

MSc

2.º
CICLO

FCUP
INSA

U.
PORTO

Development of an antisense-mediated exon
skipping therapeutic strategy for Mucopolipidosis II

Regina Filipa Machado Vilela

FC

U. PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Instituto Nacional de Saúde
Doutor Ricardo Jorge



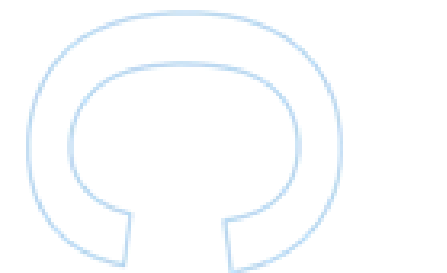
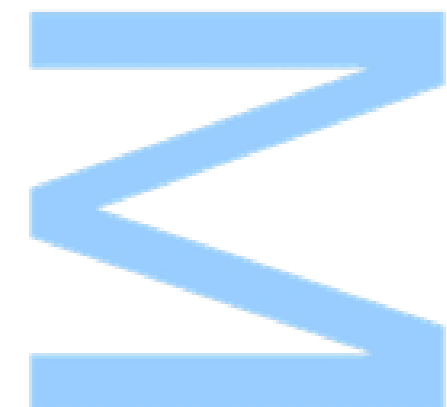
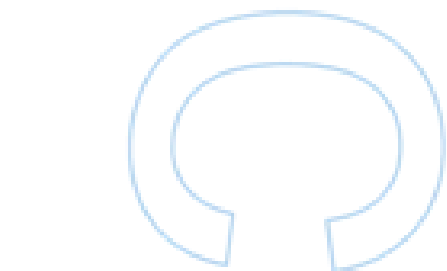
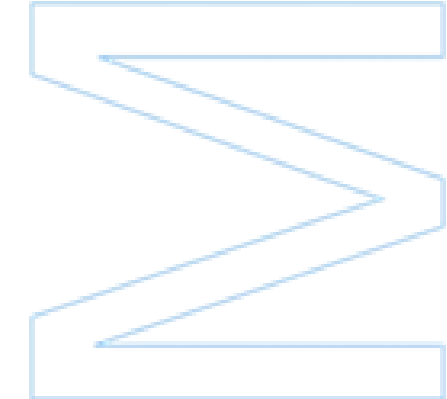
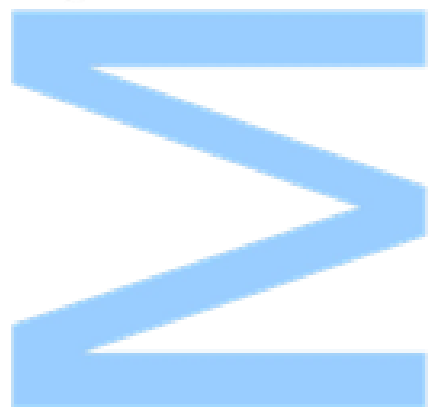
Development of an antisense-mediated exon skipping therapeutic strategy for Mucopolipidosis II

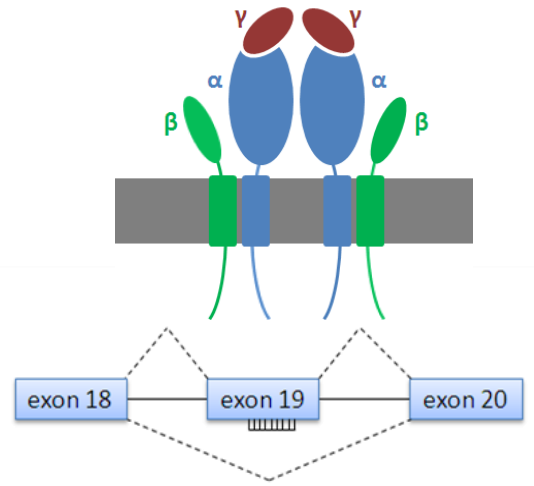
Regina Filipa Machado Vilela

Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto em
Biologia Celular e Molecular

2017

U. PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO





Development of an antisense-mediated exon skipping therapeutic strategy for Mucopolipidosis II

Regina Filipa Machado Vilela

Mestrado em Biologia Celular e Molecular

Departamento de Biologia

2017

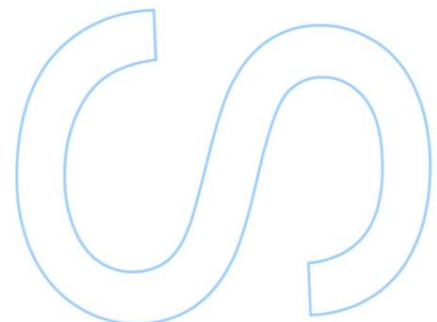
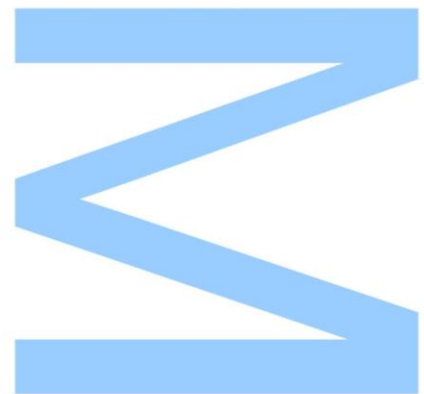
Orientadora

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

Coorientadoras

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

Doutora Liliana da Silva Matos,
Investigadora Pós-Doutoramento,
Instituto Nacional de Saúde Doutor Ricardo Jorge



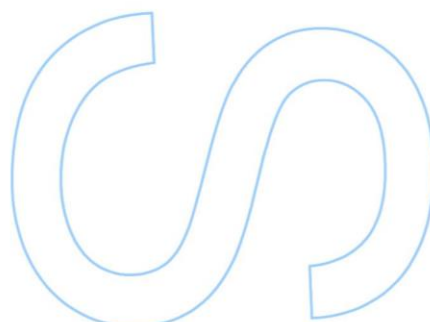
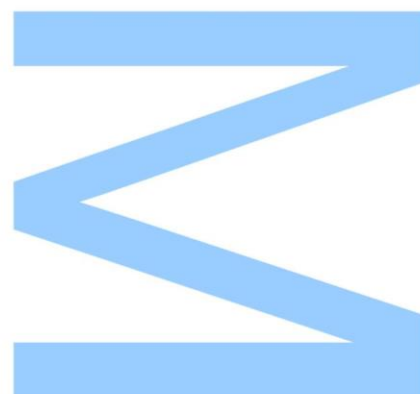
Instituto **Nacional de Saúde**
Doutor Ricardo Jorge



Todas as correções determinadas
pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____ / ____ / ____



Eu, Regina Filipa Machado Vilela, aluna com o número 201201434 do Mestrado de Biologia Celular e Molecular da edição de 2016/2017, declaro que sou a autora da totalidade do trabalho experimental descrito, assim como do texto apresentado, não apresentando texto plagiado e tendo tomado conhecimento das consequências de uma situação de plágio.

Assinatura

Dissertação apresentada à Faculdade de
Ciências da Universidade do Porto
para candidatura ao grau de Mestre em
Biologia Celular e Molecular.

Dissertation presented to the Faculty of
Sciences of the University of Porto
for application for the Master's Degree
in Cell and Molecular Biology.

Agradecimentos

Porque este trabalho é fruto da participação de várias pessoas, gostaria de aproveitar esta oportunidade para expressar o meu agradecimento a todos que de alguma forma contribuíram para que este projeto e toda a minha formação fossem possíveis:

À Doutora Sandra Alves, minha orientadora, por me ter aceitado no seu grupo de investigação e por toda a dedicação ao longo deste ano, permitindo que desenvolvesse este projeto da melhor forma possível e melhorasse os meus conhecimentos e capacidades na área da Biologia Molecular. Obrigada por todas as sugestões e conselhos e por sempre se ter preocupado comigo e com o meu percurso académico.

À Professora Doutora Maria João Prata, minha co-orientadora, por me ter inicialmente apresentado à Doutora Sandra e por se ter mostrado sempre disponível para me ajudar em todas as etapas deste Mestrado, especialmente na elaboração desta dissertação. Obrigada também por todo o conhecimento transmitido durante as aulas das diversas disciplinas que lecionou e que em muito contribuíram para a minha escolha desta área de estudo.

À Liliana, também minha co-orientadora, por me ter sempre acompanhado a nível laboratorial, por me ter ensinado todas as técnicas necessárias para a realização deste projeto e por estar sempre disponível para esclarecer qualquer dúvida que tivesse. Um obrigada muito especial por toda a confiança e por me deixares aprender ao meu próprio ritmo e sem pressões.

A todo o grupo de investigação em Doenças Lisossomais de Sobrecarga do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), por me terem acolhido da melhor forma possível e por me terem sempre apoiado ao longo de todo este percurso. Obrigada por todos os conselhos e partilha de experiências que tanto me ajudaram a crescer ao longo deste ano, tanto a nível pessoal como profissional.

Ao Paulo Gaspar, por aceitar contribuir para este projeto ao nível da realização dos ensaios enzimáticos e por se mostrar sempre disponível para esclarecer qualquer dúvida sobre este assunto.

Ao Professor Doutor José Pissarra e à Professora Doutora Conceição Santos, diretores do Mestrado de Biologia Celular e Molecular nestes dois anos letivos, assim como a todos os professores deste Mestrado, por todo o esforço e dedicação e pelo apoio dado a todos os alunos ao longo deste percurso. Obrigada por todo o conhecimento transmitido e por nos ajudarem a tornarmo-nos melhores cientistas e melhores pessoas.

A todos os professores que de alguma forma fizeram parte da minha vida escolar e académica, obrigada por terem contribuído para a pessoa que sou hoje e por me ajudarem a alargar os meus horizontes cada vez mais. Um agradecimento especial à educadora Sara e à professora Cecília que foram incansáveis nos meus primeiros anos e tanto se esforçaram por compreender e adaptar os programas escolares às minhas necessidades para que nunca perdesse a curiosidade e a vontade de aprender sempre mais.

À Dra. Edna Carvalho, minha psicóloga, um obrigada muito especial por me ter ajudado a ultrapassar medos e ansiedades e por ter contribuído para que hoje seja uma pessoa mais confiante e feliz. Lembrarei sempre os seus conselhos com carinho e amizade.

A toda a minha família, por todo o incentivo e força que sempre me deram e por terem sido as pessoas mais orgulhosas de todas as minhas conquistas. Um agradecimento especial aos meus dois avôs, Afonso e Adelino, por se revezarem em inúmeras boleias de casa para o comboio e do comboio para casa, permitindo que continuasse a morar junto da minha família apesar de estudar a várias dezenas de quilómetros de casa. Sem o vosso esforço, tudo isto não teria sido possível.

E por último, aos primeiros, os meus pais, Alda e Pedro. Eles que sempre acreditaram nas minhas capacidades e me apoiaram em todas as minhas escolhas a nível académico, incentivando-me a dar sempre o meu máximo e tentar fazer mais e melhor. Obrigada por inventarem mil e um planos em família para me distrair quando estava mais stressada e ansiosa e obrigada por me ensinarem a ver sempre o lado positivo de todas as coisas. O meu mais sincero obrigada por todo o vosso amor e carinho desde o meu primeiro dia e por estarem sempre presentes em todos os momentos da minha vida. E por fim, obrigada por me ensinarem que o mais importante da vida é simplesmente ser feliz.

Abstract

Lysosomal Storage Disorders (LSDs) are a group of rare inherited metabolic diseases caused by the malfunction of the lysosomal system, resulting in the accumulation of undegraded substrates inside the lysosomes and leading to severe and progressive pathology. Among them is Mucopolipidosis II (ML II), one of the most severe LSDs, which is caused by the total or near total deficiency of the GlcNAc-1-phosphotransferase, a key enzyme for the correct trafficking of lysosomal hydrolases to the lysosome. One of the most frequent ML II causal mutations is a dinucleotide deletion on exon 19 of the *GNPTAB* gene (c.3503_3504delTC), which disrupts the reading frame and prevents the production of an active GlcNAc-1-phosphotransferase. Despite broad understanding of the molecular causes behind this and other LSDs, the same progress has not been observed in the development of new therapies, with current treatments being mostly symptomatic and presenting several limitations.

In recent years, the RNA molecule became one of the most promising targets for therapeutic intervention in diseases associated with genetic mutations. One strategy that has been developed involves the use antisense oligonucleotides (AOs) to induce the skipping of specific exons bearing disease causing mutations with the purpose of altering the splicing pattern, the mature RNA and ultimately the final protein product. This has previously been applied to several disorders with encouraging results but has yet to be thoroughly explored for the treatment of LSDs and specifically ML II.

Acknowledging this, this project intended to design and develop an RNA-based therapeutic approach through the use of AOs capable of inducing the skipping of exon 19 of the *GNPTAB* gene and consequently circumvent the effects of the most common ML II causal mutation and ultimately ameliorate patients' phenotype.

An initial *in silico* analysis was conducted to study the effects of this therapy on the *GNPTAB* transcript and protein product to ensure it would result in an in-frame protein and that no important domains would be removed. The target sequence was also explored to select the most promising locations for AOs aimed at inducing the skipping of exon 19 and four molecules were designed. ML II patient fibroblasts were then transfected with the selected AOs at different concentrations and their *GNPTAB* transcripts were analyzed. All four molecules produced successful exon-skipping to some extent, with the best results observed with the AO directed at a region of exonic enhancers in the second half of the exon and the one targeting the 5' splice site. These two molecules were then tested at higher doses and the first proved to be the most effective, generating large amounts of the transcript

missing exon 19 and also reducing the levels of full-length transcript at the highest concentrations.

Following RNA analysis, and because this enzyme is involved in the transport of lysosomal hydrolases to the lysosome, we evaluated the effects of this therapeutic approach at biochemical level by testing the effects of the most efficient AO on the levels of activity of some of these enzymes inside fibroblasts. Fibroblasts were again transfected with this molecule and enzymatic assays were conducted to measure lysosomal enzyme activity. However, no significant changes were observed. Still, this is merely an indirect measurement of the GlcNAc-1-phosphotransferase activity and therefore it is not possible by now to affirm whether or not its activity was restored, at least partially, after removal of exon 19.

Globally, this work explored the possibility of an innovative therapeutic strategy for ML II through the skipping of the mutation harboring exon 19 of the *GNPTAB* gene and confirmed the ability of these AOs for RNA splicing modulation. Further studies are needed to verify the effect of this approach on the levels of GlcNAc-1-phosphotransferase activity and in the amelioration of the patients' phenotype, but these results provide the basis for future research regarding the treatment of this disorder through the targeting of the aberrant transcript and may ultimately contribute to better quality of life for ML II patients and their families.

Keywords: Lysosomal Storage Disorders; Mucopolipidosis II; GlcNAc-1-phosphotransferase; splicing modulation; antisense oligonucleotides; exon skipping

Resumo

As Doenças Lisossomais de Sobrecarga (DLS) são um grupo de doenças metabólicas hereditárias raras, crônicas e progressivas causadas por um mau funcionamento do sistema lisossomal, resultando na acumulação de substratos não degradados no interior dos lisossomas. A Mucopolipidose tipo II (ML II) é uma das patologias mais graves deste grupo, sendo causada pela perda de atividade da GlcNAc-1-fosfotransferase, uma enzima essencial para o correto transporte das hidrolases lisossomais para o lisossoma. Uma das mutações mais frequentes em doentes com ML II é a deleção de um dinucleótido no exão 19 do gene *GNPTAB* (c.3503_3504delTC) que altera a grelha de leitura do mRNA e impede a produção de uma GlcNAc-1-fosfotransferase ativa, comprometendo assim o correto transporte das enzimas lisossomais. Apesar do vasto conhecimento das causas moleculares responsáveis por esta e outras DLS, o mesmo progresso não tem sido observado no desenvolvimento de novas terapias, sendo os tratamentos atualmente disponíveis principalmente sintomáticos e com várias limitações.

Nos últimos anos, a molécula de RNA tornou-se um alvo promissor para a intervenção terapêutica em inúmeras doenças genéticas. Uma das estratégias aplicadas recorre ao uso de oligonucleótidos *antisense* para induzir o *skipping* de exões contendo mutações patogénicas tendo por objetivo alterar o padrão de *splicing*, o mRNA maduro e por último o produto proteico. Esta abordagem foi já aplicada a outras doenças com resultados promissores mas ainda não foi explorada para o tratamento das DLS e mais especificamente da ML II.

O objetivo principal deste projeto foi, portanto, o desenvolvimento de uma abordagem terapêutica envolvendo a manipulação da molécula de RNA, através do uso de oligonucleótidos *antisense* capazes de induzir o *skipping* do exão 19 do gene *GNPTAB* e, consequentemente, evitar os efeitos desta mutação e melhorar o fenótipo dos doentes.

Inicialmente foi realizada uma análise *in silico* para estudar os efeitos desta terapia ao nível do transcrito e do produto proteico do gene *GNPTAB* para garantir que esta abordagem resultaria numa proteína *in-frame* e que não seriam removidos domínios relevantes. A sequência alvo foi também explorada de forma a escolher as localizações mais promissoras para os oligonucleótidos *antisense* com o objetivo de induzir o *skipping* do exão 19 e quatro moléculas foram selecionadas. Fibroblastos de doentes foram então transfetados com os oligonucleótidos *antisense* escolhidos a diferentes concentrações e os seus transcritos do gene *GNPTAB* foram analisados. Todas as quatro moléculas produziram o *skipping* do exão 19 com sucesso, sendo os melhores resultados observados com o

oligonucleótido dirigido à região de *enhancers* exónicos na segunda metade do exão e com a molécula direcionada ao 5' *splice site*. Estes dois oligonucleótidos *antisense* foram então testados a doses mais elevadas e o primeiro provou ser o mais eficiente, originando quantidades elevadas de transcrito sem o exão 19 e também reduzindo os níveis do transcrito completo nas concentrações superiores.

Após a análise de RNA, e uma vez que esta enzima participa na via de transporte das hidrolases lisossomais, foram avaliados os efeitos desta abordagem terapêutica ao nível bioquímico, testando os efeitos do oligonucleótido *antisense* mais eficiente nos níveis de atividade de algumas destas enzimas no interior dos fibroblastos. Os fibroblastos foram novamente transfectados com este oligonucleótido e foram realizados ensaios enzimáticos para quantificar a atividade de diversas enzimas lisossomais, não tendo sido observadas variações significativas. Contudo, estes ensaios são apenas uma análise indireta, não sendo por isso possível concluir se existe ou não recuperação, pelo menos parcial, da atividade da enzima GlcNAc-1-fosfotransferase após a remoção do exão 19.

Globalmente, este trabalho explorou a possibilidade de uma estratégia terapêutica inovadora para ML II através do *skipping* do exão 19 do gene *GNPTAB* e confirmou a capacidade destes oligonucleótidos *antisense* para a modulação do *splicing* de RNA. De futuro, são necessários mais estudos para verificar o impacto desta abordagem nos níveis de atividade da GlcNAc-1-fosfotransferase e na recuperação do fenótipo dos doentes com esta mutação. No entanto, os resultados obtidos oferecem uma base para novos estudos relativos ao tratamento desta doença através da manipulação do transcrito aberrante e podem, em última instância, contribuir para uma melhor qualidade de vida dos doentes com ML II e das suas famílias.

Palavras-chave: Doenças Lisossomais de Sobrecarga; Mucopolidose II; GlcNAc-1-fosfotransferase; modulação do *splicing*; oligonucleótidos *antisense*; *skipping* exónico

Table of contents

Agradecimentos	9
Abstract.....	11
Resumo.....	13
Table of contents.....	15
Table Index	19
Figure Index	21
Abbreviations	27
Chapter 1 – Introduction.....	31
1.1. Lysosomal Storage Disorders	33
1.1.1. Mucopolipidosis II.....	36
1.1.2. Biochemical diagnosis and molecular diagnosis	39
1.1.3. Available treatments for LSDs and ML II.....	40
1.2. Splicing therapies.....	41
1.2.1. Splicing reaction	41
1.2.2. Splicing signals and regulatory elements.....	42
1.2.3. Alternative Splicing	44
1.3. Antisense Oligonucleotides	45
1.3.1. Chemical modifications: three generations of oligonucleotides	45
1.3.2. Oligonucleotides for splicing modulation.....	47
Chapter 2 – Objectives	51
Chapter 3 – Materials and Methods	55
3.1. Bioinformatical analysis.....	57
3.1.1. Sequence alignment of <i>GNPTAB</i> protein products	57
3.1.2. Prediction of protein domains within the <i>GNPTAB</i> protein product	57
3.1.3. Prediction of glycosylation sites within the <i>GNPTAB</i> protein product	58
3.1.4. Scoring of 5' and 3' splice sites and prediction of enhancers within <i>GNPTAB</i> exon 19.....	58
3.2. Cell culture and establishment of stable fibroblast cell lines	58
3.3. Confirmation of the reported mutation in patient fibroblast cells	59

3.3.1. Extraction of genomic DNA.....	59
3.3.2. Amplification of exon 19 of the <i>GNPTAB</i> gene	60
3.3.3. Evaluation of PCR amplification by agarose gel electrophoresis.....	60
3.3.4. Sequencing of amplification products	60
3.4. Analysis of endogenous splicing pattern in patient fibroblast cells.....	61
3.4.1. Extraction of total RNA	61
3.4.2. Synthesis of cDNA.....	62
3.4.3. Amplification of <i>GNPTAB</i> transcript cDNA.....	62
3.5. Design and synthesis of antisense oligonucleotides for the skipping of <i>GNPTAB</i> exon 19.....	63
3.6. Transfection of healthy control and ML II patient fibroblasts with antisense oligonucleotides	65
3.6.1. Cell counting and plating	65
3.6.2. Transfection.....	66
3.7. Evaluation of <i>in vitro</i> <i>GNPTAB</i> gene exon 19 splicing pattern modulation with antisense oligonucleotides	67
3.8. Quantification of lysosomal enzyme activity in healthy control and ML II patient fibroblasts after splicing pattern modulation with antisense oligonucleotides.....	68
Chapter 4 – Results	71
4.1. Bioinformatical analysis.....	73
4.1.1. Sequence alignments of <i>GNPTAB</i> protein products	73
4.1.2. Prediction of protein domains within the <i>GNPTAB</i> protein product	75
4.1.3. Prediction of glycosylation sites within the <i>GNPTAB</i> protein product	78
4.1.4. Scoring of 5' and 3' splice sites of <i>GNPTAB</i> gene exon 19.....	81
4.1.5. Prediction of enhancers within <i>GNPTAB</i> gene exon 19	81
4.2. Confirmation of the presence of the <i>GNPTAB</i> gene exon 19 c.3503_3504delTC mutation in patient fibroblast cells	84
4.3. Analysis of <i>GNPTAB</i> gene exon 19 splicing pattern in fibroblast cells.....	85
4.4. Exon-skipping antisense oligonucleotide <i>in vitro</i> therapeutic approach for patients affected by ML II	88
4.4.1. Evaluation of transfection efficiency with <i>eGFP</i> plasmid in control fibroblasts.....	88
4.4.2. Splicing modulation of exon 19 with antisense oligonucleotides in healthy control and ML II patient fibroblasts	89
4.4.2.1. Experiments in healthy control fibroblasts	89
4.4.2.2. Experiments in ML II Patient 1 fibroblasts	92

4.4.2.3. Experiments in ML II Patient 2 fibroblasts	94
4.5. Quantification of the activity of several lysosomal enzymes in fibroblasts.....	97
4.5.1. Enzyme levels prior to <i>GNPTAB</i> exon 19 splicing modulation with antisense oligonucleotides	97
4.5.2. Enzyme levels following <i>GNPTAB</i> exon 19 splicing modulation with antisense oligonucleotides	98
Chapter 5 – Discussion	101
Chapter 6 – Conclusion and Future Perspectives	113
References.....	119
Appendix	137

Table Index

Table 1 – Classification of Lysosomal Storage Disorders (LSDs). Adapted from Filocamo & Morrone, 2011; Boustany, 2013.	34
Table 2 – Sequences and annealing temperatures of primers used in the cDNA amplification of <i>GNPTAB</i> gene.....	63
Table 3 – Sequences of antisense oligonucleotides selected for transfection in ML II patient fibroblasts. Nomenclature as described by Mann et al., 2002.	65
Table 4 – Composition of each enzyme activity assay and respective incubation period at 37°C.	69
Table 5 – Protein domains predicted by Pfam 30.0 for the <i>GNPTAB</i> protein product. n/a – not applied.....	75
Table 6 – Protein domains predicted by UniProt for the <i>GNPTAB</i> protein product.	76
Table 7 – Transmembrane domains predicted by UniProt for the <i>GNPTAB</i> protein product.	77
Table 8 – Metal binding sites predicted by UniProt for the <i>GNPTAB</i> protein product.....	77
Table 9 – Calcium binding region predicted by UniProt for the <i>GNPTAB</i> protein product.	77
Table 10 – Secondary structure of the <i>GNPTAB</i> protein predicted by UniProt software product.	78
Table 11 – <i>N</i> -glycosylation sites predicted by NetNGlyc 1.0 Server for the <i>GNPTAB</i> protein product.	79
Table 12 – <i>N</i> -glycosylation sites predicted by GlycoEP for the <i>GNPTAB</i> protein product. Only positions 922-1252 are shown.....	80

Table 13 – O-glycosylation sites predicted by GlycoEP for the *GNPTAB* protein product. Only positions 1066-1250 are shown.....80

Table 14 – Putative enhancers predicted by ESEfinder 3.0 for exon 19 of the *GNPTAB* gene containing the c.3503_3504delTC mutation and 50bp of the flanking introns. Positions start at the beginning of the input sequence and not the beginning of exon 19. 1-50: intron 18; 51-216: exon 19; 217-266: intron 19.....82

Table 15 – Exons of the *GNPTAB* transcript covered by each primer set and corresponding amplicon size. 86

Table 16 – Enzyme activity of lysosomal enzymes in fibroblasts of controls and LSDs patients. Due to small sample amount of the Fabry cell line, some enzymatic assays could not be performed for these fibroblasts. 97

Table 17 – Enzyme activity of lysosomal enzymes in fibroblasts of controls and ML II Patient 1 after transfection with the H19A(+119+138) AO. Plus sign (+) indicates the treatment with Lipofectamine® 2000 + AO, while minus sign (-) indicates the treatment with just Lipofectamine® 2000. 98

Figure Index

Figure 1 – Newly synthesized lysosomal enzymes are transported from the endoplasmic reticulum into the Golgi apparatus where a mannose-6-phosphate (M6P) marker is added through the action of two enzymes: GlcNAc-1-phosphotransferase transfers GlcNAc-1-P to mannose residues of the peptide, while *N*-acetylglucosamine-1-phosphodiester α -*N*-acetylglucosaminidase (“uncovering enzyme”) removes the GlcNAc residue to expose the M6P modifications. Subsequently, the lysosomal enzymes bind to M6P receptors and are targeted to the lysosomes. Adapted from Coutinho et al., 2012.....36

Figure 2 – Schematic representation of the GlcNAc-1-phosphotransferase. The enzyme is composed of two α subunits (blue), two β subunits (green) and two γ subunits (red). α and β subunits are responsible for the catalytic activity of the enzyme, while the γ subunits appear to facilitate and increase their action. Adapted from Pohl et al., 2011.....37

Figure 3 – Motifs and domains of the α/β -precursor protein obtained from the *GNPTAB* gene. Blue line indicates the position for cleavage by the site-1 protease, while red lines indicate the location of N-terminal dileucine motif and C-terminal arginine/isoleucine/arginine (RIR) motif, respectively. While in the wild-type protein product (A) both motifs are present, the C-terminal motif is missing from the mutant protein (B) due to the frameshift caused by the c.3503_3504delTC mutation, preventing its exit from the endoplasmic reticulum and subsequent cleavage into mature α and β subunits. Adapted from Franke et al., 2013; Qian et al., 2013. 38

Figure 4 – Schematic representation of the spliceosome assembly and splicing process. U1 snRNP initially binds to the 5' splice site (ss), while the splicing factor 1 (SF1) binds to the branch point (A) and the U2 auxiliary factor heterodimer (U2AF 65/35) binds to the 3' ss and polypyrimidine tract (Py). Next, U2 snRNP binds to the branch point, excluding SF1 from the complex and interacts with U1 snRNP. U5, U4 and U6 snRNPs are then recruited as a preassembled complex with U6 snRNP binding to the 5'ss and interacting closely with U2 snRNP, displacing the U1 and U4 snRNPs and bringing the 5'ss and the branch point into close proximity for the first step of the splicing reaction. The spliceosome complex undergoes further conformational changes, resulting in the release of the intronic sequence and junction of the two exons, to produce the mature mRNA. Adapted from Chen et al., 2009; Will & Lührmann, 2011; Kornblihtt et al., 2013.....42

Figure 5 – Core splicing signals within the pre-mRNA molecule. The consensus sequences are as indicated with N being any nucleotide, R being a purine and Y being a pyrimidine. The 5' splice site (ss) is located at the upstream intron junction with the consensus sequence AGGURAGU, while the 3'ss includes the polypyrimidine tract and the dinucleotide AG at the downstream intron junction. The branch point is defined by a highly conserved adenosine surrounded by seven poorly conserved nucleotides located 20-40 nucleotides upstream of the 3'ss. Adapted from Matlin et al., 2005; Kornblihtt et al., 2013. 43

Figure 6 – *Cis*-acting splicing regulatory elements. Intronic splicing silencers (ISSs) and exonic splicing silencers (ESSs) repress the inclusion of the target exon in the mature mRNA, while intronic splicing enhancers (ISEs) and exonic splicing enhancers (ESEs) promote exon inclusion. Adapted from Matlin et al., 2005; Kornblihtt et al., 2013. 43

Figure 7 – Representation of the most common forms of alternative splicing. A: cassette exon; B: mutually exclusive exons; C: alternative 5' splice sites; D: alternative 3' splice sites; E: intron retention; F: alternative promoters; G: alternative polyadenylation. Adapted from Matlin et al., 2005. 45

Figure 8 – Chemical modifications of antisense oligonucleotides. First generation oligonucleotides are characterized by a phosphorothioate (PS) backbone, obtained through the replacement of an oxygen atom with a sulfur atom. Second generation oligonucleotides contain a methyl or methoxyethyl group attached to 2' position of the ribose. Finally, third generation oligonucleotides carry modifications in the furanose ring structure. Adapted from Chan et al., 2006. 46

Figure 9 – Different methods of antisense-mediated splicing modulation. A: blocking of cryptic splice site within the intronic sequence; B: blocking of cryptic splice site within the exonic sequence; C: inclusion of alternatively spliced exon; D: exclusion of exon or “exon skipping”. Adapted from Hammond & Wood, 2011. 48

Figure 10 – Location of antisense oligonucleotide targets on exon 19 of the *GNPTAB* gene and flanking intronic sequences. 65

Figure 11 – Example of Neubauer chamber used for cell counting. The Neubauer chamber contains two distinct chambers, composed of 9 quadrants, each with a volume of 0,1 mm³ or 10⁻⁴ mL. 66

Figure 12 – Alignment in Clustal Omega software of amino acid sequences for the full length *GNPTAB* protein product and truncated protein resulting from the skipping of exon 19. Due to the long size of the protein sequence, only exons 14 to 21 are shown in this picture. protein: full length protein product; skip19: truncated protein.....74

Figure 13 – Alignment in Clustal Omega software of amino acid sequences encoded by exon 19 of the *GNPTAB* gene in different animal species. 74

Figure 14 – Protein domains predicted by Pfam 30.0 for the *GNPTAB* protein product. Color codes and exact positions are depicted in Table 5. 75

Figure 15 – Protein domains predicted by InterPro for the *GNPTAB* protein product. 76

Figure 16 – Secondary structure of the *GNPTAB* protein predicted by UniProt software. Green: beta strand; pink: turn; blue: helix. 77

Figure 17 – N-glycosylation sites predicted by NetNGlyc 1.0 Server for the *GNPTAB* protein product. 79

Figure 18 – Graphic representation of the putative enhancers predicted by ESEfinder 3.0 for exon 19 *GNPTAB* gene containing the c.3503_3504delTC mutation and 50bp of the flanking introns. Arrows indicate the location of the donor and acceptor splice sites.....83

Figure 19 – Putative enhancers predicted by RESCUE-ESE on exon 19 of the *GNPTAB* gene containing the c.3503_3504delTC mutation and 50bp of the flanking introns. Positions start at the beginning of the input sequence and not the beginning of exon 19. 1-50: intron 18; 51-216: exon 19; 217-266: intron 19.....83

Figure 20 – Schematic figure of primer annealing for amplification of exon 19 of the *GNPTAB* gene and flanking intronic regions.....84

Figure 21 – Agarose gel electrophoresis of PCR amplification of *GNPTAB* gene exon 19 and flanking intronic regions in control and patient fibroblasts. L: ladder; B: negative control (blank); HC: healthy control; P1: ML II Patient 1; P2: ML II Patient 2. 85

Figure 22 – Schematic representation of annealing locations of all six primer sets for amplification of different fragments of the *GNPTAB* transcript. 86

Figure 23 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript in patient and control fibroblasts. A: primer set 1; B: primer set 2; C: primer set 3; D: primer set 4; E: primer set 5; F: primer set 6. L: ladder; B: negative control (blank); HC: healthy control; P1: ML II Patient 1; P2: ML II Patient 2. 87

Figure 24 – Fluorescence microscopy images of healthy control fibroblasts transfected with eGFP. The strong green fluorescence visible in the cells confirms the efficiency of Lipofectamine® 2000 in transfecting cultured fibroblasts. 88

Figure 25 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in control fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19A(-18+4): directed to the 3'ss; H19D(+12-9): directed to the 5'ss; H19A(+45+64): directed to ESEs in the first half of exon 19; H19A(+119+138): directed to ESEs in the second half of exon 19.....89

Figure 26 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in control fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19D(+12-9): directed to the 5'ss; H19A(+119+138): directed to ESEs in the second half of exon 19. 91

Figure 27 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 1 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19A(-18+4): directed to the 3'ss; H19D(+12-9): directed to the 5'ss; H19A(+45+64): directed to ESEs in the first half of exon 19; H19A(+119+138): directed to ESEs in the second half of exon 19.....92

Figure 28 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 1 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19D(+12-9): directed to the 5'ss; H19A(+119+138): directed to ESEs in the second half of exon 19.....93

Figure 29 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 2 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19A(-18+4): directed to the 3'ss; H19D(+12-9): directed to the 5'ss; H19A(+45+64): directed to ESEs in the first half of exon 19; H19A(+119+138): directed to ESEs in the second half of exon 19.....94

Figure 30 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 2 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19D(+12-9): directed to the 5'ss; H19A(+119+138): directed to ESEs in the second half of exon 19.....96

Abbreviations

2'-OMe	2'-O-methyl
2'-MOE	2'-O-methoxyethyl
α-MM	α-methylmannoside
AG	adenine/guanine
AO	antisense oligonucleotide
ARSB	arylsulfatase B
AS	alternative splicing
ATP	adenosine triphosphate
BBB	blood-brain barrier
BLAST	Basic Local Alignment Search Tool
cDNA	complementary DNA
CNS	central nervous system
CR4	conserved region 4
del	deletion
DLS	Doenças Lisossomais de Sobrecarga
DMD	Duchenne muscular dystrophy
DMAP	DNA methyltransferase-associated protein
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid
ERT	enzyme replacement therapy
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
ExPASy	Expert Protein Analysis System
FBS	fetal bovine serum
GAG	glycosaminoglycan
GB	γ-subunit binding region
GC	guanine/cytosine
GlcNAc	N-acetylglucosamine
GlcNAc-1-P	N-acetylglucosamine-1-phosphate
HC	healthy control
hnRNP	heterogeneous nuclear ribonucleoprotein
HSCT	hematopoietic stem cell transplantation
ISE	intronic splicing enhancer

ISS	intronic splicing silencer
LNA	locked nucleic acid
LSD	Lysosomal Storage Disorders
M6P	mannose 6-phosphate
ML	mucopolipidosis
MPS	mucopolysaccharidosis
mRNA	messenger RNA
MU	methyumbelliferyl
NCBI	National Center for Biotechnology Information
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PenStrep	penicillin/streptomycin
PMO	phosphoroamidate morpholino oligomer
PNA	peptide nucleic acid
pre-mRNA	precursor mRNA
PS	phosphorothioate
Py	polypyrimidine tract
RIR	arginine/isoleucine/arginine
RNA	ribonucleic acid
RNase	ribonuclease
RT-PCR	reverse transcriptase polymerase chain reaction
SF1	splicing factor 1
SMA	spinal muscular atrophy
SNP	single nucleotide polymorphism
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SR	serine/arginine
SRSF	serine/arginine-rich splicing factor
SRT	substrate reduction therapy
ss	splice site
TC	thymine/cytosine
T_m	melting temperature
TTR	transthyretin
U2AF	U2 auxiliary factor

UDP-GlcNAc	uridine diphosphate <i>N</i> -acetylglucosamine
US FDA	United States Food and Drug Administration

Chapter 1

Introduction

1.1. Lysosomal Storage Disorders

Lysosomal Storage Disorders (LSDs) are a group of rare inherited metabolic diseases caused by the malfunction of the lysosomal system, ultimately resulting in the accumulation of undegraded substrates inside the lysosomes (Filocamo & Morrone, 2011; Lieberman et al., 2012). As a consequence, multiple organelles and processes might be affected, such as signaling, vesicle trafficking, autophagy, calcium homeostasis and inflammatory response (Futerman & van Meer, 2004; Ballabio & Gieselmann, 2009; Platt et al., 2012), leading to cell death and tissue damage. These damages translate into a wide variety of clinical manifestations ranging from nearly asymptomatic individuals to serious and progressive clinical manifestations involving several organs and systems and generally resulting in death in the first few years of life (Platt et al., 2012). Despite being rare disorders individually, LSDs present a significant frequency when considered as a whole ranging from 7.6 cases per 100 000 births in British Columbia (Applegarth et al., 2000) to 26.9 cases per 100 000 births in the United Arab Emirates (Al-Jasmi et al., 2012), showcasing a strong variability in incidence between different ethnicities and populations. In the Portuguese population, the frequency of these disorders is estimated to be 25 cases per 100 000 births (Pinto et al., 2004), whereas other European countries have revealed a LSD incidence around 12-14 cases per 100 000 births (Poorthuis et al., 1999; Dionisi-Vici et al., 2002; Poupětová et al., 2010).

Encompassing about 50 distinct disorders, the classification of LSDs can be quite challenging and complex. While it has been traditionally associated with the nature of the accumulated substrate, in recent years, the molecular defect responsible for each disorder has taken on a more important role for the grouping of these diseases, allowing for the inclusion of other disorders that could not be properly organized in the traditional system (Table 1) (Filocamo & Morrone, 2011; Boustany, 2013). Nonetheless, the traditional classification is still widely used for disorders such as mucopolysaccharidoses (accumulation of mucopolysaccharides), sphingolipidoses (accumulation of sphingolipids) and oligosaccharidoses (accumulation of oligosaccharides) (Ballabio & Gieselmann, 2009), and the measurement of accumulated substrate is still quite useful as a biomarker for both diagnosis and therapy follow-up in LSDs (Cox, 2005).

Table 1 – Classification of Lysosomal Storage Disorders (LSDs). Adapted from Filocamo & Morrone, 2011; Boustany, 2013.

Disorder	Gene	Enzyme deficiency	Accumulated substrates
Sphingolipidoses			
Fabry disease	<i>GLA</i>	α -Galactosidase A	Globotriasylceramide
Farber lipogranulomatosis	<i>ASAHI</i>	Acid ceramidase	Ceramide
Gaucher disease	<i>GBA</i>	Glucosylceramidase	Glucosylceramide
Niemann-Pick disease (A/B)	<i>SMPD1</i>	Sphingomyelinase	Sphingomyelin
GM1-gangliosidosis (I, II and III)	<i>GLB1</i>	GM1- β -galactosidase	GM1 ganglioside, keratan sulphate, oligos, glycolipids
GM2-gangliosidosis (Sandhoff)	<i>HEXAB</i>	β -Hexosaminidase A + B	GM2 ganglioside, oligos
GM2-gangliosidosis (Tay-Sachs)	<i>HEXA</i>	β -Hexosaminidase A	GM2 ganglioside, oligos, glycolipids
GM3-gangliosidosis	<i>ST3GAL5</i>	GM3 synthase	GM3 ganglioside
Krabbe disease	<i>GALC</i>	β -Galactosylceramidase	Galactosylceramide
Metachromatic leucodystrophy	<i>ARSA</i>	Arylsulphatase A	Sulphatides
Sphingolipid-activator deficiency	<i>PSAP</i>	Sphingolipid activator	Glycolipids
Mucopolysaccharidoses			
MPS I (Hurler, Scheie, Hurler/Scheie)	<i>IDUA</i>	α -Iduronidase	Dermatan sulphate, heparan sulphate
MPS II (Hunter)	<i>IDS</i>	Iduronate sulphatase	Dermatan sulphate, heparan sulphate
MPS IIIA (Sanfilippo A)	<i>SGSH</i>	Heparan sulphamidase	Heparan sulphate
MPS IIIB (Sanfilippo B)	<i>NAGLU</i>	Acetyl α -glucosaminidase	Heparan sulphate
MPS IIIC (Sanfilippo C)	<i>HGSNAT</i>	Acetyl CoA; α -glucosaminide N-acetyltransferase	Heparan sulphate
MPS IIID (Sanfilippo D)	<i>GNS</i>	N-acetyl glucosamine-6-sulphatase	Heparan sulphate
MPS IVA (Morquio A)	<i>GALNS</i>	Acetyl galactosamine-6-sulphatase	Keratan sulphate, chondroitin 6-sulphate
MPS IVB (Morquio B)	<i>GLB1</i>	β -Galactosidase	Keratan sulphate
MPS VI (Maroteaux-Lamy)	<i>ARSB</i>	Arylsulphatase B	Dermatan sulphate
MPS VII (Sly)	<i>GUSB</i>	β -Glucuronidase	Dermatan sulphate, heparan sulphate, chondroitin 6-sulphate
MPS IX (Natowicz)	<i>HYAL1</i>	Hyaluronidase	Hyluronan
Oligosaccharidoses (glycoproteinoses)			
Aspartylglucosaminuria	<i>AGA</i>	Glycosylasparaginase	Aspartylglucosamine
Fucosidosis	<i>FUCA1</i>	α -Fucosidase	Glycoproteins, glycolipids, Fucoside-rich oligos
α -mannosidosis	<i>MAN2B1</i>	α -Mannosidase	Mannose-rich oligos
β -mannosidosis	<i>MANBA</i>	β -Mannosidase	Man ($\beta 1 \rightarrow 4$) GlnNAc
Schindler disease	<i>NAGA</i>	N-acetylgalactosaminidase	Sialylated/asialoglycopeptides, glycolipids
Sialidosis (mucopolipidosis I)	<i>NEU1</i>	Neuraminidase	Oligos, glycopeptides
Glycogenoses			
Glycogenosis II (Pompe)	<i>GAA</i>	α -Glucosidase	Glycogen
Lipidoses			
Wolman disease	<i>LIPA</i>	Acid lipase	Cholesterol esters
Non-enzymatic lysosomal protein defect			
GM2 gangliosidosis (GM2 activator deficiency)	<i>GM2A</i>	GM2 activator protein	GM2 ganglioside, oligos
Krabbe disease	<i>PSAP</i>	Saposin A	Galactosylceramide

Metachromatic leucodystrophy	<i>PSAP</i>	Saposin B	Sulphatides
Gaucher disease	<i>PSAP</i>	Saposin C	Glucosylceramide
Transmembrane protein defect			
Sialic acid storage disease	<i>SLC17A5</i>	Sialin	Sialic acid
Cystinosis	<i>CTNS</i>	Cystinosin	Cystine
Niemann-Pick type C1	<i>NPC1</i>	Niemann-Pick type C1 protein (proton-driven transporter)	Cholesterol and sphingolipids
Niemann-Pick type C2	<i>NPC2</i>	Niemann-Pick type C2 protein (proton-driven transporter)	Cholesterol and sphingolipids
Danon disease	<i>LAMP2</i>	Lysosome-associated membrane protein 2	Cytoplasmatic debris and glycogen
Mucopolipidosis IV	<i>MCOLN1</i>	Mucolipin	Lipids
Lysosomal enzyme protection defect			
Galactosialidosis	<i>CTSA</i>	Protective protein cathepsin A (PPCA)	Sialyloligosaccharides
Post-translational processing defect			
Multiple sulphatase deficiency	<i>SUMF1</i>	Multiple sulphatase	Sulphatides, glycolipids, GAGs ¹
Trafficking defect in lysosomal enzymes			
Mucopolipidosis II alpha/beta Mucopolipidosis III alpha/beta	<i>GNPTAB</i>	GlcNAc-1-phosphotransferase (α/β subunits)	Oligos, GAGs ¹ , lipids
Mucopolipidosis III gamma	<i>GNPTG</i>	GlcNAc-1-phosphotransferase (γ subunit)	Oligos, GAGs ¹ , lipids
Polypeptide degradation defect			
Pycnodysostosis	<i>CTSK</i>	Cathepsin K	Bone proteins
Neuronal ceroid lipofuscinoses (NCLs)			
NCL 1	<i>PPT1</i>	Palmitoyl protein thioesterase (PPT1)	Lipofuscin, saposins A and D
NCL 2	<i>TPP1</i>	Tripeptidyl peptidase 1 (TPP1)	Lipofuscin, subunit c of ATP synthase
NCL 3	<i>CLN3</i>	CLN3, lysosomal transmembrane protein	Lipofuscin, subunit c of ATP synthase
NCL 4	<i>DNAJC5</i>	DnaJ homologue subfamily C member 5	Lipofuscin, subunit c of ATP synthase
NCL 5	<i>CLN5</i>	CLN5, soluble lysosomal protein	Lipofuscin, subunit c of ATP synthase
NCL 6	<i>CLN6</i>	CLN6, transmembrane protein of ER	Lipofuscin, subunit c of ATP synthase
NCL 7	<i>MFSD8</i>	CLN7, lysosomal chloride channel	Lipofuscin, subunit c of ATP synthase
NCL 8	<i>CLN8</i>	CLN8, transmembrane protein of ER	Lipofuscin, subunit c of ATP synthase
NCL 10	<i>CTSD</i>	Cathepsin D	Lipofuscin, saposins A and D
NCL 11	<i>GRN</i>	Progranulin	—
NCL 12	<i>ATP13A2</i>	P type ATPase	—
NCL 13	<i>CTSF</i>	Cathepsin F	—
NCL 14	<i>KCTD7</i>	Potassium channel tetramerization domain-containing protein 7	—

¹ GAGs: Glycosaminoglycans

1.1.1. Mucopolipidosis II

Like many other glycoproteins, soluble lysosomal proteins are synthesized in the endoplasmic reticulum and are co-translationally glycosylated on certain asparagine residues. After synthesis and glycosylation, lysosomal enzymes enter the Golgi apparatus where a mannose 6-phosphate (M6P) marker is added to the peptide, allowing for its binding to the M6P receptors in the Golgi and subsequent targeting to the lysosomes (Kornfeld, 1986). This modification is performed by two Golgi enzymes in sequential action. Initially, the enzyme *N*-acetylglucosamine-1-phosphotransferase (GlcNAc-1-phosphotransferase) transfers *N*-acetylglucosamine-1-phosphate (GlcNAc-1-P) to mannose residues in the peptide, after which a second enzyme, *N*-acetylglucosamine-1-phosphodiester α -*N*-acetylglucosaminidase (also known as “uncovering enzyme”) removes the GlcNAc residue to expose the M6P modifications, which can then be recognized by M6P receptors (Figure 1) (Qian et al., 2010).

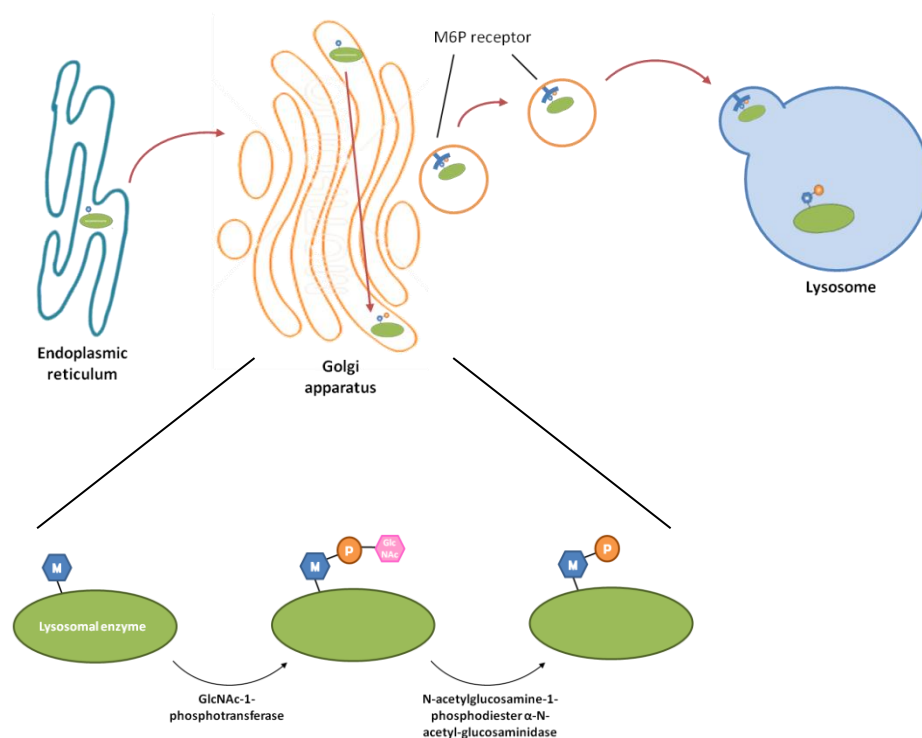


Figure 1 – Newly synthesized lysosomal enzymes are transported from the endoplasmic reticulum into the Golgi apparatus where a mannose-6-phosphate (M6P) marker is added through the action of two enzymes: GlcNAc-1-phosphotransferase transfers GlcNAc-1-P to mannose residues of the peptide, while *N*-acetylglucosamine-1-phosphodiester α -*N*-acetylglucosaminidase (“uncovering enzyme”) removes the GlcNAc residue to expose the M6P modifications. Subsequently, the lysosomal enzymes bind to M6P receptors and are targeted to the lysosomes. Adapted from Coutinho et al., 2012.

Total or near-total deficiency of GlcNAc-1-phosphotransferase impairs the formation of this M6P marker, resulting in the nearly complete absence of lysosomal targeting in many cell types and tissues and the secretion of most lysosomal enzymes (Braulke et al., 2008). This defect is clinically recognized as mucopolidosis II (ML II or I-cell disease; MIM# 252500), an autosomal recessive disease often resulting in death in the first decade of life (Reitman et al., 1981b). Another type of mucopolidosis, mucopolidosis III (ML III or pseudo-Hurler polydystrophy; MIM# 252600), is a milder disease in which GlcNAc-1-phosphotransferase activity is reduced rather than absent (Honey et al., 1982).

The GlcNAc-1-phosphotransferase enzyme is an hexameric complex composed of two α , two β and two γ subunits, with α and β exerting the catalytic activity of the enzyme, whereas γ chains facilitate the folding of the catalytic subunits and maintain them in the right configuration, while also regulating their activity (Figure 2) (Bao et al., 1996; Coutinho et al., 2012).

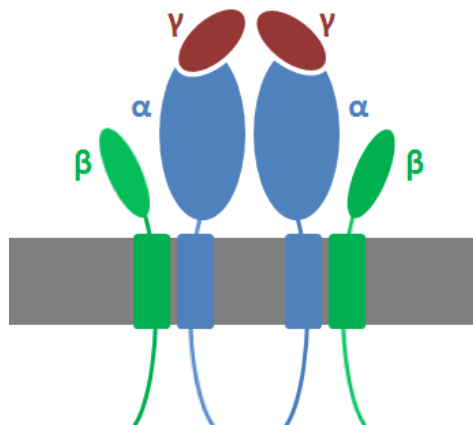


Figure 2 – Schematic representation of the GlcNAc-1-phosphotransferase. The enzyme is composed of two α subunits (blue), two β subunits (green) and two γ subunits (red). α and β subunits are responsible for the catalytic activity of the enzyme, while the γ subunits appear to facilitate and increase their action. Adapted from Pohl et al., 2011.

The α and β subunits are encoded by *GNPTAB*, a gene located on chromosome 12, composed of 21 exons (Tiede et al., 2005), while the γ subunits are encoded by *GNPTG*, a gene located on chromosome 16, composed of 11 exons (Raas-Rothschild et al., 2000). Defects in *GNPTAB* result in ML II alpha/beta and ML III alpha/beta whereas mutations in *GNPTG* have only been found in ML III gamma patients. Being encoded by the same gene, the α and β subunits are synthesized as a single precursor protein in the endoplasmic reticulum, which is later cleaved in the Golgi apparatus by the site-1 protease to produce mature α and β subunits (Marschner et al., 2011; De Pace et al., 2013). Transport of the α/β -precursor protein to the Golgi depends on the presence of two sorting motifs in the protein

sequence: an N-terminal dileucine motif and a C-terminal dibasic arginine/isoleucine/arginine (RIR) motif (Figure 3A) (Franke et al., 2013). The absence of one or both of these domains inhibits the exit of the α/β -precursor from the endoplasmic reticulum and its subsequent cleavage, preventing the formation of mature α and β subunits and therefore impairing GlcNAc-1-phosphotransferase activity.

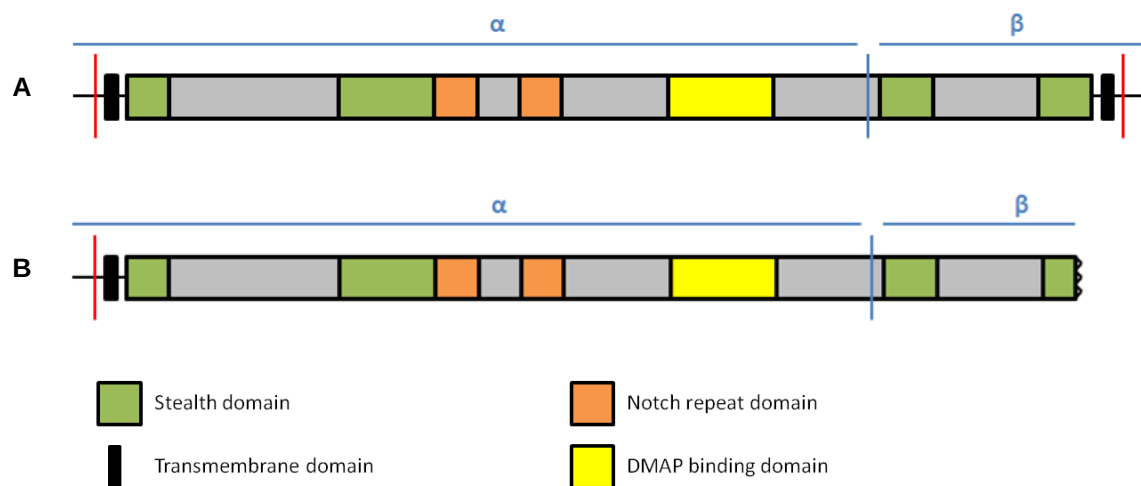


Figure 3 – Motifs and domains of the α/β -precursor protein obtained from the *GNPTAB* gene. Blue line indicates the position for cleavage by the site-1 protease, while red lines indicate the location of N-terminal dileucine motif and C-terminal arginine/isoleucine/arginine (RIR) motif, respectively. While in the wild-type protein product (A) both motifs are present, the C-terminal motif is missing from the mutant protein (B) due to the frameshift caused by the c.3503_3504delTC mutation, preventing its exit from the endoplasmic reticulum and subsequent cleavage into mature α and β subunits. Adapted from Franke et al., 2013; Qian et al., 2013.

To date 142 pathogenic mutations in *GNPTAB* that cause ML II and ML III alpha/beta are listed in The Human Gene Mutation Database (Stenson et al., 2017). Almost all ML II patients have nonsense, frameshift, or splice-site mutations in *GNPTAB*, and these mutations are usually associated with a more severe clinical course. Conversely, ML III alpha/beta patients carry missense mutations or at least one hypomorphic allele that is not associated with a complete loss of functionality and the clinical signs and symptoms of ML III alpha/beta patients are milder and more slowly progressive than in ML II (Braulke et al., 2013). Among these multiple mutations described in the *GNPTAB* gene, the two-nucleotide deletion c.3503_3504delTC on exon 19 is considered the most common among ML II patients, having been previously studied by our research group (Encarnação et al., 2009; Coutinho et al., 2011). Since it involves the deletion of two nucleotides, this mutation causes a frameshift in the transcript, leading to an altered amino acid sequence, as well as to a premature stop codon (Plante et al., 2008). The synthesized α/β -precursor peptide lacks the C-terminal domain, preventing its exit from the endoplasmic reticulum and subsequent cleavage into mature α and β subunits (Figure 3B) (De Pace et al., 2013). This impairs the

assembly of a functional GlcNAc-1-phosphotransferase enzyme and consequently the formation of the M6P marker in lysosomal enzymes, which results in an almost complete absence of hydrolases inside the lysosomes and accumulation of undegraded substrates (Kudo et al., 2006). This leads to a phenotype characterized by severe psychomotor retardation, coarse facial features, gingival hypertrophy, short stature, hepatosplenomegaly, joint contractures, osteopenia and cardiopulmonary complications that usually lead to death in the first decade of life (De Pace et al., 2013; Velho et al., 2015).

1.1.2. Biochemical and molecular diagnosis

Since the measurement of GlcNAc-1-phosphotransferase activity usually requires a very laborious assay that involves the radioactive [^{32}P]UDP-GlcNAc substrate and is only available in a very small number of facilities, an indirect diagnosis of ML II and ML III is more commonly established by the assessment of lysosomal hydrolases (Alegra et al., 2014). Given the impaired sorting of multiple enzymes to the lysosomes in these patients, the biochemical confirmation of these disorders is usually attained indirectly through the detection of activity of lysosomal enzymes in plasma or in cultured fibroblasts. In general, a ten to twentyfold increase in serum lysosomal enzymes, together with decreased intralysosomal levels in fibroblasts, is indicative of those disorders (Kornfeld & Sly, 2001). In addition, the measurement of lysosomal enzymes activity may be used in the research field for the evaluation of the efficiency of therapeutic approaches that are being investigated both *in vitro* and *in vivo*.

Nonetheless, molecular diagnosis is essential to confirm the presence of these disorders, with sequencing of the *GNPTAB* and *GNPTG* coding regions detecting disease-causing mutations in over 95% of patients. In general, ML II and ML III can be distinguished based on clinical criteria and disease progression but the identification of mutations in the *GNPTAB* and *GNPTG* genes constitutes an important tool for the correct classification of these disorders. In fact, the distinction between ML III alpha/beta and ML III gamma is only possible based on molecular diagnosis, since the two forms of ML III have the same clinical and biochemical characteristics. Moreover, prenatal diagnosis of at-risk pregnancies is dependent on prior identification of disease-causing mutations in the family (Bräulke et al., 2013).

1.1.3. Available treatments for LSDs and ML II

Although the biochemical and molecular causes behind most LSDs have been largely disclosed, similar advances have not been made in the field of therapy. Currently, no definitive cure is available for this kind of disorders, so most available treatments focus on improving life expectancy and quality of life through multiple strategies (Boustany, 2013; Parenti et al., 2013). An obvious approach for treatment of these disorders is the restoration of enzyme activity, which can be achieved by different kinds of procedures. One possibility is hematopoietic stem cell transplantation (HSCT), where the patient's tissues can be repopulated with the donor's healthy cells, which will then produce the functional enzyme at normal levels and hopefully ameliorate the symptoms of the disease (Parenti et al., 2013). However, this sort of treatment has only proved helpful for a limited number of conditions, patients must be in their early childhood for this procedure to be effective, and the rate of associated complications and mortality is significant (Malatack et al., 2003; Cox, 2015). Another therapeutic approach is enzyme replacement therapy (ERT), in which a recombinant form of the target enzyme is administered to patients to be then taken up by the patient's cells, where it is targeted to the lysosome to replace the defective enzyme and degrade the accumulated substrate (Parenti et al., 2013). Nonetheless, despite good results in improving the function of multiple organs, this form of therapy is unable to have any effect on the central nervous system (CNS) since the exogenous enzyme cannot cross the blood–brain barrier (BBB), limiting its application (Boustany, 2013; Parenti et al., 2013). Yet another common option for LSD treatment is substrate reduction therapy (SRT), which aims to reduce the production of the accumulated substrate so that patients with residual enzyme activity can successfully degrade it and avoid excessive storage in the lysosome (Bruni et al., 2007). For that purpose, this therapy uses small-molecule inhibitors to target specific steps in the biosynthetic pathways of the accumulated substrate and block or decrease the synthesis of these materials. Since the drugs used here are composed of very small molecules, they are able to cross the BBB and successfully reach the CNS, making it a better option for multiple disorders that involve the CNS (Parenti et al., 2013).

Unfortunately, in the specific case of ML II, many of these therapies are unsuitable given the nature of the affected enzyme and accumulation of multiple substrates inside the lysosomes. On one hand, ERT is not an option as the defective enzyme is not a lysosomal enzyme that can be internalized by the cells and targeted to the lysosome, and SRT also does not fit this disorder since various substrates are accumulated within the lysosomes due to a near total absence of hydrolytic enzymes in this organelle. Furthermore, although HSCT

has been reported in the past for the treatment of ML II (Grewal et al., 2003), no benefits were observed in the musculoskeletal system and the survival rate is quite low (Lund et al., 2014).

Acknowledging this, some innovative therapeutic approaches have been developed in recent years, including the usage of pharmacological chaperones that promote the correct folding and trafficking of defective proteins (Sawkar et al., 2006), the combination of these chaperones with ERT (Boustany, 2013), the induction of ribosomal read-through of nonsense mutations allowing for the production of a full-length protein (Parenti et al., 2013), and even gene therapy through the introduction of the correct gene sequence into the affected cells so that the functional enzyme can be produced (Bruni et al., 2007; Cox, 2015). Considering the impact of these disorders in both patients and families, there is an urgent need to invest in the development of new therapeutic strategies for their treatment in order to ameliorate the disease phenotype, offering to patients the best possible clinical outcome together with minimal side effects and improved quality of life.

1.2. Splicing therapies

1.2.1. Splicing reaction

The expression of eukaryotic genes requires a sequential series of finely regulated steps, amongst which one of the most important is splicing. In the process of splicing, non-coding introns are removed from the precursor messenger RNA (pre-mRNA) molecule and the coding sequences (exons) are joined together to produce the mature mRNA that will be exported to the cytoplasm and translated into protein (Matera & Wang, 2014). The splicing reaction is carried out by the spliceosome, a large RNA and protein complex, constituted by five small ribonucleoprotein complexes (snRNPs) and more than a hundred proteins (Matera & Wang, 2014). The first step of spliceosome assembly is the recognition of the 5' splice site (ss) by the U1 snRNP, followed by the binding of splicing factor 1 (SF1) to the branch point and of the U2 auxiliary factor heterodimer (U2AF 65/35) to the polypyrimidine tract (Py) and 3'ss, originating the E complex (Chen & Manley, 2009; De Conti et al., 2013). SF1 is then replaced by the U2 snRNP at the branch point, originating the A complex and allowing for the interaction between U1 snRNP and U2 snRNP across the exon (Chen & Manley, 2009; Matera & Wang, 2014). Afterwards, the U4, U5 and U6 snRNPs are recruited as a preassembled complex, leading to the formation of the B complex. The interaction between

U4 and U6 is then disrupted and the U6 snRNP base pairs with the 5'ss, displacing U1 snRNP from its initial location and releasing it from the complex, along with the U4 snRNP (Will & Lührmann, 2011). At the same time, U6 snRNP also interacts extensively with U2 snRNP, bringing the 5'ss and the branch point into close proximity and allowing for the first step of splicing to take place, which originates the C complex containing the free upstream exon and the intron-exon lariat intermediate (Matera & Wang, 2014). This complex undergoes further ATP-dependent conformational changes and completes the second step of the splicing reaction, releasing the lariat intron and joining the coding exons to form the mature mRNA, while U2, U5 and U6 snRNPs are also released from the complex and recycled for future splicing reactions (Figure 4) (Matera & Wang, 2014).

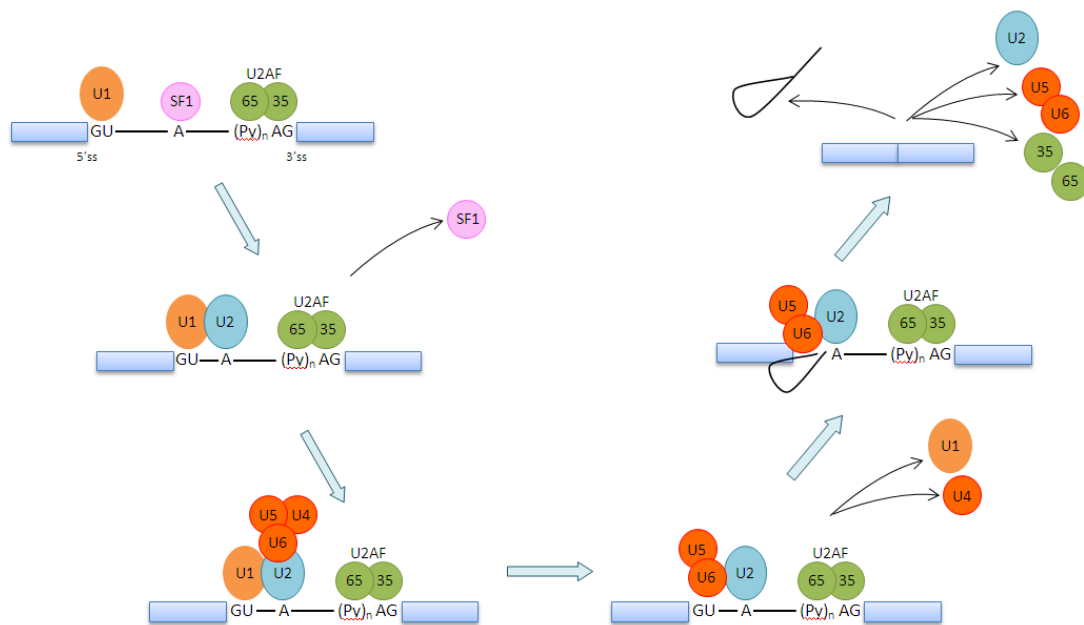


Figure 4 – Schematic representation of the spliceosome assembly and splicing process. U1 snRNP initially binds to the 5' splice site (ss), while the splicing factor 1 (SF1) binds to the branch point (A) and the U2 auxiliary factor heterodimer (U2AF 65/35) binds to the 3' ss and polypyrimidine tract (Py). Next, U2 snRNP binds to the branch point, excluding SF1 from the complex and interacts with U1 snRNP. U5, U4 and U6 snRNPs are then recruited as a preassembled complex with U6 snRNP binding to the 5'ss and interacting closely with U2 snRNP, displacing the U1 and U4 snRNPs and bringing the 5'ss and the branch point into close proximity for the first step of the splicing reaction. The spliceosome complex undergoes further conformational changes, resulting in the release of the intronic sequence and junction of the two exons, to produce the mature mRNA. Adapted from Chen et al., 2009; Will & Lührmann, 2011; Kornblihtt et al., 2013.

1.2.2. Splicing signals and regulatory elements

Although the spliceosome orchestrates pre-mRNA processing with great complexity and fidelity, this is a fairly flexible mechanism under strong regulation by both *cis*-acting and *trans*-acting elements. The 5' and 3' splice sites and the branch point are present in every

intron and therefore are known as core splicing signals (Wang & Burge, 2008). The 5'ss (donor site) is located at the upstream intron junction with the consensus sequence AGGURAGU, while the 3'ss (acceptor site) includes the 10-20 nucleotide long polypyrimidine tract and the dinucleotide AG at the downstream intron junction (Reed, 1996; Faustino & Cooper, 2003; Kornblihtt et al., 2013). The branch point is located 20-40 nucleotides upstream the 3'ss and consists of a highly conserved adenosine surrounded by a poorly conserved sequence of 7 nucleotides (Figure 5) (Reed, 1996; Kornblihtt et al., 2013).

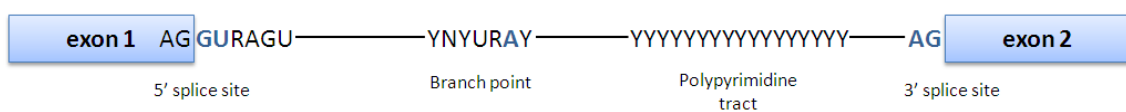


Figure 5 – Core splicing signals within the pre-mRNA molecule. The consensus sequences are as indicated with N being any nucleotide, R being a purine and Y being a pyrimidine. The 5' splice site (ss) is located at the upstream intron junction with the consensus sequence AGGURAGU, while the 3'ss includes the polypyrimidine tract and the dinucleotide AG at the downstream intron junction. The branch point is defined by a highly conserved adenosine surrounded by seven poorly conserved nucleotides located 20-40 nucleotides upstream of the 3'ss. Adapted from Matlin et al., 2005; Kornblihtt et al., 2013.

Although this core splicing signals are highly conserved in less complex organisms, they tend to be less conserved in higher eukaryotes (Warf & Berglund, 2010), thus requiring the involvement of other *cis*-elements for the correct recognition of introns and exons. These elements can be enhancers or silencers depending on their role in promoting or inhibiting exon inclusion, respectively, and can also be divided into intronic or exonic elements depending on their location. As such, intronic splicing enhancers (ISEs) and exonic splicing enhancers (ESEs) favor the inclusion of the exon into the mature mRNA molecule, while intronic splicing silencers (ISSs) and exonic splicing silencers (ESSs) repress exon inclusion (Figure 6) (De Conti et al., 2013).

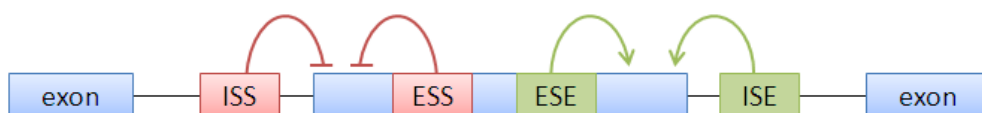


Figure 6 – *Cis*-acting splicing regulatory elements. Intronic splicing silencers (ISSs) and exonic splicing silencers (ESSs) repress the inclusion of the target exon in the mature mRNA, while intronic splicing enhancers (ISEs) and exonic splicing enhancers (ESEs) promote exon inclusion. Adapted from Matlin et al., 2005; Kornblihtt et al., 2013.

ESEs are for the most part recognized by serine/arginine (SR) proteins, which also contain domains for protein-protein interaction (Wang & Burge, 2008; De Conti et al., 2013), while proteins associated with ISEs are still poorly characterized. In the case of silencer elements, ESSs and ISSs are recognized by members of the heterogeneous nuclear ribonucleoprotein (hnRNP) protein family, which act through different mechanisms, such as blocking interactions between spliceosome elements (Spellman & Smith, 2006) or promoting inhibitory RNA structures (Fisette et al., 2010). Nonetheless, the activity of a certain regulatory element often depends on its location within the pre-mRNA, a property called “context dependence” (Wang & Burge, 2008). Additionally, splicing can be affected in multiple aspects by other factors, such as phosphorylation, subcellular localization, ratio between SR proteins and hnRNPs and RNA secondary structure (Matlin et al., 2005; Jensen et al., 2009; Warf & Berglund, 2010).

1.2.3. Alternative Splicing

Aside from the so-called constitutive splicing process, alternative splicing (AS) also occurs, with some sequences altering between inclusion and exclusion into the mature mRNA (Faustino & Cooper, 2003; Revil et al., 2010; Liu et al., 2012). Since most AS events occur in the coding sequences of the mRNA transcript, they have consequences in the sequence of encoded proteins, affecting their structure and biological activity on multiple levels (Stamm et al., 2005). In that sense, AS events can be classified into five basic groups: cassette exons, mutually exclusive exons, alternative 5' splice sites, alternative 3' splice sites and intron retention (Figure 7A to Figure 7E) (Gamazon & Stranger, 2014; Sveen et al., 2016). Aside from these main types of AS, alternative transcripts can also arise from the usage of alternative promoters or alternative polyadenylation, which provide additional variability to mRNAs produced from a single gene (Figure 7F and Figure 7G) (Matlin et al., 2005; Roy et al., 2013). Furthermore, AS can also affect gene expression through removal or addition of regulatory elements that control mRNA stability, localization and translation (Faustino & Cooper, 2003). Lastly, not only the sequence of the transcript can be altered, but also its abundance, which is associated with several development processes, such as sex determination, tissue-differentiation and muscle contraction (Black, 2003), as well as with a number of different diseases (Love et al., 2015; Sveen et al., 2016).

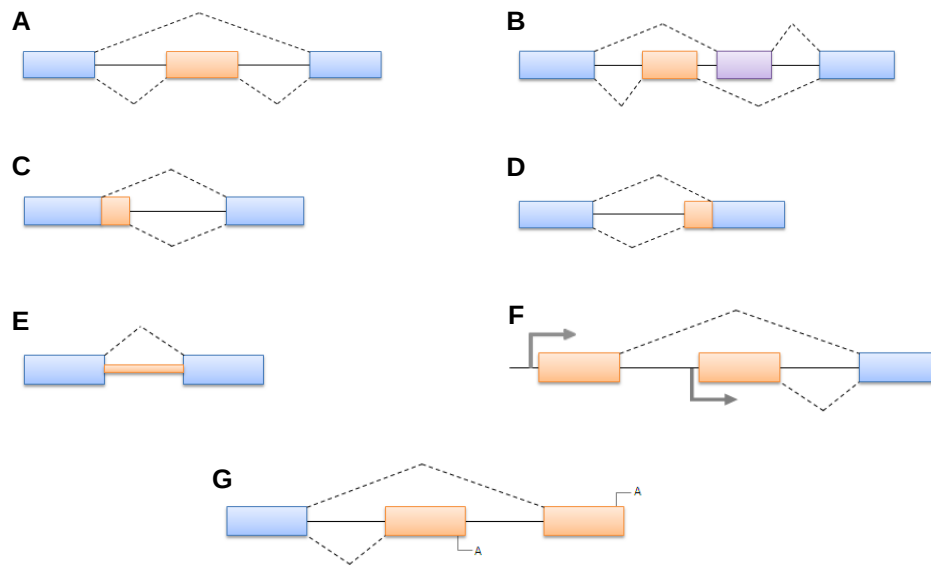


Figure 7 – Representation of the most common forms of alternative splicing. A: cassette exon; B: mutually exclusive exons; C: alternative 5' splice sites; D: alternative 3' splice sites; E: intron retention; F: alternative promoters; G: alternative polyadenylation. Adapted from Matlin et al., 2005.

1.3. Antisense Oligonucleotides

In the last two decades, many approaches have been explored in the field of RNA-based therapeutics in order to manipulate splicing of pre-mRNA or promote degradation of a mutant mRNA. Concerning the modulation of splicing, although multiple methods can be employed, the vast majority involves the usage of antisense oligonucleotides (AOs). AOs are short synthetic oligonucleotides that bind to target RNAs through standard Watson-Crick base pairing (Bennett & Swayze, 2010). AOs can act via different mechanisms and, in the case of splicing modulation, they are designed to target splice sites or splicing regulatory elements, blocking the binding of splicing factors and spliceosome components, which results in alterations in the splicing pattern and the final protein product (Hammond & Wood, 2011; Evers et al., 2015).

1.3.1. Chemical modifications: three generations of oligonucleotides

Initially, AOs were just synthetic unmodified DNA or RNA, which despite delivering some promising results, would prove to be quite ineffective in biological systems due to their susceptibility to degradation by nucleases, poor affinity for their target mRNA, multiple off-

target effects, inability to cross the cell membrane given their negative charge and weak binding to plasma proteins leading to rapid clearance by the kidney (Bennett & Swayze, 2010; Evers et al., 2015; Chery & Näär, 2016). Therefore, chemical modifications were made in the oligonucleotides, generating three categories of AOs, commonly known as generations, with different chemical and pharmacological properties (Figure 8). The first generation of oligonucleotides is characterized by alterations in the backbone, the most common being the phosphorothioate (PS) backbone, accomplished by the replacement of one of the non-bridging oxygen atoms with a sulphur atom. This modification allows for an improved nuclease resistance, as well as strong binding to plasma proteins, reducing renal clearance, but still presents poor binding affinity, low specificity and poor cellular uptake (Chan et al., 2006; Bennett & Swayze, 2010; Evers et al., 2015; Chery & Näär, 2016). Despite these disadvantages, PS oligonucleotides are still the most commonly used AOs and were the first antisense-based drug approved for clinical use in 1998 with fomivirsen (Vitravene®) (Roehr, 1998).

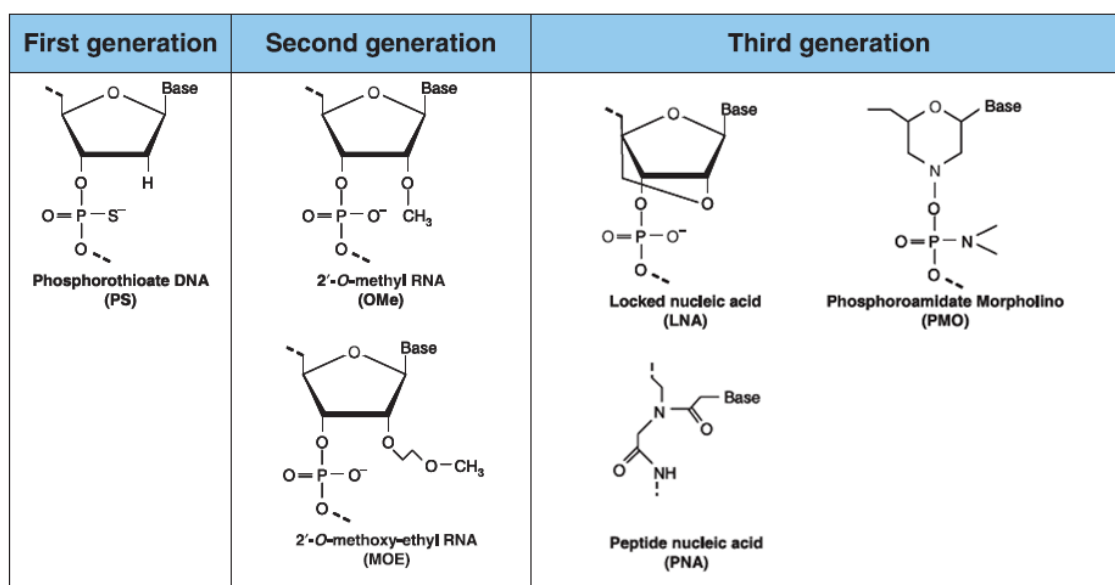


Figure 8 – Chemical modifications of antisense oligonucleotides. First generation oligonucleotides are characterized by a phosphorothioate (PS) backbone, obtained through the replacement of an oxygen atom with a sulfur atom. Second generation oligonucleotides contain a methyl or methoxyethyl group attached to 2' position of the ribose. Finally, third generation oligonucleotides carry modifications in the furanose ring structure. Adapted from Chan et al., 2006.

In order to surpass the downsides of the first generation oligonucleotides, a second generation was developed through modifications at the 2' position of the ribose (Figure 8). The most widely studied second generation AOs are 2'-O-methyl (2'-OMe) and 2'-O-methoxyethyl (2'-MOE), which present higher nuclease resistance and higher affinity for the target RNA, while also reducing non-specific protein binding and toxicity (Chan et al., 2006;

Bennett & Swayze, 2010; Chery & Nääär, 2016). These second generation AOs, however, do not support RNase H-mediated cleavage of the target mRNA, which impairs their usage for purposes of gene downregulation (Chan et al., 2006; Bennett & Swayze, 2010). To minimize this limitation, a chimeric ‘gapmer’ AO was developed with a central ‘gap’ region of approximately 10 PS-modified nucleotides flanked on both sides by ‘wings’ composed of five 2'-modified nucleotides. This allows for the binding of RNase H in the central gap region to promote cleavage and target degradation, while the external ‘wings’ provide higher nuclease resistance and higher target affinity (Chan et al., 2006; Bennett & Swayze, 2010). So far, only one second generation oligonucleotide drug – mipomersen (Kynamro®) – has been approved for clinical use in the treatment of familial hypercholesterolemia (Crooke & Geary, 2013; Chery & Nääär, 2016), while other second generation drugs are currently under clinical trials for transthyretin (TTR) amyloid cardiomyopathy, type 2 Diabetes and gastrointestinal autoimmune disease (Chery & Nääär, 2016).

Finally, the third generation of oligonucleotides is characterized by modifications of the furanose ring of the nucleotide, with the most common being peptide nucleic acids (PNAs), locked nucleic acids (LNAs) and phosphoroamidate morpholino oligomers (PMOs) (Figure 8) (Chan et al., 2006). These modifications further increase nuclease and protease resistance, target affinity, specificity and *in vivo* stability of antisense drugs, while also reducing non-specific interactions with proteins (Dias & Stein, 2002; Chan et al., 2006; Bennett & Swayze, 2010; Evers et al., 2015). Nevertheless, PNAs and PMOs present poor cellular uptake, low water solubility and are rapidly cleared from the blood due to their uncharged nature (Bennett & Swayze, 2010; Aartsma-Rus, 2012; Evers et al., 2015), whereas LNAs appear to generate higher toxicity than other chemical modifications, questioning their safety for therapeutic applications (Bennett & Swayze, 2010; Evers et al., 2015).

1.3.2. Oligonucleotides for splicing modulation

For the purpose of splicing modulation, AOs are designed to base-pair to their target pre-mRNA in regions that influence splicing to ultimately alter the sequence of the protein-coding mRNA (Havens et al., 2013). As such, second or third generation oligonucleotides are required in order to avoid RNase H-mediated cleavage and consequent mRNA degradation, which is promoted by unmodified and PS-oligonucleotides (Bauman et al., 2009). Splicing modulating AOs can then act through multiple mechanisms (Figure 9).

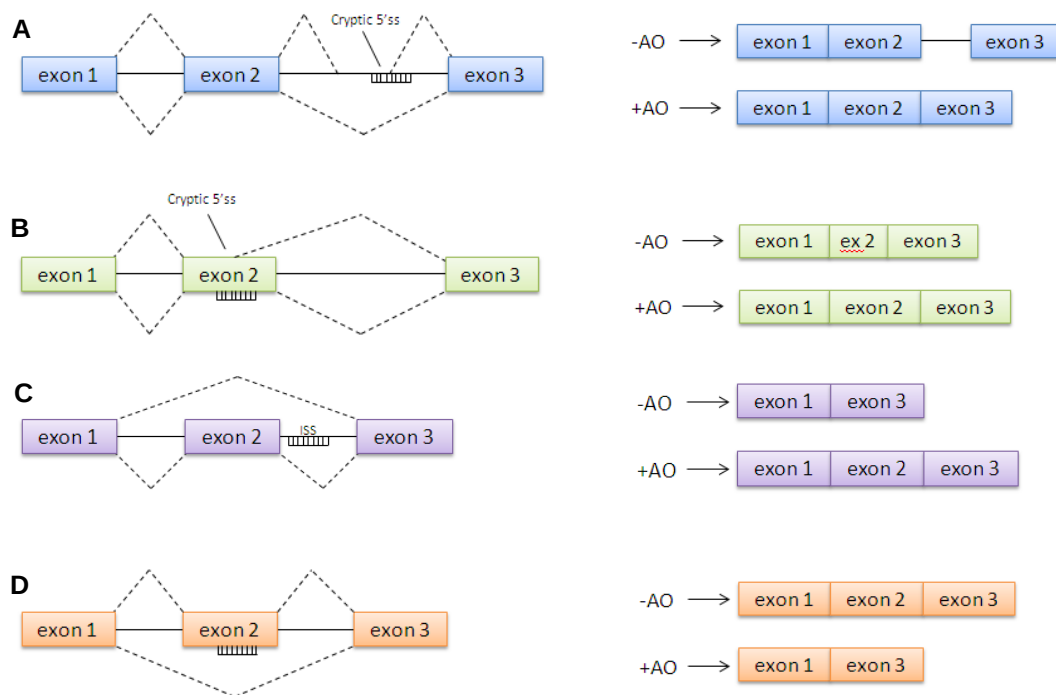


Figure 9 – Different methods of antisense-mediated splicing modulation. A: blocking of cryptic splice site within the intronic sequence; B: blocking of cryptic splice site within the exonic sequence; C: inclusion of alternatively spliced exon; D: exclusion of exon or "exon skipping". Adapted from Hammond & Wood, 2011.

On one hand, AOs can be used to block cryptic splice sites in the pre-mRNA, redirecting the splicing machinery to the canonical splice sites and restoring correct splicing of the transcript (Figure 9A and Figure 9B) (Havens et al., 2013; Arechavala-Gomez et al., 2014). This was actually the strategy employed in the first demonstration of splicing modulation using antisense mechanisms for the correction of causative mutations of β -thalassemia (Dominski & Kole, 1993), and later of ataxia-telangiectasia (Du et al., 2007), Niemann-Pick type C (Rodríguez-Pascau et al., 2009) and Pompe disease (Bergsma et al., 2016), among many other disorders. A different strategy in splicing modulation is isoform switching, where the goal is to alter the relative amounts of alternatively spliced isoforms; it is usually employed when there is a disruption in the balance of normally spliced transcripts, originating a pathological condition (Aartsma-Rus & van Ommen, 2007; Havens et al., 2013). Since disruption of isoform balance is often associated with tumorigenesis and tumor progression, this approach has shown promise for treatment of certain cancers (Renshaw et al. 2004; Williams & Kole, 2006; Zammarchi et al., 2011; Rigo et al., 2012). But in addition, it revealed to be useful for inflammatory diseases (Karras et al., 2000; Vickers et al., 2006; Graziewicz et al., 2008), fronto-temporal dementia and Parkinsonism associated with chromosome 17 (Kalbfuss et al., 2001; Peacey et al., 2012; Rigo et al., 2012) and

Hutchinson-Gilford progeria syndrome (Scaffidi & Misteli, 2005; Osorio et al., 2011). Moreover, AOs have also been used to promote the inclusion of exons that are normally excluded from the final transcript, with the targeting of the *SMN2* gene in spinal muscular atrophy (SMA) being the most common example (Figure 9C) (Hammond & Wood, 2011; Douglas & Wood, 2013). In this case, the 2'-MOE oligonucleotide nusinersen (Spinraza™) has been subject to multiple phase 2 and phase 3 clinical trials for many years, proving to be both safe and effective, and was recently approved by the US FDA for treatment of mild and severe SMA (Chiriboga et al., 2016; Ottesen, 2017).

Lastly, antisense technology has been employed to remove one or more exons from the pre-mRNA in order to eliminate a disease-causing mutation or to restore a transcript open reading frame (ORF), in a process known as “exon skipping” (Figure 9D). The process results in an internally deleted in-frame mature mRNA that will originate a truncated but functional protein, which can ameliorate the phenotype of multiple diseases (Hammond & Wood, 2011; Siva et al., 2014). One of the disorders amenable to be targeted with this strategy is Duchenne muscular dystrophy (DMD), caused by deletions or duplications in one or more exons of the dystrophin gene, disrupting the ORF and leading to a deficit of its encoded protein (Arechavala-Gomez et al., 2014). In this approach, AOs target specific exons to conceal them from the splicing machinery, causing them to be excluded from the mature transcript, which is later translated into an internally truncated but functional dystrophin similar to those associated with the much milder Becker muscular dystrophy (van Deutekom et al., 2007; Verhaart & Aartsma-Rus, 2012). The use of antisense technology as a treatment for DMD patients has experienced great development since its first application and rapidly progressed through human clinical trials (van Deutekom et al., 2007; Cirak et al., 2011; Mendell et al., 2013), culminating in the approval of the PMO eteplirsen (EXONDYS 51™) by the U. S. FDA in September 2016 (Aartsma-Rus & Krieg, 2017).

Despite antisense technology having emerged as a promising therapeutic strategy for multiple disorders caused by very distinct genetic defects, it has been insufficiently explored for the treatment of LSDs. Still, given the heavy burden of LSDs for patients and families and the lack of adequate treatments in the current medical practice, research and investment in new forms of therapy are crucial for the improvement of their condition. In this line of thought, AOs directed to specific mutations responsible for these disorders represent an important field of investigation and the main focus of this study.

Chapter 2

Objectives

While current knowledge on the mutations responsible for LSDs is already quite vast, allowing for early diagnosis and better disease management, similar progress has not been achieved in the field of treatment. In the specific case of ML II, treatments are still mostly symptomatic and unable to prevent death in the first decade of life.

Having as perspective to pave the way for future therapies, this project focused on the design and development of a specific RNA-based therapeutic strategy for the correction of the c.3503_3504delTC mutation, the most commonly found in ML II patients. The main purpose was the modulation of the pre-mRNA splicing pattern in order to restore the production of a partially functional GlcNAc-1-phosphotransferase enzyme and achieve positive therapeutic effects.

More specifically, the following aims were set:

- Establish stable cell lines of skin fibroblasts obtained from ML II patients harboring the c.3503_3504delTC mutation on exon 19 of the *GNTFAB* gene.
- Design AOs able to promote the skipping of the mutation-containing exon 19 in patient fibroblasts.
- Induce effectively the skipping of exon 19 in patient fibroblasts and consequently modulate the splicing pattern.
- Increase the activity of GlcNAc-1-phosphotransferase in patient fibroblasts, reducing the excretion of lysosomal enzymes and increasing their levels inside the lysosome.
- Contribute to a better global understanding of RNA therapies and splicing modulation in LSDs, consequently helping to reduce the current void in this kind of treatment for these disorders.

Chapter 3

Materials and Methods

3.1 Bioinformatical analysis

A bioinformatical analysis was conducted on exon 19 of the *GNPTAB* gene (harboring the mutation), as well as on flanking intronic regions from introns 18 and 19 and final protein product. It was performed to assess whether an exon skipping strategy could be applied to this protein without significantly affecting its activity and also to select the most appropriate sequence to be targeted by oligonucleotides with the purpose of inducing the skipping on exon 19. For this analysis, all nucleotide and peptidic sequences were obtained from the Ensembl database (<http://www.ensembl.org/>) (Yates et al., 2016) and the ExPASy Translate Tool (<http://web.expasy.org/translate/>) (Artimo et al., 2012) was used for the translation of nucleotide sequences into amino acid sequences.

3.1.1. Sequence alignment of *GNPTAB* protein products

In order to study whether the exclusion of exon 19 from the final *GNPTAB* transcript would result in an in-frame product and maintain the amino acid sequence encoded by the last two exons of the gene, the expected protein product after skipping of this target exon was compared with the normal protein product through the alignment of both amino acid sequences on Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (Sievers et al., 2011). This same software was also used for the alignment of amino acid sequences encoded by exon 19 in different animal species including birds, primates and other mammals, in order to assess the conservation of the amino acid sequence across different species which can be correlated with the presence of an important domain.

3.1.2. Prediction of protein domains within the *GNPTAB* protein product

Aside from considering the information already established in the literature, the presence of protein domains within the amino acid sequence encoded by exon 19 was further analyzed using three distinct bioinformatic tools: Pfam 30.0 (<http://pfam.xfam.org/>) (Finn et al., 2016), InterPro (<https://www.ebi.ac.uk/interpro/>) (Mitchell et al., 2015) and the UniProt database (<http://www.uniprot.org/>) (The Uniprot Consortium, 2015). Besides protein domains, the secondary structure and other important features of this protein product with and without exon 19 were also analyzed with these softwares, with special attention to the region encoded by the target exon.

3.1.3. Prediction of glycosylation sites within the *GNPTAB* protein product

To assess the possibility of post-translational modification sites within the amino acid sequence encoded by exon 19, glycosylation sites were predicted using two bioinformatical softwares: NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) (Gupta et al., 2004) and GlycoEP (<http://www.imtech.res.in/raghava/glycoep/submit.html>) (Chauhan et al., 2013), which provided valuable information regarding potential *N*-glycosylation, *O*-glycosylation and *C*-glycosylation sites and helped complement what had previously been identified in the literature.

3.1.4. Scoring of 5' and 3' splice sites and prediction of enhancers within *GNPTAB* exon 19

For the selection of possible targets for AO action, 5' and 3' splice sites of exon 19 were scored according to their strength, using MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) (Yeo & Burge, 2004) and NNSplice: Splice Site Prediction by Neural Network (http://www.fruitfly.org/seq_tools/splice.html) (Reese et al., 1997). In both cases, splice sites received a score that reflects their probability of being selected by the splicing machinery, with higher scores suggesting a stronger splice site and therefore a higher likelihood of being chosen.

As for the prediction of enhancer elements, the softwares ESEfinder 3.0 (<http://krainer01.cshl.edu/cgi-bin/tools/ESE3/ese finder.cgi?process=home>) (Cartegni et al., 2003; Smith et al., 2006) and RESCUE-ESE (<http://genes.mit.edu/burgelab/rescue-ese/>) (Fairbrother et al., 2002) were used. Despite the fact that these bioinformatical tools are specifically designed to predict exonic enhancers, 50bp of the flanking intronic sequences from intron 18 and 19 were also analyzed together with the sequence of exon 19 to allow for the identification of possible intronic enhancers in close proximity to the target exon.

3.2. Cell culture and establishment of stable fibroblast cell lines

Skin fibroblasts of two distinct patients with an established ML II diagnosis, both homozygous for the c.3503_3504delTC mutation, were obtained from the Cell Line and DNA Biobank from Patients Affected by Genetic Diseases (Istituto Giannina Gaslini, Genova,

Italy), while control cells were retrieved from healthy donors. Both control and patient cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco® Life Technologies, Carlsbad, California, USA) supplemented with 10% fetal bovine serum (FBS), 5% penicillin/streptomycin (PenStrep) antibiotic and 5% Fungizone® (amphotericin B), occasionally increasing the supplementation to 20% FBS when cells exhibited very low growth rates. Cultured cells were kept in T-75 culture flasks in an incubator set at the temperature of 37°C and 5% CO₂ atmosphere, receiving regular renovation of the culture medium.

When 80-90% confluence was reached, culture medium was removed from the flask, cells were washed with phosphate buffered saline (PBS) and trypsin was added to dissociate cells from the culture flask. After 5 minutes at 37°C for optimal trypsin activity, cell dissociation was observed with an inverted microscope and cells were sub-cultured in new flasks with the appropriate volume of fresh medium. To save fibroblast cells for future use, several samples of each were cryopreserved in the first sub-culture passages. Also, the sub-culture procedure was conducted a minimum of three times for each cell line before initiating any transfection experiments.

3.3. Confirmation of the reported mutation in patient fibroblast cells

Both patient fibroblast cells were subject to molecular confirmation of the reported c.3503_3504delTC mutation through amplification and sequencing of the target region to ensure the mutation was indeed present in the cultured cells. Healthy control fibroblasts genomic pattern was also screened for comparison with that observed in patient cells.

3.3.1. Extraction of genomic DNA

Firstly, the medium in the fibroblast cells containing flasks was removed by aspiration and cells were washed with PBS. Trypsin was then added to promote cell dissociation from the flask surface followed by a short period of incubation at 37°C for optimal trypsin activity. After incubation, 4mL of DMEM medium were added to inactivate the enzyme and the entire volume was transferred to a 15mL tube for centrifugation at 1500 rpm for 5 minutes. Following centrifugation, the supernatant was removed and the resulting pellet was resuspended in 1mL of PBS and transferred to a 1.5mL tube for a second centrifugation at 13 200 rpm for 5 minutes and supernatant removal.

Genomic DNA was then extracted automatically from the cell pellets using the BioRobot EZ1[®] System and the EZ1 DNA Tissue Kit (QIAGEN, Hilden, Germany) with a final elution volume of 200µL. Following extraction, the amount of DNA in the samples was quantified using NanoDrop[®] ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

3.3.2. Amplification of exon 19 of the *GNPTAB* gene

After extraction, genomic DNA was used for the amplification of exon 19 of the *GNPTAB* gene through PCR using ImmoMix[™] Red (Bioline, London, UK) and two primers previously described by Kudo et al., 2006: CCCATAGCTAAAAGGCCATCTACC (forward) and GTATACACTCACCCACACACATGC (reverse). Following the manufacturer's instructions, the PCR amplification was carried out using 1µL of genomic DNA, 1X ImmoMix[™] Red reaction mix and 0,5µM of each primer. The amplification program was composed of an initial denaturation step at 95°C for 7 minutes, followed by 35 cycles of denaturation at 94°C for 45 seconds, primer annealing at 60°C for 45 seconds and DNA strand extension at 72°C for 90 seconds, and finished with a final extension at 72°C for 10 minutes.

3.3.3. Evaluation of PCR amplification by agarose gel electrophoresis

Following amplification, products were separated through agarose gel electrophoresis in a 2% agarose gel stained with 5µL of ethidium bromide. Volumes of 8µL of each sample and 6µL of 100 bp DNA Ladder Ready to Load (Solis BioDyne, Tartu, Estonia) were loaded into the gel and submitted to an 115V voltage for 55 minutes. After separation, the gel was visualized with ultraviolet light and images were obtained using the Molecular Imager[®] Gel Doc XR+ System and Image Lab[™] software (BioRad, Hercules, California, USA).

3.3.4. Sequencing of amplification products

After confirmation of the success of the amplification reaction through gel electrophoresis, the *GNPTAB* PCR products from healthy control and the two patients were purified using the illustra[™] ExoStar[™] 1-Step Kit (GE Healthcare Life Sciences, Chicago, Illinois, USA) so that they could be subsequently sequenced. Adapting the manufacturer's protocol, the reactions were composed of 3µL of the amplification product and 1µL of the

ExoStar™ mix and were incubated at 37°C for 30 minutes for optimal enzyme activity, followed by 15 minutes at 80°C to inactivate the enzyme.

Immediately after purification, the purified amplification products underwent two separate sequencing reactions (forward and reverse) using the BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California, USA). Both reactions were composed of 0,5µL of the BigDye® reaction mix, 0,5µL of the BigDye® reaction buffer, 0,5µL of the corresponding primer (5µM), 2µL of the purified amplification product and 6,5µL of sterile water to complete a final volume of 10µL. PCR conditions consisted in an initial denaturation step of 10 minutes at 96°C, followed by 25 cycles of denaturation at 96°C for 10 seconds, primer annealing at 50°C for 5 seconds and strand extension at 60°C for 4 minutes. The sequencing products of the patient fibroblast cells were then purified and separated through capillary electrophoresis in an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, Foster City, California, USA) and the resulting electropherograms were analyzed using the Chromas Lite 2.01 software (Technelysium, South Brisbane, Australia). Sequences were then compared with the *GNPTAB* reference sequence ENSG00000111670 (retrieved from the Ensembl database) using the Clustal Omega bioinformatic tool.

3.4. Analysis of endogenous splicing pattern in patient fibroblast cells

Aside from the confirmation of the presence of the c.3503_3504delTC mutation in the cultured patient fibroblasts, the endogenous splicing pattern of the *GNPTAB* gene in these cells was also examined to ensure that no alterations were present prior to the modulation of splicing by the AOs. Once again, healthy control fibroblasts were also analyzed for comparison.

3.4.1. Extraction of total RNA

Fibroblast cell pellets were collected through a similar process to the one used for genomic DNA extraction described in 3.3.1. Therefore, culture medium was removed and cells were washed with PBS, followed by dissociation from the flask through trypsinization. Afterwards, cells were centrifuged and supernatant was removed to obtain a final cell pellet.

Afterwards, total RNA was extracted from cell pellets with the GRS Total RNA Kit – Blood & Cultured cells (GRiSP, Porto, Portugal) with a final elution volume of 50µL. The

amount of RNA in the samples was then assessed using NanoDrop® ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

3.4.2. Synthesis of cDNA

After RNA extraction and quantification, cDNA synthesis was carried out using the Ready-To-Go™ You-Prime First-Strand Beads kit (GE Healthcare Life Sciences, Chicago, Illinois, USA) with an initial sample volume containing 0,5µg of total RNA to which RNase-free deionized water was added to complete a final volume of 30µL. RNA samples were then incubated at 65°C for 10 minutes to disrupt any RNA secondary structure and placed on ice for a minimum of 2 minutes afterwards. Next, the RNA samples were transferred to the bead-containing tubes included in the kit, as well as 2µL of RNase-free deionized water and 1µL of oligo(dT)₁₈ primer mix (NZY Tech, Lisbon, Portugal) at 50µM for a final volume of 33µL. After 1 minute of room temperature incubation, the tube contents were vortexed and incubated at 37°C for 60 minutes for the synthesis of the cDNA strand. The resultant cDNA was stored at -20°C until future use.

3.4.3. Amplification of *GNPTAB* transcript cDNA

After synthesis, cDNA was used for the amplification of the 21 exons of the *GNPTAB* transcript through PCR using ImmoMix™ Red (Bioline, London, UK) and the primers described in Table 2. The PCR reaction was carried out using 2µL of cDNA, 1X ImmoMix™ Red reaction mix and 0,5µM of each primer. For reactions displaying large amounts of unspecific products, 5% of betaine was added to the reaction corresponding to a volume of 1,25µL, and for those with low intensity levels of the specific amplicon, the cDNA volume was increased to 3µL. The amplification program was composed of an initial denaturation step at 95°C for 7 minutes, followed by 35 cycles of denaturation at 94°C for 45 seconds, primer annealing at the appropriate temperature for each primer set for 45 seconds and DNA strand extension at 72°C for 60 seconds, finishing with a final extension at 72°C for 10 minutes. Primer sequences and annealing temperature are described in Table 2. Following amplification, the success of the PCR reaction was assessed through agarose gel electrophoresis as described in 3.3.3 and amplification products were sequenced as detailed in 3.3.4. Afterwards, sequences were compared with the *GNPTAB* transcript reference sequence ENST00000299314.11 (retrieved from the Ensembl database) using Clustal Omega.

Table 2 – Sequences and annealing temperatures of primers used in the cDNA amplification of *GNPTAB* gene.

Primer	Sequence (5' – 3')	Exons	T _{annealing} (°C)
cDNA 1F	AGCTGCAATGAGCGGCGC	1-6	59
cDNA 1R	AGCAGTCCAGAGTGGGCA		
cDNA 2F	CCCTTCTACCAATGTCTCAG	5-10	55
cDNA 2R	CCTTCCCAAACATGACATCA		
cDNA 3F	AAGGGCTGTCCCAGAAGTTT	10-13	55
cDNA 3R	TTGGTCCCTCCCTTGTGTC		
cDNA 4F	CAGATAACAGTGGAGGTGGA	13	58
cDNA 4R	TGTGTAATGCTGCAGCTTTC		
cDNA 5F	GCTGAAAATCACATAGGCGT	13-18	55
cDNA 5R	CTTCTTCTCCCATGATTTC		
cDNA 6F	CCAGTAACTGACAAAATCCA	17-21	55
cDNA 6R	ACAGGTCCATGAGCAAATTC		

3.5. Design and synthesis of antisense oligonucleotides for the skipping of *GNPTAB* exon 19

For the design and selection of the most promising AOs, we followed the guidelines and parameters previously described by Aartsma-Rus (Aartsma-Rus, 2012). Secondary structure of the target exon and flanking intronic regions was assessed using the mfold software (<http://unafold.rna.albany.edu/?q=mfold>) (Zuker, 2003). From the most stable structures obtained here, partially open and partially closed regions were selected for further analysis as these are the most appropriate locations for AO binding. Next, this information regarding the secondary structure of the transcript was combined with the data previously collected from ESEfinder 3.0 software concerning the presence of enhancers along the sequence of exon 19 in order to select those regions where both analysis overlap, that is, the partially open and partially closed regions that also contain the highest/biggest number of predicted enhancer elements. After this process, two internal regions within exon 19, as well as the acceptor and donor splice sites, were selected to be the targets for AO binding.

With the target locations established, several AOs spanning around 20 nucleotides were designed for each region and tested for parameters like GC percentage, which should be between 40% and 60%, and melting temperature (T_m), which should be above 48°C (Aartsma-Rus, 2012). Free energy of the AO and AO-AO complexes was calculated using the software RNAstructure (<http://rna.urmc.rochester.edu/RNAstructureWeb/index.html>)

(Reuter & Mathews, 2010). Ideally, the free energy of the AO should be above -4, while the free energy of the AO-AO complexes should be above -15 (Aartsma-Rus, 2012) and the AOs failing these numbers were discarded. Using the same software, the free energy of the exon and the free energy of the AO-target complex were obtained, allowing for the calculation of the binding energy between the AO and the target sequence, which results from the following formula:

$$B = \Delta G_e - \Delta G_t$$

where B is the binding energy between the AO and the target sequence, ΔG_e is the free energy of the exon and ΔG_t is the free energy of the AO-target complex. This binding energy is a measure of the strength of the annealing between the AO and its target sequence and should ideally be between 21 and 28 for optimal results (Aartsma-Rus, 2012).

Finally, after narrowing the number of AO to the most promising candidates, those were subject to a BLAST analysis using the nucleotide BLAST tool hosted by the NCBI database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Altschul et al., 1990), in order to ensure their uniqueness and avoid the possibility of off-target effects. Stretches of 15 or more homologous nucleotides should be avoided as much as possible, even in the presence of mismatches, as oligonucleotides with up to five mismatches have been reported to successfully bind to their targets and induce exon skipping (Errington et al., 2003). Although total uniqueness was impossible to find among the tested oligonucleotides, those with the least amount of BLAST hits spanning 15 or more nucleotides saw those hits analyzed regarding their pattern of expression according to The Human Protein Atlas project (<http://www.proteinatlas.org/>) (Pontén et al, 2008). Using this webtool, we were able to examine the expression of the genes with the highest levels of similarity to our oligonucleotides and select those AOs which affected the least amount of tissues, reducing the possibility of off-target effects in multiple organs. Also, oligonucleotides with strong similarity to genomic regions but not transcript sequences were preferred as the RNA structure of these antisense molecules makes them less likely to interact with the genomic DNA.

At last, four AOs were selected for transfection in patient fibroblasts, targeting both splice sites and two internal regions containing enhancer elements as predicted by ESEfinder 3.0 (Figure 10 and Table 3). The selected oligonucleotides were then synthesized and purified by EuroGentec (Seraing, Belgium) as second generation 2'-OMe RNA molecules containing a full-length phosphorothioate backbone.

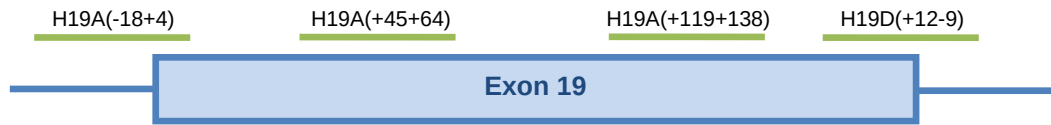


Figure 10 – Location of antisense oligonucleotide targets on exon 19 of the *GNPTAB* gene and flanking intronic sequences.

Table 3 – Sequences of antisense oligonucleotides selected for transfection in ML II patient fibroblasts. Nomenclature as described by Mann et al., 2002.

Name	Sequence (5' – 3')	Length (nt)	GC content (%)	T _m	Binding energy
H19A(-18+4)	CUUCCUGAAUAACAGAGAGCC	22	45.45	53	25.4
H19A(+45+64)	AGCCUUCACUGUCUGAGCAU	20	50	51.8	27.5
H19A(+119+138)	CGGUUUCGAUACUCUCUUGG	20	50	51.8	31.1
H19D(+12-9)	AGAAUUUACCAUCCUGCAGC	21	42.86	50.5	22.2

3.6. Transfection of healthy control and ML II patient fibroblasts with antisense oligonucleotides

After selection and synthesis of the AOs and the establishment of stable cell lines from ML II patient fibroblasts, cells were plated and transfected with each oligonucleotide in order to promote the exclusion of exon 19 from the final *GNPTAB* transcript. Fibroblasts from healthy individuals were also transfected as controls.

3.6.1. Cell counting and plating

Cells were cultured in 6-well (10 cm²) plates at a density of 275 000 to 300 000 cells per well with DMEM medium supplemented with 10% FBS, 5% PenStrep and 5% Fungizone®. To calculate the appropriate volume of cell suspension to be plated in each well, cells were counted using a Neubauer chamber as seen in Figure 11.

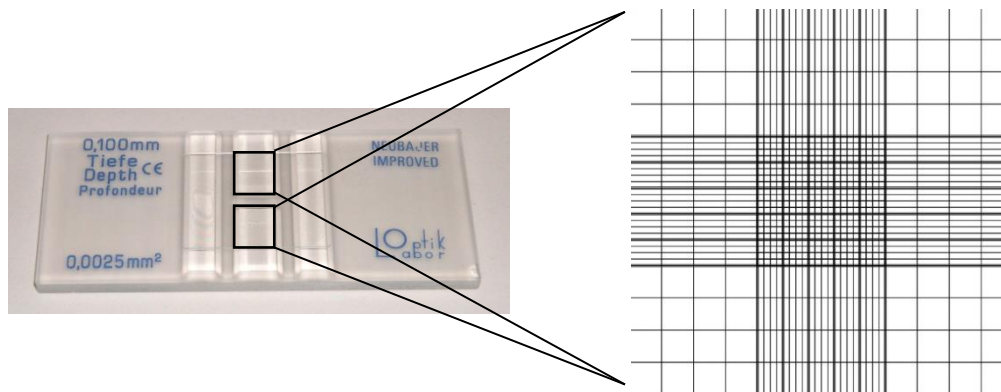


Figure 11 – Example of Neubauer chamber used for cell counting. The Neubauer chamber contains two distinct chambers, composed of 9 quadrants, each with a volume of 0,1 mm³ or 10⁻⁴ mL.

A small volume of cell suspension was deposited on each side of the chamber and cells present in all four external quadrants of both sides of the chamber were counted in an inverted microscope. Considering the dimensions of the chamber (1 mm x 1 mm x 0,1 mm), each quadrant has a total volume of 0,1 mm³, which equals 10⁻⁴ mL and the total cell concentration was calculated with the following equation:

$$N = \frac{\sum n}{8} \times 10^4$$

where N is the total number of cells per milliliter, n is the number of cells counted in each quadrant of the Neubauer chamber and multiplication by 10⁴ allows for the correction to the total number of cells in 1mL of cell suspension.

After being plated with the appropriate volume of cell suspension and growth medium, the 6-well plates were incubated at 37°C with a 5% CO₂ atmosphere for 16-18 hours prior to transfection.

3.6.2. Transfection

After reaching 80-90% confluence, patient and healthy control fibroblasts were transfected with the four AOs detailed in Table 3 and Figure 10, each in four different concentrations: 10nM, 25nM, 50nm and 100nM. Using the Lipofection method for transfection, the appropriate amount of each AO was initially diluted in Opti-MEM® medium to complete a final volume of 250μL, while 10μL of the transfection reagent Lipofectamine® 2000 were diluted in 240μL of Opti-MEM® medium for a final volume of 250μL. After an

incubation period of 10 minutes at room temperature, the lipofectamine solution was added to the oligonucleotides to form the AO-lipid complexes, completing a total of 500µL, and this final mix was incubated a second time for 20 minutes at room temperature.

Meanwhile, growth medium was removed from the plate wells and fibroblast cells were washed with 2mL PBS before receiving 1mL of fresh antibiotic/fungizone-free DMEM medium. After the second incubation period was finished, the total volume of each AO-lipid complex mix was added dropwise to the corresponding well and cells were incubated at 37°C with a 5% CO₂ atmosphere for 5 hours. After this time period, culture medium was removed and 2mL of fresh DMEM medium supplemented with 10% FBS, 5% PenStrep and 5% Fungizone® were added to the cells, which were incubated once again at 37°C, 5% CO₂ until completing a 24 hour period after transfection.

After this initial transfection with lower concentrations, the two most effective AOs – H19D(+12-9), targeting the 5'ss, and H19A(+119+138), directed at ESEs within the second half of exon 19 – were tested at higher doses to further analyze their potential in the splicing modulation of the target exon. Patient and healthy control fibroblasts were once again transfected and five different concentrations, ranging from 100nM to 500nM, were used, with all transfection and culture procedures identical to those employed for the lowest concentrations.

Additionally, a plasmid containing the reporter gene *eGFP*, which encodes for the enhanced green fluorescent protein, was used as a control to assess the efficacy of the transfection reagent through the analysis of the fluorescence observed among the transfected fibroblasts. In this case, 2,5µL of plasmid solution, corresponding to 2,5µg of DNA, were transfected with 10µL of Lipofectamine® 2000 used and fluorescence was observed in a fluorescence microscope 24 hours after.

3.7. Evaluation of *in vitro* GNPTAB gene exon 19 splicing pattern modulation with antisense oligonucleotides

Following transfection with AOs, the splicing pattern of the *GNPTAB* transcript was once again analyzed to verify whether the oligonucleotides successfully promoted the exclusion of this target exon from the final mRNA product. All procedures were identical to those described in 3.4, with only a few adjustments regarding the collection of cell pellets, using 500µL of trypsin and 1mL of DMEM medium for that purpose given the small volume of

the plate wells. Also, for cDNA synthesis, the specific primer GNPTAB cDNA 6R from Table 2 was used instead of the oligo(dT)₁₈ primer mix previously chosen. After amplification, when a shorter fragment had been observed through gel electrophoresis, suggesting the successful skipping of exon 19, products were run once again on agarose gel and the shorter fragments were cut off for purification with the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, Wisconsin, USA) followed by sequencing as described before.

3.8. Quantification of lysosomal enzyme activity in healthy control and ML II patient fibroblasts after splicing pattern modulation with antisense oligonucleotides

Following the observation of successful skipping of exon 19 from the *GNPTAB* final transcript, the effectiveness of this therapeutic approach regarding the restoration of GlcNAc-1-phosphotransferase activity was assessed indirectly through the quantification of a number of lysosomal enzymes. For that purpose, fibroblasts from Patient 1 were plated in four wells of 6-well plates and once again transfected with the H19A(+119+138) AO at the highest concentration tested (500nM) as described in 3.6.2. After the incubation periods of 24 and 48 hours, cells were dissociated and collected as previously detailed, with the volume of three wells used for the enzyme assays, while the volume of one well was separated for RNA analysis. The same experiment was done with fibroblasts from healthy individuals as a control. Additionally, the two cell lines were treated only with Lipofectamine® 2000 in order to account for the influence of this transfection reagent in the experiments.

Afterwards, pellets of each experiment were diluted in 50µL Triton 0,1% and sonicated in order to lyse the cells and release their internal components. Samples were then centrifuged at 13 000rpm for 10 minutes to collect all cellular debris at the bottom of the tube and the amount of total protein in the supernatant was quantified with NanoDrop® ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). From here, the appropriate volume of each sample for a final concentration of 0,5mg/mL was transferred into a new 1,5mL tube, completing the final volume with sterile water. Next, the reaction mixture for each enzyme was prepared and incubated at 37°C, with all assays conducted in duplicate to ensure the reproducibility of the results. Reaction composition and incubation times are detailed on Table 4.

Table 4 – Composition of each enzyme activity assay and respective incubation period at 37°C.

Enzyme	Protein sample	Substrate	Incubation time
β-glucosidase^a	10μL	25μL 4-MU-β-D-glucopyranoside 10mM	2h
α-galactosidase	10μL	50μL 4-MU-α-D-galactopyranoside 5mM	1h30
Arylsulfatase B (ARSB)	50μL	150μL 4-MU-sulfate 10mM	1h
β-glucuronidase	10μL	50μL 4-MU-β-D-glucuronide 4mM	45min
β-galactosidase	10μL	50μL 4-MU-β-D-galactopyranoside 1mM	30min
α-mannosidase	10μL	50μL 4-MU-α-D-mannopyranoside 2mM	30min
Total hexosaminidase	10μL ^b	50μL 4-MU-2-acetamido-2-deoxy-β-D-glucopyranoside 2mM	15min

^a 25μL of phosphate/citrate buffer pH 5,2 were also added and two different assays were conducted: one using a substrate solution with sodium taurocholate (T+) and a second without sodium taurocholate (T-).

^b Samples were diluted in albumin in a ratio of 10μL of sample to 30μL of albumin, and from this dilution 10μL were used for the assay.

After incubation was completed, 1mL of glycine was added to each test tube to stop the enzyme reaction and 300μL of each sample were transferred to a fluorometer plate well for fluorescence reading using the VICTOR 2TM D fluorometer (PerkinElmer, Waltham, Massachusetts, USA) at the wavelengths of 355 and 460nm. Results were then compared between ML II patients, healthy controls and other LSD patients, as well as before and after AO treatment.

Chapter 4

Results

4.1. Bioinformatical analysis

Several bioinformatics tools for the analysis of sequence information were used to explore thoroughly i) exon 19 of the *GNPTAB* gene where is located the c.3503_3504delTC mutation, the target molecular defect in this study from the therapeutic perspective; ii) its flanking intronic regions and iii) the final protein product.

4.1.1. Sequence alignments of *GNPTAB* protein products

We first sought to compare the amino acid sequences of both full length and truncated protein products in order to confirm whether the skipping of an exon in the final transcript would originate an in frame protein product, maintaining intact the remaining amino acid sequence. For this purpose, the amino acid sequence of the *GNPTAB* transcript minus exon 19 was aligned with the sequence of the complete *GNPTAB* protein product using Clustal Omega software to assess whether the exclusion of this exon from the mature mRNA would result in an in frame protein or if instead it would alter more drastically the final protein product, rendering any exon skipping approach unfit for the correction of the mutation. Following this alignment, it was possible to confirm that the exclusion of exon 19 from the final *GNPTAB* transcript originates an internally truncated protein with a total of 1200 amino acids, 56 amino acids shorter than the complete protein product, but that the reading frame remains intact, since the amino acid sequence without the excluded region presents a complete match with the normal *GNPTAB* protein (Figure 12). Furthermore, the amino acid arginine encoded across exon boundaries is also maintained, despite being the most likely to be altered during an exon skipping strategy even if the reading frame is unaltered and the amino acids beyond match the original protein sequence. There are codons for arginine across the boundaries between exons 18 and 19, as well as between exons 19 and 20, and thus the amino acid remains unchanged in the truncated protein since exon 19 and exon 20 both start with a guanine nucleotide, maintaining the AGG codon, which codes for arginine, even after removal of exon 19. As such, the truncated protein obtained after the skipping of exon 19 exhibits a complete match with the full length *GNPTAB* protein product, maintaining the reading frame and amino acid sequence but missing the 56 amino acids spanning positions 1146 through to 1201 encoded by the nucleotide sequence of exon 19 which has been removed from the final mRNA transcript.

```

protein      IDRIVMQLQDMFPEEFDKTSFHKVRHSEDMQFAFSYFYLLMSAVQPLNISQVFDVDDTD
skip19      IDRIVMQLQDMFPEEFDKTSFHKVRHSEDMQFAFSYFYLLMSAVQPLNISQVFDVDDTD
*****

protein      QSGVLSREIRTLATRIHELPLSLQDLTGLEHMLINCSKMLPADITQLNNIPPTQESTYD
skip19      QSGVLSREIRTLATRIHELPLSLQDLTGLEHMLINCSKMLPADITQLNNIPPTQESTYD
*****

protein      PNLPPVTKSLVTNCKPVTDKIHKATKDKNKYRFEIMGEBEIAFKMIRTNVSHVVGQLDDI
skip19      PNLPPVTKSLVTNCKPVTDKIHKATKDKNKYRFEIMGEBEIAFKMIRTNVSHVVGQLDDI
*****

protein      RKNPBKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
skip19      RKNPR-----
*****

protein      WRAYRDKLKFVTHCVLATLIMFTIPSFFAEQIALKPKIFPRRIHKEASPNRIRV
skip19      -RAYRDKLKFVTHCVLATLIMFTIPSFFAEQIALKPKIFPRRIHKEASPNRIRV
*****

```

Figure 12 – Alignment in Clustal Omega software of amino acid sequences for the full length *GNPTAB* protein product and truncated protein resulting from the skipping of exon 19. Due to the long size of the protein sequence, only exons 14 to 21 are shown in this picture. protein: full length protein product; skip19: truncated protein.

After confirming this basic pre-condition to implement an exon skipping strategy, the amino acid sequence encoded by exon 19 was further analyzed to ensure that its absence would not severely impact the activity of the final protein. Firstly, amino acid sequences encoded by exon 19 in different animal species were compared using Clustal Omega software in order to assess their conservation across different taxa and obtain insights on the presence of important protein domains in the region. Results from this alignment, shown in Figure 13, reveal a great level of conservation of this sequence in different animal species, with most amino acids identical across all species studied and only minor discrepancies in some residues in mouse and zebrafish. This degree of conservation suggests the presence of an important domain in this region of the protein, which could represent a serious obstacle to the exon skipping approach for the modulation of this mutation without impairing the proper activity of the final protein product. Nonetheless, although important protein domains usually show their sequences highly conserved across species, the opposite is not always true, given that there are known instances of conserved amino acid sequences despite not containing any important domain. Therefore, further analysis was conducted to check whether or not the sequence encoded by exon 19 corresponded to a crucial protein region and consequently whether or not an exon skipping approach could be considered for exon 19 without severely impacting the final protein product.

```

human      RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
cat        RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
chicken    RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
chimpanzee RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
cow        RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
duck       RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
elephant   RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
monkey     RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
horse      RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
rabbit     RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
rat        RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
sheep      RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
mouse      RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
zebrafish  RKFVCLNDNIDHNHKAQTVKAVLRDFTESMFPFPSQFELPREYRNRFLLHMHLEQW
***-*****  **** *****-*****

```

Figure 13 – Alignment in Clustal Omega software of amino acid sequences encoded by exon 19 of the *GNPTAB* gene in different animal species.

4.1.2. Prediction of protein domains within the *GNPTAB* protein product

To predict the presence of important domains within the protein region of *GNPTAB* encoded by exon 19, the Pfam 30.0 database was used, revealing the presence of multiple protein domains with a variety of functions across the entire sequence of the *GNPTAB* protein as seen in Table 5 and Figure 14.



Figure 14 – Protein domains predicted by Pfam 30.0 for the *GNPTAB* protein product. Color codes and exact positions are depicted in Table 5.

The expected N-terminal and C-terminal transmembrane domains essential for the translocation of the α/β -precursor from the endoplasmic reticulum to the Golgi apparatus where it is later cleaved into the α and β subunits of GlcNAc-1-phosphotransferase were clearly identified at both ends of the protein sequence (red boxes in Figure 14), while two Notch domains were also found to be present within the protein (dark blue boxes in Figure 14). Furthermore, this protein has one DNA methyltransferase-associated protein (DMAP) domain (yellow in Figure 14), and four stealth conserved regions (CR) across its amino acid sequence (green, red, purple and orange in Figure 14), with CR4 (positions 1138-1194) covering almost the entire sequence encoded by exon 19 (1146-1201).

Table 5 – Protein domains predicted by Pfam 30.0 for the *GNPTAB* protein product. n/a – not applied.

Source	Domain	Start	End
transmembrane	n/a	20	38
Pfam	Stealth CR1	73	101
coiled_coil	n/a	86	113
disorder	n/a	104	117
disorder	n/a	195	196
disorder	n/a	280	284
Pfam	Stealth CR2	322	429
Pfam	Notch	435	469
low_complexity	n/a	461	474
low_complexity	n/a	469	489
Pfam	Notch	502	536
disorder	n/a	628	629
disorder	n/a	636	637
disorder	n/a	640	642
disorder	n/a	677	683
Pfam	DMAP binding	699	814
disorder	n/a	750	751
disorder	n/a	757	762
disorder	n/a	764	769
disorder	n/a	773	776
disorder	n/a	795	869
low_complexity	n/a	846	857
Pfam	Stealth CR3	955	1003
disorder	n/a	1076	1079
Pfam	Stealth CR4	1138	1194
transmembrane	n/a	1210	1235

The bioinformatic analysis with InterPro database also predicted the presence of two Notch domains, one DMAP domain and four stealth regions within the *GNPTAB* protein product, and confirmed that stealth CR4 encompassed the amino acid sequence encoded by exon 19 (Figure 15). Additionally, this software pinpointed the presence of an EF-hand domain responsible for calcium binding, which was not predicted by Pfam 30.0. Aside from these annotated domains, other matches were identified among unintegrated signatures (shown in grey), indicating the N-terminal and C-terminal transmembrane domains, the cytoplasmic and non-cytoplasmic regions, the two GlcNAc-1-phosphotransferase subunits and some structural information.

Domains and repeats

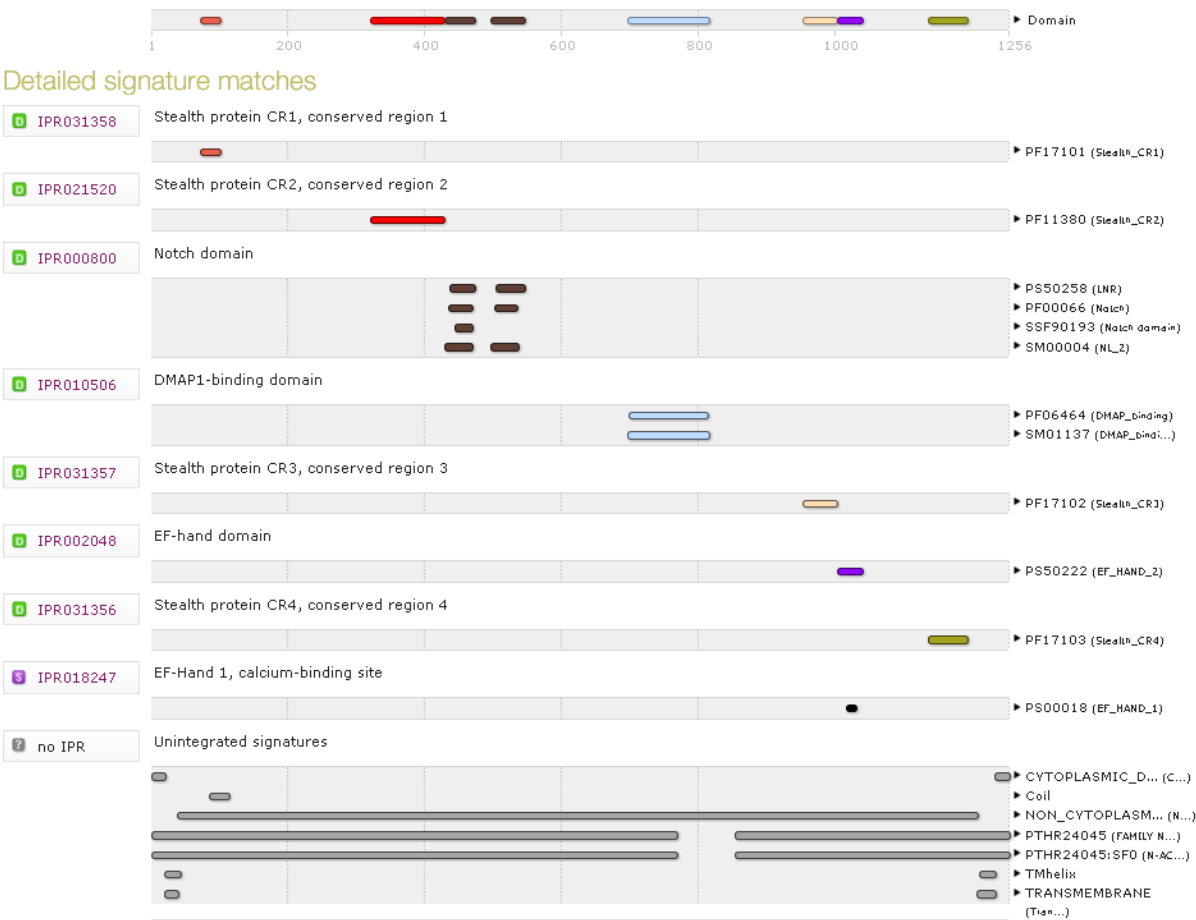


Figure 15 – Protein domains predicted by InterPro for the *GNPTAB* protein product.

Finally, the GNPTAB protein was also analyzed in the UniProt database concerning protein domains, functions and secondary structure. In the case of protein domains, the UniProt resource provided results coincident with the two previous bioinformatic tools, including this protein within the stealth protein family and confirming the presence of two Notch domains, one DMAP domain and one EF-hand domain (Table 6).

Table 6 – Protein domains predicted by UniProt for the *GNPTAB* protein product.

Feature key	Position(s)	Length	Description	Graphical view
Repeat ⁱ	438 – 473	36	LNR 1	
Repeat ⁱ	505 – 545	41	LNR 2	
Domain ⁱ	700 – 813	114	DMAP-interaction	
Domain ⁱ	1005 – 1040	36	EF-hand	

Similarly to InterPro and Pfam 30.0, UniProt also predicted both N-terminal and C-terminal transmembrane domains, as well as their helical conformation (Table 7).

Table 7 – Transmembrane domains predicted by UniProt for the *GNPTAB* protein product.

Feature key	Position(s)	Length	Description	Graphical view
Transmembrane ⁱ	22 – 42	21	Helical Sequence analysis	
Transmembrane ⁱ	1215 – 1235	21	Helical Sequence analysis	

Conversely, this database anticipated that GNPTAB harbors multiple metal binding sites for calcium (Table 8) which had not been detected by the previous tools, as well as a calcium binding region which had been suggested by InterPro but not Pfam 30.0 (Table 9). Nonetheless, none of these sites and regions is located in the amino acid sequence encoded by exon 19 (1146-1201).

Table 8 – Metal binding sites predicted by UniProt for the *GNPTAB* protein product.

Feature key	Position(s)	Length	Description	Graphical view
Metal binding ⁱ	449 – 449	1	Calcium PROSITE-ProRule annotation	
Metal binding ⁱ	464 – 464	1	Calcium PROSITE-ProRule annotation	
Metal binding ⁱ	467 – 467	1	Calcium PROSITE-ProRule annotation	
Metal binding ⁱ	516 – 516	1	Calcium PROSITE-ProRule annotation	
Metal binding ⁱ	531 – 531	1	Calcium PROSITE-ProRule annotation	
Metal binding ⁱ	534 – 534	1	Calcium PROSITE-ProRule annotation	

Table 9 – Calcium binding region predicted by UniProt for the *GNPTAB* protein product.

Feature key	Position(s)	Length	Description	Graphical view
Calcium binding ⁱ	1018 – 1029	12	PROSITE-ProRule annotation	

The UniProt database further revealed some relevant information regarding the secondary structure of GNPTAB, predicting several beta strands, helices and turns, all between positions 138 and 280 (Figure 16 and Table 10). Once again, no remarkable features were detected for the region encoded by exon 19.

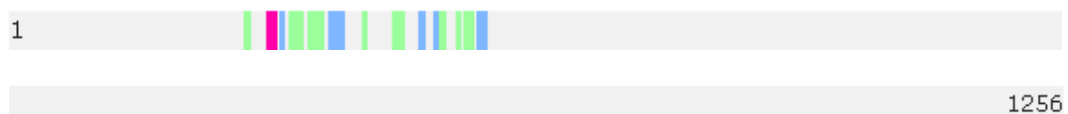
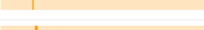
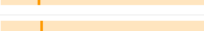
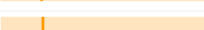
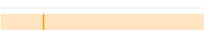
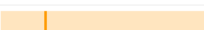
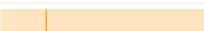


Figure 16 – Secondary structure of the GNPTAB protein predicted by UniProt software. Green: beta strand; pink: turn; blue: helix.

Table 10 – Secondary structure of the GNPTAB protein predicted by UniProt software product.

Feature key	Position(s)	Length	Description	Graphical view
Beta strand ⁱ	138 – 141	4	 Combined sources ▾	
Turn ⁱ	151 – 153	3	 Combined sources ▾	
Turn ⁱ	155 – 157	3	 Combined sources ▾	
Helix ⁱ	159 – 161	3	 Combined sources ▾	
Beta strand ⁱ	165 – 172	8	 Combined sources ▾	
Beta strand ⁱ	175 – 184	10	 Combined sources ▾	
Helix ⁱ	188 – 196	9	 Combined sources ▾	
Beta strand ⁱ	207 – 210	4	 Combined sources ▾	
Beta strand ⁱ	225 – 231	7	 Combined sources ▾	
Helix ⁱ	240 – 244	5	 Combined sources ▾	
Helix ⁱ	249 – 251	3	 Combined sources ▾	
Beta strand ⁱ	252 – 256	5	 Combined sources ▾	
Beta strand ⁱ	262 – 265	4	 Combined sources ▾	
Beta strand ⁱ	267 – 272	6	 Combined sources ▾	
Helix ⁱ	274 – 280	7	 Combined sources ▾	

Overall, none of these three bioinformatic tools evidenced the presence of other important domains or relevant secondary structure features within the amino acid sequence encoded by exon 19, besides the CR4 stealth domain, which should be carefully studied given its role in the catalytic activity of the GlcNAc-1-phosphotransferase enzyme.

4.1.3. Prediction of glycosylation sites within the *GNPTAB* protein product

Lastly, the presence of glycosylation sites within the amino acid sequence encoded by exon 19 of the *GNPTAB* gene was also predicted to ensure that this modification would not be affected by the removal of exon 19 from the mature transcript. Glycosylation is a co-translational and post-translational modification in which a glycan is attached to a specific amino acid residue within a protein, playing an important role for its structure and function (Moremen et al., 2012). Therefore, if glycosylation sites were present within the amino acid sequence encoded by exon 19, the exclusion of those sites could severely affect the tridimensional conformation and activity of the protein produced after skipping of exon 19. In that sense, three different types of glycosylation were predicted for that amino acid sequence: *N*-glycosylation, where the glycan is linked to a nitrogen atom of asparagine or arginine residues; *O*-glycosylation, where the glycan is linked to the hydroxyl oxygen of

serine, threonine, tyrosine, hydroxylysine or hydroxyproline residues; and C-glycosylation, where a mannose sugar is added to a carbon of a tryptophan residue (Spiro, 2002).

Using the software NetNGlyc 1.0 Server, *N*-glycosylation sites were predicted for the entire amino acid sequence of the *GNPTAB* protein product using a threshold of 0.5, revealing that multiple locations showcased a strong potential for this modification, with results reaching the 0.7 range (Table 11 and Figure 17). Nonetheless, the amino acid sequence encoded by exon 19 (positions 1146-1201) did not present any predicted *N*-glycosylation sites, suggesting that the removal of this protein region would not influence the state of glycosylation of the final protein product.

Table 11 – *N*-glycosylation sites predicted by NetNGlyc 1.0 Server for the *GNPTAB* protein product.

(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
Sequence	83 NGTD	0.7798	(9/9)	+++
Sequence	114 NTTE	0.5938	(9/9)	++
Sequence	148 NITL	0.7364	(9/9)	++
Sequence	175 NPST	0.5060	(4/9)	+
Sequence	179 NVSV	0.7481	(9/9)	++
Sequence	250 NLSS	0.6923	(9/9)	++
Sequence	287 NMTI	0.4626	(6/9)	-
Sequence	376 NLSH	0.4324	(6/9)	-
Sequence	457 NNSA	0.5427	(7/9)	+
Sequence	551 NQTH	0.6196	(6/9)	+
Sequence	607 NATT	0.6048	(7/9)	+
Sequence	614 NLTF	0.6661	(9/9)	++
Sequence	642 NSTA	0.5606	(7/9)	+
Sequence	684 NSTR	0.5932	(7/9)	+
Sequence	699 NISL	0.5602	(7/9)	+
Sequence	729 NLSK	0.7184	(9/9)	++
Sequence	759 NDSL	0.5054	(4/9)	+
Sequence	829 NVTK	0.7694	(9/9)	+++
Sequence	1009 NISQ	0.6173	(9/9)	++
Sequence	1056 NCSK	0.5410	(6/9)	+
Sequence	1129 NVSH	0.6610	(9/9)	++

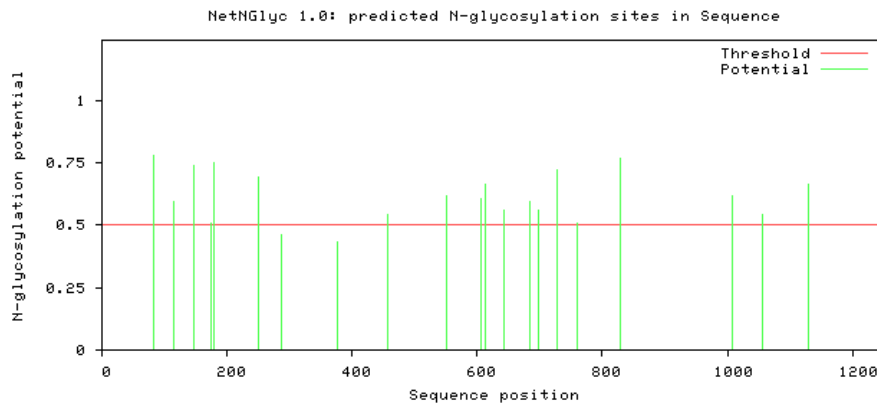


Figure 17 – *N*-glycosylation sites predicted by NetNGlyc 1.0 Server for the *GNPTAB* protein product.

A second software used for glycosylation predictions, GlycoEP, besides the presence of the more common *N*-glycosylation sites, also predicts the sites for *O*-glycosylation and *C*-glycosylation, providing a broader appraisal of possible post-translational modifications within the GNPTAB protein. Regarding *N*-glycosylation sites, the predictions were similar to those provided by NetNGlyc 1.0, with some positions showcasing potential for glycosylation but none within the amino acid sequence encoded by exon 19 (positions 1146-1201), indicating the very low probability of glycosylation in that portion of the protein (Table 12). In the case of *O*-glycosylation sites, predictions once again showed negative scores for positions within the amino acid sequence encoded by exon 19, further supporting the absence of glycosylation in this location, while other positions outside of exon 19 presented potential for glycosylation (Table 13). Finally, predictions of *C*-glycosylation sites showed no positive scores in the entire GNPTAB protein sequence, indicating that this protein does not undergo this specific modification (results not shown).

Table 12 – *N*-glycosylation sites predicted by GlycoEP for the GNPTAB protein product. Only positions 922-1252 are shown.

922	NTG	-1.0051085	Non-glycosylated
939	NKI	-1.0537177	Non-glycosylated
943	NSK	-1.0506794	Non-glycosylated
1009	NIS	0.43475891	Potential Glycosylated
1056	NCS	0.96727476	Potential Glycosylated
1069	NHI	-1.046069	Non-glycosylated
1070	NIP	-1.3998707	Non-glycosylated
1082	NLP	-1.3911925	Non-glycosylated
1093	NCK	-0.98069839	Non-glycosylated
1109	NKY	-1.0607582	Non-glycosylated
1129	NVS	0.81223484	Potential Glycosylated
1143	NPR	-1.6211505	Non-glycosylated
1151	NDH	-0.9463451	Non-glycosylated
1153	NDI	-1.4877977	Non-glycosylated
1157	NHK	-1.0297803	Non-glycosylated
1190	NRF	-0.91318196	Non-glycosylated
1252	NRI	-1.3453223	Non-glycosylated

Table 13 – *O*-glycosylation sites predicted by GlycoEP for the GNPTAB protein product. Only positions 1066-1250 are shown.

1066	I	-0.59911711	Non-glycosylated
1074	I	-0.18961227	Non-glycosylated
1077	S	0.21172554	Potential Glycosylated
1087	I	-0.34229457	Non-glycosylated
1089	S	-1.1425374	Non-glycosylated
1092	I	-0.6472745	Non-glycosylated
1098	I	-1.2835187	Non-glycosylated
1128	I	-0.38036664	Non-glycosylated
1131	S	-0.75539179	Non-glycosylated
1163	I	-0.41423619	Non-glycosylated
1174	S	-0.38561511	Non-glycosylated
1180	S	-0.40767196	Non-glycosylated
1212	I	-0.38188939	Non-glycosylated
1218	I	-0.75375684	Non-glycosylated
1223	I	-0.23741771	Non-glycosylated
1226	S	0.24954717	Potential Glycosylated
1250	S	-0.47598332	Non-glycosylated

As a whole, these bioinformatical predictions indicate that the amino acid sequence encoded by exon 19 of the GNPTAB gene does not contain any positions with strong potential for any kind of glycosylation. Therefore, its absence likely would not disrupt the normal pattern of this post-translational modification in the GNPTAB protein. Additionally, this amino acid sequence does not include other important protein domains or secondary structures, besides the CR4 stealth domain, which raises concerns in respect to the

therapeutic strategy in mind. In fact, given that the stealth domains together form the catalytic activity core of the GlcNAc-1-phosphotransferase enzyme (Liu et al., 2017), the removal of exon 19 from the *GNPTAB* transcript and the consequent absence of the CR4 amino acid sequence from the final protein product might have a negative impact on the activity of the enzyme, which should be carefully investigated. Nonetheless, the remaining results are not unfavorable to the implementation of this exon-skipping therapeutic approach to ameliorate the effects of the c.3503_3504delTC mutation in the exon 19 of the *GNPTAB* gene, ultimately expecting to produce a truncated but functional enzyme with the majority of its essential domains and structures.

4.1.4. Scoring of 5' and 3' splice sites of *GNPTAB* gene exon 19

Initially, 5' and 3' splice sites of exon 19 were scored according to their strength to check whether they could be useful targets for oligonucleotides with the intent of promoting the skipping of the exon. For this purpose, two distinct softwares were used, MaxEntScan and NNSplice, producing concordant results despite utilizing different algorithms for their predictions. In the case of MaxEntScan, whose ideal scores are 13.59 for the 3'ss and 11.81 for the 5'ss, the 3'ss received a final score of 9.32, while the 5'ss was scored at just 6.41, suggesting a lower level of strength of the donor splice site, but a strong acceptor site. Similarly, NNSplice scored 3'ss at 0.94 (on a scale of 0 to 1), while 5'ss was given a 0.81. Taken together, these results indicate a very strong acceptor site on the boundary between intron 18 and exon 19 of the *GNPTAB* gene, which is expected to be robustly recognized by the splicing machinery, ensuring the proper splicing of this exon. Acknowledging the importance of this splice site for the recognition of the exon, this location emerges as a promising target for AOs in exon skipping experiments, since the blocking of this 3'ss alone could possibly alter the recognition of the exon by the splicing machinery and induce its exclusion from the final transcript.

4.1.5. Prediction of enhancers within *GNPTAB* gene exon 19

Aside from splice sites, enhancers are also an important component for the correct functioning of the splicing reaction, promoting the inclusion of the exon under their regulation. Therefore, targeting these elements with AOs can also be useful for the purpose of exon skipping and so the exon 19 sequence of *GNPTAB* gene was analyzed by two bioinformatic tools: ESEfinder 3.0 and RESCUE-ESE. Using the software ESEfinder 3.0, enhancers were

predicted on the mutant variant of exon 19 of the *GNPTAB* gene, as well as on 50bp of each flanking intron, through the search for binding motifs for five human SR proteins, revealing the presence of multiple enhancers on various sites across the exon and the flanking intronic regions. Results are shown on Table 14 and Figure 18 and multiple locations have been identified as putative enhancers by each SR protein matrix, with SRSF2 and SRSF6 containing the largest number of predicted enhancer binding motifs, ten and twelve, respectively. Among these predictions, five locations, three targeted by SRSF2 and two targeted by SRSF5, present a score above four, showcasing strong potential for enhancer activity in these sites and therefore presenting themselves as clear candidates for targeting by AOs for the purpose of promoting the skipping of exon 19. Furthermore, analyzing the graphic in Figure 18, it is possible to observe a great number of putative enhancers in the downstream end of exon 19 and the beginning of intron 19, suggesting this could also be a good location for AO targeting since multiple enhancers could be disrupted with a single AO, increasing the probability of excluding exon 19 from the final transcript. Finally, it is easily noticeable in the graphic that the highest scoring location overlaps with the 3'ss, for which a high score was also predicted in the MaxEntScan and NNSplice analysis concerning the two splice sites flanking exon 19. Taken together with previous experiments regarding exon skipping for the treatment of DMD, which revealed acceptor sites and the first half of the exon as the most efficient annealing sites for AOs (Fletcher et al., 2012), these results suggest that this region could be a very strong target for AOs intended to promote the skipping of exon 19 by simultaneously disrupting the acceptor splice site and a powerful splicing enhancer.

Table 14 – Putative enhancers predicted by ESEfinder 3.0 for exon 19 of the *GNPTAB* gene containing the c.3503_3504delTC mutation and 50bp of the flanking introns. Positions start at the beginning of the input sequence and not the beginning of exon 19. 1-50: intron 18; 51-216: exon 19; 217-266: intron 19.

SRSF1 threshold: 1.956	SRSF1 (lgM-BRCA1) threshold: 1.867	SRSF2 threshold: 2.383	SRSF5 threshold: 2.67	SRSF6 threshold: 2.676
Position*/Site/Score	Position*/Site/Score	Position*/Site/Score	Position*/Site/Score	Position*/Site/Score
1 (-266) gtcagga3.28514	100 (-167) CAGACAG2.80462	26 (-241) taaccctg3.03048	45 (-222) ttccagG4.96507	11 (-256) tccata2.85801
100 (-167) CAGACAG3.03834	164 (-103) CTGCCAA2.40293	34 (-233) gctctctg3.08741	69 (-198) TGACAAC4.09239	15 (-252) tatgta3.61028
171 (-96) GAGAGTA2.04141	207 (-60) TGCAGGA2.38615	79 (-188) GACCCACA4.14538	81 (-186) CCACAAT3.01814	130 (-137) TGAATC3.38755
207 (-60) TGCAGGA2.58726	224 (-43) ctccatgt2.93017	122 (-145) GACTTCTA4.65083	206 (-61) CTGCAAG3.04686	192 (-75) TTCATA2.85801
224 (-43) ctccatgt2.58960		132 (-135) AATCCATG2.49301	223 (-44) tctcatg3.25458	198 (-69) TGCATG3.76391
		139 (-128) GTTCCCA4.14071	260 (-7) atacatg2.94018	207 (-60) TGCAGG2.95003
		186 (-81) GTTCCTT2.74404		233 (-34) tatata3.00162
		204 (-63) AGCTGCAG2.54795		239 (-28) tgcagtg3.76391
		212 (-55) GAATGgtg2.58640		243 (-24) tgtgtg2.82087
		229 (-38) gttttata2.51827		245 (-22) tgtgtg2.82087
				257 (-10) tgtata3.61028
				261 (-6) tccatg3.15525

* both positions from 5'end (through 1) and 3'end (through -1) are given

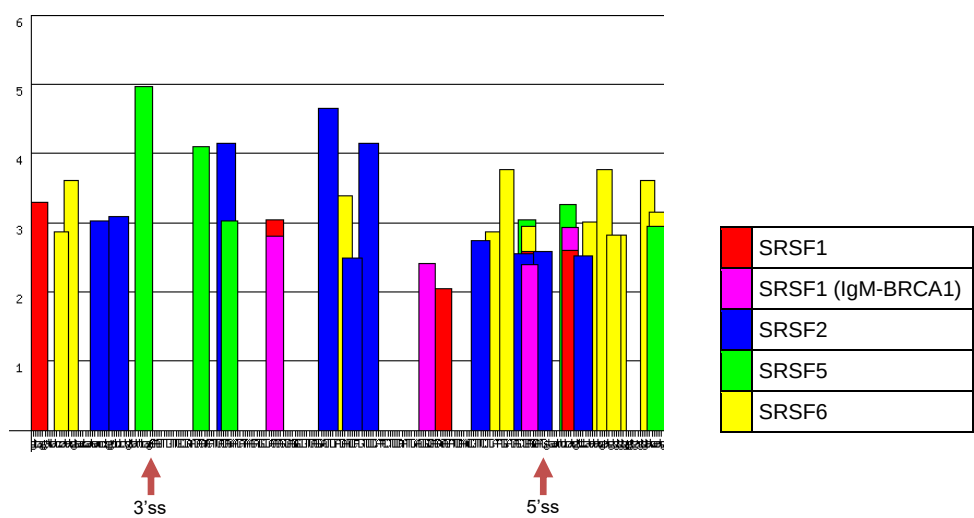


Figure 18 – Graphic representation of the putative enhancers predicted by ESEfinder 3.0 for exon 19 *GNPTAB* gene containing the c.3503_3504delTC mutation and 50bp of the flanking introns. Arrows indicate the location of the donor and acceptor splice sites.

With RESCUE-ESE, putative enhancers are predicted according to their frequency on exons versus introns, as well as their frequency on exons with weak splice sites versus exons with strong splice sites where enhancers are expected to play a more subtle role in the splicing reaction. Once again, analysis with RESCUE-ESE included the mutant nucleotide sequence of exon 19 of the *GNPTAB* gene, as well as 50bp of introns 18 and 19, to predict enhancers within the exon but also in the flanking intronic regions (Figure 19).

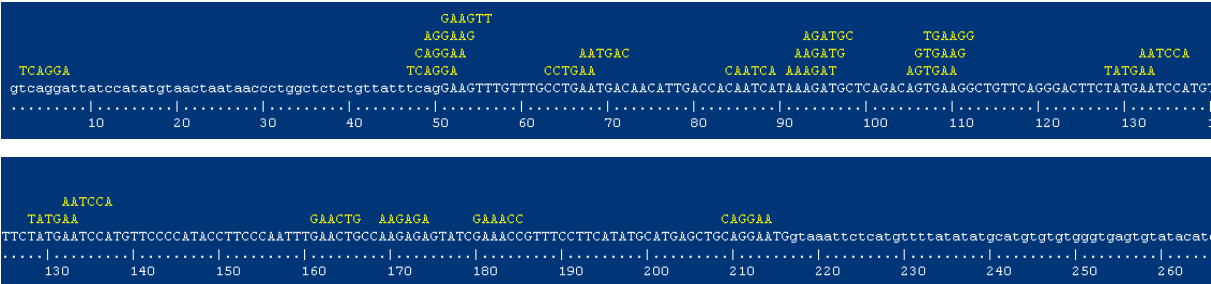


Figure 19 – Putative enhancers predicted by RESCUE-ESE on exon 19 of the *GNPTAB* gene containing the c.3503_3504delTC mutation and 50bp of the flanking introns. Positions start at the beginning of the input sequence and not the beginning of exon 19. 1-50: intron 18; 51-216: exon 19; 217-266: intron 19.

Similar to ESEfinder 3.0, RESCUE-ESE also identified multiple putative enhancers within the input sequence, as well as some overlap between certain predicted elements. Moreover, the region around the acceptor site (position 50) emerges once again as the location with the most predicted enhancers, with the first half of the exon also showing multiple predictions. However, the cluster of putative enhancers located at the downstream

end of the exon in the ESEfinder analysis is not replicated by this software and the sequence and position of many predicted elements are also very distinct from those observed with ESEfinder. This divergence of results highlights the variability and diversity in the prediction of enhancer sequences when performed by different bioinformatic tools, which can only be validated as splicing regulatory elements after functional splicing analysis.

Overall, the 3'ss and the downstream end of exon 19 appear to be the most promising targets for AOs with the purpose of inducing the skipping of this exon, given the high scoring of this splice site when analyzed by the bioinformatic tools selected, as well as the presence of numerous putative enhancers in these regions that can be targeted to alter the splicing pattern and promote the exclusion of exon 19 from the final mRNA transcript.

4.2. Confirmation of the presence of the *GNPTAB* gene exon 19 c.3503_3504delTC mutation in patient fibroblast cells

After the establishment of stable fibroblast cell lines from two distinct ML II patients, it was necessary to confirm the presence of the c.3503_3504delTC mutation in the cells before advancing to any further experiments. For that purpose, genomic DNA was extracted from both patient fibroblast cells, as well as from healthy control fibroblasts and the target region including exon 19 of the *GNPTAB* gene and its flanking intronic regions was amplified through PCR and subsequently sequenced to search for the ML II causing mutation.

A single pair of primers was selected, showing specific annealing to the last portion of intron 18 (forward) and the beginning of intron 19 (reverse), allowing the amplification of a final product of 436bp that included the entire sequence of exon 19 (Figure 20).



Figure 20 – Schematic figure of primer annealing for amplification of exon 19 of the *GNPTAB* gene and flanking intronic regions.

Following amplification, products were separated through agarose gel electrophoresis with ethidium bromide staining and visualized with ultraviolet light. As seen in Figure 21, all samples were successfully amplified and produced a single amplicon of around 450bp, as expected given the primer annealing locations. Furthermore, no unspecific products or

contamination were detected in any of the samples, which allowed their use for subsequent purification and sequencing.

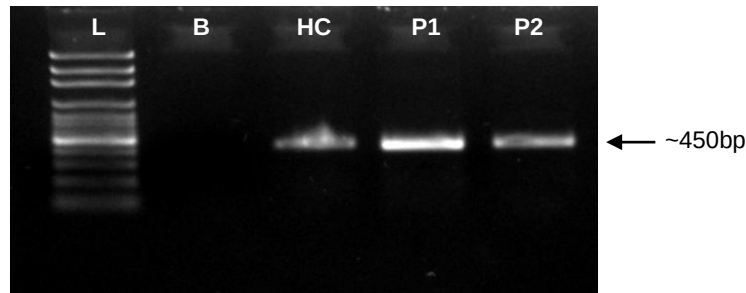


Figure 21 – Agarose gel electrophoresis of PCR amplification of *GNPTAB* gene exon 19 and flanking intronic regions in control and patient fibroblasts. L: ladder; B: negative control (blank); HC: healthy control; P1: ML II Patient 1; P2: ML II Patient 2.

Then, the amplification products of patient fibroblasts were sequenced and results were compared with the *GNPTAB* reference sequence ENSG00000111670 obtained from the Ensembl database using the Clustal Omega web tool. In both cases, sequence analysis revealed the deletion of a TC dinucleotide, in homozygosity, at positions c.3503 and c.3504. As such, it was possible to confirm that this deletion corresponds to the expected c.3503_3504delTC mutation responsible for the ML II phenotype presented by these patients and the most frequent pathogenic mutation associated with this disorder (Coutinho et al., 2011).

4.3. Analysis of *GNPTAB* gene exon 19 splicing pattern in fibroblast cells

After confirming the presence of the c.3503_3504delTC mutation in exon 19 of the *GNPTAB* gene in the cultured patient fibroblast cells, it was then needed to study the endogenous splicing pattern of these fibroblasts before starting the splicing modulation experiments in order to assure that no previous disruption of the splicing process occurred in the cells prior to transfection with AOs and that the *GNPTAB* transcript was produced as expected.

Viewing that, total RNA was extracted from the two patient cell lines, as well as from healthy control fibroblasts, and cDNA was synthesized using an oligo(dT)₁₈ primer, therefore covering all the transcripts present in the samples. For the PCR reaction, six different sets of primers were employed, each of them directed at a specific location of the *GNPTAB* transcript and ensuring that the total coding sequence of the mRNA molecule was amplified, as shown schematically on Figure 22. Exons covered by each of the primer sets and corresponding amplicon sizes are detailed in Table 15.

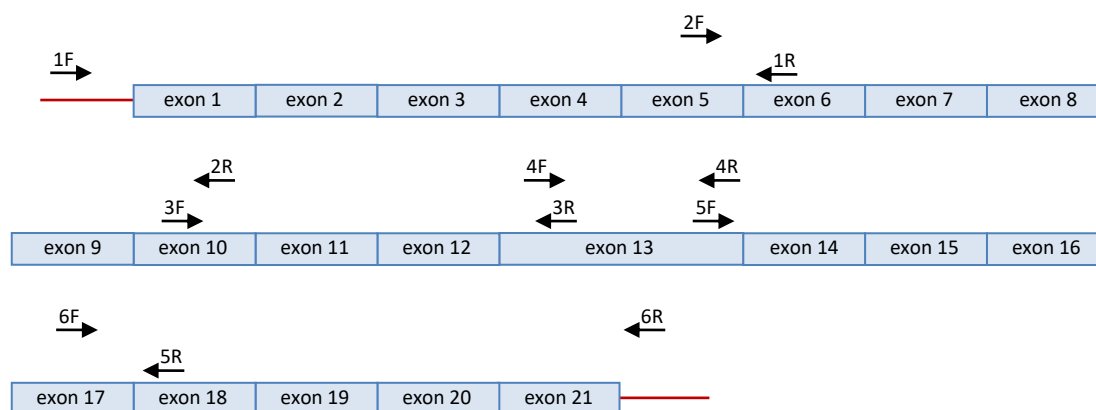


Figure 22 – Schematic representation of annealing locations of all six primer sets for amplification of different fragments of the *GNPTAB* transcript.

Table 15 – Exons of the *GNPTAB* transcript covered by each primer set and corresponding amplicon size.

Primer set	Exons	Amplicon size (bp)
cDNA 1F / cDNA 1R	1-6	736
cDNA 2F / cDNA 2R	5-10	716
cDNA 3F / cDNA 3R	10-13	732
cDNA 4F / cDNA 4R	13	774
cDNA 5F / cDNA 5R	13-18	763
cDNA 6F / cDNA 6R	17-21	734

Following amplification, the products were separated through agarose gel electrophoresis with ethidium bromide staining and visualized under ultraviolet light. As expected, all samples revealed a band between 700bp and 750bp, confirming the success and specificity of the PCR reaction (Figure 23).

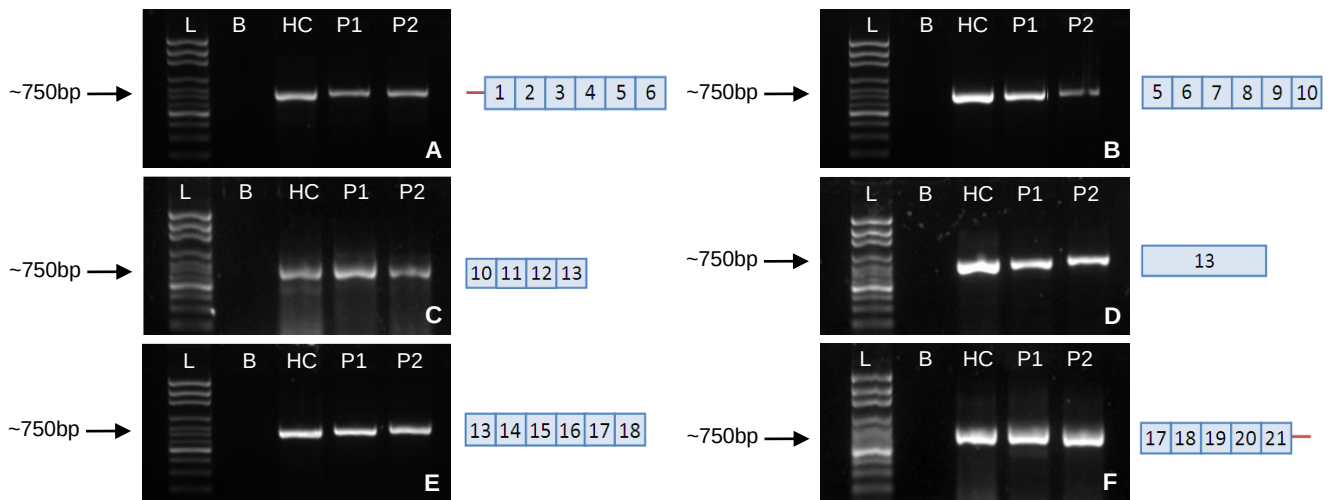


Figure 23 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript in patient and control fibroblasts. A: primer set 1; B: primer set 2; C: primer set 3; D: primer set 4; E: primer set 5; F: primer set 6. L: ladder; B: negative control (blank); HC: healthy control; P1: ML II Patient 1; P2: ML II Patient 2.

As so, the amplification products with primer set 6 (Figure 23F), spanning exons 17 to 21, were subsequently purified and sequenced. The obtained sequences were next compared with the *GNPTAB* transcript reference sequence ENST00000299314.11 retrieved from the Ensembl database to confirm the absence of any aberrant splicing surrounding the c.3503_3504delTC mutation and attest their availability for further experiments. Using the online software Clustal Omega, an almost complete match was observed between the two patients' amplified sequences and the reference sequence, as well the presence of all expected exons without any alteration in their order or length. Aside from the expected c.3503_3504delTC mutation in exon 19 responsible for the ML II phenotype and a small number of single nucleotide polymorphisms (SNPs), the cDNA sequence of these two patients proved to be identical to the reference sequence and showed no evidence of any kind of aberrant splicing in this region of the gene.

Overall, this analysis revealed a regular endogenous splicing pattern in the *GNPTAB* gene of the ML II patients, favoring the potential of these cell lines for use in the exon-skipping experiments viewing the development of a novel therapeutic approach.

4.4. Exon-skipping antisense oligonucleotide *in vitro* therapeutic approach for patients affected by ML II

Having confirmed the presence of the c.3503_3504delTC mutation in the two patient cells and the absence of any splicing alterations in the *GNPTAB* transcript, the fibroblasts were plated in 6-well plates and transfected with the four selected AOs at different concentrations in an attempt to induce the skipping of exon 19 from the final mRNA molecule. Following transfection, the cells were incubated for 24 hours after which total RNA was extracted for molecular analysis. The transcript region surrounding the target exon was then amplified through RT-PCR and the resultant products were separated through agarose gel electrophoresis, where full-length amplicons were expected to stretch for 734bp, while the shorter transcripts without exon 19 should only reach 566bp, allowing for a clear distinction between the two transcripts. Next, after this initial experiment with all four AOs at concentrations between 10nM and 100nM, the two most effective molecules were selected for additional testing and transfected into cells at higher doses (between 100nM and 500nM) to further study their effects on the splicing pattern of the *GNPTAB* gene. Lastly, when a possible pattern of exon-skipping was detected, the corresponding product was cut out from the gel and sequenced to confirm the exclusion of this exon from the final *GNPTAB* transcript. Below these experiments and their results are described in more detail.

4.4.1. Evaluation of transfection efficiency with *eGFP* plasmid in control fibroblasts

Before initiating the experiments of transfection with AOs, a plasmid containing the reporter gene *eGFP* was transfected into control fibroblasts in order to assess the efficiency of the Lipofectamine® 2000 transfection reagent. Following transfection and a 24 hours incubation period, cells were observed under a fluorescence microscope and green fluorescence was clearly visible in multiple cells (Figure 24), confirming the effectiveness of this reagent in transfecting these cultured fibroblasts and supporting its use for transfection of this kind of cells with the selected AOs.

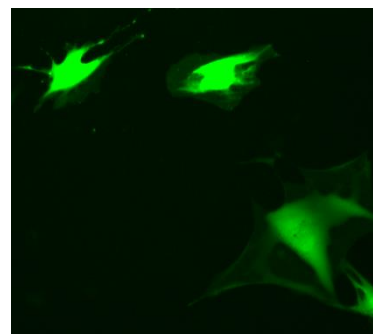


Figure 24 – Fluorescence microscopy images of healthy control fibroblasts transfected with *eGFP*. The strong green fluorescence visible in the cells confirms the efficiency of Lipofectamine® 2000 in transfecting cultured fibroblasts.

4.4.2. Splicing modulation of exon 19 with antisense oligonucleotides in healthy control and ML II patient fibroblasts

4.4.2.1. Experiments in healthy control fibroblasts

After confirming the effectiveness of the Lipofectamine® 2000 reagent in transfecting cultured fibroblasts and before testing the selected AOs in ML II fibroblasts cells, these molecules were transfected into fibroblasts of a healthy control. Here, the skipping of exon 19 was also to be expected, since the target sequences for these AOs do not encompass the ML II causing mutation, which means they should have the same effect in both patient and control fibroblasts. Accordingly, all four oligonucleotides have induced some level of exon-skipping in the control fibroblasts cells at the initially tested concentrations. As shown in Figure 25, an amplification product of around 550bp is present in most of the samples, matching the length of the *GNPTAB* transcript without exon 19 and indicating this exon has been successfully removed from the final mRNA molecule.

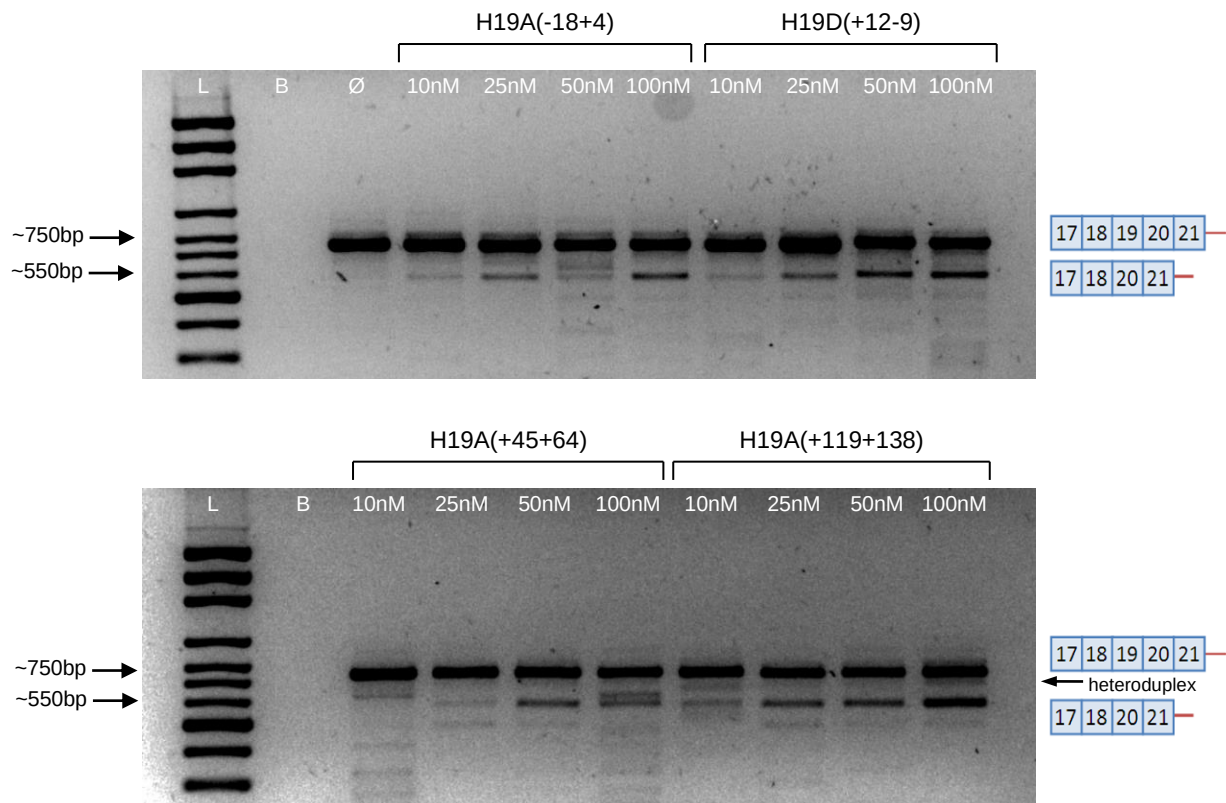


Figure 25 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in control fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19A(-18+4): directed to the 3'ss; H19D(+12-9): directed to the 5'ss; H19A(+45+64): directed to ESEs in the first half of exon 19; H19A(+119+138): directed to ESEs in the second half of exon 19.

Out of the four AO tested, H19A(-18+4), which is directed to the 3'ss of exon 19, induced the skipping of this exon regardless of the concentration used, but showing higher efficiency at 25nM and 100nM. At 50nM, the shorter transcript appears with lower intensity than at 25nM but a third fragment is visible between the full-length mRNA and the shorter transcript without exon 19, which could be caused by partial skipping of the exon, an event that would also explain the lower amount of the transcript missing exon 19. Nonetheless, the amount of full-length transcript at the two highest concentrations, 50nM and 100nM, seems to be lower than at the two lowest doses, suggesting an increase in the ratio of transcript missing exon 19 versus full-length mRNA, which is one of the desired conditions to make this therapeutic strategy possible. Using the H19D(+12-9) AO, targeting the 5'ss, it is possible to observe a clear dose-dependent pattern concerning the amount of the shorter transcript, with a gradual increase in its levels from the lowest concentration of 10nM to 100nM.

Among the two internal AOs, targeting ESEs within exon 19, the first molecule, H19A(+45+64), directed at the first half of the exon, shows some level of exon-skipping in all of the concentrations. However, at the lowest dose, 10nM, the visible product is longer than the expected transcript without exon 19, spanning over 600bp, which could suggest another case of partial skipping of the exon sequence. For the three other concentrations, the expected product is present, exhibiting an increase between 25nM and 50nM but decreasing once more at 100nM. Similar to H19A(-18+4), this decrease could be explained by the presence of a third fragment, with approximately the same length as the one present at 10nM, possibly produced by the skipping of only part of exon 19. Furthermore, at 25nM and 50nM, an amplification product around 500bp is tenuously visible, suggesting the possibility of the opposite situation, that is, the skipping of more than just exon 19, leading to an even shorter final transcript. As will be described in the next section, a similar product was observed in the experiments with the patient fibroblasts in which subsequent sequencing revealed a situation of skipping of both exon 19 and 20, which most likely also occurred in the control cells. At last, the H19A(+119+138) AO, directed at ESEs present in the second half of the exon, seems to achieve the skipping of exon 19 at all concentrations tested, displaying a clear dose-dependent response with a gradual increase in the amount of the shorter transcript across all four doses. Furthermore, at 10nM and 100nM, a third fragment is visible between the full-length mRNA and the shorter transcript missing exon 19, this time with around 700bp, and its sequencing revealed the existence of a heteroduplex between the full-length strand and the shorter strand without exon 19.

As the best results were obtained for the H19D(+12-9) and H19A(+119+138) AOs, higher concentrations (between 100nM and 500nM) of these molecules were further tested and both AOs were able to successfully induce the skipping of exon 19 at all of the tested conditions (Figure 26).

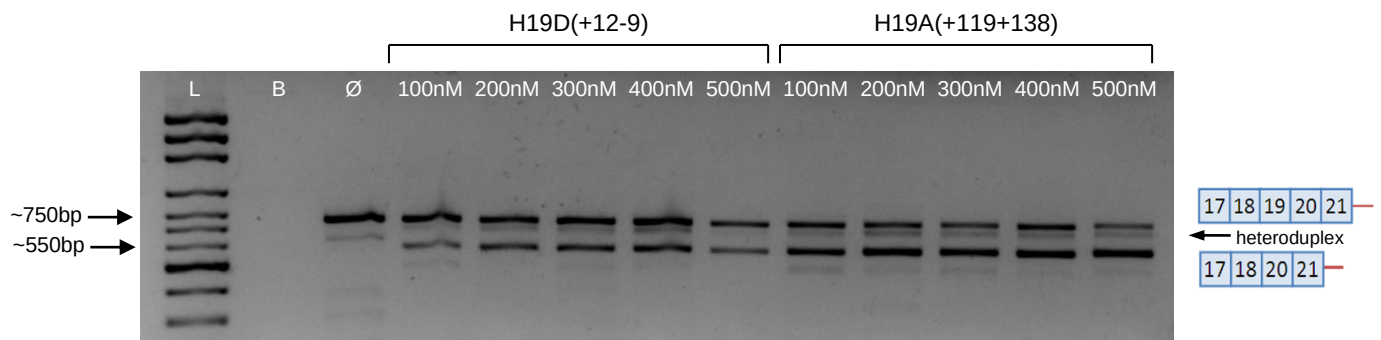


Figure 26 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in control fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19D(+12-9): directed to the 5'ss; H19A(+119+138): directed to ESEs in the second half of exon 19.

At this range of AO concentration, H19D(+12-9) AO still appears to originate a dose-dependent response, with gradual increases in the amount of the shorter transcript missing exon 19 as the concentrations tested also increase. The highest dose of 500nM was the exception here, with comparatively low levels of both transcripts when balanced against the other samples, possibly due to a very low amount of initial RNA for this amplification. Nonetheless, the amount of both transcripts obtained with this dose is nearly equivalent, reinforcing the positive results observed for the other four concentrations regarding the increase in the amount of shorter transcript to levels close to the full-length mRNA molecule.

On the other hand, the amount of shorter transcript produced by action of the H19A(+119+138) AO only seemed to suffer a very small increase from 100nM to 200nM, with all the other concentrations, between 300nM and 500nM, exhibiting similar increasing levels of this transcript. Furthermore, the levels of full-length transcript in all of the tested concentrations appeared to be considerably lower than those of its shorter variant, which is most notorious at 500nM where the full-length mRNA is clearly reduced compared to the other conditions, leading to an increase in the ratio between the shorter transcript and the full-length variant and supporting the potential of this AO as an exon-skipping therapeutic agent for the treatment of ML II patients with mutations on exon 19.

4.4.2.2. Experiments in ML II Patient 1 fibroblasts

For the first ML II patient, the electrophoresis gel seen in Figure 27 also revealed the existence of a shorter amplification product at about 550bp that most likely corresponds to the transcript missing exon 19 and appears to be present in the tests with all four AOs under analysis.

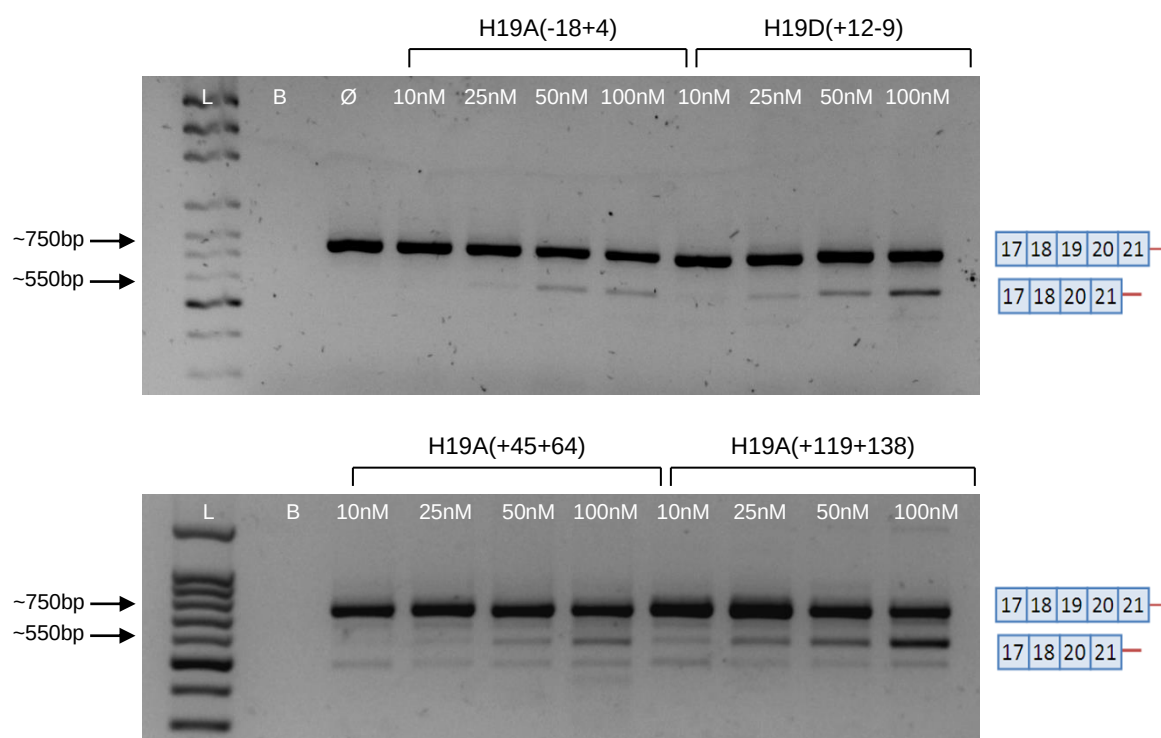


Figure 27 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 1 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19A(-18+4): directed to the 3'ss; H19D(+12-9): directed to the 5'ss; H19A(+45+64): directed to ESEs in the first half of exon 19; H19A(+119+138): directed to ESEs in the second half of exon 19.

The first AO, H19A(-18+4) binding the 3'ss region, seems to produce successful exon-skipping at the two highest concentrations, 50nM and 100nM, with a very faint band also visible at 25nM, which could suggest the presence of the shorter transcript at this dose, as well. Nonetheless, there is no clear increase in the amount of shorter transcript from 50nM to 100nM, preventing us from establishing a direct relation between the dosage used and the resultant level of exon-skipping. In the case of the second AO, H19D(+12-9) for the 5'ss region, the exon-skipping is visible at the three highest concentrations in a clear dose-dependent pattern, with just 10nM, the lowest dose tested here, not revealing the presence of the shorter transcript. Furthermore, a fragment below 500bp also seems to be visible at 25nM and 100nM, which is probably the result of the exclusion of exon 20 together with the target exon from the final transcript, as was already observed in the control fibroblasts.

Concerning the two exonic internal AOs, H19A(+45+64) for the first half of exon 19, induced the production of a shorter transcript at the concentrations of 25nM, 50nM and 100nM, albeit in very low levels compared to the full-length transcript. Still, a distinct increase in the amount of the shorter mRNA is visible across the three doses, suggesting a dose-dependent response to this treatment and that higher concentrations could offer even better exon-skipping levels. The AO for the second half of exon 19, H19A(+119+138), was considerably more successful than H19A(+45+64), with exon-skipping products in all concentrations tested and exhibiting a clear dose-dependent response. In fact, the highest tested dose, 100nM, revealed a considerable large amount of the shorter transcript, above what was observed for any of the other AOs, at any of the tested concentrations, nearly reaching the level of the full-length transcript and further supporting the efficiency of this AO in inducing the skipping of exon 19 from the *GNPTAB* transcript.

In the same way, for the experiments with higher doses (between 100nM and 500nM) of H19D(+12-9) and H19A(+119+138), successful exon-skipping was achieved in all of the tested concentrations. A fragment below 500bp was also present in every sample, including the negative control without AO, which could result from the exclusion of the adjacent exon 20 together with the target exon 19 from the *GNPTAB* transcript, as observed for the control fibroblasts and Patient 2, but could also be a naturally occurring alternative transcript only visible under certain conditions (Figure 28).

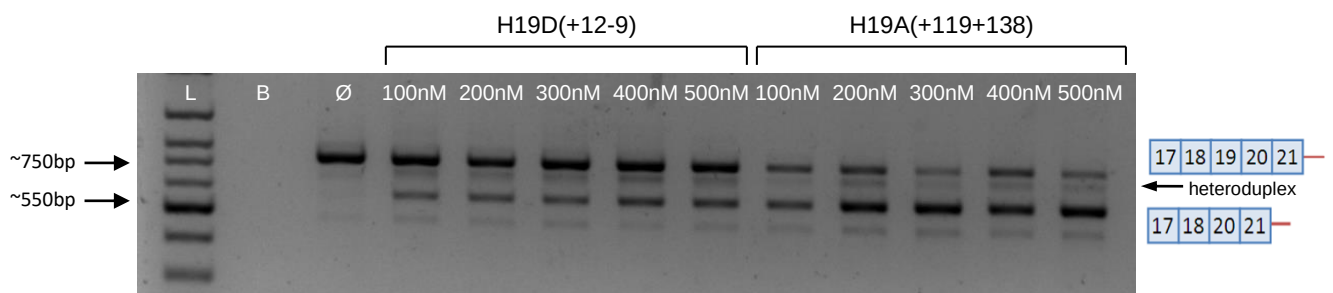


Figure 28 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 1 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19D(+12-9): directed to the 5'ss; H19A(+119+138): directed to ESEs in the second half of exon 19.

In the case of the H19D(+12-9) AO for the 3'ss region, a clear dose-dependent pattern is observed, with a gradual increase in the amount of shorter transcript missing exon 19 across the five doses, while the full-length transcript appears to remain unchanged. The H19A(+119+138) molecule, however, proved to be much more effective in these

experiments, with strong production of the shorter fragment in all tested concentrations except 100nM, where the amount of this transcript is a bit lower. Additionally, the full-length transcript appears to be massively reduced in all samples, showing levels below those observed for the shorter variant without exon 19 in all of the conditions. Just as previously observed for this molecule in lower concentrations and other fibroblast cells, these results strongly support its capability of effectively inducing the skipping of exon 19, producing a shorter *GNPTAB* transcript, while also significantly reducing the amount of full-length transcript, ultimately inverting the ratio between these two molecules.

4.4.2.3. Experiments in ML II Patient 2 fibroblasts

In the case of Patient 2, the electrophoresis gel revealed the presence of some unspecific fragments, but still it was possible to identify the transcripts where the skipping of exon 19 was apparently successful. As seen in Figure 29, all four AOs exhibited positive results in all of the initially tested concentrations (10nM-100nM), with a dose-dependent response visible in most cases. Additionally, as seen in healthy control and Patient 1 fibroblasts, a shorter fragment of about 500bp, resulting from the removal of both exon 19 and 20, is also visible in most situations, albeit with much lower intensity than both the full-length and the transcript without exon 19.

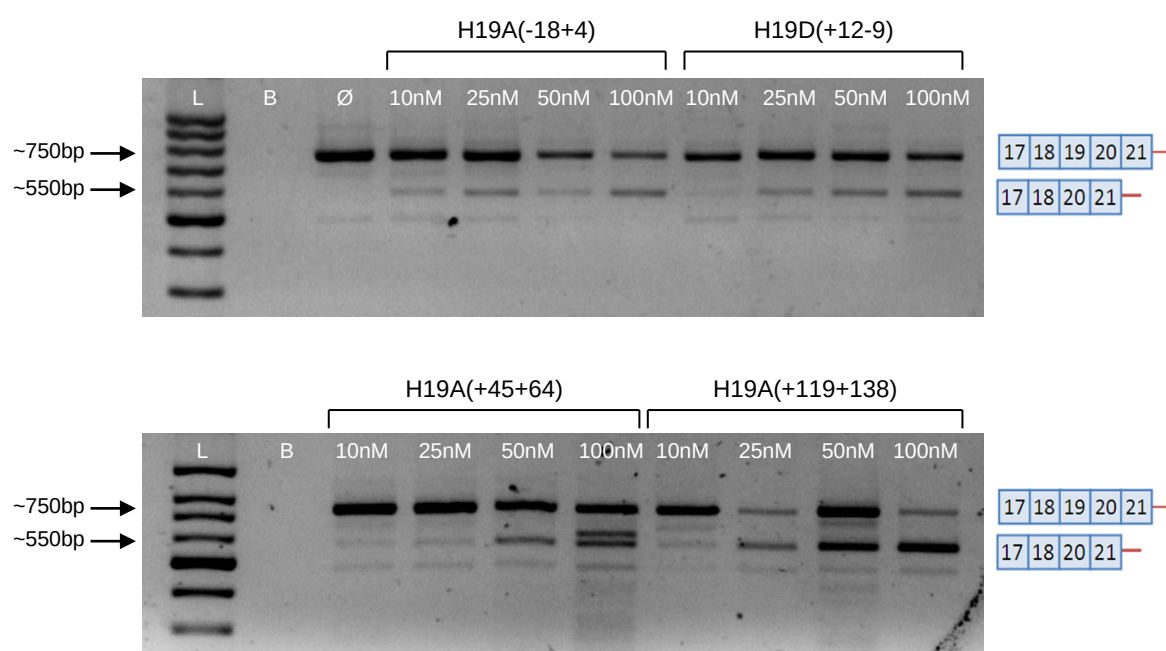


Figure 29 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 2 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19A(-18+4): directed to the 3'ss; H19D(+12-9): directed to the 5'ss; H19A(+45+64): directed to ESEs in the first half of exon 19; H19A(+119+138): directed to ESEs in the second half of exon 19.

In the specific case of the H19A(-18+4) AO binding the 3'ss region, there appears to be a decrease in the amount of the shorter fragment between the concentrations of 25nM and 50nM that can likely be attributed to the use of a smaller amount of initial RNA rather than to a reduced effect of the AO at this dosage. However, at the concentration of 100nM, the full-length transcript and the shorter variant seem to be present in equal amounts, proving the capacity of this molecule to induce exon 19 skipping. On the other hand, the H19D(+12-9) AO for the 5'ss, showcases a clear dose-dependent response, with regular increases in the amount of shorter transcript as the concentration of the therapeutic agent increases. Noticeably, the full-length mRNA is produced at a higher level than the transcript missing exon 19 and does not seem to undergo major changes across the four tested concentrations.

As for the AOs targeting ESEs, the H19A(+45+64) molecule lead to a dose-dependent response to the treatment, with the 10nM and 25nM concentrations generating very low levels of exon-skipping but producing a visible increase of the shorter transcript missing exon 19 at 50nM and 100nM. Moreover, a third fragment of about 600bp is visible at 100nM, suggesting the possibility of a partial exon-skipping rather than the exclusion of the entire exon 19 from the final transcript. Lastly, in the case of the H19A(+119+138) AO, a dose-dependent response is once again visible with increasing amounts of the shorter transcript missing exon 19 accompanying the increasing AO concentrations. Simultaneously, the full-length transcript appears to be reduced at 25nM and 100nM, showing lower intensity than its shorter counterpart, which is particularly noticeable at the highest dose tested. In general, an increase in the ratio of transcript without exon 19 versus the full-length mRNA was observed, which once again supports the feasibility of the planned AO therapeutic approach.

As for the experiments with the higher AO concentrations, both H19D(+12-9) and H19A(+119+138) successfully induced the skipping of exon 19 from the final *GNPTAB* transcript in all five concentrations tested (Figure 30). Once more, fragments of approximately 500bp and 700bp are also visible in the gel aside from the full-length and the truncated transcript, which most likely result from the skipping of exons 19 and 20 in the former and a partial skipping of exon 19 in the latter.

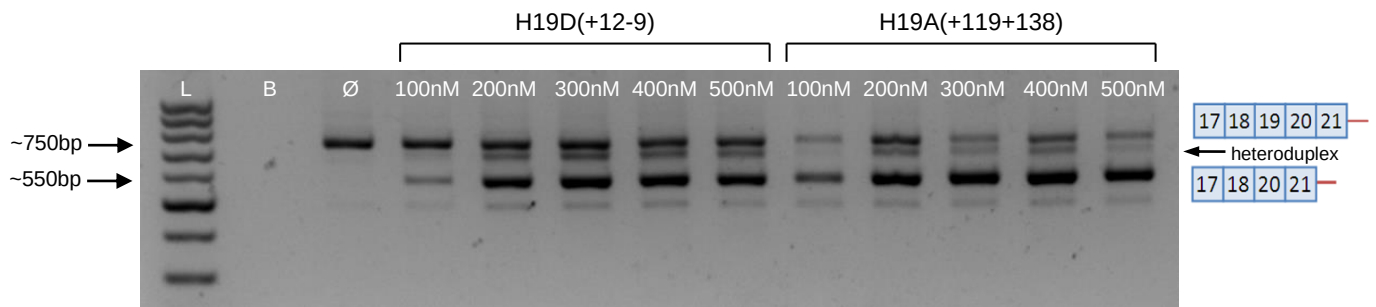


Figure 30 – Agarose gel electrophoresis of amplification products of the *GNPTAB* transcript (exons 17-21) in ML II Patient 2 fibroblasts after transfection with AOs. L: ladder; B: negative control (blank). Ø: control sample without AO. H19D(+12-9): directed to the 5'ss; H19A(+119+138): directed to ESEs in the second half of exon 19.

In the case of H19D(+12-9), there is an increase in the amount of transcript missing exon 19 between the concentrations of 100nM and 200nM, with the remaining doses showing strong and steady levels of this product. Additionally, with the exception of the lowest concentration, the shorter transcript levels are always superior to the full-length transcript, revealing an inversion in the ratio between these two transcript variants. Conversely, an increase in the levels of the transcript missing exon 19 is also visible between 100nM and 200nM of the H19A(+119+138) AO, followed by solid levels of this transcript in all the other doses. Furthermore, as observed for Patient 1, the amount of full-length transcript is greatly reduced with this molecule, especially at the concentrations of 300nM and 500nM, where it is only visible as a faint band. Once more, this shows a massive increase in the ratio of shorter transcript versus full-length transcript, combining the successful skipping of exon 19 from the *GNPTAB* transcript with a considerable reduction of the full-length mRNA.

As a whole, these results confirm the ability of these AOs to induce the skipping of exon 19 from the *GNPTAB* transcript, therefore circumventing the ML II-causing mutation. Among the four molecules tested, H19A(+119+138) revealed to be the most promising, with great levels of transcript missing exon 19 while also reducing massively the amount of full-length transcript, especially at the higher concentrations.

4.5. Quantification of the activity of several lysosomal enzymes in fibroblasts

After obtaining preliminary data on the effectiveness of the therapeutic approach at RNA level, with the successful exclusion of exon 19 from the final *GNPTAB* transcript, quantification of some lysosomal enzymes within the cells was undertaken in order to indirectly evaluate the activity of GlcNAc-1-phosphotransferase following splicing modulation by AOs. To do that, healthy control and Patient 1 fibroblasts were transfected again with H19A(+119+138) – the most promising of the four AOs – at the concentration of 500nM and their pellets were collected for enzymatic assays to measure the amount of different enzymes before and after treatment.

4.5.1. Enzyme levels prior to *GNPTAB* exon 19 splicing modulation with antisense oligonucleotides

Prior to transfection with AOs, the amount of lysosomal hydrolases was measured through enzymatic assays to establish their endogenous levels in the different fibroblast cells. Aside from cells from the two ML II patients, two controls (HC1 and HC2) were also used, as well as cells from a Fabry patient (P3) and a MPS II (Hunter syndrome) patient (P4), in order to discriminate the pattern of ML II regarding the levels of multiple lysosomal enzymes. Results are shown on Table 16.

Table 16 – Enzyme activity of lysosomal enzymes in fibroblasts of controls and LSDs patients. Due to small sample amount of the Fabry cell line, some enzymatic assays could not be performed for these fibroblasts.

Enzymes	Enzyme activity (nmol/h/mg protein)					
	Cell lines					
	HC1	HC2	P1 (ML II)	P2 (ML II)	P3 (Fabry)	P4 (MPS II)
β-Glucosidase T+	80,21	109,47	242,68	178,34	125,02	158,94
β-Glucuronidase	78,42	77,44	13,39	29,64	—	233,98
β-Galactosidase	366,01	450,06	5,84	6,64	519,96	410,69
α-Galactosidase	17,16	14,31	2,19	0,32	17,20	28,62
Arylsulfatase B (ARSB)	57,40	95,90	21,77	25,46	111,75	203,72
Hexosaminidase T	35561,88	20973,83	6026,56	7112,47	—	43130,98
α-Mannosidase	113,07	84,08	6,44	10,04	70,30	68,06

Undoubtedly, both ML II patients present reduced levels of activity in all tested enzymes except β -glucosidase T+, whose activity in patients is actually higher than in the controls or the two other LSD patients. In reality, this is commonly observed in ML II patients, given that this enzyme is not transported through the M6P pathway and therefore is not affected by the absence of active GlcNAc-1-phosphotransferase (Reczek et al., 2007; Liou et al., 2014). Contrarily, the remaining six hydrolases, all rely on the M6P pathway for lysosomal targeting, which explains their reduced amounts in both ML II patients, with α -galactosidase only showing residual activity in Patient 2. This is in line with results previously reported for ML II patients, who present lysosomal enzyme activity levels considerably below the normal range observed in control fibroblasts (Otomo et al., 2011; Alegra et al., 2014).

Overall, the ML II fibroblast cells exhibit a unique pattern of general reduction in the levels of lysosomal enzymes, with the exception of β -glucosidase, which, in opposition, is quantitatively augmented. Thus, a clear distinction between ML II and other LSDs is feasible through the analysis of the activity of certain lysosomal enzymes, which is very important for both the diagnosis and the study of this lysosomal dysfunction.

4.5.2. Enzyme levels following *GNPTAB* exon 19 splicing modulation with antisense oligonucleotides

Following the initial assessment of enzyme activity, Patient 1 and control fibroblasts were transfected with 500nM of H19A(+119+138) AO and their enzyme activity was measured to assess whether the skipping of exon 19 could be associated with an increase in the levels of lysosomal hydrolases transported through the M6P-dependent pathway. Results are shown on Table 17.

Table 17 – Enzyme activity of lysosomal enzymes in fibroblasts of controls and ML II Patient 1 after transfection with the H19A(+119+138) AO. Plus sign (+) indicates the treatment with Lipofectamine® 2000 + AO, while minus sign (-) indicates the treatment with just Lipofectamine® 2000.

Enzymes	Enzyme activity (nmol/h/mg protein)					
	HC1 24 hours		P1 (ML II) 24 hours		P1 (ML II) 48 hours	
	-	+	-	+	-	+
β-Glucosidase T+	143,89	155,17	1243,77	1148,11	320,66	410,58
β-Glucuronidase	80,92	79,89	21,72	18,82	10,10	12,16
β-Galactosidase	663,02	671,20	15,91	14,04	13,90	17,34
α-Galactosidase	50,98	54,02	13,78	12,26	8,62	8,99
Arylsulfatase B (ARSB)	121,66	125,60	52,41	58,37	44,08	47,26
Hexosaminidase T	37774,81	53463,49	12588,29	11000,06	5818,07	6240,44
α-Mannosidase	216,24	236,39	33,91	27,12	17,54	20,02

Comparatively with the results observed prior to AO transfection, the levels of the enzymes transported through the M6P-dependent pathway remains quite reduced compared to the control fibroblasts and very little change is observed in the presence of the AO despite RNA analysis confirming the successful skipping of exon 19. In fact, contrary to expectations, all enzymes in Patient 1 seem to present a slight decrease in their activity 24 hours after AO transfection, with the exception of ARSB which is slightly augmented in treated cells. On the other hand, 48 hours following transfection, enzyme activity appears to suffer a small increase in the patient fibroblasts for all enzymes tested, even if the results are still much lower than those observed for the healthy controls. It should be noted, however, that all mentioned variations appear to be too small to be significative and most probably cannot be considered relevant changes in the levels of lysosomal enzyme activities in the cells.

As a whole, the results of the enzymatic assays suggest that the H19A(+119+138) AO was unable to produce significant changes in the levels of the lysosomal enzymes transported through the M6P-dependent pathway in ML II Patient 1 cells, despite effectively inducing the skipping of exon 19 from the final *GNPTAB* transcript.

Chapter 5

Discussion

Among the multiple mutations that can result in an ML II phenotype, the most frequent is c.3503_3504delTC, a dinucleotide deletion on exon 19 of the *GNPTAB* gene responsible for encoding the α/β -precursor peptide (Kudo et al., 2006; Coutinho et al., 2011). Broadly studied at our research group in the past, this mutation causes a frameshift in the transcript reading frame, as well as premature stop codon, and eventually leads to the production of a shorter inactive protein product (Plante et al., 2008). Included in the lost segment of the protein is the C-terminal arginine/isoleucine/arginine (RIR) motif, which together with the N-terminal dileucine motif, is essential for the transport of the α/β -precursor from the endoplasmic reticulum to the Golgi apparatus where it can be cleaved into mature α and β subunits of the GlcNAc-1-phosphotransferase enzyme (Franke et al., 2013). Therefore, in ML II patients, this mutant peptide cannot exit the endoplasmic reticulum and consequently impairs its cleavage and the formation of mature α and β subunits, preventing GlcNAc-1-phosphotransferase activity (De Pace et al., 2013).

Considering the nature of this ML II causing mutation, the possibility of using AOs as therapeutic agents arose. Among the multiple mechanisms for which AOs can be employed, the most appropriate in this case was exon-skipping, which implies a strategy where AOs bind to specific splicing regulatory elements in the pre-mRNA in an attempt to exclude the target exon from the final transcript, therefore removing the disease causing mutation and restoring the reading frame (Hammond & Wood, 2011; Siva et al., 2014). This approach has previously been applied with encouraging results for the treatment of DMD, where AOs were used to induce the skipping of exon 51 (harboring the mutation) and produced an internally truncated but partially functional dystrophin protein, ameliorating the patients' condition quite significantly (van Deutekom et al., 2007; Mendell et al., 2013). Similarly, we thought to use AOs in ML II patient fibroblasts to promote the exclusion of exon 19 from the final *GNPTAB* transcript, in order to circumvent the c.3503_3504delTC mutation and restore the reading frame, hoping to produce a truncated but functional α/β -precursor that could then be cleaved into mature α and β subunits and achieve GlcNAc-1-phosphotransferase activity.

Considering the implications of removing a reasonable segment from the final protein product, our first step in this project was to analyze *in silico* the α/β -precursor peptide to ensure that the removal of exon 19 would actually produce an in-frame protein and that no important domains would be removed due to this exon-skipping approach. For that purpose, we performed a thorough analysis of this peptide, using multiple bioinformatic tools and softwares. The alignment of the amino acid sequence of the *GNPTAB* protein product minus exon 19 with the normal *GNPTAB* protein product showed that removing exon 19 indeed produced an in-frame protein, 56 amino acids shorter than the regular protein product, but

with the amino acid sequence following the excluded region presenting a complete match to the normal α/β -precursor peptide. When the amino acid sequence encoded by exon 19 was aligned with from multiple animal species, a great level of conservation was detected between the sequences, with just small discrepancies in the sequences from mouse and zebrafish. However, while important protein domains are usually conserved across species, the opposite is not always true, with several regions presenting high levels of conservation without any apparent function.

Acknowledging this, protein domains were predicted for the *GNPTAB* protein product to confirm the absence of essential elements within the region encoded by exon 19 that would naturally be excluded by the planned exon-skipping approach. Predictions obtained from three different databases all confirmed the presence of several domains previously described for this protein product: the stealth domain, spread across four regions, two Notch repeats and a DMAP domain (Qian et al., 2015; van Meel et al., 2016). The Notch repeats are characteristic of Notch receptor family members and regulate the cleavage of the Notch receptor, preventing its activation in the absence of a proper ligand (Sanchez-Irizarry et al., 2004), while the DMAP domain functions as an interaction domain between proteins and is responsible for the selective binding of this enzyme to lysosomal hydrolases (Qian et al., 2013). Together, these domains are essential for the recognition and binding of lysosomal enzymes to GlcNAc-1-phosphotransferase, with mutants on either of these locations maintaining the catalytic activity of this enzyme and its location in the Golgi apparatus but losing the ability to phosphorylate any lysosomal hydrolases (van Meel et al., 2016). On the other hand, the stealth domain is highly conserved in bacteria and metazoans and appears to play an important role in the synthesis of extracellular polysaccharides and evasion of detection by the host's immune system, providing great advantage to pathogenic bacteria (Sperisen et al., 2005). In the GlcNAc-1-phosphotransferase enzyme, however, this domain has been shown to harbor the catalytic site, as mutations in these regions have severely reduced the activity of the enzyme, without affecting its Golgi localization and proteolytic cleavage (Qian et al., 2015). Given this, the presence of stealth domain CR4 within the amino acid sequence encoded by exon 19 could appeared to be quite problematic, since the removal of this domain from the final protein product could lead to a decrease or even absence of GlcNAc-1-phosphotransferase catalytic activity, rendering this therapeutic approach unfit for the treatment of ML II. Still, since only one out of the four stealth domains would be excluded, we hypothesized that some residual activity could still be present after skipping of exon 19. It is, therefore, imperative that enzyme activity is carefully analyzed following the application of this innovative therapy in order to ensure it produces a functional

enzyme, even if only partially, rather than a Golgi-located but inactive protein, which would do very little to ameliorate the phenotype of this disorder. Aside from those domains, the γ -subunit binding (GB) region has also been identified in a previously uncharacterized region of the α -subunit between amino acid positions 535 and 698 (De Pace et al., 2015) but was not predicted by any of the three consulted databases. More recently, Velho and colleagues were able to distinguish two subdomains within this region – GB1 (aa 535-588) and GB3 (aa 645-698) – which play an essential role in the interaction between the α -subunits and the γ -subunits, as indicates the observation that the simultaneous deletion of both subdomains leads to a complete loss of interaction between the subunits, while the deletion of each of them alone causes a reduction in protein interaction (Velho et al., 2016). While these subdomains do not interfere directly with our target region, the very recent identification of this γ -subunit binding region testifies the constant progress on research and knowledge about this enzyme, as well as multiple others, and reinforces the necessity for permanent updates of online databases in order to provide users with the most recent discoveries and correct information in the field.

Regarding protein glycosylation, despite the fact that the α/β -precursor is a highly *N*-glycosylated peptide (Velho et al., 2016), bioinformatic predictions did not indicate any sites with strong glycosylation potential within the amino acids encoded by exon 19, which is consistent with information previously published in the literature (Kollmann et al., 2010). Thus, it is unlikely that the removal of exon 19 from the final *GNPTAB* transcript would have some impact on the glycosylation of this protein product, since all of the glycosylation sites would remain intact and available for this post-translational modification. In summary, according to the combined bioinformatic analyses, the main contrariety of correcting the c.3503_3504delTC mutation through an exon-skipping approach was that a considerable portion of the stealth domain CR4 would be removed from the α/β -precursor, which could severely impact the catalytic activity of the GlcNAc-1-phosphotransferase enzyme. Nonetheless, the fact of such impairment being a mere prediction obtained from bioinformatic tools, having not yet been experimentally confirmed, together with the remaining bioinformatic indications being quite encouraging, prompted us to further continue our work in the development of this innovative therapeutic approach for the treatment of ML II.

Bioinformatic tools were then used to design AOs and select the most appropriate targets to induce the skipping of exon 19 of the *GNPTAB* gene. For that purpose, the acceptor and donor splice sites of the exon were scored according to their strength, that is, their similarity to the consensus sequence, to verify whether they could be useful targets for AO binding. Two distinct bioinformatic tools provided concordant scores, with the 3'ss

consistently outscoring the 5'ss on both platforms, which suggests the presence of a very strong acceptor site on exon 19, expected to play a crucial role in the splicing reaction. In light of these findings, this location came forward as a very promising target for AOs, as the blocking of the 3'ss alone could prevent the recognition of the exon by the splicing machinery and result in its exclusion from the final transcript. Additionally, the presence of enhancers within exon 19 and its flanking intronic regions (50bp) was predicted with bioinformatic tools as these could equally be potential targets for AOs. In this case, results revealed the presence of several enhancers across the sequence of the exon, with a dense cluster around the 5'ss. However, the highest scoring enhancers were not present in this cluster, but rather in the first half of the exon and surrounding the 3'ss, suggesting this region could also be target by AOs for the purpose of inducing the skipping of exon 19. In fact, previous studies associated with exon-skipping in the dystrophin gene indicate a tendency for AOs targeting the acceptor site or enhancers present in the first half of the exon to show higher levels of efficiency than those targeting other locations (Wilton et al., 2007; Fletcher et al., 2012). Therefore, we decided to also explore the design of AOs in this region to verify whether the efficient outcomes would hold true for the *GNPTAB* gene and if AOs targeting these positions would successfully promote the exclusion of exon 19 from the final transcript. At the end, four AOs were designed: one for each splice site and two for exonic internal regions with high density of predicted enhancers, expecting that at least one of these molecules would efficiently induce the skipping of exon 19 of the *GNPTAB* gene.

Moving next to the experimental framework, fibroblast cells from two distinct ML II patients were cultured and characterized to confirm the presence of the c.3503_3504delTC mutation, while their endogenous splicing pattern was also assessed to ensure that no splicing abnormalities were present in these cells prior to transfection with AOs. As expected, the results from these two molecular analyses confirmed the presence of the c.3503_3504delTC mutation in the patient fibroblasts, as well as a regular endogenous splicing pattern, proving that these cells are suitable for use in the planned exon-skipping experiments.

Since the c.3503_3504delTC mutation and an unaltered splicing pattern were successfully confirmed in the patient fibroblasts, we moved on to their transfection with the selected AOs. Overall, all four AOs achieved successful skipping of exon 19, at least to some extent. However, contrary to what has previously been published (Arechavala-Gomez et al., 2007; Harding et al., 2007; Wilton et al., 2007), the most effective molecules were not the AOs targeting the acceptor splice site or the first half of the target exon. In fact, among the four molecules tested, the AOs directed at the donor splice site, H19D(+12-9), and the

second half of the exon, H19A(+119+138), managed to accomplish higher levels of exon-skipping and displayed a very consistent dose-dependent response, with a gradual increase in the amount of the shorter transcript across the four tested concentrations.

In the case of H19D(+12-9) AO, the efficiency of exon-skipping can be explained by the crucial role the donor splice site plays in the spliceosome assembly, since its recognition and binding by U1 snRNP represents the first step in the complex cellular process of splicing (Chen & Manley, 2009; De Conti et al., 2013). In that sense, the blocking of this location alone by the H19D(+12-9) AO might have inhibited the assembly of the spliceosome complex in this region and consequently lead to the exclusion of exon 19 from the final *GNPTAB* transcript in the treated cells. Furthermore, the 5'ss of this particular exon is surrounded by a considerably high number of enhancer elements, both within the exon and in the intronic sequence, and these are equally blocked by H19D(+12-9), which greatly diminishes their action and ultimately prevents the inclusion of exon 19 in the final mRNA.

Another important factor is the secondary RNA structure of the target region, which can greatly influence the binding and efficiency of AOs. As described by Aartsma-Rus (Aartsma-Rus, 2012), immediately following transcription, the pre-mRNAs form secondary structures, leaving some sequences strongly bound to each other in a closed structure, while others originate loops, where they are in a more accessible and open structure. Considering this, AOs must have as target a balanced structure, partially open and partially closed, in order to bind successfully to it and significantly alter the RNA secondary structure to induce the intended splicing modulation. Otherwise, if the target structure is too closed, the AO will probably fail to bind to the sequence, whereas if the structure is completely open, the AO might bind to the sequence but most likely will not alter the RNA structure in a substantial manner, producing little to no effect (Aartsma-Rus, 2012). The results observed here could also reflect these differences in target structure, with the binding of AOs to some locations showing higher levels of exon-skipping and more consistent outcomes probably due to targeting a more favorable RNA structure, whereas other RNA segments may present excessively open or closed structures, hindering the action of AOs.

Aside from these factors associated with RNA structure and AO location, the transfection reagent selected for the experiment may also play a significant role in its efficiency, leading to different levels of exon-skipping. In fact, when the skipping of exon 19 from the *GNPTAB* transcript was experimentally tested using Lipofectamine® LTX no exon-skipping was observed in any of the conditions tested (data not shown). However, when Lipofectamine® 2000 was used for the transfection, exon-skipping was successfully achieved in most of the conditions and a shorter transcript without the target exon was clearly visible in

most of the samples. It is, therefore, important to weight the choice of the transfection reagent to be used in this sort of experiments and test different alternatives in order to attain a more accurate perception on the efficiency of specific AOs, and consequently on the feasibility of this technique.

Additionally, besides the full-length *GNPTAB* transcript and its shorter variant missing exon 19, some samples also revealed the presence of other amplification products between these two fragments. As previously observed in studies carrying out the skipping of exons in DMD (van Deutekom et al., 2007; Kinali et al., 2009, Adkin et al., 2012), intermediate fragments appear with relative frequency in the course of exon-skipping techniques. Possibly, they can result from a partial skipping of the target exon sequence, producing a transcript shorter than its endogenous variant but still longer than the desired molecule, which in this work was a sequence missing the full exon 19. In turn, an even shorter transcript, spanning about 500bp, was visible in some of the samples, including from transfection controls without any AOs, and its sequencing revealed the opposite situation, with both exon 19 and 20 absent in the final *GNPTAB* transcript. While the unexpected splicing of extra exons has also been reported as a common side event in exon-skipping procedures (Adkin et al., 2012; Fletcher et al., 2012), here the detection of such a product in the control sample is quite intriguing since no AO was administered to justify the splicing alteration. Still, this transcript could be a minor endogenous alternative splicing product only observed under certain experimental conditions. Although these amplification products do not compromise the success of this exon-skipping approach, they serve as a reminder that this sort of technique still requires strong refinement before it can proceed to *in vivo* experiments in order to prevent the possibility of adverse secondary effects.

Given the positive results observed for the H19A(+119+138) and H19D(+12-9) AOs, targeting the ESEs within the second half of the exon and the donor splice site, respectively, these two molecules were selected for further evaluation with higher concentrations. As observed at lower AO concentrations, both molecules induced a successful skipping of exon 19 in patient and control fibroblasts. However, while H19D(+12-9) revealed a clear dose-dependent response and a gradual increase in the amount of shorter transcript missing the target exon, H19A(+119+138) did not, instead resulting in stable but solid levels of shorter transcript and only small amounts of the full-length mRNA. This was especially clear in the patient fibroblasts where the levels of shorter transcript far surpassed those observed for the full-length mRNA, effectively inverting the ratio between these two variants of the *GNPTAB* gene. Once again, these results oppose the data previously reported (Arechavala-Gomez et al., 2007; Harding et al., 2007; Wilton et al., 2007), according to which the most effective

AOs are those directed to the acceptor splice site or the first half of the target exon. Here we observed that, despite blocking an essential location for the assembly of the spliceosome, the H19D(+12-9) AO turned out to be slightly less efficient than H19A(+119+138), the one binding to an internal location in the second half of the exon, without particularly strong ESEs, which, based on the bioinformatic predictions would make it the least likely of the two AOs to induce successful exon-skipping. Actually, H19A(+119+138) showed the best results and appears to be the most promising AO among the four tested, strongly inducing exon-skipping and leading to the production of very satisfactory levels of *GNPTAB* transcript missing exon 19. Additionally, it also effectively reduced the full-length out-of-frame mRNA to levels below those observed for its shorter variant, demonstrating the potential of this molecule as a future innovative therapeutic agent for ML II.

At last, the impact of this therapeutic approach at protein level was assessed indirectly by quantification of lysosomal enzymes and their activity. Since in ML II patients harboring the c.3503_3504delTC mutation, GlcNAc-1-phosphotransferase activity is almost totally absent, the removal of exon 19 from the final *GNPTAB* transcript was expected to restore, at least to some degree, the activity of the enzyme, to attain an improvement of lysosomal trafficking and reduction of excreted lysosomal enzymes. Unfortunately, in patient fibroblasts transfected once again with the H19A(+119+138) AO for 24 and 48 hours, no significant change in activity in a number of lysosomal hydrolases was detected, which was a sign that no GlcNAc-1-phosphotransferase activity had been restored. Nonetheless, minor increases in enzyme activity were observed at 48 hours, suggesting this period of incubation was slightly more effective in improving GlcNAc-1-phosphotransferase activity and that extending the experiment for a longer period of time could produce overall better results.

Soon in the work, we anticipated that, given the role of stealth domains for the catalytic activity of GlcNAc-1-phosphotransferase (Qian et al., 2015), the presence of a considerable portion of stealth domain CR4 within the amino acid sequence encoded by exon 19 could represent a serious issue for the planned therapeutic approach. However, the presence of three other stealth domains in the α/β -precursor allowed us to hypothesize that enzyme activity would not be totally lost after removing the CR4 domain. Also, only two missense mutations are listed in The Human Gene Mutation Database for this domain, one associated with a very mild ML III phenotype (Otomo et al., 2009) and the other associated with a case of stuttering (Kang et al., 2010), while in other locations of this protein missense mutations have been linked to a severe ML II phenotype (Otomo et al., 2009; Zarghooni et al., 2009; Qian et al., 2015). Therefore, based on these first results it is not possible to conclude for sure that the activity of the GlcNAc-1-phosphotransferase was not partially

recovered, since we only used indirect methods for checking its status. In fact, GlcNAc-1-phosphotransferase levels are much more reduced in ML II patients than in ML III patients and well correlated with the more severe phenotype that is observed in ML II patients. However, these two ML forms cannot be distinguished through the indirect quantification of the lysosomal enzyme activities since their levels are similar in both diseases, which supports the need for a direct evaluation of GlcNAc-1-phosphotransferase activity.

As corollary, while this work showed encouraging results at RNA level, further research is still needed to obtain a better assessment of the activity of the resultant protein product. As explained in previous chapters, direct measurement of GlcNAc-1-phosphotransferase activity is not routinely used for diagnosis because it requires a cumbersome method involving the radioactive substrate [^{32}P]UDP-GlcNAc only feasible to be carried out in specialized facilities (Reitman & Kornfeld, 1981a), with most routine evaluations performed through indirect measurements. In this research project, however, direct evaluation should necessarily be conducted, because it is much more sensitive than the indirect alternative, constituting the best means to verify whether the developed exon-skipping strategy manages to restore some level of protein activity, despite the absence of the 56 amino acids encoded by exon 19. In addition, expression analysis of a mutant *GNPTAB* construct without exon 19 by Western blotting and immunofluorescence microscopy could help verify whether the α/β -precursor is retained in the endoplasmic reticulum or is correctly transported to the Golgi apparatus and proteolytically processed, supporting or not the feasibility of this approach.

As previously reported (Steet et al., 2005), patients diagnosed with the milder ML III disorder also exhibit levels of GlcNAc-1-phosphotransferase activity below the normal range (2-20%) but their symptoms and overall phenotype are considerably more favorable than in ML II patients. One patient in particular, carrying a missense mutation within the region encoded by exon 19 of the *GNPTAB* gene, exhibited very low levels of GlcNAc-1-phosphotransferase activity (<2%) but still presented with a very mild ML III phenotype (Otomo et al., 2009; Qian et al., 2015). Expectedly, an increase in the levels of enzyme activity to about 6-10% of normal would probably allow for a serious improvement in the condition of these patients and extend their life expectancy (Steet et al., 2005). Moreover, past studies have reported on asymptomatic individuals with very reduced levels of enzyme activity for other lysosomal enzymes (Leinekugel et al., 1992), as well as for enzymes associated with cystic fibrosis (Rowntree & Harris, 2003). More recently, an heterozygous ML III carrier was demonstrated to only exhibit 36,2% of normal GlcNAc-1-phosphotransferase activity despite presenting with a normal clinical phenotype (Cathey et al., 2010). These

findings reinforce that it is possible to achieve a normal asymptomatic phenotype with levels of enzyme activity considerably lower than those observed in the normal range and that even small increases in enzyme activity can have a major impact in ameliorating the symptoms associated with these disorders.

Nonetheless, further research and investment is still needed in the field of ML II for the development and implementation of innovative therapeutic strategies, opening the way to offer one day personalized treatments to patients that can ultimately improve their overall quality of life.

Chapter 6

Conclusion and Future Perspectives

In recent years, as new information on RNA biology and chemistry arose, this molecule became one of the most promising targets for therapeutic intervention in multiple diseases associated with genetic mutations present in the transcriptome. Unlike other forms of therapy focused on ameliorating the symptoms of a disorder, RNA therapeutics act directly upon the disease causing mutation, allowing for a very precise and specific treatment with reduced secondary effects. One of the most notorious examples of the utility of RNA-based therapies is the recent developments in the field of DMD, where AOs have successfully been used for the skipping of specific exons and have been approved for clinical use since the end of last year after years of intensive research (Aartsma-Rus & Krieg, 2017).

Considering this and other achievements in disorders where AOs were approved for clinical use (Crooke & Geary, 2013; Chery & Näär, 2016; Ottesen, 2017), this project set out to develop a similar exon-skipping strategy for the treatment of patients harboring the c.3503_3504delTC mutation responsible for ML II. After testing *in vitro* different AOs and concentrations through transfection, we can safely affirm that all four AOs tested were able to induce the skipping of exon 19 to some extent, confirming that the exclusion of this sequence from the *GNPTAB* transcript is indeed feasible through this sort of technique. Among the selected AOs, H19A(+119+138), directed at a region of ESEs within the second half of the target exon, clearly stood out as the most successful, having produced a considerable amount of the shorter transcript missing exon 19, even at relatively low dosage. At the same time, this AO also promoted a significant reduction in the amount of full-length mRNA in the treated fibroblast cells, consequently increasing the ratio between the shorter transcript variant and the full-length molecule. Given the ultimate goal of this therapeutic approach – the complete replacement of the natural occurring *GNPTAB* transcript harboring the c.3503_3504delTC mutation with the shorter mRNA missing exon 19 resultant from this exon-skipping strategy –, this AO appears quite promising and should therefore be further explored as a potential treatment for ML II patients carrying exon 19 mutations. Unfortunately, these encouraging results did not replicate at protein level, as the assessment of enzyme activity for several lysosomal hydrolases following transfection with H19A(+119+138) revealed very small changes when compared to untreated fibroblasts. However, considering these are only indirect measurements of GlcNAc-1-phosphotransferase activity, for the moment it cannot be definitely excluded that some level of enzyme activity has been restored, since that can only be confirmed with direct assays. Nonetheless, the mRNA results are very encouraging and further research should be conducted regarding this exon-skipping strategy.

On grounds of these results, the next logical step would be to thoroughly study the expression levels of both *GNPTAB* transcripts through Real-Time PCR techniques in order to measure accurately the action of the AOs used. Additionally, the protein product resultant from the skipping of exon 19 should also be carefully analyzed. As previously noted (Velho et al., 2015; De Pace et al., 2013; Franke et al., 2013; van Meel et al., 2016), it is crucial to study the effects of such an impactful alteration of the transcript at protein level, using Western Blot techniques to compare the expression patterns of both protein variants (full-length and artificially truncated), as well as verify the length of the α/β -precursor following the exclusion of exon 19 and ultimately confirm the successful cleavage of the α and β subunits, allowing for the assembly of an active GlcNAc-1-phosphotransferase enzyme.

Additionally, experiments for quantification of lysosomal enzyme activity should be carried out for a longer period of time and using higher AO concentrations, and enzyme levels should also be measured on culture media since enzymes that fail to reach the lysosome are subsequently secreted to the extracellular milieu (Kornfeld & Sly, 2001; Braulke et al., 2008). Furthermore, GlcNAc-1-phosphotransferase activity should also be assessed directly to accurately determine its level of restoration. The use of radioactive [^{32} P]UDP-GlcNAc as the donor substrate together with the phosphate acceptor α -methylmannoside (α -MM) as first described by Reitman and Kornfeld allows for a direct measurement of GlcNAc-1-phosphotransferase activity (Reitman & Kornfeld, 1981a), which could provide very important information regarding the effects of this therapeutic approach at enzyme level (Steet et al., 2005; Kudo et al., 2006; Qian et al., 2010; Velho et al., 2015). In alternative, given the role of this enzyme in marking lysosomal enzymes with M6P, its activity could also be estimated through the immunoprecipitation of M6P-containing proteins using a M6P-specific antibody in order to verify if any changes in the amount of M6P-containing proteins are visible following AO transfection (Muller-Loennies et al., 2010).

Lastly, the location of GlcNAc-1-phosphotransferase is also an important point to address in the scope of this therapeutic approach, since the endogenous α/β -precursor in ML II is retained in the endoplasmic reticulum and is unable to progress to the Golgi apparatus, where the cleavage of the α and β subunits occurs and the assembly of a functional GlcNAc-1-phosphotransferase takes place (De Pace et al., 2013; Franke et al., 2013). In that sense, it would be interesting to examine the subcellular localization of both the α/β -precursor and the GlcNAc-1-phosphotransferase enzyme, using specific antibodies and fluorescence microscopy, in order to assess whether the altered protein product can successfully reach the Golgi apparatus. Similarly, the same methods could also be used to identify the subcellular localization of some lysosomal enzymes to confirm their presence within the

lysosomal compartment and not in the extracellular milieu, as commonly happens in ML II patients (De Pace et al., 2013; Franke et al., 2013; Velho et al., 2015; van Meel et al., 2016).

The success of this therapeutic approach also depends of other major challenges, such as the successful delivery of this AO to the target tissues, which is still one of the most complex issues in the field of RNA therapeutics. Given the difficulty of cellular uptake for naked AOs, several delivery mechanisms have been tested in an attempt to improve the success of this process and increase the levels of AO inside the cell. However, some of the agents used, like cell-penetrating peptides, cationic lipids or polymers, still present some risks and associated toxicity, requiring careful testing and refinement before they can be considered for clinical use (Amantana et al., 2007; Moreno & Pêgo, 2014; Gurav & Srinivasan, 2017). Furthermore, viral vectors, which have previously been used for the incorporation and delivery of AOs due to their natural ability to penetrate cells, can also face some obstacles when used *in vivo* because they can activate the host's immune system with the consequent production of specific antibodies that neutralize the action of these viral vectors in future systemic injections (Douglas & Wood, 2011; Imbert et al., 2017). Additionally, even in the presence of delivery mechanisms, most AOs fail to reach the central nervous system when administered peripherally due to their inability to cross the blood-brain barrier (Chery & Näär, 2016). Considering most of the disorders targeted by AO therapies, including ML II, present severe neurological symptoms, the nervous system must be an essential focus for the treatment with AOs and alternative delivery methods that ensure efficient distribution of these agents into the brain must be developed and refined.

Finally, before reaching widespread medical use, AOs must face one last hurdle: clinical approval and regulation by drug administrations and medicine agencies. While all medicines go through this phase of the drug development pipeline before entering the market, the process can be quite more complex for AOs, especially those targeting rare diseases. Not only is the number of patients with these disorders obviously reduced, those who carry a specific mutation feasible to be treated with the same AO are an even smaller group, making it very difficult to conduct robust clinical trials (Douglas & Wood, 2011). Taking this into account, the scientific community has been working towards considering AOs as a single drug class in order to facilitate their approval and ultimately accelerate the progress in the development of new therapies (Hammond & Wood, 2011). This would give a completely different opportunity to these drugs, opening the possibility to provide patients, even those with private mutations, with the necessary AO treatment for their disorders, bringing everyone closer to the prospects of truly personalized medicine.

Overall, this study was pioneer in the development of antisense-mediated exon-skipping strategies for the treatment of ML II, achieving encouraging results and reinforcing the role of AOs as very promising therapeutic agents for this and other severe disorders.

References

A

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

Aartsma-Rus A (2012). Overview on AON Design. In: Aartsma-Rus A, eds, *Exon Skipping: Methods and Protocols*, Methods in Molecular Biology, vol. 867. Humana Press, New York, USA, pp 117-129.

Aartsma-Rus A, Krieg AM (2017). FDA Approves Eteplirsen for Duchenne Muscular Dystrophy: The Next Chapter in the Eteplirsen Saga. *Nucleic Acid Therapeutics*, 27 (1): 1-3.

Adkin CF, Meloni PL, Fletcher S, Adams AM, Muntoni F, Wong B, Wilton SD (2012). Multiple exon skipping strategies to by-pass dystrophin mutations. *Neuromuscular Disorders*, 22 (4): 297-305.

Alegra T, Koppe T, Acosta A, Sarno M, Burin M, Kessler RG, Sperb-Ludwig F, Cury G, Baldo G, Matte U, Giugliani R, Schwartz IV (2014). Pitfalls in the prenatal diagnosis of mucopolipidosis II alpha/beta: A case report. *Meta Gene*, 2: 403-406.

Al-Jasmi FA, Tawfig N, Berniah A, Ali BR, Taleb M, Hertecant, JL, Bastaki F, Souid AK (2012). Prevalence and Novel Mutations of Lysosomal Storage Disorders in United Arab Emirates. *JIMD reports*, 10: 1-9.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215: 403-410.

Amantana A, Moulton HM, Cate ML, Reddy MT, Whitehead T, Hassinger JN, Youngblood DS, Iversen PL (2007). Pharmacokinetics, biodistribution, stability and toxicity of a cell-penetrating peptide-morpholino oligomer conjugate. *Bioconjugate Chemistry*, 18 (4): 1325-1331.

Applegarth DA, Toone JR, Lowry RB (2000). Incidence of Inborn Errors of Metabolism in British Columbia, 1969–1996. *Pediatrics*, 105 (1): e10.

Arechavala-Gomez V, Graham IR, Popplewell LJ, Adams AM, Aartsma-Rus A, Kinali M, Morgan JE, van Deutekom JC, Wilton SD, Dickson G, Muntoni F (2007). Comparative analysis of antisense oligonucleotide sequences for targeted skipping of exon 51 during dystrophin pre-mRNA splicing in human muscle. *Human Gene Therapy*, 18 (9): 798-810.

Arechavala-Gomez V, Khoo B, Aartsma-Rus A (2014). Splicing modulation therapy in the treatment of genetic diseases. *The Application of Clinical Genetics*, 7: 245-252.

Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D,

Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research*, 40: W597–W603.

B

Ballabio A, Gieselmann V (2009). Lysosomal disorders: From storage to cellular damage. *Biochimica et Biophysica Acta*, 1793 (4): 684-696.

Bao M, Booth JL, Elmendorf BJ, Canfield WM (1996). Bovine UDP-N-acetylglucosamine:lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure. *The Journal of Biological Chemistry*, 271 (49): 31437-31445.

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

Bennett CF, Swayze EE (2010). RNA Targeting Therapeutics: Molecular Mechanisms of Antisense Oligonucleotides as a Therapeutic Platform. *Annual Review of Pharmacology and Toxicology*, 50: 259-293.

Bergsma AJ, in't Groen SL, Verheijen FW, van der Ploeg AT, Pijnappel WP (2016). From Cryptic Toward Canonical Pre-mRNA Splicing in Pompe Disease: a Pipeline for the Development of Antisense Oligonucleotides. *Molecular Therapy: Nucleic Acids*, 5 (9): e361.

Black DL (2003). Mechanisms of Alternative pre-Messenger RNA Splicing. *Annual Review of Biochemistry*, 72: 291–336.

Boustany RMN (2013). Lysosomal storage diseases – the horizon expands. *Nature reviews. Neurology*, 9 (10): 583-598.

Braulke T, Pohl S, Storch S (2008). Molecular analysis of the GlcNac-1-phosphotransferase. *Journal of inherited metabolic disease*, 31 (2): 253-257.

Braulke T, Raas-Rothschild A, Kornfeld S (2013). I-cell disease and pseudo-Hurler polydystrophy: Disorders of lysosomal enzyme phosphorylation and localization. In: Valle D, Vogelstein B, Kinzler K, Antonarakis S, Ballabio A, Scriver C, Sly W, Bunz F, Gibson K, Mitchell G, eds., *The online metabolic and molecular basis of inherited diseases*. McGraw-Hill Education, New York, USA.

Bruni S, Loschi L, Incerti C, Gabrielli O, Coppa GV (2007). Update on treatment of lysosomal storage diseases. *Acta myologica: myopathies and cardiomyopathies: official journal of the Mediterranean Society of Myology*, 26 (1): 87-92.

C

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003). ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acid Research*, 31 (13): 3568-3571.

Cathey SS, Leroy JG, Wood T, Eaves K, Simensen RJ, Kudo M, Stevenson RE, Friez MJ (2010). Phenotype and genotype in mucopolipidoses II and III alpha/beta: a study of 61 probands. *Journal of Medical Genetics*, 47 (1): 38-48.

Chan JH, Lim S, Wong WS (2006). Antisense oligonucleotides: from design to therapeutic application. *Clinical and Experimental Pharmacology & Physiology*, 33 (5-6): 533-540.

Chauhan JS, Rao A, Raghava GPS (2013). *In silico* Platform for Prediction of N-, O- and C-Glycosites in Eukaryotic Protein Sequences. *PloS One*, 8 (6): e67008.

Chen M, Manley JL (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature Reviews Molecular Cell Biology*, 10 (11): 741–754.

Chery J, Näär A (2016). RNA therapeutics: RNAi and antisense mechanisms and clinical applications. *Postdoc Journal: a Journal of Postdoctoral Research and Postdoctoral Affairs*, 4 (7): 35-50.

Chiriboga CA, Swoboda KJ, Darras BT, Iannaccone ST, Montes J, De Vivo DC, Norris DA, Bennett CF, Bishop KM (2016). Results from a phase 1 study of nusinersen (ISIS-SMN_{Rx}) in children with spinal muscular atrophy. *Neurology*, 86 (10): 890-897.

Cirak S, Arechavala-Gomez V, Guglieri M, Feng L, Torelli S, Anthony K, Abbs S, Garralda ME, Bourke J, Wells DJ, Dickson G, Wood MJA, 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 (9791): 595-605.

Crooke ST, Geary RS (2013). Clinical pharmacological properties of mipomersen (Kynamro), a second generation antisense inhibitor of apolipoprotein B. *British Journal of Clinical Pharmacology*, 76 (2): 269-276.

Coutinho MF, Encarnação M, Gomes R, da Silva Santos L, Martins S, Sirois-Gagnon D, Bargal R, Filocamo M, Raas-Rothschild A, Tappino B, Laprise C, Cury GK, Schwartz IV, Artigalás O, Prata MJ, Alves S (2011). Origin and spread of a common deletion causing mucopolipidosis type II: insights from patterns of haplotypic diversity. *Clinical Genetics*, 80 (3): 273-80.

Coutinho MF, Prata MJ, Alves S (2012). Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction. *Molecular Genetics & Metabolism*, 105 (4): 542-550.

Cox TM (2005). Biomarkers in lysosomal storage diseases: a review. *Acta Pædiatrica*, 94 (447): 39-42.

Cox TM (2015). Innovative Treatments for Lysosomal Diseases. *Best practice & research. Clinical Endocrinology & Metabolism*, 29 (2): 275-311.

D

De Conti L, Baralle M, Buratti E (2013). Exon and intron definition in pre-mRNA splicing. *Wiley Interdisciplinary reviews. RNA*, 4 (1): 49-60.

De Pace R, Coutinho MF, Koch-Nolte F, Haag F, Prata MJ, Alves S, Bräulke T, Pohl S (2013). Mucopolidosis II-Related Mutations Inhibit the Exit from the Endoplasmic Reticulum and Proteolytic Cleavage of GlcNAc-1-Phosphotransferase Precursor Protein (GNPTAB). *Human mutation*, 35 (3): 368-376.

De Pace R, Velho RV, Encarnação M, Marschner K, Bräulke T, Pohl S (2015). Subunit interactions of the disease-related hexameric GlcNAc-1-phosphotransferase complex. *Human Molecular Genetics*, 24 (23): 6826-6835.

Dias N, Stein CA. (2002). Antisense oligonucleotides: basic concepts and mechanisms. *Molecular Cancer Therapeutics*, 1 (5): 347-355.

Dionisi-Vici C, Rizzo C, Burlina AB, Caruso U, Sabetta G, Uziel G, Abeni D (2002). Inborn errors of metabolism in the Italian pediatric population: A national retrospective survey. *The Journal of Pediatrics*, 140 (3): 321-327.

Dominski Z, Kole R (1993). Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proceedings of the National Academy of Sciences of the United States of America*, 90 (18): 8673-8677.

Douglas AGL, Wood MJA (2011). RNA splicing: disease and therapy. *Briefings in Functional Genomics*, 10 (3): 151-164.

Douglas AG, Wood MJ (2013). Splicing therapy for neuromuscular disease. *Molecular and Cellular Neurosciences*, 56: 169-185.

Du L, Pollard JM, Gatti RA (2007). Correction of prototypic ATM splicing mutations and aberrant ATM function with antisense morpholino oligonucleotides. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (14): 6007-6012.

E

Encarnação M, Lacerda L, Costa R, Prata MJ, Coutinho MF, Ribeiro H, Lopes L, Pineda M, Ignatius J, Galvez H, Mustonen A, Vieira P, Lima MR, Alves S (2009). Molecular analysis of the GNPTAB and GNPTG genes in 13 patients with mucopolipidosis type II or type III - identification of eight novel mutations. *Clinical Genetics*, 76 (1): 76-84.

Errington SJ, Mann CJ, Fletcher S, Wilton SD (2003). Target selection for antisense oligonucleotide induced exon skipping in the dystrophin gene. *The Journal of Gene Medicine*, 5 (6): 518-527.

Evers MM, Toonen LJA, van Roon-Mom WMC (2015). Antisense oligonucleotides in therapy for neurodegenerative disorders. *Advanced Drug Delivery Reviews*, 87: 90-103.

F

Fairbrother WG, Yeh RF, Sharp PA, Burge CB (2002). Predictive identification of exonic splicing enhancers in human genes. *Science*, 297 (5583): 1007-1013.

Faustino NA, Cooper TA (2003). Pre-mRNA splicing and human disease. *Genes & development*, 17 (4): 419-437.

Filocamo M, Morrone A (2011). Lysosomal storage disorders: Molecular basis and laboratory testing. *Human Genomics*, 5 (3): 156-169.

Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*, 44: D279–D285.

Fisette JF, Toutant J, Dugré-Brisson S, Desgroseillers L, Chabot B (2010). hnRNP A1 and hnRNP H can collaborate to modulate 5' splice site selection. *RNA*, 16 (1): 228-238.

Fletcher S, Adkin CF, Meloni P, Wong B, Muntoni F, Kole R, Fragall C, Greer K, Johnsen R, Wilton SD (2012). Targeted Exon Skipping to Address “Leaky” Mutations in the Dystrophin Gene. *Molecular Therapy – Nucleic Acids*, 1 (10): e48.

Franke M, Bräulke T, Storch S (2013). Transport of the GlcNAc-1-phosphotransferase α/β -Subunit Precursor Protein to the Golgi apparatus Requires a Combinatorial Sorting Motif. *The Journal of Biological Chemistry*, 288 (2): 1238-1249.

Futerman AH, van Meer G (2004). The cell biology of lysosomal storage disorders. *Nature reviews. Molecular cell biology*, 5 (7): 554-565.

G

Gamazon ER, Stranger BE (2014). Genomics of alternative splicing: evolution, development and pathophysiology. *Human Genetics*, 133 (6): 679–687.

Graziewicz MA, Tarrant TK, Buckley B, Roberts J, Fulton L, Hansen H, Ørum H, Kole R, Sazani P (2008). An endogenous TNF- α antagonist induced by splice-switching oligonucleotides reduces inflammation in hepatitis and arthritis mouse models. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 16 (7): 1316-1322.

Grewal S, Shapiro E, Braunlin E, Charnas L, Krivit W, Orchard P, Peters C (2003). Continued neurocognitive development and prevention of cardiopulmonary complications after successful BMT for I-cell disease: a long-term follow-up report. *Bone Marrow Transplantation*, 32 (9): 957–960.

Gupta R, Jung E, Brunak S (2004). Prediction of N-glycosylation sites in human proteins. In preparation.

Gurav B, Srinivasan G (2017). Antisense oligonucleotides as therapeutics and their delivery. *Current Science*, 112 (3): 490-498.

H

Hammond SM, Wood MJA (2011). Genetic therapies for RNA mis-splicing diseases. *Trends in Genetics*, 27 (5): 196-205.

Harding PL, Fall AM, Honeyman K, Fletcher S, Wilton SD (2007). The influence of antisense oligonucleotide length on dystrophin exon skipping. *Molecular Therapy: the Journal of the American Society of Gene Therapy*, 15 (1): 157-166.

Havens MA, Duelli DM, Hastings ML (2013). Targeting RNA splicing for disease therapy. *Wiley Interdisciplinary Reviews: RNA*, 4 (3): 247-266.

Honey NK, Mueller OT, Little LE, Miller AL, Shows TB (1982). Mucopolipidosis III is genetically heterogeneous. *Proceedings of the National Academy of Sciences of the United States of America*, 79 (23): 7420-7424.

I

Imbert M, Dias-Florencio G, Goyenvallé A (2017). Viral Vector-Mediated Antisense Therapy for Genetic Diseases. *Genes*, 8 (2): 51.

J

Jensen CJ, Oldfield BJ, Rubio JP (2009). Splicing, *cis* genetic variation and disease. *Biochemical Society Transactions*, 37 (6): 1311-1315.

K

Kalbfuss B, Mabon SA, Misteli T (2001). Correction of alternative splicing of tau in frontotemporal dementia and parkinsonism linked to chromosome 17. *The Journal of Biological Chemistry*, 276 (46): 42986-42993.

Kang C, Riazuddin S, Mundorff J, Krasnewich D, Friedman P, Mullikin JC, Drayna D (2010). Mutations in the Lysosomal Enzyme-Targeting Pathway and Persistent Stuttering. *The New England Journal of Medicine*, 362 (8): 677–685.

Karras JG, McKay RA, Dean NM, Monia BP (2000). Deletion of Individual Exons and Induction of Soluble Murine Interleukin-5 Receptor- α Chain Expression through Antisense Oligonucleotide-Mediated Redirection of Pre-mRNA Splicing. *Molecular Pharmacology*, 58 (2): 380-387.

Kinali M, Arechavala-Gomeza V, Feng L, Cirak S, Hunt D, Adkin C, Guglieri M, Ashton E, Abbs S, Nihoyannopoulos P, Garralda ME, Rutherford M, McCulley C, Popplewell L, Graham IR, Dickson G, Wood MJ, Wells DJ, Wilton SD, Kole R, Straub V, Bushby K, Sewry C, Morgan JE, Muntoni F (2009). Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *The Lancet: Neurology*, 8 (10): 918-928.

Kollmann K, Pohl S, Marschner K, Encarnação M, Sakwa I, Tiede S, Poorthuis BJ, Lübke T, Müller-Loennies S, Storch S, Bräulke T (2010). Mannose phosphorylation in health and disease. *European Journal of Cell Biology*, 89 (1): 117-123.

Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nature Reviews Molecular Cell Biology*, 14 (3): 153-165.

Kornfeld S (1986). Trafficking of Lysosomal Enzymes in Normal and Disease States. *The Journal of Clinical Investigation*, 77 (1): 1-6.

Kornfeld S, Sly WS (2001). I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localisation. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. Childs B, Kinzler KW, Vogelstein B, assoc. eds. *The metabolic and molecular bases of inherited disease*, 8th ed. McGraw-Hill Education, New York, USA, pp 3421–3442.

Kudo M, Brem MS, Canfield WM (2006). Mucopolidosis II (I-Cell Disease) and Mucopolidosis IIIA (Classical Pseudo-Hurler Polydystrophy) Are Caused by Mutations in the GlcNAc-Phosphotransferase α/β -Subunits Precursor Gene. *American Journal of Human Genetics*, 78 (3): 451-463.

L

Leinekugel P, Michel S, Conzelmann E, Sandhoff K (1992). Quantitative correlation between the residual activity of β -hexosaminidase A and arylsulfatase A and the severity of the resulting lysosomal storage disease. *Human Genetics*, 88 (5): 513-523.

Lieberman AP, Puertollano R, Raben N, Slaugenhaupt S, Walkley SU, Ballabio A (2012). Autophagy in lysosomal storage disorders. *Autophagy*, 8 (5): 719-730.

Liou B, Haffey WD, Greis KD, Grabowski GA (2014). The LIMP-2/SCARB2 binding motif on acid β -glucosidase: basic and applied implications for Gaucher disease and associated neurodegenerative diseases. *The Journal of Biological Chemistry*, 289 (43): 30063-30074.

Liu H, He L, Tang L (2012). Alternative Splicing Regulation and Cell Lineage Differentiation. *Current Stem Cell Research & Therapy*, 7 (6): 400-406.

Liu L, Lee WS, Doray B, Kornfeld S (2017). Engineering of GlcNAc-1-Phosphotransferase for Production of Highly Phosphorylated Lysosomal Enzymes for Enzyme Replacement Therapy. *Molecular Therapy. Methods & Clinical Development*, 5: 59-65.

Love JE, Hayden EJ, Rohn TT (2015). Alternative Splicing in Alzheimer's Disease. *Journal of Parkinson's Disease and Alzheimer's Disease*, 2 (2): 6.

Lund TC, Cathey SS, Miller WP, Eapen M, Andreansky M, Dvorak CC, Davis JH, Dalal JD, Devine SM, Eames GM, Ferguson WS, Giller RH, He W, Kurtzberg J, Krance R, Katsanis E, Lewis VA, Sahdev I, Orchard PJ (2014). Outcomes after Hematopoietic Stem Cell Transplant for Children with I-Cell Disease. *Biology of Blood and Marrow Transplantation*, 20 (11): 1847-1851.

M

Malatack JJ, Consolini DM, Bayever E (2003). The Status of Hematopoietic Stem Cell Transplantation in Lysosomal Storage Disease. *Pediatric Neurology*, 29 (5): 391-403.

Mann CJ, Honeyman K, McClorey G, Fletcher S, Wilton SD (2002). Improved antisense oligonucleotide induced exon skipping in the mdx mouse model of muscular dystrophy. *The Journal of Gene Medicine*, 4 (6): 644–654.

Marschner K, Kollmann K, Schweizer M, Braulke T, Pohl S (2011). A key enzyme in the biogenesis of lysosomes is a protease that regulates cholesterol metabolism. *Science*, 333 (6038): 87-90.

Matera AG, Wang Z (2014). A day in the life of the spliceosome. *Nature reviews Molecular Cell Biology*, 15 (2): 108-121.

Matlin AJ, Clark F, Smith CWJ (2005). Understanding alternative splicing: towards a cellular code. *Nature reviews Molecular Cell Biology*, 6 (5): 386-398.

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, the Eteplirsen Study Group (2013). Eteplirsen for the Treatment of Duchenne Muscular Dystrophy. *Annals of Neurology*, 74 (5): 637-647.

Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK, Sigrist CJA, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD (2015). The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Research*, 43: D213–D221.

Moremen KW, Tiemeyer M, Nairn AV (2012). Vertebrate protein glycosylation: diversity, synthesis and function. *Nature reviews: Molecular Cell Biology*, 13 (7): 448-462.

Moreno PMD, Pêgo AP (2014). Therapeutic antisense oligonucleotides against cancer: hurdling to the clinic. *Frontiers in Chemistry*, 2: 87.

Muller-Loennies S, Galliciotti G, Kollmann K, Glatzel M, Braulke T (2010). A novel single-chain antibody fragment for detection of mannose 6-phosphate-containing proteins: application in mucopolipidosis type II patients and mice. *The American Journal of Pathology*, 177 (1): 240-247.

O

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. *Science Translational Medicine*, 3 (106): 106ra107.

Otomo T, Muramatsu T, Yorifuji T, Okuyama T, Nakabayashi H, Fukao T, Ohura T, Yoshino M, Tanaka A, Okamoto N, Inui K, Ozono K, Sakai N (2009). Mucopolipidosis II and III alpha/beta: mutation analysis of 40 Japanese patients showed genotype-phenotype correlation. *Journal of Human Genetics*, 54 (3): 145-151.

Otomo T, Higaki K, Nanba E, Ozono K, Sakai N (2011). Lysosomal storage causes cellular dysfunction in mucopolipidosis II skin fibroblasts. *The Journal of Biological Chemistry*, 286 (40): 35283-35290.

Ottesen E (2017). ISS-N1 makes the first FDA-approved drug for spinal muscular atrophy. *Translational Neuroscience*, 8 (1): 1-6.

P

Parenti G, Pignata C, Vajro P, Salerno M (2013). New strategies for the treatment of lysosomal storage diseases (review). *International Journal of Molecular Medicine*, 31 (1): 11-20.

Peacey E, Rodriguez L, Liu Y, Wolfe MS (2012). Targeting a pre-mRNA structure with bipartite antisense molecules modulates tau alternative splicing. *Nucleic Acids Research*, 40 (19): 9836-9849.

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. *European Journal of Human Genetics*, 12 (2): 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. *The Journal of Cell Biology*, 199 (5): 723-734.

Plante M, Claveau S, Lepage P, Lavoie EM, Brunet S, Roquis D, Morin C, Vézina H, Laprise C (2008). Mucopolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (*GNPTAB*) in a French Canadian founder population. *Clinical Genetics*, 73 (3): 236-244.

Pohl S, Marschner K, Thies B, Darnell LL, Bräulke T. *Proteolyse als regulatorisches Ereignis in der Pathophysiologie*. Poster presented at the IRTG-Symposium 2011. Hamburg, Germany.

Pontén F, Jirström K, Uhlen M (2008). The Human Protein Atlas — a tool for pathology. *Journal of Pathology*, 216 (4): 387-393.

Poorthuis BJHM, Wevers RA, Kleijer WJ, Groener JEM, de Jong JGN, van Weely S, Niezen-Koning KE, van Diggelen OP (1999). The frequency of lysosomal storage diseases in The Netherlands. *Human Genetics*, 105 (1-2): 151-156.

Poupětová H, Ledvinová J, Berná L, Dvořáková L, Kožich V, Elleder M (2010). The birth prevalence of lysosomal storage disorders in the Czech Republic: comparison with data in different populations. *Journal of Inherited Metabolic Disease*, 33 (4): 387-396.

Q

Qian Y, Lee I, Lee WS, Qian M, Kudo M, Canfield WM, Lobel P, Kornfeld S (2010). Functions of the α , β , and γ Subunits of UDP-GlcNAc:Lysosomal Enzyme N-Acetylglucosamine-1-phosphotransferase. *The Journal of Biological Chemistry*, 285 (5): 3360-3370.

Qian Y, Flanagan-Steet H, van Meel E, Steet R, Kornfeld SA (2013). The DMAP interaction domain of UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase is a substrate recognition module. *Proceedings of the National Academy of Sciences of the United States of America*, 110(25): 10246-10251.

Qian Y, van Meel E, Flanagan-Steet H, Yox A, Steet R, Kornfeld S (2015). Analysis of mucopolipidosis II/III GNPTAB missense mutations identifies domains of UDP-GlcNAc:lysosomal enzyme GlcNAc-1-phosphotransferase involved in catalytic function and lysosomal enzyme recognition. *The Journal of Biological Chemistry*, 290 (5): 3045-3056.

R

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 (mucopolipidosis IIIC). *The Journal of Clinical Investigation*, 105 (5): 673–681.

Reczek D, Schwake M, Schröder J, Hughes H, Blanz J, Jin X, Brondyk W, Van Patten S, Edmunds T, Saftig P (2007). LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell*, 131 (4): 770-783.

Reed R (1996). Initial splice-site recognition and pairing during pre-mRNA splicing. *Current Opinion in Genetics & Development*, 6 (2): 215-220.

Reese MG, Eeckman FH, Kulp D, Haussler D (1997). Improved splice site detection in Genie. *Journal of Computational Biology*, 4 (3): 311-323.

Reitman ML, Kornfeld S (1981a). UDP-N-acetylglucosamine:glycoprotein N-acetylglucosamine-1-phosphotransferase. Proposed enzyme for the phosphorylation of the high mannose oligosaccharide units of lysosomal enzymes. *The Journal of Biological Chemistry*, 256 (9): 4275-4281.

Reitman ML, Varki A, Kornfeld S (1981b). Fibroblasts from patients with I-cell disease and pseudo-Hurler polydystrophy are deficient in uridine 5'-diphosphate-N-acetylglucosamine: glycoprotein N-acetylglucosaminylphosphotransferase activity. *The Journal of Clinical Investigation*, 67 (5): 1574-1579.

Renshaw J, Orr RM, Walton MI, te Poele R, Williams RD, Wancewicz EV, Monia BP, Workman P, Pritchard-Jones K (2004). Disruption of WT1 gene expression and exon 5 splicing following cytotoxic drug treatment: Antisense down-regulation of exon 5 alters target gene expression and inhibits cell survival. *Molecular Cancer Therapeutics*, 3 (11): 1467-1484.

Reuter JS, Mathews DH (2010). RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, 11: 129.

Revil T, Gaffney D, Dias C, Majewski J, Jerome-Majewska LA (2010). Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics*, 11: 399.

Rigo F, Hua Y, Chun SJ, Prakash TP, Krainer AR, Bennett CF (2012). Synthetic oligonucleotides recruit ILF2/3 to RNA transcripts to modulate splicing. *Nature Chemical Biology*, 8: 555–561.

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. *Human mutation*, 30 (11): E993-E1001.

Roehr B (1998). Fomivirsen approved for CMV retinitis. *Journal of the International Association of Physicians in AIDS care*, 4 (10): 14-16.

Rowntree RK, Harris A (2003). The Phenotypic Consequences of *CFTR* Mutations. *Annals of Human Genetics*, 67 (Pt 5): 471-485.

Roy B, Haupt LM, Griffiths LR (2013). Alternative Splicing (AS) of Genes As An Approach for Generating Protein Complexity. *Current Genomics*, 14 (3): 182-194.

S

Sanchez-Irizarry C, Carpenter AC, Weng AP, Pear WS, Aster JC, Blacklow SC (2004). Notch subunit heterodimerization and prevention of ligand-independent proteolytic activation depend, respectively, on a novel domain and the LNR repeats. *Molecular and Cell Biology*, 24 (21): 9265-9273.

Sawkar AR, D'Haese W, Kelly JW (2006). Therapeutic strategies to ameliorate lysosomal storage disorders – a focus on Gaucher disease. *Cell and Molecular Life Sciences*, 63 (10): 1179-92.

Scaffidi P, Misteli T (2005). Reversal of the cellular phenotype in the premature aging disease Hutchinson-Gilford progeria syndrome. *Nature Medicine*, 11 (4): 440-445.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7: 539.

Siva K, Covello G, Denti MA (2014). Exon-Skipping Antisense Oligonucleotides to Correct Missplicing in Neurogenetic Diseases. *Nucleic Acid Therapeutics*, 24 (1): 69–86.

Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR (2006). An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Human Molecular Genetics*, 15 (16): 2490-2508.

Spellman R, Smith CWJ (2006). Novel modes of splicing repression by PTB. *Trends in Biochemical Sciences*, 31 (2): 73-76.

Sperisen P, Schmid CD, Bucher P, Zilian O (2005). Stealth proteins: *in silico* identification of a novel protein family rendering bacterial pathogens invisible to host immune defense. *PLoS Computational Biology*, 1 (6): e63.

Spiro RG (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, 12 (4): 43R-56R.

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.

Steet RA, Hullin R, Kudo M, Martinelli M, Bosshard NU, Schaffner T, Kornfeld S, Steinmann B. (2005). A splicing mutation in the alpha/beta GlcNAc-1-phosphotransferase gene results in an adult onset form of mucopolipidosis III associated with sensory neuropathy and cardiomyopathy. *American Journal of Medical Genetics*, 132A (4): 369-375.

Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, Hussain M, Phillips AD, Cooper DN (2017). The Human Gene Mutation Database: towards a comprehensive

repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Human Genetics*, 136 (6): 665-677.

Sveen A, Kilpinen S, Ruusulehto A, Lothe RA, Skotheim RI (2016). Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene*, 35 (19): 2413-2427.

T

The UniProt Consortium (2015). UniProt: a hub for protein information. *Nucleic Acids Research*, 43.

Tiede S, Storch S, Lübke T, Henrissat B, Bargal R, Raas-Rothschild A, Braulke T (2005). Mucopolidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase. *Nature Medicine*, 11 (10): 1109-1112.

V

van Deutekom JC, Janson AA, Ginjaar IB, Frankhuizen WS, Aartsma-Rus A, Bremmer-Bout M, den Dunnen JT, Koop K, van der Kooi AJ, Goemans NM, de Kimpe SJ, Ekhardt PF, Venneker EH, Platenburg GJ, Verschuuren JJ, van Ommen GJ (2007). Local Dystrophin Restoration with Antisense Oligonucleotide PRO051. *The New England Journal of Medicine*, 357 (26): 2677-2686.

van Meel E, Lee WS, Liu L, Qian Y, Doray B, Kornfeld S (2016). Multiple Domains of GlcNAc-1-phosphotransferase Mediate Recognition of Lysosomal Enzymes. *The Journal of Biological Chemistry*, 291(15): 8295-8307.

Velho RV, De Pace R, Klünder S, Sperb-Ludwig F, Lourenço CM, Schwartz IVD, Braulke T, Pohl S (2015). Analyses of disease-related GNPTAB mutations define a novel GlcNAc-1-phosphotransferase interaction domain and an alternative site-1 protease cleavage site. *Human molecular genetics*, 24 (12): 3497-3505.

Velho RV, De Pace R, Tidow H, Braulke T, Pohl S (2016). Identification of the interaction domains between α - and γ -subunits of GlcNAc-1-phosphotransferase. *FEBS Letters*, 590: 4287-4295.

Verhaart IEC, Aartsma-Rus A (2012). Gene therapy for Duchenne muscular dystrophy. *Current Opinion in Neurology*, 25 (5): 588–596.

Vickers TA, Zhang H, Graham MJ, Lemonidis KM, Zhao C, Dean NM (2006). Modification of MyD88 mRNA splicing and inhibition of IL-1 β signaling in cell culture and in mice with a 2'-O-methoxyethyl-modified oligonucleotide. *Journal of Immunology*, 176 (6): 3652-3661.

W

Wang Z, Burge CB (2008). Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA*, 14 (5): 802-813.

Warf MB, Berglund JA (2010). The role of RNA structure in regulating pre-mRNA splicing. *Trends in Biochemical Sciences*, 35 (3): 169-178.

Will CL, Lührmann R (2011). Spliceosome Structure and Function. *Cold Spring Harbor Perspectives in Biology*, 3 (7): a003707.

Williams T, Kole R (2006). Analysis of Prostate-Specific Membrane Antigen Splice Variants in LNCap Cells. *Oligonucleotides*, 16 (2): 186–195.

Wilton SD, Fall AM, Harding PL, McClorey G, Coleman C, Fletcher S (2007). Antisense oligonucleotide-induced exon skipping across the human dystrophin gene transcript. *Molecular Therapy*, 15 (7): 1288-1296.

Y

Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L, García Girón C, Gordon L, Hourlier T, Hunt SE, Janacek SH, Johnson N, Juettemann T, Keenan S, Lavidas I, J. Martin FJ, Maurel T, McLaren W, Murphy DN, Nag R, Nuhn M, Parker A, Patricio M, Pignatelli M, Rahtz M, Riat HS, Sheppard D, Taylor K, Thormann A, Vullo A, Wilder SP, Zadissa A, Birney E, Harrow J, Muffato M, Perry E, Ruffier M, Spudich G, Trevanion SJ, Cunningham F, Aken BL, Zerbino DR, Flicek P (2016). Ensembl 2016. *Nucleic Acids Research*, 44 (Database issue): D710–D716.

Yeo G, Burge CB (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of Computational Biology*, 11 (2-3): 377-394.

Z

Zammarchi F, Stanchina E, Bournazou E, Supakorndej T, Martires K, Riedel E, Corben AD, Bromberg JF, Cartegni L (2011). Antitumorigenic potential of STAT3 alternative splicing modulation. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (43): 17779–17784.

Zarghooni M, Dittakavi SS (2009). Molecular analysis of cell lines from patients with mucopolipidosis II and mucopolipidosis III. *American Journal of Medical Genetics. Part A*, 149A (12): 2753-2761.

Zuker M (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 31 (13), 3406-3415.

Appendix

Poster presented at the 21st ESGLD Workshop, Lyon, France (September 2017) and at the 13th Annual Meeting of the Oligonucleotide Therapeutics Society (OTS), Bordeaux, France (September 2017). **Recipient of Best Poster Award at the 13th OTS Annual Meeting**

Development of an antisense-mediated exon skipping approach as a therapeutic option for the ML II causing mutation c.3503_3504delTC

Liliana Matos¹, Regina Vilela^{1,2}, Francisca Coutinho¹, Paulo Gaspar³, Sandra Alves¹

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

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

³Newborn Screening, Metabolism and Genetics Unit, Department of Human Genetics, INSA, Porto, Portugal

Background

Mucopolipidosis type II is one of the most severe LSDs and is caused by the deficiency of GlcNAc-1-phosphotransferase activity, a hexameric enzyme which plays an essential role in the trafficking of lysosomal hydrolases to the lysosomes (Kornfeld & Sly, 2001). Despite the severity of this disorder, which often leads to death within the first decade of life, therapeutic approaches are scarce and there is no definitive cure currently available, with most treatments simply tackling the symptoms and not the genetic origin of ML II. Among the multiple mutations that can result in an ML II phenotype, the most frequent is c.3503_3504delTC, a dinucleotide deletion on exon 19 of the *GNPTAB* gene. In recent years, the RNA molecule became one of the most promising targets for therapeutic intervention in multiple diseases associated with genetic mutations. Some of the most notorious examples of this utility of the RNA molecule as a target for therapy are the recent developments in the field of Duchenne Muscular Dystrophy (DMD), where antisense oligonucleotides (AOs) have successfully been used for the skipping of specific exons and have been approved for clinical use (Aartsma-Rus & Krieg, 2017). Acknowledging this, the present study intends to design and develop an RNA-based therapeutic agent through the use of antisense oligonucleotides capable of inducing the skipping of exon 19 of the *GNPTAB* gene and consequently circumvent the effects the most common ML II causal mutation.

Methodology

1. Several bioinformatics tools were used to explore thoroughly i) exon 19 of the *GNPTAB* gene where is located the c.3503_3504delTC mutation, the target molecular defect in this study from the therapeutic perspective; ii) its flanking intronic regions and iii) the final protein product.

2. Skin fibroblasts of two distinct patients with an established ML II diagnosis, both homozygous for the c.3503_3504delTC mutation, were obtained from the Cell Line and DNA Biobank from Patients Affected by Genetic Diseases (Istituto Giannina Gaslini, Genova, Italy), while control cells were retrieved from healthy donors.

3. Design and selection of the most promising AOs were done using the guidelines and parameters previously described by Aartsma-Rus, 2012.

4. The fibroblasts were plated in P6 plates and transfected using Lipofectamine 2000 with the four selected antisense oligonucleotides at different concentrations. Fibroblasts from healthy individuals were also transfected as control experiments.

5. The effectiveness of this therapeutic approach regarding the restoration of GlcNAc-1-phosphotransferase activity was assessed indirectly through the quantification of some lysosomal enzymes.

Results

1. Sequence alignments of *GNPTAB* protein products

First, we aligned the amino acid sequence of the *GNPTAB* protein product minus exon 19 with the normal *GNPTAB* protein product and verified that removing exon 19 indeed produced an in-frame protein, 56 amino acids shorter than the regular protein product.

2. Prediction of protein domains within the *GNPTAB* protein product

Predictions obtained from three different databases (Pfam 30.0, UniProt and InterPro) confirmed the presence of several domains previously described for this protein product: the stealth domain, spread across four regions, two Notch repeats and a DMAP domain (Qian et al., 2015; Velho et al., 2015; van Meel et al., 2016).

3. Splicing modulation of *GNPTAB* exon 19 with antisense oligonucleotides in healthy control and ML II patient fibroblasts

After testing different concentrations, we can safely affirm that all four oligonucleotides tested were able to induce the skipping of this target exon to some extent. Among the selected AOs, H19A(+119+138), directed at a region of exonic enhancer elements within the second half of the target exon, clearly stood out as the most successful of the four.

4. Quantification of the activity of several lysosomal enzymes in fibroblasts

The assessment of enzyme activity for several lysosomal hydrolases following transfection with H19A(+119+138) aimed at the skipping of exon 19 of the *GNPTAB* gene revealed very small changes when compared to untreated fibroblasts.

Conclusion

Given the ultimate goal of this therapeutic approach – replacement of the natural occurring *GNPTAB* transcript harboring the c.3503_3504delTC mutation with the shorter mRNA missing exon 19 resultant from this exon-skipping strategy – , this oligonucleotide appears quite promising in this pursuit and should therefore be further explored. However, results showed little to no change in enzyme activity following transfection with AOs. Given the role of stealth domains for the catalytic activity of GlcNAc-1-phosphotransferase, the presence of a portion of stealth domain CR4 within the amino acid sequence encoded by exon 19 was expected to be an issue for this therapeutic approach. However, the presence of three other stealth domains in the α/β -precursor could allow for some enzyme activity after removing the CR4 domain. Indeed, based on these first results, it is not possible to conclude for sure that the activity of GlcNAc-1-phosphotransferase was not partially recovered, once we are using an indirect way of checking its status. In fact, ML II cannot be distinguished from ML III on the basis of residual lysosomal enzymes activities as they are similar in both disorders even though the activity of the GlcNAc-1-phosphotransferase is much more reduced in ML II and the clinical symptoms are much milder in ML III patients. In summary, while this work showed encouraging results at RNA level, further research is still needed regarding the activity of the resultant protein product.

Acknowledgments

This work was supported by FCT (project FCT/PTDC/BBB-BMD(6301/2014) and Asociación Nour de Mucopolipidosis.

References

- Aartsma-Rus A (2012). Overview on AON Design. In A Aartsma-Rus, eds, *Exon Skipping: Methods and Protocols*, Methods in Molecular Biology, vol. 867. Humana Press, New York, USA, pp 117-129.
- Aartsma-Rus A, Krieg AM (2017). FDA Approves Eteplirsen for Duchenne Muscular Dystrophy: The Next Chapter in the Eteplirsen Saga. *Nucleic Acid Therapeutics*, 27 (1): 1-3.
- Kornfeld S, Sly WS (2001). I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localisation. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. Childs B, Kinzler KW, Vogelstein B, assoc. eds. *The metabolic and molecular bases of inherited disease*, 8th ed. McGrawHill, New York, pp 3421–3442.
- Qian Y, van Meel E, Flanagan-Steet H, Yox A, Steet R, Kornfeld S (2015). Analysis of mucopolipidosis II/III GNPTAB missense mutations identifies domains of UDP-GlcNAc:lysosomal enzyme GlcNAc-1-phosphotransferase involved in catalytic function and lysosomal enzyme recognition. *The Journal of Biological Chemistry*, 290 (5): 3045-3056.
- van Meel E, Lee WS, Liu L, Qian Y, Doray B, Kornfeld S (2016). Multiple Domains of GlcNAc-1-phosphotransferase Mediate Recognition of Lysosomal Enzymes. *The Journal of Biological Chemistry*, 291 (15): 8295-8307.
- Velho RV, De Pace R, Klünder S, Sperb-Ludwig F, Lourenço CM, Schwartz IVD, Bräulke T, Pohl S (2015). Analyses of disease-related GNPTAB mutations define a novel GlcNAc-1-phosphotransferase interaction domain and an alternative site-1 protease cleavage site. *Human molecular genetics*, 24 (12): 3497-3505.