snRNA with nucleotides at

positions +4 to +6 of the SDSs (Kandels-Lewis & Séraphin, 1993; Lesser & Guthrie, 1993). In the recent study of Schmid et al. (Schmid et al., 2013), involving a mutation affecting the +5 position of the donor site, the partial correction of aberrant splicing was only achieved using modified U1 vectors. The authors evaluated the effect of several adaptations of the U6 snRNA sequence for its correction, having found out that only a co-application of adapted U1 and U6 isoforms corrected the splice defect caused by the mutation at SDS position +5. The sequence complementarity between U6 and the three SDS positions (+4, +5, and +6) showed to be relevant for the outcome of the therapy.

Extrapolating these results for the mutations at positions +5 and +6 that we are studying, it would be important to also test the effect of mutation-adapted U6 isoforms in the correction of the splicing defects, given its importance in the interaction with the affected donor site positions during the splicing process.

CHAPTER 4

CONCLUSIONS AND FUTURE PERSPECTIVES

4.1 Functional analysis and splicing mechanisms comprehension of exonic variants on *IDS* gene

Exonic point mutations can create or disrupt *cis*-acting sequence elements (e.g. ESEs or ESSs) thus affecting splicing. A significant proportion of these mutations may consequently have a wide range of effects, which, however, are very difficult to predict.

In this study, the analysis of *IDS* splicing patterns in the presence of the mutation c.241C>T, previously classified as a conventional nonsense mutation (Brusius-Facchin et al., 2014), allowed us to verify that this mutation also affects the normal splicing process of the *IDS* gene. Furthermore, the functional analysis of reporter minigenes for this mutation and for two other exonic mutations on *IDS* gene, c.257C>T and c.1122C>T, previously characterized at the cDNA level, further confirmed its involvement on the deregulation of splicing in the *IDS* gene. Therefore, an accurate characterization of genetic mutations able to clarify the link between genotype and phenotype cannot be restricted to the examination of gDNA. Unless cDNA analysis is also conducted, simple classifications as synonymous, missense or nonsense mutations can mitigate the fact that a mutation may also generate severe, unconventional splicing alterations. In this way, the results obtained in this study reinforce the importance of analysing all exonic variations that predictably can induce any effect on splicing, not only through cDNA analysis but also through other analytical tools such as the use of minigenes.

In addition, exonic mutations that cause defective splicing can give important lessons for understanding exon-identity determinants and AS mechanisms. In this study, the *in silico* analysis of the splicing mutation c.257C>T located on *IDS* exon 3 allowed to identify changes in putative binding motifs for some SR and hnRNP *trans*-acting proteins. The functional analysis through protein splicing factors overexpression or depletion assays and through *cis*-acting motifs loss of function minigene studies validated the bioinformatic predictions, confirming that the detected changes in *cis*-acting motifs are related to the splicing pattern alterations observed. Further, the functional experiments also demonstrated that the SRSF2 and hnRNP E1 proteins are involved in the use and repression of the constitutive 3' ss of exon 3, respectively.

Globally, this study contributed to a deeper understanding of the molecular basis of *IDS* gene splicing regulation through the identification of critical SREs with effects on splicing, as also provided the basis to propose an explanatory model for the splicing regulation around the *IDS* exon 3 region. This kind of exploratory studies encourages

similar approaches in other genes, which hopefully will further increase the knowledge about the pre-mRNA splicing.

4.2 Development of antisense oligonucleotides and U1 snRNA-mediated therapeutic strategies

Synonymous mutations that alter splicing are suitable targets for AO therapy given that once corrected the splicing defect the restored transcript will allow normal protein production. In this work, for the synonymous mutations c.1122C>T and c.66G>A that impair the normal splicing process of *IDS* and *CSTB* genes respectively, we have investigated the applicability of RNA-based therapies to revert the caused splicing defects.

For the c.1122C>T change present in a MPS II patient, which creates a new 5' ss within exon 8 giving rise to a transcript shorter than usual, the use of four different AOs (three AMOs and one LNA) did not correct the aberrant splicing in patient cells. Instead, it led to the production of a new abnormal splicing product with the total skipping of exon 8. Given this result, we reasoned that the AOs could be masking some exonic motifs potentially involved in the inclusion of exon 8 in the mature transcript. In fact, an *in silico* analysis predicted the existence of several putative *cis*-binding motifs for *trans*-acting proteins in the mutated region. Additionally, AOs transfection experiments performed in control fibroblasts resulted in a change of the normal splicing pattern towards aberrant forms, including the exon 8 skipping observed in patient cells after treatment, further demonstrating that the AOs are indeed masking relevant motifs for the 5' ss regulation of exon 8. Altogether these observations support the conclusion that the efficacy of AO therapy greatly depends on the genomic environment.

Given the "silent" effect of this splicing mutation and the potential of AOs for the correction of this kind of variants, it will be important to continue the search for an optimal AO that effectively corrects the aberrant splicing. A suitable approach to find that AO involves the design of an AO microwalk at a single nucleotide resolution along the entire length of *IDS* exon 8, with the subsequent test of all resulting AOs in patient fibroblasts. If the correction is successfully achieved at RNA and protein levels, the further development of *in vivo* studies using an animal model will also be mandatory for the translation of this approach into the clinic.

The c.66G>A mutation was described in *CSTB* gene in a patient with ULD. ULD is a lysosome-related progressive myoclonic epilepsy for which the pathophysiologic mechanism is still not completely understood, even though it might be underlain by loss

or gain of function mechanisms (Lehtinen et al., 2009; Polajnar et al., 2012). This synonymous mutation, leads to missplicing of *CSTB* pre-mRNA, generating a normal transcript with the synonymous G>A change at the last nucleotide of exon 1 and a mutant one with a partial inclusion of intron 1, due to the activation of a cryptic 5' ss. This aberrant transcript does not seem to be a target of the NMD mechanism as shown by cycloheximide experiments, indicating that a truncated protein may be produced. So, in this case the disease can result from the reduced expression of the normal CSTB protein in combination with the presence of a mutant truncated protein with a possible toxic role. To try to overcome the consequences of this splicing defect, we used a specific LNA oligonucleotide that effectively blocked the activated cryptic splice site in intron 1, resulting in the recovery of the normal splicing pattern of a single transcript with the synonymous change G>A. Hence, this work represents the *in vitro* proof of concept that AOs can be used to overcome the effect of this *CSTB* splicing mutation. Currently, only symptomatic pharmacologic and rehabilitative management are available for the treatment of ULD patients. So, a therapy based on mRNA correction would be a desirable alternative, although only applicable to a small number of patients, because, despite the high frequency (~31%) of splicing mutations described in *CSTB* (HGMD® professional release 2015.3), only two of the reported mutations, including the one here studied (Kagitani-Shimono et al., 2002; Pinto et al., 2012), would profit from AOs mutation-based therapy, since the others affect canonical splice sites. It is possible that more research on innovative therapeutic options, as well as more reports on *in vitro/in vivo* studies to overcome this pathology, help draw further attention to ULD, encouraging clinicians to refer additional cases of epilepsies of unknown aetiology to be screened for *CSTB* defects, potentially increasing the number of diagnosed mutations suitable for correction through this kind of therapy.

Having been here demonstrated the mutation correction at the RNA level, future studies addressing the clarification of its impact at protein level will be fundamental. The use of an antibody specifically designed for the detection of the mutant *CSTB* protein will lead to a more accurate analysis of the CSTB protein profile before and after the AO treatment, which certainly would afford a more comprehensive picture about the effect of the c.66G>A mutation at the protein level, and, consequently, about the effect of the treatment strategy. Again, for the effective development of the therapy it will also be essential to conduct *in vivo* studies.

Another major issue in this work was to explore the potential of U1 snRNA-mediated therapies. Once this approach had been previously used to rescue splicing mutations in the last base of an exon, we also tried to apply it to the just before mentioned c.66G>A in *CSTB*, given its location in the last nucleotide of exon 1 of the

gene. Still, in this case, the correction of the abnormal transcript in the patient cells was not achieved. The data obtained from the *in silico* analyses actually explains this negative result once it showed that the 5' canonical splice site presents a lower score than the 5' downstream cryptic splice site and the predicted ESE binding motifs at the end region of the WT exon 1 were altered in the presence of the c.66G>A mutation. These predictions support the hypothesis that the proper recognition of the *CSTB* exon 1 may depend, not only from the U1 complementarity but also from other splicing factors interacting with the pre-mRNA, which probably compromised the success of the U1 approach for this mutation.

Modified U1 snRNAs were also used as an effort to correct donor site splicing defects in *HGSNAT*, namely those which concerned the 5' ss mutations c.234+1G>A, c.633+1G>A and c.1542+4dupA affecting four different MPS IIIC patients.

In a first optimization approach where the effect of the overexpression of different designed U1 snRNAs on the splicing process was tested on minigenes bearing each specific donor site mutation, no correction of the aberrant splicing was observed. In fact, despite the skipping of each respective exon was corrected, the transcript generated after treatment was still aberrant due to the use of cryptic "gt" donor sites situated at intronic positions +5 and +6 (c.234+1G>A and c.633+1G>A mutations) or +6 and +7 (c.1542+4dupA mutation) which was promoted, partially or completely, by some adapted U1 snRNAs. Despite the absence of correction, and given that minigenes do not entirely reproduce the human cellular environment, the different U1 variants were also tested directly into patients' fibroblasts. In the case of the c.633+1G>A and c.1542+4dupA mutations, the normal endogenous splicing process was not recovered after the overexpression of the different modified U1 snRNAs. For the c.234+1G>A mutation present in homozygosity in two patients, the overexpression of a modified U1 that completely matched the mutated 5' ss allowed to achieve a partial recovery (almost 50%) of the splicing process, together with the use of the cryptic "gt" site at intronic positions + 5 and +6. This was an unexpected and surprising positive result, since the majority of the mutations localized in the high conserved nucleotide positions +1 and +2 of the SDS that were submitted to correction with modified U1 snRNAs were not successfully rescued (Fernandez Alanis et al., 2012; Schmid et al., 2013). To the best of our knowledge, the partial rescue of a +1 splice site mutation was reported only once, but in a case in which the mutated allele already produced some degree of the normal spliced transcript before treatment (Hartmann et al., 2010). So, in our study, we present the first case where a partial recovery of the normal splicing was achieved for a +1 mutation that did not produce any WT spliced mRNA when untreated. In this case, given the partial recovery of the correctly spliced mRNA, a

measurement of the enzymatic activity in patients' cells after treatment was performed to analyse the effect of the partial splicing correction in protein function. However, no increase in enzyme activity level was observed, indicating that possibly a higher level of recovery of the splicing process is required to enhance the enzymatic activity.

Whilst in this study we have added evidence that modified U1 snRNAs hold the potential to be used in the treatment of +1 donor site mutations, the results also pointed out that the success of the therapy would depend on the presence of other "gt" dinucleotides in the donor site region which, depending on the context, may be used as an alternative splice site interfering with the correction of the splicing process.

Concerning the c.234+1G>A mutation, given that the overexpression of a totally complementary U1 snRNA promoted the utilization of an alternative splice site at positions +5 and +6, and taking into account the role of U6 snRNA in the recognition of the 5' ss, in the near future it would be interesting to check whether the effect of modified U6 snRNAs alone or in combination with modified U1 snRNAs could specifically improve the use of the mutated site and not the alternative one, allowing the total correction of the 5' ss defect. Once this is achieved, *in vivo* approaches should be designed. Recently, Balestra and colleagues (Balestra et al., 2014) described a novel methodology to generate mice expressing a 5' ss defect for the coagulation factor VII, through liver-directed expression by plasmid or recombinant AAV vector administration. For those mice, subsequent injection of a modified U1 vector allowed the correction of an exogenous mutant construct. Thus, for this *HGSNAT* mutation it should be analysed the possibility of developing a similar approach in order to provide an *in vivo* proof of concept that this kind of therapy may be used for the correction of this MPS IIIC donor site splicing defect.

Finally, we also have begun the development of U1 snRNA therapeutic approaches to correct two other mutations that affect the recognition of the 5' ss, the c.1650+5G>A in the *IDUA* gene and c.3335+6T>G in the *GNPTAB* gene, causing MPS I and ML III alpha/beta, respectively. Different modified U1 snRNAs were transfected in patients' fibroblasts, but no rescue of the normal splicing process was observed. Despite these preliminary results, this study is still ongoing and further experiments such as the co-transfection of the different U1 modifications with WT and +5 and +6 mutant minigenes, the transduction in patients' cells of the U1 variants using viral vectors as well as the study of the effect of different U6 snRNA modifications are already planned to find out the potential of antisense-U1snRNA therapy for these donor site mutations.

The implementation in our laboratory of the methodologies for the correction of splicing mutations using RNA-based therapies was also one of the important achievements of this work, once it opens the possibility of their application in future studies.

Overall, this work allowed the identification of specific SREs involved in the splicing regulation of a particular region of the *IDS* gene, contributing not only to extend the knowledge about the mechanisms underlying the splicing of this gene but also of the splicing process in general. Additionally, the *in vitro* development of specific antisense therapeutic strategies to correct different splicing defects in LSDs was successfully achieved, leading to a better understanding of mutation-based therapies, and pointing them out as a feasible way to reach the personalized treatment of LSDs patients. These findings further encourage the investigation of similar approaches in other rare genetic diseases.

To conclude, we would like to remember that science is a constant step-by-step construction. Our hope is that studies like this might seed the grounds to others that taking advantage on small lessons can proceed to *in vivo* approaches and use additional analytical methods to gain a comprehensive picture of the potential application of splicing therapeutics not only to LSDs but also to other genetic disorders.

CHAPTER 5

REFERENCES

Aartsma-Rus A. (2012) Overview on AON design. *Methods Mol Biol* 867:117-129.

Aartsma-Rus A. (2014) Dystrophin Analysis in Clinical Trials. *J Neuromuscul Dis* 1:41-53.

Aartsma-Rus A, Singh KH, Fokkema IF, Ginjaar IB, van Ommen GJ, den Dunnen JT, van der Maarel SM. (2010) Therapeutic exon skipping for dysferlinopathies? *Eur J Hum Genet* 18:889-894.

Aartsma-Rus A, van Ommen GJ. (2007) Antisense-mediated exon skipping: a versatile tool with therapeutic and research applications. *RNA* 13:1609-1624.

Aartsma-Rus A, van Vliet L, Hirschi M, Janson AA, Heemskerk H, de Winter CL, de Kimpe S, van Deutekom JC, 't Hoen PA, van Ommen GJ. (2009) Guidelines for antisense oligonucleotide design and insight into splice-modulating mechanisms. *Mol Ther* 17:548-553.

Alekseyenko AV, Kim N, Lee CJ. (2007) Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes. *RNA* 13:661-670.

Alter J, Lou F, Rabinowitz A, Yin H, Rosenfeld J, Wilton SD, Partridge TA, Lu QL. (2006) Systemic delivery of morpholino oligonucleotide restores dystrophin expression bodywide and improves dystrophic pathology. *Nat Med* 12:175-177.

Alves S, Mangas M, Prata MJ, Ribeiro G, Lopes L, Ribeiro H, Pinto-Basto J, Lima MR, Lacerda L. (2006) Molecular characterization of Portuguese patients with mucopolysaccharidosis type II shows evidence that the IDS gene is prone to splicing mutations. *J Inherit Metab Dis* 29:743-754.

Aoki Y, Yokota T, Nagata T, Nakamura A, Tanihata J, Saito T, Duguez SM, Nagaraju K, Hoffman EP, Partridge T, Takeda S. (2012) Bodywide skipping of exons 45-55 in dystrophic mdx52 mice by systemic antisense delivery. *Proc Natl Acad Sci U S A* 109:13763-13768.

Arechavala-Gomeza V, Khoo B, Aartsma-Rus A. (2014) Splicing modulation therapy in the treatment of genetic diseases. *Appl Clin Genet* 7:245-252.

Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, Bernardi F, Arruda VR, Pinotti M. (2014) An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice. *J Thromb Haemost* 12:177-185.

Baralle D, Baralle M. (2005) Splicing in action: assessing disease causing sequence changes. *J Med Genet* 42:737-748.

Baralle D, Lucassen A, Buratti E. (2009) Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep* 10:810-816.

Bauman J, Jearawiriyapaisarn N, Kole R. (2009) Therapeutic potential of splice-switching oligonucleotides. *Oligonucleotides* 19:1-13.

Bauman JA, Kole R. (2011) Modulation of RNA splicing as a potential treatment for cancer. *Bioeng Bugs* 2:125-128.

Bauman JA, Li SD, Yang A, Huang L, Kole R. (2010) Anti-tumor activity of splice-switching oligonucleotides. *Nucleic Acids Res* 38:8348-8356.

Bennett CF, Swayze EE. (2010) RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu Rev Pharmacol Toxicol* 50:259-293.

Berget SM. (1995) Exon recognition in vertebrate splicing. *J Biol Chem* 270:2411-2414.

Berget SM, Moore C, Sharp PA. (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* 74:3171-3175.

Betts C, Saleh AF, Arzumanov AA, Hammond SM, Godfrey C, Coursindel T, Gait MJ, Wood MJ. (2012) Pip6-PMO, A New Generation of Peptide-oligonucleotide Conjugates With Improved Cardiac Exon Skipping Activity for DMD Treatment. *Mol Ther Nucleic Acids* 1:e38.

Boustany RM. (2013) Lysosomal storage diseases--the horizon expands. *Nat Rev Neurol* 9:583-598.

Brusius-Facchin AC, Schwartz IV, Zimmer C, Ribeiro MG, Acosta AX, Horovitz D, Monlleó IL, Fontes MI, Fett-Conte A, Sobrinho RP, Duarte AR, Boy R, Mabe P, Ascurra M, de Michelena M, Tylee KL, Besley GT, Garreton MC, Giugliani R, Leistner-Segal S. (2014) Mucopolysaccharidosis type II: identification of 30 novel mutations among Latin American patients. *Mol Genet Metab* 111:133-138.

Buratti E, Baralle D. (2010) Novel roles of U1 snRNP in alternative splicing regulation. *RNA Biol* 7:412-419.

Buratti E, Chivers M, Královicová J, Romano M, Baralle M, Krainer AR, Vorechovsky I. (2007) Aberrant 5' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res* 35:4250-4263.

Burdick AD, Sciabola S, Mantena SR, Hollingshead BD, Stanton R, Warneke JA, Zeng M, Martsen E, Medvedev A, Makarov SS, Reed LA, Davis JW, Whiteley LO. (2014) Sequence motifs associated with hepatotoxicity of locked nucleic acid--modified antisense oligonucleotides. *Nucleic Acids Res* 42:4882-4891.

Burge C, Karlin S. (1997) Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 268:78-94.

Busch A, Hertel KJ. (2012) Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip Rev RNA* 3:1-12.

Cartegni L, Chew SL, Krainer AR. (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 3:285-298.

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

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. (2003) ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31:3568-3571.

Chan JH, Lim S, Wong WS. (2006) Antisense oligonucleotides: from design to therapeutic application. *Clin Exp Pharmacol Physiol* 33:533-540.

Chaudhury A, Chander P, Howe PH. (2010) Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. *RNA* 16:1449-1462.

Chen L, Tovar-Corona JM, Urrutia AO. (2012) Alternative splicing: a potential source of functional innovation in the eukaryotic genome. *Int J Evol Biol* 2012:596274.

Chen M, Manley JL. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol* 10:741-754.

Chow LT, Gelinas RE, Broker TR, Roberts RJ. (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12:1-8.

Cirak S, Arechavala-Gomeza V, Guglieri M, Feng L, Torelli S, Anthony K, Abbs S, Garralda ME, Bourke J, Wells DJ, Dickson G, Wood MJ, Wilton SD, Straub V, Kole R, Shrewsbury SB, Sewry C, Morgan JE, Bushby K, Muntoni F. (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.

Cline A, Gao N, Flanagan-Steet H, Sharma V, Rosa S, Sonon R, Azadi P, Sadler KC, Freeze HH, Lehrman MA, Steet R. (2012) A zebrafish model of PMM2-CDG reveals altered neurogenesis and a substrate-accumulation mechanism for N-linked glycosylation deficiency. *Mol Biol Cell* 23:4175-4187.

Consortium IHGS. (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431:931-945.

Cooper GM, Hausman RE. (2013) *The Cell: A Molecular Approach*. Sinauer Associates Sunderland, Massachusetts U.S.A.

Cooper TA, Wan L, Dreyfuss G. (2009) RNA and disease. *Cell* 136:777-793.

Coutinho MF, Lacerda L, Alves S. (2012) Glycosaminoglycan storage disorders: a review. *Biochem Res Int* 2012:471325.

Coutinho MF, Lacerda L, Prata MJ, Ribeiro H, Lopes L, Ferreira C, Alves S. (2008) Molecular characterization of Portuguese patients with mucopolysaccharidosis IIIC: two novel mutations in the HGSNAT gene. *Clin Genet* 74:194-195.

Cox T. (2012) Current treatments. In Mehta A, Winchester B (Eds) *Lysosomal storage disorders: a practical guide*, Wiley-Blackwell, UK, pp. 153-165.

Cox TM. (2015) Innovative treatments for lysosomal diseases. *Best Pract Res Clin Endocrinol Metab* 29:275-311.

Crick F. (1958) On protein synthesis. *Symp Soc Exp Biol* 12:138-163.

Crick F. (1970) Central dogma of molecular biology. *Nature* 227:561-563.

Crooke ST, Geary RS. (2013) Clinical pharmacological properties of mipomersen (Kynamro), a second generation antisense inhibitor of apolipoprotein B. *Br J Clin Pharmacol* 76:269-276.

Cáceres JF, Kornblihtt AR. (2002) Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet* 18:186-193.

Dal Mas A, Fortugno P, Donadon I, Levati L, Castiglia D, Pagani F. (2015a) Exon-Specific U1s Correct SPINK5 Exon 11 Skipping Caused by a Synonymous Substitution that Affects a Bifunctional Splicing Regulatory Element. *Hum Mutat* 36:504-512.

Dal Mas A, Rogalska ME, Bussani E, Pagani F. (2015b) Improvement of SMN2 pre-mRNA processing mediated by exon-specific U1 small nuclear RNA. *Am J Hum Genet* 96:93-103.

Dauksaite V, Akusjärvi G. (2002) Human splicing factor ASF/SF2 encodes for a repressor domain required for its inhibitory activity on pre-mRNA splicing. *J Biol Chem* 277:12579-12586.

De Conti L, Baralle M, Buratti E. (2013) Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip Rev RNA* 4:49-60.

Dembowski JA, An P, Scoulos-Hanson M, Yeo G, Han J, Fu XD, Grabowski PJ. (2012) Alternative Splicing of a Novel Inducible Exon Diversifies the CASK Guanylate Kinase Domain. *J Nucleic Acids* 2012:816237.

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, Bozzoni I. (2006) Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model. *Proc Natl Acad Sci U S A* 103:3758-3763.

Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. (2009) Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 37:e67.

Desviat L, Pérez B, Ugarte M. (2012) Minigenes to Confirm Exon Skipping Mutations. In Aartsma-Rus A (Ed) *Exon Skipping*, Humana Press, New York, pp. 37-47.

DeVos SL, Miller TM. (2013) Antisense oligonucleotides: treating neurodegeneration at the level of RNA. *Neurotherapeutics* 10:486-497.

Dhir A, Buratti E. (2010) Alternative splicing: role of pseudoexons in human disease and potential therapeutic strategies. *FEBS J* 277:841-855.

Disset A, Bourgeois CF, Benmalek N, Claustres M, Stevenin J, Tuffery-Giraud S. (2006) An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements. *Hum Mol Genet* 15:999-1013.

Dominski Z, Kole R. (1993) Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proc Natl Acad Sci U S A* 90:8673-8677.

Douglas AG, Wood MJ. (2011) RNA splicing: disease and therapy. *Brief Funct Genomics* 10:151-164.

Douglas AG, Wood MJ. (2013) Splicing therapy for neuromuscular disease. *Mol Cell Neurosci* 56:169-185.

Du L, Gatti RA. (2009) Progress toward therapy with antisense-mediated splicing modulation. *Curr Opin Mol Ther* 11:116-123.

Du L, Pollard JM, Gatti RA. (2007) Correction of prototypic ATM splicing mutations and aberrant ATM function with antisense morpholino oligonucleotides. *Proc Natl Acad Sci U S A* 104:6007-6012.

Eisen JS, Smith JC. (2008) Controlling morpholino experiments: don't stop making antisense. *Development* 135:1735-1743.

Erickson MA, Niehoff ML, Farr SA, Morley JE, Dillman LA, Lynch KM, Banks WA. (2012) Peripheral administration of antisense oligonucleotides targeting the amyloid-β protein precursor reverses AβPP and LRP-1 overexpression in the aged SAMP8 mouse brain. *J Alzheimers Dis* 28:951-960.

Erkelenz S, Mueller WF, Evans MS, Busch A, Schöneweis K, Hertel KJ, Schaal H. (2013) Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. *RNA* 19:96-102.

Evers MM, Pepers BA, van Deutekom JC, Mulders SA, den Dunnen JT, Aartsma-Rus A, van Ommen GJ, van Roon-Mom WM. (2011) Targeting several CAG expansion diseases by a single antisense oligonucleotide. *PLoS One* 6:e24308.

Evers MM, Toonen LJ, van Roon-Mom WM. (2015) Antisense oligonucleotides in therapy for neurodegenerative disorders. *Adv Drug Deliv Rev* 87:90-103.

Fairbrother WG, Yeh RF, Sharp PA, Burge CB. (2002) Predictive identification of exonic splicing enhancers in human genes. *Science* 297:1007-1013.

Faustino NA, Cooper TA. (2003) Pre-mRNA splicing and human disease. *Genes Dev* 17:419-437.

Ferguson DP, Schmitt EE, Lightfoot JT. (2013) Vivo-morpholinos induced transient knockdown of physical activity related proteins. *PLoS One* 8:e61472.

Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, Rogalska ME, Bernardi F, Pagani F. (2012) An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. *Hum Mol Genet* 21:2389-2398.

Filocamo M, Morrone A. (2011) Lysosomal storage disorders: molecular basis and laboratory testing. *Hum Genomics* 5:156-169.

Fuller M, Meikle P, Hopwood J. (2006) Epidemiology of lysosomal storage diseases: an overview. In Mehta A, Beck M, Sunder-Plassmann G (Eds) *Fabry disease: Perspectives from 5 years of FOS*, Oxford PharmaGenesis, Oxford, pp. 9-20.

Futerman AH, van Meer G. (2004) The cell biology of lysosomal storage disorders. *Nat Rev Mol Cell Biol* 5:554-565.

Gagnon KT, Pendergraff HM, Deleavey GF, Swayze EE, Potier P, Randolph J, Roesch EB, Chattopadhyaya J, Damha MJ, Bennett CF, Montaillier C, Lemaitre M, Corey DR. (2010) Allele-selective inhibition of mutant huntingtin expression with antisense oligonucleotides targeting the expanded CAG repeat. *Biochemistry* 49:10166-10178.

Gallego-Villar L, Viecelli HM, Pérez B, Harding CO, Ugarte M, Thöny B, Desviat LR. (2014) A sensitive assay system to test antisense oligonucleotides for splice suppression therapy in the mouse liver. *Mol Ther Nucleic Acids* 3:e193.

Gamazon ER, Stranger BE. (2014) Genomics of alternative splicing: evolution, development and pathophysiology. *Hum Genet* 133:679-687.

Geary RS, Norris D, Yu R, Bennett CF. (2015) Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Adv Drug Deliv Rev* 87:46-51.

Geller BL, Deere J, Tilley L, Iversen PL. (2005) Antisense phosphorodiamidate morpholino oligomer inhibits viability of Escherichia coli in pure culture and in mouse peritonitis. *J Antimicrob Chemother* 55:983-988.

Glaus E, Schmid F, Da Costa R, Berger W, Neidhardt J. (2011) Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells. *Mol Ther* 19:936-941.

Goemans NM, Tulinius M, van den Akker JT, Burm BE, Ekhart PF, Heuvelmans N, Holling T, Janson AA, Platenburg GJ, Sipkens JA, Sitsen JM, Aartsma-Rus A, van Ommen GJ, Buyse G, Darin N, Verschuuren JJ, Campion GV, de Kimpe SJ, van Deutekom JC. (2011) Systemic administration of PRO051 in Duchenne's muscular dystrophy. *N Engl J Med* 364:1513-1522.

Gonçalves V, Theisen P, Antunes O, Medeira A, Ramos JS, Jordan P, Isidro G. (2009) A missense mutation in the APC tumor suppressor gene disrupts an ASF/SF2 splicing enhancer motif and causes pathogenic skipping of exon 14. *Mutat Res* 662:33-36.

Goren A, Ram O, Amit M, Keren H, Lev-Maor G, Vig I, Pupko T, Ast G. (2006) Comparative analysis identifies exonic splicing regulatory sequences--The complex definition of enhancers and silencers. *Mol Cell* 22:769-781.

Gottlieb S. (2003) The Splice of Life *Horizon Symposia: Understanding the RNAissance*, pp. 1-4.

Goyenvalle A, Davies KE. (2011) Challenges to oligonucleotides-based therapeutics for Duchenne muscular dystrophy. *Skelet Muscle* 1:8.

Goyenvalle A, Griffith G, Babbs A, El Andaloussi S, Ezzat K, Avril A, Dugovic B, Chaussenot R, Ferry A, Voit T, Amthor H, Bühr C, Schürch S, Wood MJ, Davies KE, Vaillend C, Leumann C, Garcia L. (2015) Functional correction in mouse models of muscular dystrophy using exon-skipping tricyclo-DNA oligomers. *Nat Med* 21:270-275.

Graziewicz MA, Tarrant TK, Buckley B, Roberts J, Fulton L, Hansen H, Ørum H, Kole R, Sazani P. (2008) An endogenous TNF-alpha antagonist induced by splice-switching oligonucleotides reduces inflammation in hepatitis and arthritis mouse models. *Mol Ther* 16:1316-1322.

Guiro J, O'Reilly D. (2015) Insights into the U1 small nuclear ribonucleoprotein complex superfamily. *Wiley Interdiscip Rev RNA* 6:79-92.

Gupta N, Fisker N, Asselin MC, Lindholm M, Rosenbohm C, Ørum H, Elmén J, Seidah NG, Straarup EM. (2010) A locked nucleic acid antisense oligonucleotide (LNA) silences PCSK9 and enhances LDLR expression in vitro and in vivo. *PLoS One* 5:e10682.

Hammond SM, Wood MJ. (2011) Genetic therapies for RNA mis-splicing diseases. *Trends Genet* 27:196-205.

Hartmann L, Neveling K, Borkens S, Schneider H, Freund M, Grassman E, Theiss S, Wawer A, Burdach S, Auerbach AD, Schindler D, Hanenberg H, Schaal H. (2010) Correct mRNA processing at a mutant TT splice donor in FANCC

ameliorates the clinical phenotype in patients and is enhanced by delivery of suppressor U1 snRNAs. *Am J Hum Genet* 87:480-493.

Hartmann L, Theiss S, Niederacher D, Schaal H. (2008) Diagnostics of pathogenic splicing mutations: does bioinformatics cover all bases? *Front Biosci* 13:3252-3272.

Havens MA, Duelli DM, Hastings ML. (2013) Targeting RNA splicing for disease therapy. *Wiley Interdiscip Rev RNA* 4:247-266.

Hertel KJ. (2008) Combinatorial control of exon recognition. *J Biol Chem* 283:1211-1215.

Hopwood J. (2012) Genetics of lysosomal storage disorders and counselling. In Mehta A, Winchester B (Eds) *Lysosomal storage disorders: a practical guide*, Wiley-Blackwell, UK, pp. 29-36.

House AE, Lynch KW. (2008) Regulation of alternative splicing: more than just the ABCs. *J Biol Chem* 283:1217-1221.

Hrebícek M, Mrázová L, Seyrantepe V, Durand S, Roslin NM, Nosková L, Hartmannová H, Ivánek R, Cízkova A, Poupetová H, Sikora J, Urinovská J, Stranecký V, Zeman J, Lepage P, Roquis D, Verner A, Ausseil J, Beesley CE, Maire I, Poorthuis BJ, van de Kamp J, van Diggelen OP, Wevers RA, Hudson TJ, Fujiwara TM, Majewski J, Morgan K, Kmoch S, Pshezhetsky AV. (2006) Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *Am J Hum Genet* 79:807-819.

Hu J, Matsui M, Gagnon K, Schwartz J, Gabillet S, Arar K, Wu J, Bezprozvanny I, Corey D. (2009) Inhibiting Expression of Mutant Huntingtin and Ataxin-3 by Targeting Expanded CAG Repeat RNAs *Nat Biotechnol* 27:478-484.

Hua Y, Sahashi K, Hung G, Rigo F, Passini MA, Bennett CF, Krainer AR. (2010) Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev* 24:1634-1644.

Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. (2008) Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *Am J Hum Genet* 82:834-848.

Jelen N, Ule J, Zivin M, Darnell RB. (2007) Evolution of Nova-dependent splicing regulation in the brain. *PLoS Genet* 3:1838-1847.

Jensen CJ, Oldfield BJ, Rubio JP. (2009) Splicing, cis genetic variation and disease. *Biochem Soc Trans* 37:1311-1315.

Joensuu T, Lehesjoki AE, Kopra O. (2008) Molecular background of EPM1-Unverricht-Lundborg disease. *Epilepsia* 49:557-563.

Kagitani-Shimono K, Imai K, Okamoto N, Ono J, Okada S. (2002) Unverricht-Lundborg disease with cystatin B gene abnormalities. *Pediatr Neurol* 26:55-60.

Kalbfuss B, Mabon SA, Misteli T. (2001) Correction of alternative splicing of tau in frontotemporal dementia and parkinsonism linked to chromosome 17. *J Biol Chem* 276:42986-42993.

Kalsotra A, Cooper TA. (2011) Functional consequences of developmentally regulated alternative splicing. *Nat Rev Genet* 12:715-729.

Kandels-Lewis S, Séraphin B. (1993) Involvement of U6 snRNA in 5' splice site selection. *Science* 262:2035-2039.

Kaur H, Babu BR, Maiti S. (2007) Perspectives on chemistry and therapeutic applications of Locked Nucleic Acid (LNA). *Chem Rev* 107:4672-4697.

Kelemen O, Convertini P, Zhang Z, Wen Y, Shen M, Falaleeva M, Stamm S. (2013) Function of alternative splicing. *Gene* 514:1-30.

Keren H, Lev-Maor G, Ast G. (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* 11:345-355.

Kim E, Goren A, Ast G. (2008) Alternative splicing: current perspectives. *Bioessays* 30:38-47.

Kim E, Magen A, Ast G. (2007) Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res* 35:125-131.

Kole R, Krainer AR, Altman S. (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* 11:125-140.

Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ. (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol* 14:153-165.

Kornfeld S, Sly W. (2001) I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localization. In Scriver C, Beaudet A, Sly W, Valle D (Eds) *The metabolic and molecular bases of inherited disease*, McGraw-Hill, New York, pp. 3469–3505.

Kudo M, Brem MS, Canfield WM. (2006) Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha / beta -subunits precursor gene. *Am J Hum Genet* 78:451-463.

Kumar VB, Farr SA, Flood JF, Kamlesh V, Franko M, Banks WA, Morley JE. (2000) Site-directed antisense oligonucleotide decreases the expression of amyloid precursor protein and reverses deficits in learning and memory in aged SAMP8 mice. *Peptides* 21:1769-1775.

Kurreck J. (2003) Antisense technologies. Improvement through novel chemical modifications. *Eur J Biochem* 270:1628-1644.

Lachmann R. (2010) Treatments for lysosomal storage disorders. *Biochem Soc Trans* 38:1465-1468.

Ladd AN, Cooper TA. (2002) Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol* 3:reviews0008.

Lee JJ, Yokota T. (2013) Antisense therapy in neurology. *J Pers Med* 3:144-176.

Leger AJ, Mosquea LM, Clayton NP, Wu IH, Weeden T, Nelson CA, Phillips L, Roberts E, Piepenhagen PA, Cheng SH, Wentworth BM. (2013) Systemic delivery of a Peptide-linked morpholino oligonucleotide neutralizes mutant RNA toxicity in a mouse model of myotonic dystrophy. *Nucleic Acid Ther* 23:109-117.

Lehesjoki A, Gardiner M. (2012) Progressive myoclonus epilepsy: Unverricht-Lundborg disease and Neuronal ceroid lipofuscinoses. In Noebels J, Avoli M, Rogawski M, Olsen R, Delgado-Escueta A (Eds) *Jasper's Basic Mechanisms of the Epilepsies*, National Center for Biotechnology Information, Bethesda MD, USA.

Lehtinen MK, Tegelberg S, Schipper H, Su H, Zukor H, Manninen O, Kopra O, Joensuu T, Hakala P, Bonni A, Lehesjoki AE. (2009) Cystatin B deficiency sensitizes neurons to oxidative stress in progressive myoclonus epilepsy, EPM1. *J Neurosci* 29:5910-5915.

Lentz JJ, Jodelka FM, Hinrich AJ, McCaffrey KE, Farris HE, Spalitta MJ, Bazan NG, Duelli DM, Rigo F, Hastings ML. (2013) Rescue of hearing and vestibular function by antisense oligonucleotides in a mouse model of human deafness. *Nat Med* 19:345-350.

Lerner MR, Boyle JA, Mount SM, Wolin SL, Steitz JA. (1980) Are snRNPs involved in splicing? *Nature* 283:220-224.

Lesser CF, Guthrie C. (1993) Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* 262:1982-1988.

Lewandowska MA. (2013) The missing puzzle piece: splicing mutations. *Int J Clin Exp Pathol* 6:2675-2682.

Li Q, Lee JA, Black DL. (2007) Neuronal regulation of alternative pre-mRNA splicing. *Nat Rev Neurosci* 8:819-831.

Li YF, Morcos PA. (2008) Design and synthesis of dendritic molecular transporter that achieves efficient in vivo delivery of morpholino antisense oligo. *Bioconjug Chem* 19:1464-1470.

Lu QL, Mann CJ, Lou F, Bou-Gharios G, Morris GE, Xue SA, Fletcher S, Partridge TA, Wilton SD. (2003) Functional amounts of dystrophin produced by skipping the mutated exon in the mdx dystrophic mouse. *Nat Med* 9:1009-1014.

Lund E, Dahlberg JE. (1984) True genes for human U1 small nuclear RNA. Copy number, polymorphism, and methylation. *J Biol Chem* 259:2013-2021.

López-Bigas N, Audit B, Ouzounis C, Parra G, Guigó R. (2005) Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett* 579:1900-1903.

Magen I, Hornstein E. (2014) Oligonucleotide-based therapy for neurodegenerative diseases. *Brain Res* 1584:116-128.

Mahmood F, Fu S, Cooke J, Wilson SW, Cooper JD, Russell C. (2013) A zebrafish model of CLN2 disease is deficient in tripeptidyl peptidase 1 and displays progressive neurodegeneration accompanied by a reduction in proliferation. *Brain* 136:1488-1507.

Mansoor M, Melendez AJ. (2008) Advances in antisense oligonucleotide development for target identification, validation, and as novel therapeutics. *Gene Regul Syst Bio* 2:275-295.

Matera AG, Wang Z. (2014) A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* 15:108-121.

Mattioli C, Pianigiani G, De Rocco D, Bianco AM, Cappelli E, Savoia A, Pagani F. (2014) Unusual splice site mutations disrupt FANCA exon 8 definition. *Biochim Biophys Acta* 1842:1052-1058.

McManus CJ, Graveley BR. (2011) RNA structure and the mechanisms of alternative splicing. *Curr Opin Genet Dev* 21:373-379.

Meikle PJ, Hopwood JJ, Clague AE, Carey WF. (1999) Prevalence of lysosomal storage disorders. *JAMA* 281:249-254.

Mendell JR, Rodino-Klapac LR, Sahenk Z, Roush K, Bird L, Lowes LP, Alfano L, Gomez AM, Lewis S, Kota J, Malik V, Shontz K, Walker CM, Flanigan KM, Corridore M, Kean JR, Allen HD, Shilling C, Melia KR, Sazani P, Saoud JB, Kaye EM, Group ES. (2013) Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol* 74:637-647.

Merwick A, O'Brien M, Delanty N. (2012) Complex single gene disorders and epilepsy. *Epilepsia* 53 Suppl 4:81-91.

Miller TM, Pestronk A, David W, Rothstein J, Simpson E, Appel SH, Andres PL, Mahoney K, Allred P, Alexander K, Ostrow LW, Schoenfeld D, Macklin EA, Norris DA, Manousakis G, Crisp M, Smith R, Bennett CF, Bishop KM, Cudkowicz ME. (2013) An antisense oligonucleotide against SOD1 delivered

intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study. *Lancet Neurol* 12:435-442.

Monia BP, Lesnik EA, Gonzalez C, Lima WF, McGee D, Guinosso CJ, Kawasaki AM, Cook PD, Freier SM. (1993) Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression. *J Biol Chem* 268:14514-14522.

Morcos PA. (2001) Achieving efficient delivery of morpholino oligos in cultured cells. *Genesis* 30:94-102.

Morcos PA, Li Y, Jiang S. (2008) Vivo-Morpholinos: a non-peptide transporter delivers Morpholinos into a wide array of mouse tissues. *Biotechniques* 45:613-614, 616, 618 passim.

Moro E, Tomanin R, Friso A, Modena N, Tiso N, Scarpa M, Argenton F. (2010) A novel functional role of iduronate-2-sulfatase in zebrafish early development. *Matrix Biol* 29:43-50.

Moulton HM, Hase MC, Smith KM, Iversen PL. (2003) HIV Tat peptide enhances cellular delivery of antisense morpholino oligomers. *Antisense Nucleic Acid Drug Dev* 13:31-43.

Moulton HM, Moulton JD. (2004) Arginine-rich cell-penetrating peptides with uncharged antisense oligomers. *Drug Discov Today* 9:870.

Moulton JD, Jiang S. (2009) Gene knockdowns in adult animals: PPMOs and vivo-morpholinos. *Molecules* 14:1304-1323.

Mount SM, Pettersson I, Hinterberger M, Karmas A, Steitz JA. (1983) The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell* 33:509-518.

Muntoni F, Torelli S, Ferlini A. (2003) Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *Lancet Neurol* 2:731-740.

Nakamura K, Du L, Tunuguntla R, Fike F, Cavalieri S, Morio T, Mizutani S, Brusco A, Gatti RA. (2012) Functional characterization and targeted correction of ATM mutations identified in Japanese patients with ataxia-telangiectasia. *Hum Mutat* 33:198-208.

Neufeld E, Muenzer J. (2001) The Mucopolysaccharidoses. In Scriver C, Beaudet A, Sly W, Valle D (Eds) *The metabolic and molecular bases of inherited disease*, McGraw-Hill, New York, pp. 3421-3452.

Nilsen TW, Graveley BR. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463:457-463.

Obika S, Nanbu D, Hari Y, Morio K, In Y, Ishida T, Imanishi T. (1997) Synthesis of 2'-O,4'-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C-3, -endo sugar puckering. *Tetrahedron Lett* 38:8735-8738.

Osorio FG, Navarro CL, Cadiñanos J, López-Mejía IC, Quirós PM, Bartoli C, Rivera J, Tazi J, Guzmán G, Varela I, Depetris D, de Carlos F, Cobo J, Andrés V, De Sandre-Giovannoli A, Freije JM, Lévy N, López-Otín C. (2011) Splicing-directed therapy in a new mouse model of human accelerated aging. *Sci Transl Med* 3:106ra107.

Pajares MJ, Ezponda T, Catena R, Calvo A, Pio R, Montuenga LM. (2007) Alternative splicing: an emerging topic in molecular and clinical oncology. *Lancet Oncol* 8:349-357.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* 40:1413-1415.

Parkinson-Lawrence EJ, Shandala T, Prodoehl M, Plew R, Borlace GN, Brooks DA. (2010) Lysosomal storage disease: revealing lysosomal function and physiology. *Physiology (Bethesda)* 25:102-115.

Partridge M, Vincent A, Matthews P, Puma J, Stein D, Summerton J. (1996) A simple method for delivering morpholino antisense oligos into the cytoplasm of cells. *Antisense Nucleic Acid Drug Dev* 6:169-175.

Pastores GM. (2010) Therapeutic approaches for lysosomal storage diseases. *Ther Adv Endocrinol Metab* 1:177-188.

Paterson BM, Roberts BE, Kuff EL. (1977) Structural gene identification and mapping by DNA-mRNA hybrid-arrested cell-free translation. *Proc Natl Acad Sci U S A* 74:4370-4374.

Peacey E, Rodriguez L, Liu Y, Wolfe MS. (2012) Targeting a pre-mRNA structure with bipartite antisense molecules modulates tau alternative splicing. *Nucleic Acids Res* 40:9836-9849.

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

Pinotti M, Bernardi F, Dal Mas A, Pagani F. (2011) RNA-based therapeutic approaches for coagulation factor deficiencies. *J Thromb Haemost* 9:2143-2152.

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

Pinto E, Freitas J, Duarte AJ, Ribeiro I, Ribeiro D, Lima JL, Chaves J, Amaral O. (2012) Unverricht-Lundborg disease: homozygosity for a new splicing mutation in the cystatin B gene. *Epilepsy Res* 99:187-190.

Pinto R, Caseiro C, Lemos M, Lopes L, Fontes A, Ribeiro H, Pinto E, Silva E, Rocha S, Marcão A, Ribeiro I, Lacerda L, Ribeiro G, Amaral O, Sá Miranda MC. (2004) Prevalence of lysosomal storage diseases in Portugal. *Eur J Hum Genet* 12:87-92.

Platt FM, Boland B, van der Spoel AC. (2012) The cell biology of disease: lysosomal storage disorders: the cellular impact of lysosomal dysfunction. *J Cell Biol* 199:723-734.

Platt FM, Jeyakumar M. (2008) Substrate reduction therapy. *Acta Paediatr Suppl* 97:88-93.

Polajnar M, Ceru S, Kopitar-Jerala N, Zerovnik E. (2012) Human stefin B normal and patho-physiological role: molecular and cellular aspects of amyloid-type aggregation of certain EPM1 mutants. *Front Mol Neurosci* 5:88.

Poorthuis BJ, Wevers RA, Kleijer WJ, Groener JE, de Jong JG, van Weely S, Niezen-Koning KE, van Diggelen OP. (1999) The frequency of lysosomal storage diseases in The Netherlands. *Hum Genet* 105:151-156.

Poupetová H, Ledvinová J, Berná L, Dvoráková L, Kozich V, Elleder M. (2010) The birth prevalence of lysosomal storage disorders in the Czech Republic: comparison with data in different populations. *J Inherit Metab Dis* 33:387-396.

Pérez B, Gutiérrez-Solana LG, Verdú A, Merinero B, Yuste-Checa P, Ruiz-Sala P, Calvo R, Jalan A, Marín LL, Campos O, Ruiz M, San Miguel M, Vázquez M, Castro M, Ferrer I, Navarrete R, Desviat LR, Lapunzina P, Ugarte M, Pérez-Cerdá C. (2013) Clinical, biochemical, and molecular studies in pyridoxine-dependent epilepsy. Antisense therapy as possible new therapeutic option. *Epilepsia* 54:239-248.

Pérez B, Rodríguez-Pascau L, Vilageliu L, Grinberg D, Ugarte M, Desviat LR. (2010) Present and future of antisense therapy for splicing modulation in inherited metabolic disease. *J Inherit Metab Dis* 33:397-403.

Pérez B, Ugarte M, Desviat L. (2012) RNA-based therapies for inherited metabolic diseases. In Erdmann V, Barciszewski J (Eds) *From Nucleic Acids Sequences to Molecular Medicine* Springer-Verlag, Berlin Heidelberg, pp. 358-370.

Pérez B, Vilageliu L, Grinberg D, Desviat LR. (2014) Antisense mediated splicing modulation for inherited metabolic diseases: challenges for delivery. *Nucleic Acid Ther* 24:48-56.

Raas-Rothschild A, Cormier-Daire V, Bao M, Genin E, Salomon R, Brewer K, Zeigler M, Mandel H, Toth S, Roe B, Munnich A, Canfield WM. (2000) Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). *J Clin Invest* 105:673-681.

Raas-Rothschild A, Pohl S, Braulke T. (2012) Multiple enzyme deficiencies. In Mehta A, Winchester B (Eds) *Lysosomal storage disorders: a practical guide*, Wiley-Blackwell, UK, pp. 121-126.

Ramachandran N, Girard JM, Turnbull J, Minassian BA. (2009) The autosomal recessively inherited progressive myoclonus epilepsies and their genes. *Epilepsia* 50 Suppl 5:29-36.

Raponi M, Baralle D. (2008) Can donor splice site recognition occur without the involvement of U1 snRNP? *Biochem Soc Trans* 36:548-550.

Reed R. (1996) Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr Opin Genet Dev* 6:215-220.

Reese MG, Eeckman FH, Kulp D, Haussler D. (1997) Improved splice site detection in Genie. *J Comput Biol* 4:311-323.

Regis S, Corsolini F, Grossi S, Tappino B, Cooper DN, Filocamo M. (2013) Restoration of the normal splicing pattern of the PLP1 gene by means of an antisense oligonucleotide directed against an exonic mutation. *PLoS One* 8:e73633.

Rigo F, Hua Y, Krainer AR, Bennett CF. (2012) Antisense-based therapy for the treatment of spinal muscular atrophy. *J Cell Biol* 199:21-25.

Rigo F, Seth P, Bennett C. (2014) Antisense oligonucleotide-based therapies for diseases caused by pre-mRNA processing defects. In Yeo G (Ed) *Systems Biology of RNA Binding Proteins, Advances in Experimental Medicine and Biology*, Springer, New York, pp. 303-352.

Risso G, Pelisch F, Quaglino A, Pozzi B, Srebrow A. (2012) Regulating the regulators: serine/arginine-rich proteins under scrutiny. *IUBMB Life* 64:809-816.

Roberts J, Palma E, Sazani P, Ørum H, Cho M, Kole R. (2006) Efficient and persistent splice switching by systemically delivered LNA oligonucleotides in mice. *Mol Ther* 14:471-475.

Roca X, Krainer AR, Eperon IC. (2013) Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev* 27:129-144.

Roca X, Olson AJ, Rao AR, Enerly E, Kristensen VN, Børresen-Dale AL, Andresen BS, Krainer AR, Sachidanandam R. (2008) Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics. *Genome Res* 18:77-87.

Rodríguez-Pascau L, Coll MJ, Vilageliu L, Grinberg D. (2009) Antisense oligonucleotide treatment for a pseudoexon-generating mutation in the NPC1 gene causing Niemann-Pick type C disease. *Hum Mutat* 30:E993-E1001.

Roehr B. (1998) Fomivirsen approved for CMV retinitis. *J Int Assoc Physicians AIDS Care* 4:14-16.

Rogers J, Wall R. (1980) A mechanism for RNA splicing. *Proc Natl Acad Sci U S A* 77:1877-1879.

Roy B, Haupt LM, Griffiths LR. (2013) Review: Alternative Splicing (AS) of Genes As An Approach for Generating Protein Complexity. *Curr Genomics* 14:182-194.

Rymond B. (2007) Targeting the spliceosome. *Nat Chem Biol* 3:533-535.

Sabatini D, Adesnik M. (2001) The biogenesis of membranes and organelles In Scriver C, Beaudet A, Sly W, Valle D, Childs B, Kinzler K, Vogelstein B (Eds) *The metabolic and molecular bases of inherited disease*, McGraw-Hill, New York, pp. 433-520.

Sakharkar MK, Perumal BS, Sakharkar KR, Kangueane P. (2005) An analysis on gene architecture in human and mouse genomes. *In Silico Biol* 5:347-365.

Schaub MC, Lopez SR, Caputi M. (2007) Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. *J Biol Chem* 282:13617-13626.

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

Schmid F, Hiller T, Korner G, Glaus E, Berger W, Neidhardt J. (2013) A gene therapeutic approach to correct splice defects with modified U1 and U6 snRNPs. *Hum Gene Ther* 24:97-104.

Schmucker D, Clemens JC, Shu H, Worby CA, Xiao J, Muda M, Dixon JE, Zipursky SL. (2000) Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 101:671-684.

Scott HS, Litjens T, Hopwood JJ, Morris CP. (1992) A common mutation for mucopolysaccharidosis type I associated with a severe Hurler syndrome phenotype. *Hum Mutat* 1:103-108.

Shepard PJ, Hertel KJ. (2009) The SR protein family. *Genome Biol* 10:242.

Singer RH, Green MR. (1997) Compartmentalization of eukaryotic gene expression: causes and effects. *Cell* 91:291-294.

Singh RK, Cooper TA. (2012) Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med* 18:472-482.

Singh S, Nielsen P, Koshkin A, Wengel J. (1998) LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition. *Chem Commun* 4:455-456.

Siva K, Covello G, Denti MA. (2014) Exon-skipping antisense oligonucleotides to correct missplicing in neurogenetic diseases. *Nucleic Acid Ther* 24:69-86.

Spraggon L, Cartegni L. (2013) U1 snRNP-Dependent Suppression of Polyadenylation: Physiological Role and Therapeutic Opportunities in Cancer. *Int J Cell Biol* 2013:846510.

Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z, Toiber D, Thanaraj TA, Soreq H. (2005) Function of alternative splicing. *Gene* 344:1-20.

Sterner DA, Carlo T, Berget SM. (1996) Architectural limits on split genes. *Proc Natl Acad Sci U S A* 93:15081-15085.

Sugnet CW, Kent WJ, Ares M, Haussler D. (2004) Transcriptome and genome conservation of alternative splicing events in humans and mice. *Pac Symp Biocomput*:66-77.

Summerton J. (1999) Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochim Biophys Acta* 1489:141-158.

Summerton J, Weller D. (1993) Uncharged Morpholino-based polymers having phosphorous containing chiral intersubunit linkages. In Organization WIP (Ed), US Patent 5, pp. 5,185,444.

Summerton JE. (2005) Endo-Porter: a novel reagent for safe, effective delivery of substances into cells. *Ann N Y Acad Sci* 1058:62-75.

Summerton JE. (2007) Morpholino, siRNA, and S-DNA compared: impact of structure and mechanism of action on off-target effects and sequence specificity. *Curr Top Med Chem* 7:651-660.

Susani L, Pangrazio A, Sobacchi C, Taranta A, Mortier G, Savarirayan R, Villa A, Orchard P, Vezzoni P, Albertini A, Frattini A, Pagani F. (2004) TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA. *Hum Mutat* 24:225-235.

Swayze EE, Siwkowski AM, Wancewicz EV, Migawa MT, Wyrzykiewicz TK, Hung G, Monia BP, Bennett CF. (2007) Antisense oligonucleotides containing locked nucleic acid improve potency but cause significant hepatotoxicity in animals. *Nucleic Acids Res* 35:687-700.

Sánchez-Alcudia R, Pérez B, Pérez-Cerdá C, Ugarte M, Desviat LR. (2011) Overexpression of adapted U1snRNA in patients' cells to correct a 5' splice site mutation in propionic acidemia. *Mol Genet Metab* 102:134-138.

Tabuchi K, Südhof TC. (2002) Structure and evolution of neurexin genes: insight into the mechanism of alternative splicing. *Genomics* 79:849-859.

Tanner G, Glaus E, Barthelmes D, Ader M, Fleischhauer J, Pagani F, Berger W, Neidhardt J. (2009) Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. *Hum Mutat* 30:255-263.

Tazi J, Durand S, Jeanteur P. (2005) The spliceosome: a novel multi-faceted target for therapy. *Trends Biochem Sci* 30:469-478.

Tiede S, Storch S, Lübke T, Henrissat B, Bargal R, Raas-Rothschild A, Braulke T. (2005) Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase. *Nat Med* 11:1109-1112.

Tomanin R, Zanetti A, Zaccariotto E, D'Avanzo F, Bellettato CM, Scarpa M. (2012) Gene therapy approaches for lysosomal storage disorders, a good model for the treatment of mendelian diseases. *Acta Paediatr* 101:692-701.

Ule J, Stefani G, Mele A, Ruggiu M, Wang X, Taneri B, Gaasterland T, Blencowe BJ, Darnell RB. (2006) An RNA map predicting Nova-dependent splicing regulation. *Nature* 444:580-586.

Valadkhan S, Gunawardane LS. (2013) Role of small nuclear RNAs in eukaryotic gene expression. *Essays Biochem* 54:79-90.

van Deutekom JC, Bremmer-Bout M, Janson AA, Ginjaar IB, Baas F, den Dunnen JT, van Ommen GJ. (2001) Antisense-induced exon skipping restores dystrophin expression in DMD patient derived muscle cells. *Hum Mol Genet* 10:1547-1554.

Veedu RN, Wengel J. (2009) Locked nucleic acid as a novel class of therapeutic agents. *RNA Biol* 6:321-323.

Vega AI, Pérez-Cerdá C, Desviat LR, Matthijs G, Ugarte M, Pérez B. (2009) Functional analysis of three splicing mutations identified in the PMM2 gene: toward a new therapy for congenital disorder of glycosylation type Ia. *Hum Mutat* 30:795-803.

Veltrop M, Aartsma-Rus A. (2014) Antisense-mediated exon skipping: taking advantage of a trick from Mother Nature to treat rare genetic diseases. *Exp Cell Res* 325:50-55.

Venturi N, Rovelli A, Parini R, Menni F, Brambillasca F, Bertagnolio F, Uziel G, Gatti R, Filocamo M, Donati MA, Biondi A, Goldwurm S. (2002) Molecular analysis of 30

mucopolysaccharidosis type I patients: evaluation of the mutational spectrum in Italian population and identification of 13 novel mutations. *Hum Mutat* 20:231.

Vester B, Wengel J. (2004) LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. *Biochemistry* 43:13233-13241.

Vickers TA, Zhang H, Graham MJ, Lemonidis KM, Zhao C, Dean NM. (2006) Modification of MyD88 mRNA splicing and inhibition of IL-1beta signaling in cell culture and in mice with a 2'-O-methoxyethyl-modified oligonucleotide. *J Immunol* 176:3652-3661.

Wahl MC, Will CL, Lührmann R. (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* 136:701-718.

Wahlestedt C, Salmi P, Good L, Kela J, Johnsson T, Hökfelt T, Broberger C, Porreca F, Lai J, Ren K, Ossipov M, Koshkin A, Jakobsen N, Skouv J, Oerum H, Jacobsen MH, Wengel J. (2000) Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc Natl Acad Sci U S A* 97:5633-5638.

Wally V, Murauer EM, Bauer JW. (2012) Spliceosome-mediated trans-splicing: the therapeutic cut and paste. *J Invest Dermatol* 132:1959-1966.

Wan J, Sazani P, Kole R. (2009) Modification of HER2 pre-mRNA alternative splicing and its effects on breast cancer cells. *Int J Cancer* 124:772-777.

Wang BB, Brendel V. (2006) Genomewide comparative analysis of alternative splicing in plants. *Proc Natl Acad Sci U S A* 103:7175-7180.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470-476.

Wang GS, Cooper TA. (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* 8:749-761.

Wang Z, Burge CB. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* 14:802-813.

Ward AJ, Cooper TA. (2010) The pathobiology of splicing. *J Pathol* 220:152-163.

Warren TK, Warfield KL, Wells J, Swenson DL, Donner KS, Van Tongeren SA, Garza NL, Dong L, Mourich DV, Crumley S, Nichols DK, Iversen PL, Bavari S. (2010) Advanced antisense therapies for postexposure protection against lethal filovirus infections. *Nat Med* 16:991-994.

Wein N, Avril A, Bartoli M, Beley C, Chaouch S, Laforêt P, Behin A, Butler-Browne G, Mouly V, Krahn M, Garcia L, Lévy N. (2010) Efficient bypass of mutations in

dysferlin deficient patient cells by antisense-induced exon skipping. *Hum Mutat* 31:136-142.

Wender PA, Mitchell DJ, Pattabiraman K, Pelkey ET, Steinman L, Rothbard JB. (2000) The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. *Proc Natl Acad Sci U S A* 97:13003-13008.

West S. (2012) The increasing functional repertoire of U1 snRNA. *Biochem Soc Trans* 40:846-849.

Wheeler TM, Sobczak K, Lueck JD, Osborne RJ, Lin X, Dirksen RT, Thornton CA. (2009) Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science* 325:336-339.

Winchester B. (2012) Classification of lysosomal storage disorders. In Mehta A, Winchester B (Eds) *Lysosomal storage disorders: a practical guide*, Wiley-Blackwell, UK, pp. 37-46.

Xiao X, Wang Z, Jang M, Burge CB. (2007) Coevolutionary networks of splicing cis-regulatory elements. *Proc Natl Acad Sci U S A* 104:18583-18588.

Yeo G, Burge CB. (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 11:377-394.

Yeo G, Hoon S, Venkatesh B, Burge CB. (2004) Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc Natl Acad Sci U S A* 101:15700-15705.

Yeo GW, Coufal NG, Liang TY, Peng GE, Fu XD, Gage FH. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat Struct Mol Biol* 16:130-137.

Yokota T, Lu QL, Partridge T, Kobayashi M, Nakamura A, Takeda S, Hoffman E. (2009) Efficacy of systemic morpholino exon-skipping in Duchenne dystrophy dogs. *Ann Neurol* 65:667-676.

Yuste-Checa P, Medrano C, Gámez A, Desviat LR, Matthijs G, Ugarte M, Pérez-Cerdá C, Pérez B. (2015) Antisense-mediated therapeutic pseudoexon skipping in TMEM165-CDG. *Clin Genet* 87:42-48.

Zalachoras I, Evers MM, van Roon-Mom WM, Aartsma-Rus AM, Meijer OC. (2011) Antisense-mediated RNA targeting: versatile and expedient genetic manipulation in the brain. *Front Mol Neurosci* 4:10.

Zamecnik PC, Stephenson ML. (1978) Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc Natl Acad Sci U S A* 75:280-284.

Zammarchi F, de Stanchina E, Bournazou E, Supakorndej T, Martires K, Riedel E, Corben AD, Bromberg JF, Cartegni L. (2011) Antitumorigenic potential of STAT3 alternative splicing modulation. *Proc Natl Acad Sci U S A* 108:17779-17784.

Zanetta C, Nizzardo M, Simone C, Monguzzi E, Bresolin N, Comi GP, Corti S. (2014) Molecular therapeutic strategies for spinal muscular atrophies: current and future clinical trials. *Clin Ther* 36:128-140.

Zarghooni M, Dittakavi SS. (2009) Molecular analysis of cell lines from patients with mucolipidosis II and mucolipidosis III. *Am J Med Genet A* 149A:2753-2761.

Zhou H, Janghra N, Mitrpant C, Dickinson RL, Anthony K, Price L, Eperon IC, Wilton SD, Morgan J, Muntoni F. (2013) A novel morpholino oligomer targeting ISS-N1 improves rescue of severe spinal muscular atrophy transgenic mice. *Hum Gene Ther* 24:331-342.

Zhou Z, Fu XD. (2013) Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma* 122:191-207.

Zhuang Y, Weiner AM. (1986) A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* 46:827-835.

ANNEX

Review Article

Maria Francisca Coutinho, Liliana Matos, Sandra Alves. **From bedside to cell biology: A century of history on lysosomal dysfunction.**

Gene, 2015; 555(1):50-8

Gene 555 (2015) 50–58

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

From bedside to cell biology: A century of history on lysosomal dysfunction

Maria Francisca Coutinho *, Liliana Matos, Sandra Alves

Research and Development Unit, Department of Human Genetics, INSA, Portugal

ARTICLE INFO

Article history:
Received 28 May 2014
Received in revised form 22 September 2014
Accepted 24 September 2014
Available online 29 September 2014

Keywords:
Lysosomal storage disorders (LSDs)
Mucopolysaccharidoses (MPSs)
Enzyme replacement therapy (ERT)

ABSTRACT

Lysosomal storage disorders (LSDs) are a group of rare genetic diseases, generally caused by a deficiency of specific lysosomal enzymes, which results in abnormal accumulation of undegraded substrates. The first clinical reports describing what were later shown to be LSDs were published more than a hundred years ago. In general, the history and pathophysiology of LSDs has impacted on our current knowledge of lysosomal biology.
Classically, depending on the nature of the substrates, LSDs can be divided into different subgroups. The mucopolysaccharidoses (MPSs) are those caused by impaired degradation of glycosaminoglycans (GAGs). Amongst LSDs, the MPSs are a major group of pathologies with crucial historical relevance, since their study has revealed important biological pathways and highlighted interconnecting pathological cascades which are still being unveiled nowadays. Here we review the major historical discoveries in the field of LSDs and their impact on basic cellular knowledge and practical applications. Attention will be focused on the MPSs, with occasional references to other LSDs. We will show as studies on the metabolic basis of this group of diseases have increased our knowledge of the complex degradative pathways associated with the lysosome and established the basis to the development of specific therapeutic approaches aiming at correcting or, at least ameliorating their associated phenotypes.

1. Introduction

The first clinical reports describing what were later shown to be Lysosomal Storage Disorders (LSDs) were published around a hundred years ago (Tay, 1881; Gaucher, 1882; Fabry, 1898; Niemann, 1914; Hunter, 1917). The lysosome itself, however, was only discovered five decades later, by Christian de Duve. The designation 'lysosome' is derived from the Greek words 'lysis' (destruction) and 'soma' (body), and was chosen to reflect its role as the major intracellular site for enzymatic degradation of macromolecules (de Duve et al., 1955). Another ten years were needed for the concept of LSD to be first proposed by H.G. Hers, who suggested that Pompe disease, a glycogen storage disorder, was due to deficiency of a lysosomal enzyme (acid maltase deficiency; Hers, 1963). Soon after the idea of enzymatic deficiency was established, several metabolic disorders that had been previously described by clinicians were catalogued as LSDs, and classified according to the storage material accumulated intralysosomally. Thus, disorders in which glycosaminoglycan (GAG, also called mucopolysaccharides) accumulation prevailed were classified as mucopolysaccharidoses; those dominated by lipid storage were called lipidoses; the ones in which the accumulation of sphingolipids prevailed were sphingolipidoses; those mostly characterized by the storage of oligosaccharides were oligosaccharidoses (Schultz et al., 2011).

By the same time, the concept of 'cross correction', first formulated by de Duve in 1964, was demonstrated by Elizabeth Neufeld and her group, who discovered that co-cultured fibroblasts derived from two patients with different lysosomal storage disorders mutually corrected each other, established the basis of 'enzyme replacement therapy' (ERT) (Barton et al., 1990; Brady, 2006). Then, all efforts turned to developing effective ERT for LSDs, starting with Gaucher disease (GD), the most common of these pathologies. Several teams centered efforts in purifying β-glucocerebrosidase (GCase) to supply GD patients with the enzyme in which they were deficient, in order to evaluate whether clinical advantage came from that therapeutic approach. Because of its high hydrophobicity, and of the lack of experience in its handling, it was extremely hard to get useful quantities of the requisite enzyme. Eventually, a small amount of sufficiently purified GCase was obtained and administered to two GD patients (Brady et al., 1974) leading to a reduction of GCase in the liver of both patients, together with a striking

Abbreviations: AD, Alzheimer's disease; ARSG, arylsulfatase G; AV, autophagic vacuole; CNS, central nervous system; CMA, chaperone-mediated autophagy; CS, chondroitin sulfate; DS, dermatan sulfate; EE, early endosome; ER, endoplasmic reticulum; ERT, enzyme replacement therapy; GD, Gaucher disease; GAGs, glycosaminoglycans; HS, heparan sulfate; KS, keratan sulfate; LE, late endosome; LSDs, lysosomal storage disorders; MPRs, mannose 6-phosphate receptors; MPSs, mucopolysaccharidoses; RE, recycling endosome; TGN, *trans*-Golgi network; UPS, ubiquitin–proteasome system; GCase, β-glucocerebrosidase.

* Corresponding author at: Research and Development Unit, Department of Human Genetics, INSA, Rua Alexandre Herculano, 321, 4000-055 Porto, Portugal.
E-mail addresses: francisca_coutinho@yahoo.com, francisca.coutinho@insa.min-saude.pt (M.F. Coutinho).

http://dx.doi.org/10.1016/j.gene.2014.09.054
0378-1119/

reduction in the quantity of GCase associated with circulating erythrocytes in the recipients. Nevertheless, a long way had to be walked before consistent clinical benefit of ERT was demonstrated in a cohort of patients with GD (Barton et al., 1991) and an even longest period was needed for the first therapies of such kind to reach the market and become commercially available.

Here we review the major historical discoveries in the field of LSDs and their impact on basic cellular knowledge and practical applications (Fig. 1). Attention will be focused on the MPSs, with occasional references to other LSDs. We will show as studies on the metabolic basis of this group of diseases have increased our knowledge of the complex degradative pathways associated with the lysosome and established the basis to the development of specific therapeutic approaches aiming at correcting or, at least ameliorating their associated phenotypes.

2. Early clinical descriptions: the Hurler and Hunter syndromes

To find, in the literature, the first medical report of a MPS, we have to go back in time until 1917. In that year, Charles Hunter provided the scientific community with a detailed description and radiological imaging of two brothers, aged 10 and 8 years, with multiple skeletal and somatic abnormalities, who were proposed to suffer from an endocrine disturbance. The two children were full-term and were normally delivered. Both were breast-fed and no digestive disturbances were observed in infancy; both were walking at the age of 17 months. The elder began to talk when aged around 1 year and had normal intelligence. The younger brother, however, was somewhat late in learning to talk and was making slow progresses at school by the time the original paper was published. Apart from throat trouble, both children were reported to have been healthy in every way. Both were operated on for tonsils and adenoids; both had hearing deficiencies. The brothers were described as "undersized" having "extremely large" heads, "curiously shaped, with very marked bulging of the squamous portion of the temporal bone and of the frontal bones; the hair of the head rather thin and very harsh (…)." Other features included "saddle nose, with large thick nostrils; (…) very large tongue; very short neck, with slight enlargement of right lobe of thyroid (…); knees (…) slightly flexed; [thick] knees and ankles (…) [and] hands [which could] not be clenched". Also common to both children's clinical course was the fact that their condition progressively deteriorated over time (Hunter, 1917). These early observations were later confirmed by the discovery of other patients harboring similar physical deformities, and the typical features of the severe form of this syndrome were established: coarse facial features, short stature, skeletal deformities, joint stiffness and mental retardation. A milder form of the same disorder was also recognized, with preservation of intelligence and survival into late adult life. Even though still presenting with an obvious somatic involvement, similar to the one observed in severely affected Hunter patients, this subtype of the disease presents later in life and with a much slower progression (Neufeld and Muenzer, 2001).

Two years later, in 1919, Pfaunder reported two cases to the Medical Society in Munich that were described in detail by his assistant Hurler, who at once suspected to be in the presence of another new syndrome. The patients displayed the following combination of congenital abnormalities: "clouding of the corneae, (…) oxycephaly, disproportionate dwarfism strongly resembling that of hypothyroidism and associated with some of the usual signs of that condition, [such as] saddle nose, mental defect, dry skin, inguinal and umbilical hernia, crura valga, pedes valgi". In addition, the patients presented "a contraction of the fingers, limitation of movement in other joints (shoulders, elbows,

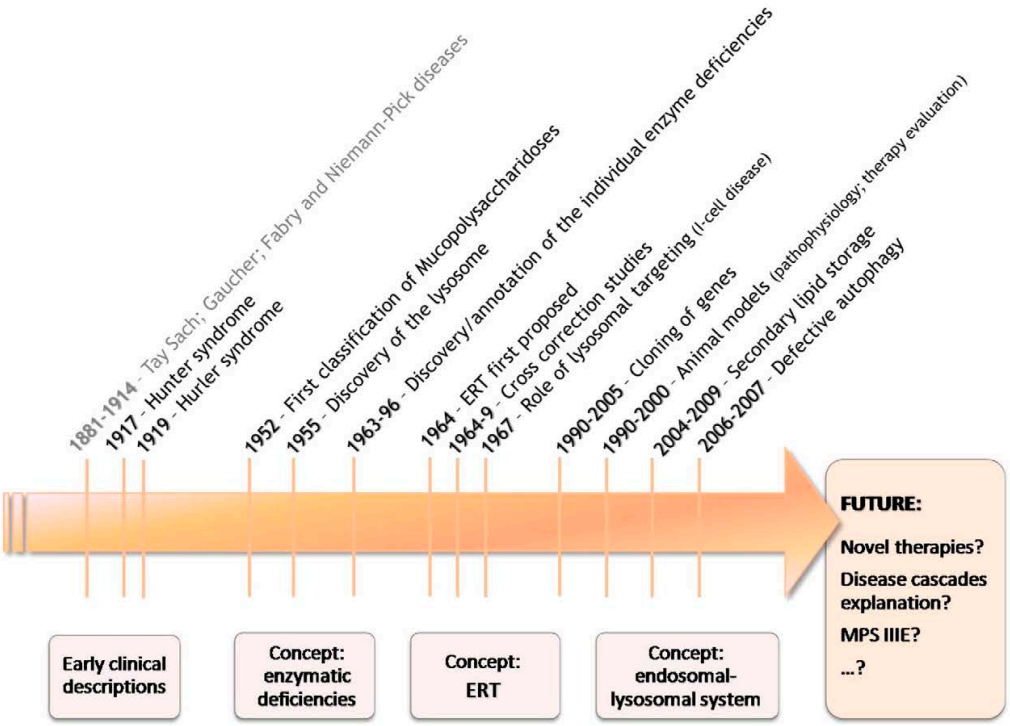


Fig. 1. Timeline for discoveries on lysosomal storage diseases and their impact on cell biology. Overview of the major discoveries concerning MPSs in particular, highlighting their impact on the establishment of some of the most important concepts on the LSD field: i) enzymatic deficiencies as the basis of disease; ii) enzyme replacement as a possible therapeutic approach and iii) the lysosome as central coordinator of a remarkably dynamic system, the *endosomal–lysosomal system*.

knees) and defective hearing, to mention only the most striking features". The syndrome was named after the clinicians who originally reported it, as Pfaundler–Hurler, but nowadays it is usually referred to as Hurler syndrome. Once again, this initial report was confirmed by the description and annotation of other patients with similar symptoms, later shown to harbor the same metabolic defect. Since this initial report, and throughout the following two decades, twenty two more cases showing the same combination of abnormalities were reported in Britain, USA, Germany and China (Engel, 1939). Presently, hallmark features of this syndrome include coarse facies, short stature, corneal clouding, joint stiffening, umbilical hernia, dysostosis multiplex and hepatosplenomegaly. Depending on the cases, it may or may not include mental retardation. Typically, the most severe cases have an early onset of symptoms, while the mild ones appear later in life. Severely affected patients usually present some degree of intellectual dysfunction associated with the above referred somatic features (Neufeld and Muenzer, 2001). As observed in several other LSDs, this wide range of phenotypic involvement is generally divided into three major clinical entities, established in accordance with the degree of severity and age of onset of the symptoms: Hurler (MPS IH; MIM#607014), Hurler–Scheie (MPS IH/S; MIM#607015), and Scheie (MPS IS; MIM#607016) syndromes. Hurler and Scheie syndromes represent phenotypes at the severe and mild ends of the MPS I clinical spectrum, respectively; and the Hurler–Scheie syndrome is intermediate in phenotypic expression (reviewed in Neufeld and Muenzer, 2001; Coutinho et al., 2012).

These were the first reports recognizing the clinical presentation of MPSs and considering the whole phenotype as a disease, even though having an unknown etiology at that time. The particularly severe skeletal deformation that led to an easily recognizable physical deficiency and facial dysmorphism, however, had not gone unnoticed until the beginning of the 20th century. In fact, it is believed that some of the early representations of 'gargoyles' were actually inspired on the observation of MPS patients by the artists who first designed those sculptures, since some of them do perfectly mimic the hallmark features of this group of diseases. When first reported, these syndromes were actually designated 'gargoylism' (Norman et al., 1959; Van Pelt, 1960; Koskenoja and Suvanto, 1959; Scheie et al., 1962). Apparently, several characters spread in both literature and oral tradition also seem to have been inspired on the observation of people with that same phenotype. One example is "Rip Van Winkle", a well known short story from the American author Washington Irving, published in 1819 and that had actually been borrowed from the folklore of Europe and rewritten with a distinct American sensibility. Van Winkle was a farmer living in a pleasant village, at the foot of New York's Catskill Mountains who, being fed up with his querulous wife, runs away into the wilderness where he finds himself surrounded by a large group of 'gargoyles': short, strange-looking people in ancient Dutch clothing. Irving's original description of these strange men was, in almost every way, consistent with our present notion of the typical MPS phenotype and its associated dysmorphism: "short, square-built (…) with thick bushy hair (…) odd-looking personages (…). Their visages, too, were peculiar; one had a large head, broad face, and small piggish eyes".

3. Establishment of the concept: MPSs as enzymatic deficiencies

When these two disorders were first described, no clue existed on their underlying defects. Therefore, apart from some obvious phenotypic similarity, no one guessed that both conditions result from defects in the same metabolic pathway and would later be recognized as clinical entities belonging to the same family of diseases: the mucopolysaccharidoses (MPSs). The term MPS was only suggested in 1952, after the identification of excessive amounts of GAGs in samples from patients with Hurler syndrome (Brante, 1952; Dorfman and Lorincz, 1957). GAGs are unbranched polysaccharide chains composed of repeating disaccharide units. They are so called because one of the two sugars in the repeating disaccharide is always an amino sugar (either *N*-acetylglucosamine or *N*-acetylgalactosamine), which, in most cases, is sulfated. Four main groups of GAGs are distinguished according to their sugars, the type of linkage between them, and the number and location of sulfate groups: (1) hyaluronan, (2) chondroitin sulfate (CS) and dermatan sulfate (DS), (3) heparan sulfate (HS) and (4) keratan sulfate (KS) (Alberts et al., 2002). The recognition that different patients, with slightly different symptoms had also different relative amounts of HS, DS, KS or CS in urine, confirmed the existence of multiple syndromes in addition to Hurler and Hunter (Harris, 1961; Lorincz, 1958; Maroteaux et al., 1963; Sanfilipo et al., 1963). Nevertheless, even then, the MPSs were thought to result from an overproduction of GAGs, other than from a defective degradation of those products. Only in 1964 did the concept of MPSs as enzymatic deficiencies appear, when van Hoof and Hers proposed that these disorders were analogous to Pompe disease and therefore, due to specific lysosomal hydrolase deficiencies. Supporting this idea were the results from these two authors, who observed dramatically enlarged lysosomes in Hurler hepatocytes (Van Hoof and Hers, 1964) and, four years later, the studies conducted by Fratantoni and co-workers, who analyzed $^{35}SO_4$ uptake on Hurler fibroblasts. Together, these experiments have demonstrated impaired GAG degradation (Fratantoni et al., 1968). Similar impairments were detected in patients with other MPSs and, gradually, the different syndromes were annotated and the diagnosis of each one established on the basis of increased urinary excretion of GAGs, taking into account the relative amounts of each one (see Table 1). From then on, researchers have focused their attention on identifying the enzyme deficiencies underlying each of the MPS syndromes. Most of them were discovered in the 1970s, thanks to the efforts of several different teams (Maroteaux et al., 1963; Kresse et al., 1971; Kresse and Neufeld, 1972; McKusick, 1972; Sly et al., 1973; Kresse et al., 1976; Kresse et al., 1980; Gitzelmann et al., 1987; Nelson et al., 1988; Natowicz et al., 1996). Then, interest turned to purification of the normal enzymes (Rome et al., 1978; Clements et al., 1985a; Clements et al., 1985b; Ohshita et al., 1989; Schuchman et al., 1984a; Schuchman et al., 1984b; Stoltzfus et al., 1992; Stahl and Touster, 1971; Himeno Mnishimura et al., 1976; Brot et al., 1978; Wasteson and Neufeld, 1982; Di Natale and Daniele, 1985; Bielicki et al., 1990; McGovern et al., 1982; Gibson et al., 1987; von Figura, 1977; Sasaki et al., 1991; Freeman and Hopwood, 1986; Freeman et al., 1987; Freeman et al., 1983; Bame and Rome, 1987; Distler and Jourdian, 1978; Glössl et al., 1979; Masue et al., 1991) and elucidation of their synthesis and transport to the lysosomes (Hanai et al., 1971; Hickman and Neufeld, 1972; Hickman et al., 1974; Hasilik and Neufeld, 1980; Varki and Kornfeld, 1980; Fischer et al., 1980; Kornfeld, 1990). Better understanding of the primary structure of the enzyme proteins and subsequent modifications also came from cloning and characterization of their coding genes (Scott et al., 1992; Wilson et al., 1993; Scott et al., 1995; Zhao et al., 1996; Fan et al., 2006; Hrebícek et al., 2006; Mok et al., 2003; Nakashima et al., 1994; Morris et al., 1994; Oshima et al., 1988; Karageorgos et al., 2007; Miller et al., 1990; Triggs-Raine et al., 1999).

Our current understanding of the normal pathways of the GAG catabolism has been closely tied to the elucidation of enzyme deficiencies in the MPSs. In fact, the role of many of those enzymes became apparent only through the consequences of their absence (Neufeld and Muenzer, 2001). Nowadays, the MPSs group comprises 11 different conditions (see Table 1), all of them caused by deficiency of enzymes catalyzing the stepwise degradation of GAGs. As a whole, MPSs are characterized by intralysosomal accumulation and increased excretion in urine of partially degraded GAGs, which ultimately results in cell, tissue, and organ dysfunction (Neufeld and Muenzer, 2001; Winchester, 2012a; Giugliani, 2012). In general terms, it is possible to say that MPSs are characterized by a chronic and progressive course, with different velocities of progression depending on the severity of each one. Typical symptoms include those already reported on the first descriptions of Hurler and Hunter patients: organomegaly, dysostosis multiplex, and a characteristic abnormal facies. Hearing, vision, and cardiovascular

Table 1
Summary table of mucopolysaccharidoses.

Pathology	Subtype	Increased GAGs	Enzyme deficiency	Gene (localization)	ERT
MPS I	Hurler (H)	DS + HS	α-L-Iduronidase	*IDUA* 4p16.3	_
	Hurler–Scheie (H/S)	DS + HS	α-L-Iduronidase		Aldurazyme®
	Scheie (S)	DS + HS	α-L-Iduronidase		Aldurazyme®
MPS II (Hunter)		DS + HS	Iduronate sulfatase	*IDS* Xq28	Elaprase®
MPS III (Sanfilippo)	III A	HS	Heparan-*N*-sulfatase	*SGSH* 17q25.3	_
	III B	HS	α-*N*-acetylglucosaminidase	*NAGLU* 17q21.2	_
	III C	HS	Heparan acetyl-CoA:α-glucosaminide *N*-acetyltransferase	*HGSNAT* 8p11.21	_
	III D	HS	*N*-acetylglucosamine 6-sulfatase	*GNS* 12q14.3	_
MPS IV (Morquio)	IV A	KS + CS	Galactose 6-sulfatase	*GALNS* 16q24.3	_
	IV B	KS	β-Galactosidase	*GLB1* 3p22.3	_
MPS VI (Maroteaux–Lamy)		DS	Arylsulfatase B (*N*-acetylglucosamine 4-sulfatase)	*ARSB* 5q14.1	Naglazyme®
MPS VII (Sly)		DS + KS + CS	β-Glucuronidase	*GUSB* 7q11.21	_
MPS IX (Natowicz)		Hyaluronan	Hyaluronidase 1	*HYAL* 3p21.31	_

function may also be affected. Additionally, joint mobility may also be compromised (reviewed in Neufeld and Muenzer, 2001; Coutinho et al., 2012). With the exception of MPS II, which is X-linked, all other MPSs are inherited in an autosomal recessive fashion. Nevertheless, all MPSs, including the X-linked syndrome, are true recessive diseases, with no detectable manifestations in carriers. A clinical suspicion of MPS is usually followed by a series of laboratory tests in both urine and blood samples that, ultimately, allow a proper and definitive diagnosis (Fig. 2). Initial suspicions are checked by a screening test for GAGs in a urine sample. If the test is positive, the same sample may be used for a quantitative and/or qualitative assay of urinary GAGs. It is the qualitative pattern of urinary GAGs, together with the clinical presentation, that provides a deeper insight to the MPS diagnosis. Taking those results into account, blood samples are retrieved and used to perform specific enzyme assays, confirming the diagnosis and assigning the specific MPS type. Finally, the blood sample may also be used to identify the underlying genetic defect (see Table 1; Filocamo and Morrone, 2011; Winchester, 2012b; Giugliani, 2012; Hopwood, 2012).

Altogether, these studies provided the background for understanding the effects of disease-causing mutations and, as we will present in

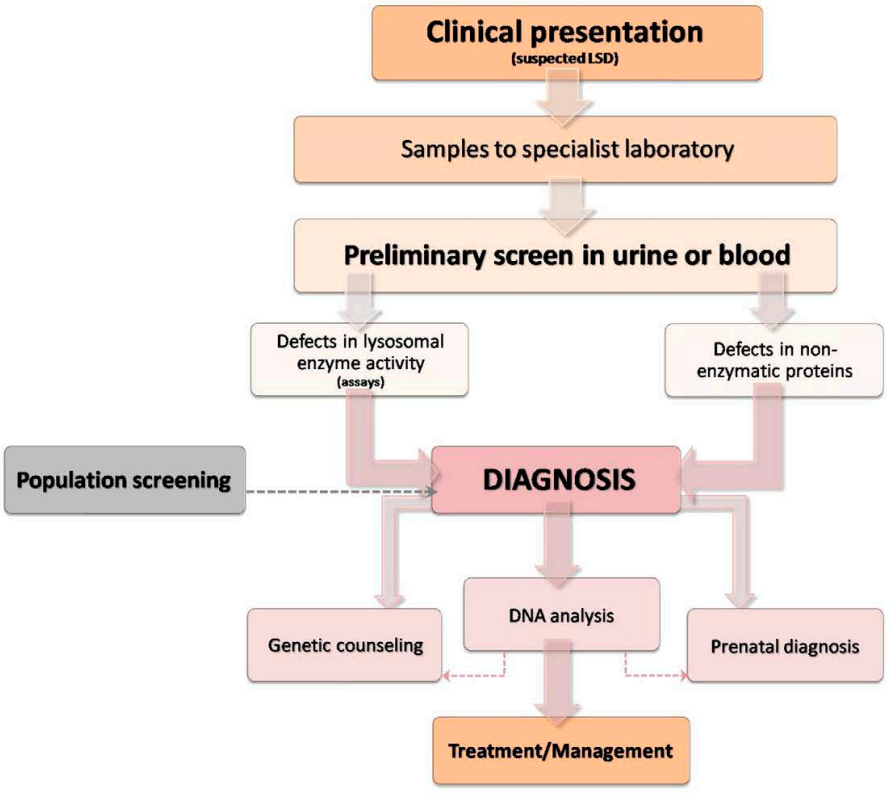


Fig. 2. Scheme for the laboratory diagnosis of lysosomal storage diseases*. To route to a definitive diagnosis is based upon clinical presentation and a panel of laboratory tests on blood and urine. Once the initial diagnosis is achieved, DNA analyses are performed to identify the genetic defect underlying pathology. This allows genetic counseling to the affected families and prenatal diagnosis, whenever necessary. In some particular cases, screening may be performed in targeted populations (e.g., screening for Krabbe disease in New York State). *Based on an original schematic diagram by Winchester, 2012b.

the next section, for designing the concept of enzymatic cross correction, which ultimately led in the development of several therapeutic approaches, not only for MPSs, but for LSDs in general (Neufeld and Muenzer, 2001).

4. Cross correction of MPSs: the enzyme as a drug?

The notion that LSDs could be treated by replacing the defective enzymes with their normal counterparts is almost as old as the lysosome itself, having been first suggested by Christian de Duve in 1964, with the following brief explanation: "In our pathogenic speculations and in our therapeutic attempts, it may be well to keep in mind that any substance which is taken up intracellularly in an endocytic process is likely to end up within lysosomes. This obviously opens up many possibilities for interaction, including replacement therapy" (de Duve, 1964). Actual evidences of this process, however, came from the pivotal studies of Elizabeth Neufeld's lab, a few years later. This group reported cross correction phenomena between cultured fibroblasts from patients with different MPSs, pulse-labeled with ^{35}S-sulfate. "The aberrant metabolism of Hurler cells [was corrected] by the secretions of fibroblasts of genotype other than Hurler and, similarly, the defect of Hunter cells [could] be corrected by secretions of fibroblasts of genotype other than Hunter. The active factors in these secretions, which [were] heat labile and associated with macromolecules" were designated as "corrective factors" and proposed to be the enzymes which were deficient in each MPS (Fratantoni et al., 1969). Subsequent experiments demonstrated that there was no need to have the genetically distinct cells in contact with each other to reach correction. In fact, when the appropriate active enzyme was added to the media of enzyme-deficient cultured fibroblasts from individuals with specific LSDs, the exogenous enzyme gained access to the lysosomes and degraded the accumulated substrates (Cantz and Kresse, 1974; O'Brien et al., 1973; Porter et al., 1971). A remarkable observation that came from these early studies was the recognition that, sometimes, values as low as 1%–5% of normal intracellular enzyme activity were enough to correct the metabolic defects in the enzyme-deficient cells (Desnick et al., 1973; Hasilik et al., 2009; von Figura, 1991).

The subsequent discovery that lysosomal enzymes are targeted to the lysosome by the mannose 6-phosphate receptor-mediated pathway (Hanai et al., 1971; Hickman and Neufeld, 1972; Sando and Neufeld, 1977; Kaplan et al., 1977; reviewed in Desnick et al., 1973; von Figura, 1991; Hasilik et al., 2009), along with the discovery that the mannose 6-phosphate receptors (MPRs) on the plasma membranes of cells mediate the cellular uptake and delivery of the intravenously administered normal enzymes to the lysosomes (Kaplan et al., 1977; reviewed in Dahms et al., 1989; Kornfeld, 1992; Ludwig et al., 1995; Ghosh et al., 2003), provided further rationale for the treatment of nonneural LSDs by ERT. Finally, the fact that most LSDs have significantly milder subtypes, which were typically associated with low levels of residual enzymatic activity, also indicated that it was not necessary to restore full activity, or even heterozygous levels of activity in the treated individuals to reach correction, provided that the enzyme effectively reached the proper sites of pathology.

Altogether, these studies established the grounds for clinical studies of ERT for LSDs and the features that the exogenous lysosomal enzyme needed to have to be safe and therapeutically useful: not only would it have to be of human origin, highly purified and available in adequate quantity, but it would also have to carry the recognition signal for the targeted cells (reviewed in Neufeld, 2006).

ERT pilot clinical studies in several LSDs (Fabry, Gaucher, Pompe, and Sandhoff diseases) began in the early 1970s, by intravenous infusion of the respective normal human enzyme. In each case, the partially purified enzyme was rapidly cleared from the circulation even though clearance of the respective accumulated substrate(s) was observed (reviewed in Desnick et al., 1973, 1976). Nevertheless, only after a deeper understanding of the basic science underlying receptor-mediated endocytosis and trafficking of the defective enzymes was achieved, was it really possible to design and establish effective and safe ERT approaches. The whole process from the concept of enzyme replacement to the commercial development of the first recombinant enzyme (GCase as a pharmaceutical for GD) took nearly three decades, and another ten years were needed for the commercial development of other enzymes presently available for ERT (reviewed in Neufeld, 2006).

Presently, ERTs are available for GD, Fabry disease, Pompe disease and several MPSs: MPS I (Aldurazyme®, recombinant human α-L-iduronidase produced by Genzyme and commercialized since 2003), MPS II (Elaprase®, recombinant iduronate sulfatase produced by Shire, available since 2005) and MPS VI (Naglazyme®, recombinant arylsulfatase B, produced by BioMarin and approved since 2006). Before the introduction of ERT, treatment of LSDs was essentially palliative and aimed at alleviating specific symptoms, without targeting the underlying pathological condition.

It is important to stress, though, that ERT is not the only therapeutic approach for LSDs (Fig. 3). It is, in general, more effective on later onset or milder forms, since the recombinant enzyme is unable to cross the blood brain barrier to reach the brain. Therefore, additional therapeutic approaches are being investigated for LSDs in general and MPSs in particular. Substrate Reduction Therapy (SRT), for example, is currently being evaluated for some of these diseases (e.g. MPS I, II, IIIA and IIIB). Also under evaluation are chaperone therapy, which holds the potential to stabilize the defective enzymes produced by patients; and/or gene enhancement and gene therapy techniques in which specific viral vectors may be used to directly transduce central nervous system (CNS) cells that would thus become an endogenous source of high levels of functional enzyme (Beck, 2010; Gritti, 2011).

5. A broader perspective: the lysosome is (much) more than a bag of enzymes

Since the discovery of the lysosome in 1955 (de Duve et al., 1955), several advances have been made to understand the key roles and functions of this organelle. Its finding led to many new questions. The most critical one was: what is the physiological function of this "bag" of enzymes?

Even before the identification of the lysosome as an individual organelle, the first clue to one of its main functions came from the work of Werner Straus, who deserves the credit for undertaking studies which would almost certainly have led to an independent discovery of lysosomes. By 1954, he had already succeeded in subfractionating what he called "droplet" fractions from kidney homogenates of normal rats and showed them to be rich in acid phosphatase and proteolytic enzymes (Straus, 1954). In a subsequent study using the same kidney homogenates of rats, which had received an intra-peritoneal injection of egg white protein, he observed that the isolated droplets contained, in addition to the enzymes, relatively high concentrations of the injected foreign proteins, obtaining good evidence that the droplets were a site of storage and breakdown of reabsorbed proteins (Straus, 1957). This early work provided the first clear link between lysosomal digestion and endocytotic uptake of extracellular materials. Despite these clues, the lysosome itself was only described as a membrane-limited component containing hydrolytic enzymes that functions as a "digestive body", in a work of de Duve and co-workers with liver cells in 1955 (de Duve et al., 1955).

As more cell lines were studied and the ubiquitous distribution of lysosomes in mammalian cells was recognized, it became clear that the lysosome is not actually only a "body", but the central coordinator of a remarkably diverse and dynamic system collectively known as the *endosomal–lysosomal system*, a term first coined by de Duve and Wattiaux that still resists the time passing (de Duve and Wattiaux, 1966). This organelle and the constituent pathways to which it is linked comprise a processing and recycling center essential to all cells (Fig. 4).

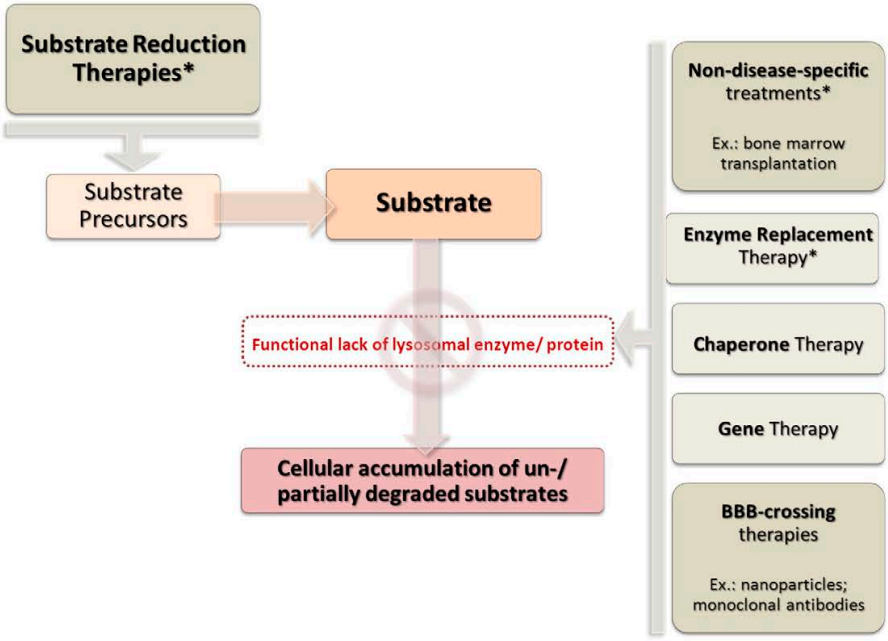


Fig. 3. Therapies for lysosomal storage disorders. LSD treatments can be divided into those that reduce the levels of biosynthesis of the accumulating substrate (left side of the image) and those aiming at directly addressing either the functional loss of the enzyme or its clinical effects (right side of the image). The later include (a) general treatments that deal mainly with the symptoms (ex.: bone marrow transplantation); (b) ERT and (c) potential new therapies such as the use of chemical chaperones, gene therapy or BBB-crossing therapies. Treatments marked with an * are in clinical use at present.

Lysosomes are now known to contain more than 50 acid hydrolases (phosphatases, nucleases, glycosidases, proteases, peptidases, sulfatases and lipases) responsible for the cellular digestion of most macromolecules. The breakdown products are then transported back to the cytosol by specific transporter proteins embedded in the lysosomal membrane — a salvage pathway which facilitates the egress of lysosomal degradation products to other cell organelles and membranes for subsequent use in biosynthetic processes (Tettamanti et al., 2003).

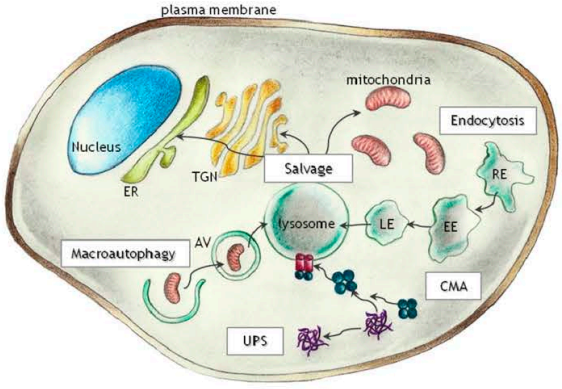


Fig. 4. The lysosome as a central element of the cell*. Lysosomes link to the endocytic pathways, to the macroautophagy stream and its close allies, the ubiquitin–proteosomal system (UPS) and the chaperone-mediated autophagy (CMA) components. What flows into this system must also leave in some form, depicted here as the salvage pathway with delivery to the *trans*-Golgi network (TGN), mitochondria, and other sites in the cell. The existence of all these interrelated components of the *endosomal–lysosomal system* is probably one of the motifs why disease cascades in lysosomal disorders are so complex and remain mostly unknown even after decades of study. (EE, early endosome; RE, recycling endosome; LE, late endosome; AV, autophagic vacuole; ER, endoplasmic reticulum). *Illustration based on an original schematic diagram by Walkley, 2009.

The *endossomal–lysosomal system* includes also the biosynthetic pathway, which is involved in the transport of newly synthesized proteins from the Golgi complex to lysosomes; the endocytic pathway where extracellular components are internalized linking cells with their external environment and the autophagic pathway, an ubiquitous cellular route by which long-lived proteins and whole organelles are sequestered for lysosomal degradation. Autophagy, which is often activated following starvation stress, can be divided into macroautophagy (organelle or cytoplasm enclosed in a double membrane), microautophagy (cytoplasmic degradation following engulfment by the lysosome), and chaperone-mediated autophagy (CMA; selective transport of cytosolic proteins one-by-one to the lysosome via specific chaperones) (Vellodi, 2005; Ballabio and Gieselmann, 2009; Schultz et al., 2011; Micsenyi and Walkley, 2012; Kaushik and Cuervo, 2012).

Operating in close parallel with these pathways is an additional mechanism — the ubiquitin–proteasome system (UPS), which is functionally allied with the lysosomal system in maintaining proteolytic quality control mechanisms in eukaryotic cells (Micsenyi and Walkley, 2012; Walkley, 2009).

The *endosomal–lysosomal system* is also intimately involved in the processes of phagocytosis and antigen presentation, which are necessary for regulation of inflammation and control of the immune response against foreign bodies (bacteria, viruses, etc.) (Saftig, 2006; Parkinson-Lawrence et al., 2010; Boustany, 2013). In addition, it plays a major role in the regulation of apoptotic processes within the cytosol via signal transduction (Saftig, 2006; Boustany, 2013) and in the secretion of cellular cargo (eg. signaling molecules as neurotransmitters or toxic cellular metabolites) by exocytosis, factors that are all involved in inflammation, oncogenesis and neurodegenerative disorders (Schultz et al., 2011; Boustany, 2013).

Apart from these central cellular functions, lysosomes can also be implied in processes such as cholesterol homeostasis, cell membrane repair (calcium regulated), fertilization, receptor recycling and regulation, cell division, skin pigmentation as well as in bone and tissue remodeling (Vellodi, 2005; Saftig, 2006; Walkley, 2009; Parkinson-Lawrence et al., 2010; Boustany, 2013). As a whole, the lysosomal system functions as a highly efficient and coordinated network essential for the metabolic homeostasis of the cell.

Disruption of such carefully orchestrated processes leads necessarily, to catastrophic consequences for cells, organs, and individuals, with nearly 60 different types of lysosomal diseases documented to date (Micsenyi and Walkley, 2012). Expansion of knowledge on LSDs over the past 50 decades reflects the progress in the field of "metabolic medicine" begun by Garrod in 1909, who first introduced the concept of inborn errors of metabolism (Garrod, 1909). Over the years the use of combined biochemical, morphological, cellular and molecular methodologies to study the lysosome and its related diseases have been crucial not only to clarify their complex pathobiology but also to expand our view of the lysosome from a degradative organelle to one intimately linked to multiple cellular processes. Nowadays, the lysosome is far from being a "stationary" compartment. Novel genes and proteins contributing to lysosome function continue to be identified and the list of lysosomal and lysosomal-associated diseases continues to grow. Further understanding of the molecular and cellular pathogenesis of lysosomal diseases will broaden the range of therapeutic targets for LSDs, as well as the opportunity to improve knowledge about the *endosomal–lysosomal system*.

6. A look forward …

As shown in the last section, our understanding of the lysosome and its role in cells has evolved significantly since its discovery by Christian de Duve more than 50 years ago. The lysosome is no longer viewed as just an end-point degradative compartment, but rather as part of a very complex and interactive set of intracellular organelles that have a wide array of specialized functions (Fig. 4). All of a sudden, the panoply of genes that code for proteins involved in any of these processes have become potential targets for screening of causal mutations in patients with clinical suspicion of LSD, but whose biochemical and/or genetic profile does not fit any known disease. With the advent of new sequencing technologies, the search for the underlying molecular basis of these patients' phenotypes had a new incentive. The future will probably surprise us with the discovery of several novel LSDs of diverse origins. In fact, this new view on the lysosome role brought the notion that defects in any one of the processes in which it is involved can be associated with lysosomal disease, leading then to the recognition that LSDs may be better viewed as conditions of molecular deprivation that result in specific pathogenic cascades (Walkley, 2009; Walkley and Vanier, 2009; Parkinson-Lawrence et al., 2010).

We have come a long way since the first descriptions of LSDs were published and their underlying biochemical basis unveiled. Nevertheless, much remains to be discovered on the processes through which excessive lysosomal storage leads to organ dysfunction and, ultimately, premature death.

And even when we consider MPSs alone, there is still much to learn on this topic. In fact, even though the majority of the affected enzymes and genes were described 20–30 years ago, there are several open questions, dealing with the pathogenetic relevance of minor structural GAG elements or on the interaction of enzymes with the lysosomal membrane. Also the pathogenic cascade which originates disease remains elusive, in MPSs as in LSDs in general. The current literature suggests a complex interplay between neuroinflammation, microglial activation and adaptive immunity in MPS disease but there is still much to be learned on this topic. The extent of adaptive immune involvement, for example remains to be elucidated (reviewed in Archer et al., 2014). Furthermore, there are still critical steps in the degradation of the GAG HS, which remain enigmatic. Recently, Kowalewski and collaborators (Kowalewski et al., 2012) have demonstrated that the lysosomal arylsulfatase G (ARSG), reported previously to be associated with an adult form of ceroid lipofuscinosis in dogs (Abitbol et al., 2010), is the long-sought glucosamine-3-O-sulfatase required to complete the degradation of heparan sulfate. These authors have generated an ARSG-deficient mouse model by targeted disruption of the *Arsg* gene and showed that *Arsg*-deficient mice accumulate HS both in visceral organs and central nervous system (CNS). Furthermore, the knockout mice develop neuronal cell death and behavioral deficits. These results strongly suggest that other MPSs may exist, which haven't been identified yet. In particular, ARSG deficiency is now thought to represent a unique, yet unknown form of MPS, tentatively termed MPS IIIE. Importantly, the HS that is accumulated in *Arsg*-deficient mice exhibits unique nonreducing end structures with terminal N-sulfoglucosamine-3-O-sulfate residues which, ultimately, may allow for a precise biochemical diagnosis of the disorder (Kowalewski et al., 2012).

So, as once the discovery of unexplained diseases prompted the search for the underlying causes of such abnormalities and led to a greater understanding of lysosomal function and physiology, now is the time to cross the opposite pathway and wonder whether additional diseases exist, caused by defects in other proteins of the greater *endosomal–lysosomal system*.

Finally, it is also important to refer that, over the last decades, lysosomal dysfunction has been extensively documented in other diseases that afflict a much higher number of individuals than all LSDs together: Alzheimer disease (AD) and other conditions associated with β-amyloidogenesis, such as parkinsonism. In fact, one of the earliest known intracellular histopathological features of AD is the presence of increased levels of lysosomal hydrolases and the accumulation of acid-hydrolase-containing compartments. Neurons from AD brains have striking alterations in the size and volume of early endosomal (EE) compartments, as well as on the localization pattern of acid hydrolases (both pro- and mature forms). Increased endocytic uptake and autophagy have also been documented, consistent with a major cellular attempt to regenerate, catabolize injured membranes, and synthesize new membranes. This leads to an increased demand for hydrolases in EE of neurons within affected regions of AD brains, which may activate alternative mechanisms of delivery to this compartment. Furthermore, in AD and similar conditions, some other specific pathological phenomena involving the *endosomal–lysosomal system* have been observed, such as the persistence of acid hydrolase deposits in the extracellular space, coming from degenerating neurons (reviewed in Cataldo et al., 1995).

Future efforts should also be focused in elucidating the temporal progression of brain disease in both LSDs and other neurodegenerative disorders such as AD. In addition, identifying sites of initial vulnerability will help to develop directed therapies where focal small molecule or gene replacement holds the most promise (Schultz et al., 2011).

References

Abitbol, M., Thibaud, J.L., Olby, N.J., Hitte, C., Puech, J.P., Maurer, M., Pilot-Storck, F., Hédan, B., Dréano, S., Brahimi, S., Delattre, D., André, C., Gray, F., Delisle, F., Caillaud, C., Bernex, F., Panthier, J.J., Aubin-Houzelstein, G., Blot, S., Tiret, L., 2010. A canine arylsulfatase G (ARSG) mutation leading to a sulfatase deficiency is associated with neuronal ceroid lipofuscinosis. Proc. Natl. Acad. Sci. U. S. A. 107, 14775–14780.

Alberts, B., Johnson, A., Lewis, J., et al., 2002. Molecular Biology of the Cell, 4th edition. , (New York).

Archer, L.D., Langford-Smith, K.J., Bigger, B.W., Fildes, J.E., 2014. Mucopolysaccharide diseases: a complex interplay between neuroinflammation, microglial activation and adaptive immunity. J. Inherit. Metab. Dis. 37, 1–12.

Ballabio, A., Gieselmann, V., 2009. Lysosomal disorders: from storage to cellular damage. Biochim. Biophys. Acta 1793, 684–696.

Bame, K.J., Rome, L.H., 1987. Acetyl-CoA: alpha-glucosaminide N-acetyltransferase from rat liver. Methods Enzymol. 138, 607–611.

Barton, N.W., Furbish, F.S., Murray, G.J., Garfield, M., Brady, R.O., 1990. Therapeutic response to intravenous infusions of glucocerebrosidase in a patient with Gaucher disease. Proc. Natl. Acad. Sci. U. S. A. 87, 1913–1916.

Barton, N.W., Brady, R.O., Dambrosia, J.M., Di Bisceglie, A.M., Doppelt, S.H., Hill, S.C., Mankin, H.J., Murray, G.J., Parker, R.I., Argoff, C.E., et al., 1991. Replacement therapy for inherited enzyme deficiency—macrophage-targeted glucocerebrosidase for Gaucher's disease. N. Engl. J. Med. 324, 1464–1470.
Beck, M., 2010. Emerging drugs for lysosomal storage diseases. Expert Opin. Emerg. Drugs 15, 495–507.
Bielicki, J., Freeman, C., Clements, P.R., Hopwood, J.J., 1990. Human liver iduronate-2-sulphatase. Purification, characterization and catalytic properties. Biochem. J. 271, 75–86.
Boustany, R.M., 2013. Lysosomal storage diseases—the horizon expands. Nat. Rev. Neurol. 9, 583–598.
Brady, R.O., 2006. Enzyme replacement for lysosomal diseases. Annu. Rev. Med. 57, 283–296.
Brady, R.O., Pentchev, P.G., Gal, A.E., Hibbert, S.R., Dekaban, A.S., 1974. Replacement therapy for inherited enzyme deficiency: use of purified glucocerebrosidase in Gaucher's disease. N. Engl. J. Med. 291, 989–993.
Brante, G., 1952. Gargoylism; a mucopolysaccharidosis. Scand. J. Clin. Lab. Invest. 4, 43–46.
Brot, F.E., Bell Jr., C.E., Sly, W.S., 1978. Purification and properties of beta-glucuronidase from human placenta. Biochemistry 17, 385–391.
Cantz, M., Kresse, H., 1974. Sandhoff disease: defective glycosaminoglycan catabolism in cultured fibroblasts and its correction by beta-N-acetylhexosaminidase. Eur. J. Biochem. 47, 581–590.
Cataldo, A.M., Barnett, J.L., Berman, S.A., Li, J., Quarless, S., Bursztajn, S., Lippa, C., Nixon, R.A., 1995. Gene expression and cellular content of cathepsin D in Alzheimer's disease brain: evidence for early up-regulation of the endosomal–lysosomal system. Neuron 14, 671–680.
Clements, P.R., Brooks, D.A., Saccone, G.T., Hopwood, J.J., 1985a. Human alpha-L-iduronidase. 1. Purification, monoclonal antibody production, native and subunit molecular mass. Eur. J. Biochem. 152, 21–28.
Clements, P.R., Muller, V., Hopwood, J.J., 1985b. Human alpha-L-iduronidase. 2. Catalytic properties. Eur. J. Biochem. 152, 29–34.
Coutinho, M.F., Lacerda, L., Alves, S., 2012. Glycosaminoglycan storage disorders: a review. Biochem. Res. Int. 2012, 471325.
Dahms, N.M., Lobel, P., Kornfeld, S., 1989. Mannose 6-phosphate receptors and lysosomal enzyme targeting. J. Biol. Chem. 264, 12115–12118.
de Duve, C., 1964. From cytases to lysosomes. Fed. Proc. 23, 1045–1049.
de Duve, C., Wattiaux, R., 1966. Functions of lysosomes. Annu. Rev. Physiol. 28, 435–492.
de Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., Appelmans, F., 1955. Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. Biochem. J. 60, 604–617.
Desnick, R.J., Bernlohr, R.W., Krivit, W., 1973. Enzyme therapy in genetic diseases. Preface. Birth Defects Orig. Artic Ser. 9, viii-x.
Desnick, R.J., Thorpe, S.R., Fiddler, M.B., 1976. Toward enzyme therapy for lysosomal storage diseases. Physiol. Rev. 56, 57–99.
Di Natale, P., Daniele, A., 1985. Iduronate sulfatase from human placenta. Biochim. Biophys. Acta 839, 258–261.
Distler, J.J., Jourdian, G.W., 1978. Beta-galactosidase from bovine testes. Methods Enzymol. 50, 514–520.
Dorfman, A., Lorincz, A.E., 1957. Occurrence of urinary acid mucopolysaccharides in the Hurler syndrome. Proc. Natl. Acad. Sci. U. S. A. 43, 443–446.
Engel, D., 1939. Dysostosis multiplex: Pfaundler–Hurler syndrome. Arch. Dis. Child. 14, 217–230.
Fabry, J., 1898. Ein Beitrag zur Kenntniss der Purpura haemorrhagica nodularis (*Purpura papulosa haemorrhagica Hebrae*). Arch. Dermatol. Res. 43, 187–200.
Fan, X., Zhang, H., Zhang, S., Bagshaw, R.D., Tropak, M.B., Callahan, J.W., Mahuran, D.J., 2006. Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). Am. J. Hum. Genet. 79, 738–744.
Filocamo, M., Morrone, A., 2011. Lysosomal storage disorders: molecular basis and laboratory testing. Hum. Genomics 5, 156–169.
Fischer, H.D., Gonzalez-Noriega, A., Sly, W.S., Morré, D.J., 1980. Phosphomannosyl-enzyme receptors in rat liver. Subcellular distribution and role in intracellular transport of lysosomal enzymes. J. Biol. Chem. 255, 9608–9615.
Fratantoni, J.C., Hall, C.W., Neufeld, E.F., 1968. The defect in Hurler's and Hunter's syndromes: faulty degradation of mucopolysaccharide. Proc. Natl. Acad. Sci. U. S. A. 60, 699–706.
Fratantoni, J.C., Hall, C.W., Neufeld, E.F., 1969. The defect in Hurler and Hunter syndromes. II. Deficiency of specific factors involved in mucopolysaccharide degradation. Proc. Natl. Acad. Sci. U. S. A. 64, 360–366.
Freeman, C., Hopwood, J.J., 1986. Human liver sulphamate sulphohydrolase. Determinations of native protein and subunit Mr values and influence of substrate agylcone structure on catalytic properties. Biochem. J. 234, 83–92.
Freeman, C., Clements, P.R., Hopwood, J.J., 1983. Acetyl CoA:alpha-glucosaminide N-acetyl transferase: partial purification from human liver. Biochem. Int. 6, 663–671.
Freeman, C., Clements, P.R., Hopwood, J.J., 1987. Human liver N-acetylglucosamine-6-sulphate sulphatase. Purification and characterization. Biochem. J. 246, 347–354.
Garrod, A.E., 1909. Inborn Errors of Metabolism: The Croonian Lectures Delivered Before the Royal College of Physicians of London in June 1908. Frowde; Hodder and Stoughton, London.
Gaucher, P., 1882. De l'epithelioma primitif de la rateMD thesis Faculté de Médicine, Paris.
Ghosh, P., Dahms, N.M., Kornfeld, S., 2003. Mannose 6-phosphate receptors: new twists in the tale. Nat. Rev. Mol. Cell Biol. 4, 202–212.
Gibson, G.J., Saccone, G.T., Brooks, D.A., Clements, P.R., Hopwood, J.J., 1987. Human N-acetylgalactosamine-4-sulphate sulphatase. Purification, monoclonal antibody production and native and subunit Mr values. Biochem. J. 248, 755–764.
Gitzelmann, R., Steinmann, B., Wiesmann, U., Spycher, M., Herschkowitz, N., Marti, H.-R., 1987. Aldersche Granulationsanomalie: Albert Alders Patienten litten nicht an M. Pfaundler–Hurler. (Abstract). Helv. Paediatr. Acta 42, 90.
Giugliani, 2012. The mucopolysaccharidoses, In: Metha, A., Winchester, B. (Eds.), Lysosomal Storage Disorders — A Practical Guide, 1th edition Wiley-Blackwell (John Wiley & Sons, Ltd), pp. 94–100.
Glössl, J., Truppe, W., Kresse, H., 1979. Purification and properties of N-acetylgalactosamine 6-sulphate sulphatase from human placenta. Biochem. J. 181, 37–46.
Gritti, A., 2011. Gene therapy for lysosomal storage disorders. Expert. Opin. Biol. Ther. 11, 1153–1167.
Hanai, J., Leroy, J., O'Brien, J.S., 1971. Ultrastructure of cultured fibroblasts in I-cell disease. Am. J. Dis. Child. 122, 34–38.
Harris, R.C., 1961. Mucopolysaccharide disorder: a possible new genotype of Hurler's syndrome. Am. J. Dis. Child. 102, 741.
Hasilik, A., Neufeld, E.F., 1980. Biosynthesis of lysosomal enzymes in fibroblasts. Phosphorylation of mannose residues. J. Biol. Chem. 255, 4946–4950.
Hasilik, A., Wrocklage, C., Schröder, B., 2009. Intracellular trafficking of lysosomal proteins and lysosomes. Int. J. Clin. Pharmacol. Ther. 47, S18–S33.
Hers, H.G., 1963. Alpha-glucosidase deficiency in generalized glycogenstorage disease (Pompe's disease). Biochem. J. 86, 11–16.
Hickman, S., Neufeld, E.F., 1972. A hypothesis for I-cell disease: defective hydrolases that do not enter lysosomes. Biochem. Biophys. Res. Commun. 49, 992–999.
Hickman, S., Shapiro, L.J., Neufeld, E.F., 1974. A recognition marker required for uptake of a lysosomal enzyme by cultured fibroblasts. Biochem. Biophys. Res. Commun. 57, 55–61.
Himeno Mnishimura, Y., Tsuji, H., Kato, K., 1976. Purification and characterization of microsomal and lysosomal beta-glucuronidase from rat liver by use of immunoaffinity chromatography. Eur. J. Biochem. 70, 349–359.
Hopwood, J., 2012. Genetics of lysosomal storage disorders and counselling, In: Metha, A., Winchester, B. (Eds.), Lysosomal Storage Disorders — A Practical Guide, 1th edition Wiley-Blackwell (John Wiley & Sons, Ltd), pp. 29–36.
Hrebícek, M., Mrázová, L., Seyrantepe, V., Durand, S., Roslin, N.M., Nosková, L., Hartmannová, H., Ivánek, R., Cízkova, A., Poupetová, H., Sikora, J., Urinovská, J., Stranecký, V., Zeman, J., Lepage, P., Roquis, D., Verner, A., Ausseil, J., Beesley, C.E., Maire, I., Poorthuis, B.J., van de Kamp, J., van Diggelen, O.P., Wevers, R.A., Hudson, T.J., Fujiwara, T.M., Majewski, J., Morgan, K., Kmoch, S., Pshezhetsky, A.V., 2006. Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Am. J. Hum. Genet. 79, 807–819.
Hunter, C., 1917. A rare disease in two brothers. Proc. R. Soc. Med. 10, 104–116.
Kaplan, A., Achord, D.T., Sly, W.S., 1977. Phosphohexosyl components of a lysosomal enzyme are recognized by pinocytosis receptors on human fibroblasts. Proc. Natl. Acad. Sci. U. S. A. 74, 2026–2030.
Karageorgos, L., Brooks, D.A., Pollard, A., Melville, E.L., Hein, L.K., Clements, P.R., Ketteridge, D., Swiedler, S.J., Beck, M., Giugliani, R., Harmatz, P., Wraith, J.E., Guffon, N., Leão Teles, E., Sá Miranda, M.C., Hopwood, J.J., 2007. Mutational analysis of 105 mucopolysaccharidosis type VI patients. Hum. Mutat. 28, 897–903.
Kaushik, S., Cuervo, A.M., 2012. Chaperone-mediated autophagy: a unique way to enter the lysosome world. Trends Cell Biol. 22, 407–417.
Kornfeld, S., 1990. Lysosomal enzyme targeting. Biochem. Soc. Trans. 18, 367–374.
Kornfeld, S., 1992. Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors. Annu. Rev. Biochem. 61, 307–330.
Koskenoja, M., Suvanto, E., 1959. Gargoylism: report of adult form with glaucoma in two sisters. Acta Ophthalmol. 37, 234–240.
Kowalewski, B., Lamanna, W.C., Lawrence, R., Damme, M., Stroobants, S., Padva, M., Kalus, I., Frese, M.A., Lübke, T., Lüllmann-Rauch, R., D'Hooge, R., Esko, J.D., Dierks, T., 2012. Arylsulfatase G inactivation causes loss of heparan sulfate 3-O-sulfatase activity and mucopolysaccharidosis in mice. Proc. Natl. Acad. Sci. U. S. A. 109, 10310–10315.
Kresse, H., Neufeld, E.F., 1972. The Sanfilippo A corrective factor: purification and mode of action. J. Biol. Chem. 247, 2164–2170.
Kresse, H., Wiesmann, U., Cantz, M., Hall, C.W., Neufeld, E.F., 1971. Biochemical heterogeneity of the Sanfilippo syndrome: preliminary characterization of two deficient factors. Biochem. Biophys. Res. Commun. 42, 892–898.
Kresse, H., Von Figura, K., Bartsocas, C., 1976. Clinical and biochemical findings in a family with Sanfilippo disease, type C. (abstract). Clin. Genet. 10, 364.
Kresse, H., Paschke, E., von Figura, K., Gilberg, W., Fuchs, W., 1980. Sanfilippo disease type D: deficiency of N-acetylglucosamine-6-sulfate sulfatase required for heparan sulfate degradation. Proc. Natl. Acad. Sci. 77, 6822–6826.
Lorincz, A.E., 1958. Acid mucopolysaccharides in the Hurler syndrome (abstract). Fed. Proc. 17, 266.
Ludwig, T., Le Borgne, R., Hoflack, B., 1995. Roles for mannose-6-phosphate receptors in lysosomal enzyme sorting, IGF-II binding and clathrin-coat assembly. Trends Cell Biol. 5, 202–206.
Maroteaux, P., Leveque, B., Marie, J., Lamy, M., 1963. Une nouvelle dysostose avec elimination urinaire de chondroitine-sulfate B. Presse Med. 71, 1849–1852.
Masue, M., Sukegawa, K., Orii, T., Hashimoto, T., 1991. N-acetylgalactosamine-6-sulfate sulfatase in human placenta: purification and characteristics. J. Biochem. 110, 965–970.
McGovern, M.M., Vine, D.T., Haskins, M.E., Desnick, R.J., 1982. Purification and properties of feline and human arylsulfatase B isozymes. Evidence for feline homodimeric and human monomeric structures. J. Biol. Chem. 257, 12605–12610.
McKusick, V.A., 1972. The mucopolysaccharidoses, Heritable Disorders of Connective Tissue4th ed. C. V. Mosby (pub.), St. Louis, pp. 556–574.
Micsenyi, M.C., Walkley, S.U., 2012. The lysosomal system: physiology and pathology, In: Metha, A., Winchester, B. (Eds.), Lysosomal Storage Disorders — A Practical Guide, 1th edition Wiley-Blackwell (John Wiley & Sons, Ltd).
Miller, R.D., Hoffmann, J.W., Powell, P.P., Kyle, J.W., Shipley, J.M., Bachinsky, D.R., Sly, W.S., 1990. Cloning and characterization of the human beta-glucuronidase gene. Genomics 7, 280–283.
Mok, A., Cao, H., Hegele, R.A., 2003. Genomic basis of mucopolysaccharidosis type IIID (MIM 252940) revealed by sequencing of GNS encoding N-acetylglucosamine-6-sulfatase. Genomics 81, 1–5.

Morris, C.P., Guo, X.H., Apostolou, S., Hopwood, J.J., Scott, H.S., 1994. Morquio A syndrome: cloning, sequence, and structure of the human N-acetylgalactosamine 6-sulfatase (GALNS) gene. Genomics 22, 652–654.
Nakashima, Y., Tomatsu, S., Hori, T., Fukuda, S., Sukegawa, K., Kondo, N., Suzuki, Y., Shimozawa, N., Orii, T., 1994. Mucopolysaccharidosis IV A: molecular cloning of the human N-acetylgalactosamine-6-sulfatase gene (GALNS) and analysis of the 5'-flanking region. Genomics 20, 99–104.
Natowicz, M.R., Short, M.P., Wang, Y., Dickersin, G.R., Gebhardt, M.C., Rosenthal, D.I., Sims, K.B., Rosenberg, A.E., 1996. Clinical and biochemical manifestations of hyaluronidase deficiency. N. Engl. J. Med. 335, 1029–1033.
Nelson, J., Broadhead, D., Mossman, J., 1988. Clinical findings in 12 patients with MPS IV A (Morquio's disease): further evidence for heterogeneity. Part I: clinical and biochemical findings. Clin. Genet. 33, 111–120.
Neufeld, E.F., 2006. Enzyme replacement therapy — a brief history. In: Mehta, A., Beck, M., Sunder-Plassmann, G. (Eds.), Fabry Disease: Perspectives From 5 Years of FOS. Oxford PharmaGenesis, Oxford (Chapter 10).
Neufeld, E.F., Muenzer, J., 2001. The mucopolysaccharidoses, In: Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D. (Eds.), The Metabolic & Molecular Bases of Inherited Disease, 8th edition vol. 3. McGraw-Hill, New York, NY, USA.
Niemann, A., 1914. Ein unbekanntes Krankheitsbild. Jahrb. Kinderheilkd. 79, 1–10.
Norman, R.M., Urich, H., Rance, N.E., 1959. Perivascular cavitation of the basal ganglia in gargoylism. J. Ment. Sci. 195, 1970–1977.
O'Brien, J.S., Miller, A.L., Loverde, A.W., Veath, M.L., 1973. Sanfilippo disease type B: enzyme replacement and metabolic correction in cultured fibroblasts. Science 181, 753–755.
Ohshita, T., Sakuda, H., Nakasone, S., Iwamasa, T., 1989. Purification, characterization and subcellular localization of pig liver alpha-L-iduronidase. Eur. J. Biochem. 179, 201–207.
Oshima, A., Tsuji, A., Nagao, Y., Sakuraba, H., Suzuki, Y., 1988. Cloning, sequencing, and expression of cDNA for human beta-galactosidase. Biochem. Biophys. Res. Commun. 157, 238–244.
Parkinson-Lawrence, E.J., Shandala, T., Prodoehl, M., et al., 2010. Lysosomal storage disease: revealing lysosomal function and physiology. Physiology (Bethesda) 25, 102–115.
Porter, M.T., Fluharty, A.L., Kihara, H., 1971. Correction of abnormal cerebroside sulfate metabolism in cultured metachromatic leukodystrophy fibroblasts. Science 172, 1263–1265.
Rome, L.H., Garvin, A.J., Neufeld, E.F., 1978. Human kidney alpha-L-iduronidase: purification and characterization. Arch. Biochem. Biophys. 189, 344–353.
Saftig, P., 2006. Physiology of the lysosome. In: Mehta, A., Beck, M., Sunder-Plassmann, G. (Eds.), Fabry Disease: Perspectives From 5 Years of FOS. Oxford PharmaGenesis, Oxford (Chapter 3).
Sando, G.N., Neufeld, E.F., 1977. Recognition and receptor-mediated uptake of a lysosomal enzyme, alpha-l-iduronidase, by cultured human fibroblasts. Cell 12, 619–627.
Sanfilippo, S.J., Podosin, R., Langer, L., Good, R.A., 1963. Mental retardation associated with acid mucopolysaccharidiuria (heparitin sulfate type). J. Pediatr. 63, 837–838.
Sasaki, T., Sukegawa, K., Masue, M., Fukuda, S., Tomatsu, S., Orii, T., 1991. Purification and partial characterization of alpha-N-acetylglucosaminidase from human liver. J. Biochem. 110, 842–846.
Scheie, H.G., Hambrick Jr., G.W., Barness, L.A., 1962. A newly recognized forme fruste of Hurler's disease (gargoylism). Am. J. Ophthalmol. 53, 753–769.
Schuchman, E.H., Guzman, N.A., Desnick, R.J., 1984a. Human alpha-L-iduronidase. I. Purification and properties of the high uptake (higher molecular weight) and the low uptake (processed) forms. J. Biol. Chem. 259, 3132–3140.
Schuchman, E.H., Guzman, N.A., Takada, G., Desnick, R.J., 1984b. Human alpha-L-iduronidase. II. Comparative biochemical and immunologic properties of the purified low and high uptake forms. Enzyme 31, 166–175.
Schultz, M.L., Tecedor, L., Chang, M., Davidson, B.L., 2011. Clarifying lysosomal storage diseases. Trends Neurosci. 34, 401–410.
Scott, H.S., Guo, X.H., Hopwood, J.J., Morris, C.P., 1992. Structure and sequence of the human alpha-L-iduronidase gene. Genomics 13, 1311–1313.
Scott, H.S., Blanch, L., Guo, X.H., Freeman, C., Orsborn, A., Baker, E., Sutherland, G.R., Morris, C.P., Hopwood, J.J., 1995. Cloning of the sulphamidase gene and identification of mutations in Sanfilippo A syndrome. Nat. Genet. 11, 465–467.
Sly, W.S., Quinton, B.A., McAlister, W.H., Rimoin, D.L., 1973. Beta-glucuronidase deficiency: report of clinical, radiologic and biochemical features of a new mucopolysaccharidosis. J. Pediatr. 82, 249–257.
Stahl, P.D., Touster, O., 1971. Beta-glucuronidase of rat liver lysosomes. Purification, properties, subunits. J. Biol. Chem. 246, 5398–5406.
Stoltzfus, L.J., Sosa-Pineda, B., Moskowitz, S.M., Menon, K.P., Dlott, B., Hooper, L., Teplow, D.B., Shull, R.M., Neufeld, E.F., 1992. Cloning and characterization of cDNA encoding canine alpha-L-iduronidase. mRNA deficiency in mucopolysaccharidosis I dog. J. Biol. Chem. 267, 6570–6575.
Straus, W., 1954. Isolation and biochemical properties of droplets from the cells of rat kidney. J. Biol. Chem. 207, 745–755.
Straus, W., 1957. Segregation of an intravenously injected protein by droplets of the cells of rat kidneys. J. Biophys. Biochem. Cytol. 3, 1037–1040.
Tay, W., 1881. Symmetrical changes in the region of the yellow apot in each eye of an infant. Trans. Ophthalmol. Soc. U. K. 1, 15–57.
Tettamanti, G., Bassi, R., Viani, P., Riboni, L., 2003. Salvage pathways in glycosphingolipid metabolism. Biochemie 85, 432–437.
Triggs-Raine, B., Salo, T.J., Zhang, H., Wicklow, B.A., Natowicz, M.R., 1999. Mutations in HYAL1, a member of a tandemly distributed multigene family encoding disparate hyaluronidase activities, cause a newly described lysosomal disorder, mucopolysaccharidosis IX. Proc. Natl. Acad. Sci. U. S. A. 96, 6296–6300.
Van Hoof, F., Hers, H.G., 1964. Ultrastructure of the hepatic cells in Hurler's disease (gargoylism). C. R. Hebdomadaires Seances Acad. Sci. 259, 1281–1283.
Van Pelt, J.F., 1960. GargoylismThesis Nijmegen (pub.).
Varki, A., Kornfeld, S., 1980. Structural studies of phosphorylated high mannose-type oligosaccharides. J. Biol. Chem. 255, 10847–10858.
Vellodi, A., 2005. Lysosomal storage disorders. Br. J. Haematol. 128, 413–431.
von Figura, K., 1977. Human alpha-N-acetylglucosaminidase. 1. Purification and properties. Eur. J. Biochem. 80, 523–533.
von Figura, K., 1991. Molecular recognition and targeting of lysosomal proteins. Curr. Opin. Cell Biol. 3, 642–646.
Walkley, S.U., 2009. Pathogenic cascades in lysosomal disease—why so complex? J. Inherit. Metab. Dis. 32, 181–189.
Walkley, S.U., Vanier, M.T., 2009. Secondary lipid accumulation in lysosomal disease. Biochim. Biophys. Acta 1793, 726–736.
Wasteson, A., Neufeld, E.F., 1982. Iduronate sulfatase from human plasma. Methods Enzymol. 83, 573–578.
Wilson, P.J., Meaney, C.A., Hopwood, J.J., Morris, C.P., 1993. Sequence of the human iduronate 2-sulfatase (IDS) gene. Genomics 17, 773–775.
Winchester, B., 2012a. Classification of lysosomal storage diseases, In: Metha, A., Winchester, B. (Eds.), Lysosomal Storage Disorders — A Practical Guide, 1th edition Wiley-Blackwell (John Wiley & Sons, Ltd).
Winchester, B., 2012b. Laboratory diagnosis of lysosomal storage diseases, In: Metha, A., Winchester, B. (Eds.), Lysosomal Storage Disorders — A Practical Guide, 1th edition Wiley-Blackwell (John Wiley & Sons, Ltd).
Zhao, H.G., Li, H.H., Bach, G., Schmidtchen, A., Neufeld, E.F., 1996. The molecular basis of Sanfilippo syndrome type B. Proc. Natl. Acad. Sci. U. S. A. 93, 6101–6105.

Glossary

Lysosomal Storage Disorders (LSDs): group of nearly 60 inborn errors of metabolism caused by defects in lysosomal function. The rate of incidence of these genetically inherited disorders is estimated to be around 1 in each 4000–9000 live births.
Lysosome: first described by Christian de Duve, in 1955, as a particle containing enzymes with lytic properties. It is a membrane-enclosed compartment containing over 60 hydrolytic enzymes that digest macromolecules. Presently, the lysosome is known to play a crucial role in a series of processes in addition to cellular digestion, including phagocytosis, endocytosis, autophagy, salvage and ubiquitin–proteasome degradation.
Mucopolysaccharidoses (MPSs): subgroup of LSD caused by the absence or malfunctioning of the lysosomal enzymes responsible for breaking down glycosaminoglycans (formerly called mucopolysaccharides) — long chains of sugar carbohydrates involved in the processes of building bone, cartilage, tendons, corneas, skin and connective tissue.
Enzyme replacement therapy (ERT): medical treatment that consists of replacing the defective enzyme in patients suffering from particular storage diseases. This is usually done by giving the patient an intravenous infusion containing the recombinant enzyme. ERT is currently available for some LSDs. This treatment does not affect the underlying genetic defect, but increases the concentration of enzyme in which the patient is deficient.