

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL



Ciências
ULisboa

**mRNA Metabolism: Nonsense Mediated mRNA Decay
as a Tool for Gene Therapy and the Role of Human
DIS3L2 in Transcript Degradation**

Mestrado em Biologia Humana e Ambiente

Gerson Leonel Asper Amaral

Dissertação orientada por:
Doutora Luísa Romão
Professora Doutora Deodália Dias

2016

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL



**mRNA Metabolism: Nonsense Mediated mRNA Decay
as a Tool for Gene Therapy and the Role of Human
DIS3L2 in Transcript Degradation**

Mestrado em Biologia Humana e Ambiente

Gerson Leonel Asper Amaral

Dissertação orientada por:
Doutora Luísa Romão (Instituto Nacional de Saúde Dr. Ricardo Jorge)
Professora Doutora Deodália Dias (Departamento de Biologia Animal,
Faculdade de Ciências da Universidade de Lisboa)

2016

“It is finished.” – Jesus Christ (The Bible, *John 19:30*)

ACKNOWLEDGEMENTS

This dissertation is the result of the very hard work, patience and resources from a lot of people. They were instrumental in the accomplishment of this project, be it through their knowledge, plain lab work, their friendship, guidance, support or sheer trust. I am sincerely thankful that all of you made part of my life at least for this year, because without you this would never see the light of day and would remain in the darkness of night. Clichéd poetry aside, honestly, thank you all.

I want to start by thanking my main advisor, Dr. Luísa Romão, for accepting me into her brilliant lab and trusting me and my work. Thank you for sharing your vast knowledge with me, helping me, guiding me, being patient and calling my attention to my mistakes, all this without ever stopping from being pleasant! I feel so honoured and thankful.

I thank my internal advisor, Dr. Deodália Dias, for being a fantastic human being and the best guide I could ever hope for. Thank you so much for your time, savvy advice, professionalism and remarkable support and understanding. You were much more than an advisor and I was truly blessed. I'm deeply grateful.

I also want to thank Dr. João Lavinha and Dr. Glória Dias for allowing me to carry out this study at Departamento de Genética Humana from Instituto Nacional de Saúde Dr. Ricardo Jorge.

I thank Juliane Menezes and Rafaela Lacerda for their close guidance, expertise and insight into all of my problems in the lab. You were always so willing to help me and so patient, without you I couldn't have done this. To say I learned a lot from you is a harsh understatement. You were the keys. I thank Paulo for his help in these last few months, our chats and your experience have guided me in my work and without them I'd probably be in trouble. I also thank my other lab mates, Nuno and Cláudia. Nuno for his companionship and help throughout this; Cláudia for her sympathy and for her very wise advice. It's all of you together that make our lab the best one. Thank you for this experience.

I want to thank all the members of the Oncobiologia laboratory for your patience with us! Whenever I needed something for my work, be it materials or advice, you were always there, and with a friendly smile on your faces even though we were kind of 'bankrupting' you. You're also the best.

I thank all of my friends; you are the family I chose. But especially Sandra, Filomena and Joana Neno, you were wonderful. Among the three of you, thank you for the silly funny moments that helped more than you think, the rides and all the fantastic support you gave

me, especially in the stressful final phase! Also, I thank Sílvia, Ricardo and the girls next floor in the lab, Mariana and Joana, with whom I lost my breath laughing and lived many unexpected episodes several times. What a year! There are others, who didn't directly contribute to this work, but unfortunately I can't mention you all, you know who you are. I'd also like to thank a very special friend, Sofia Gouveia, who taught me so much but above all offered me, without asking anything in return, her unwavering support, kindness and deep understanding, in a not so easy period of my life; I'm so fortunate for having you.

Lastly, I thank my mother, my aunt and my maternal grandparents. My aunt for her unconditional support; no matter what I do, even if you don't understand it, I can always count on you to support me. My mother and grandparents for their sacrifices and patience. I thank all of you for your never-ending trust in me.

Mother, you taught me by example how to work hard every day, not to give up, and fight for my future. These and other lessons have pushed me through this process in the hope that I'm worthy enough of your sacrifices.

As Mother Teresa once said, kind words are short and easy to speak, but their echoes are truly endless.

Thank you all for believing in me.
I hope I can make you proud.

INDEX

ACKNOWLEDGEMENTS	VII
FIGURES	XI
ABBREVIATIONS	XII
ABSTRACT	XV
RESUMO	XVII
1. INTRODUCTION	1
1.1. Gene Expression in Eukaryotes	1
1.2. mRNA Translation	1
1.2.1. Translation Initiation	2
1.2.2. Translation Elongation.....	4
1.2.3. Translation Termination.....	5
1.3. Nonsense Mediated mRNA Decay	6
1.3.1. PTC Recognition and NMD Eliciting.....	6
1.3.2. mRNA Degradation via NMD.....	8
1.3.3. NMD Targets.....	10
1.3.4. NMD and Human Disease	10
1.3.5. β -thalassemia as a Model	11
1.3.6. Nonsense Suppression Therapy.....	11
1.4. Nonstop mRNA Decay	13
1.5. Eukaryotic Exosome	14
1.6. Human Dis3L2.....	15
2. AIMS	17
2.1. Chapter I.....	17
2.2. Chapter II.....	17
3. MATERIALS AND METHODS	19
3.1. Plasmid Constructs.....	19
3.2. Cell Culture and Plasmid Transfections	19
3.3. Knockdown Experiments.....	19
3.4. RNA isolation	20
3.5. Reverse transcription-coupled quantitative PCR (RT-qPCR).....	20

3.6. SDS-PAGE and Western Blotting.....	20
3.7. SOEing PCR.....	21
3.8. Digestions and Ligation Reactions.....	21
3.9. Statistical Analysis.....	22
CHAPTER I.....	23
4. RESULTS & DISCUSSION.....	25
4.1. Human β -globin protein wasn't detectable by western blot analysis.....	25
4.2. SOEing PCR was not successful enough as an approach to clone a FLAG-tag into the C-terminal region of Human β -globin.....	26
5. FINAL CONSIDERATIONS & FUTURE DIRECTIONS.....	31
CHAPTER II.....	33
4. RESULTS & DISCUSSION.....	35
4.1. DIS3L2 knockdown correlates with increased levels of β -globin mRNAs independently from their decay mechanism.....	35
4.2. hDIS3L2 knockdown correlates with increased <i>HFE</i> mRNA levels.....	36
5. FINAL CONSIDERATIONS & FUTURE DIRECTIONS.....	38
6. REFERENCES.....	39
ANNEXES.....	47
ANNEX I.....	48

FIGURES

Figure 1 – Canonical model of cap-dependent translation initiation.	3
Figure 2 – Translation Elongation.....	5
Figure 3 – Translation Termination.	6
Figure 4 - Simplified model of mammalian EJC-dependent NMD eliciting.....	8
Figure 5 – Simplified decay pathway of human NMD	9
Figure 6 – Suppression of translation termination.....	12
Figure 7 – Eukaryotic exosome structures.	15
Figure 8 – Model for Dis3 domains’ structure.....	15
Figure 9 – mRNA degradation pathways in the cytoplasm of <i>S. pombe</i>	16
Figure 10 – Western blot analysis of cell lysates.....	25
Figure 11 – Illustrations for SOEing PCR 1, 2 and 3.....	27
Figure 12 – Model for SOEing 4 PCR and Ligation to pTRE vectors.....	28
Figure 13 – Gel Electrophoresis of the SOEing PCR reactions products.	29
Figure 14 – Gel electrophoresis of the colony PCR products	30
Figure 15 – RT-qPCR analysis of Dis3L2 knockdown experiments performed for the Human β -globin gene	35
Figure 16 – RT-qPCR analysis of Dis3L2 or UPF1 knockdown experiments performed for the <i>HFE</i> gene.....	36

ABBREVIATIONS

A: Adenine

A-site: aminoacyl-site

Bp: base pairs

CCR4: Chemokine (C-C Motif) receptor 4

cdNA: complementary deoxyribonucleic acid

C-terminal: carboxyl-terminal

CD: (any) codon

CFTR: Cystic Fibrosis Transmembrane Conductance Regulator

CSD: cold shock domain

DCP: decapping protein

DEAD: Asp-Glu-Ala-Asp motif

DECID: decay inducing complex

Dis3L: DIS3-Like Exosome 3'-5' Exoribonucleases

DMEM: Dulbecco's modified Eagle medium

DNA: deoxyribonucleic acid

eEF: eukaryotic translation elongation factors

eIF: eukaryotic initiation factors

EJC: exon junction complex

eRF: eukaryotic translation release factor

E-site: exit-site

FBS: foetal bovine serum

G: guanine

GDP: guanosine diphosphate

GTP: guanosine triphosphate

HBB: Haemoglobin beta chain

HeLa: human cervical cancer cell line

HFE: human hemochromatosis protein

MAGOH: mago-nashi homolog

Met: methionine

Met-tRNA_i^{Met}: methionine-loaded initiator transfer ribonucleic acid

min: minutes

miRNA: micro ribonucleic acid

mRNA: messenger ribonucleic acid

mRNP: messenger ribonucleoprotein particle

NGD: no-go mRNA decay

NMD: nonsense-mediated mRNA decay

NOT: negative regulator of transcription

NSD: nonstop mRNA decay

nt: (any) nucleotide

N-terminal: amino-terminus

ORF: open reading frame

PABP: poly(A)-binding protein

PABPC1: cytoplasmic poly(A)-binding protein 1

PAGE: polyacrylamide gel electrophoresis

PAN: poly(A) nuclease

P-bodies: Processing bodies

PBS: phosphate buffered saline

PCR: polymerase chain reaction

PIC: pre-initiation complex

PIN: PiIT N terminus

PLB: passive lysis buffer

PNPase: Polynucleotide phosphorylase

Poly(A): poly-adenilate

PP2A: protein phosphatase 2A

Pre-mRNA: messenger ribonucleic acid precursor

P-site: peptidyl-site

PTC: premature translation termination codon

PVDF: polyvinylidene difluoride

RNA: ribonucleic acid

RNB: Ribonuclease domain

RNase: ribonuclease

Rrp: Ribosomal RNA processing protein

RT: reverse transcription

RT-qPCR: reverse transcription quantitative polymerase chain reaction

SDS: sodium dodecyl sulphate

sec: seconds

siRNA: small interfering ribonucleic acid

SMG: suppressor of morphological defects on genitalia

SURF: SMG1-UPF1-eRFs complex

T: thymine

TBS: tris-buffered saline

tRNA: transfer ribonucleic acid

U: uracil

uORF: upstream open reading frame

UPF: up-frameshift protein

UTR: untranslated region

WT: wild type

XRN1: 5'-3' exoribonuclease 1

β15: human β-globin transcript with a PTC at codon 15

β26: human β-globin transcript with a PTC at codon 26

β39: human β-globin transcript with a PTC at codon 39

βNS: human β-globin transcript without an in-frame stop codon

βWT: normal/wild type human β-globin transcript

ABSTRACT

Eukaryotic cells have developed elaborate mechanisms of mRNA quality control that secure gene expression fidelity through the detection and degradation of abnormal transcripts. NMD (nonsense-mediated mRNA decay), which detects and degrades transcripts containing premature translation termination codons (PTCs), and NSD (nonstop mRNA decay), that detects and degrades transcripts without in-frame stop codons, are just two examples. Nonsense-mediated mRNA decay (NMD) in particular, is a conserved surveillance system in all eukaryotic cells and is also the most extensively studied. PTC-containing mRNAs could, without NMD, give rise to C-terminally truncated proteins toxic for the cell. The physiological importance of NMD is further manifested by the fact that about one third of genetic disease-associated mutations generate PTCs. Recently, some studies have shown that aminoglycosides, low molecular weight compounds, and non-aminoglycosides can suppress PTCs in cystic fibrosis, Duchenne's muscular dystrophy others, as a novel therapeutic approach, suppression therapy, which uses these compounds to induce recoding of a PTC into a sense codon. It is unclear whether β -thalassaemia would also be responsive to suppression therapy. Some recent studies show positive results for the compound PTC124 in suppressing nonsense mutations in the CFTR gene and others; also preliminary results obtained in our lab have shown that the aminoglycoside G418 can suppress a nonsense mutation at codon 39 of the human β -globin mRNA, although at low levels in cultured erythroid cells. As a first part of this work, we decided to investigate if suppression therapy can restore enough β -globin protein to correct the disease manifestations of β -thalassaemia. We intended to test whether G418 and/or PTC124 were able to induce efficient levels of suppression in a dose-dependent manner in HeLa cells transfected with plasmids containing the human β -globin wild type gene (β WT) or the other variants carrying a nonsense mutation at codon 15 (β 15) or 39 (β 39). However, we weren't successful in this approach due to difficulties in gene cloning. The next step for RNAs targeted by NMD or NSD, as well as normal transcripts, which don't accumulate indefinitely, is degradation. Generally the exosome complex, a multi-subunit ribonuclease complex, is responsible for the 3'-5' degradation of every type of RNA in the cell, with its main catalytic component being either Dis3 or Dis3L1 in humans. However, another ribonuclease has been identified: DIS3L2. This protein is thought to be a

cytoplasmic exosome-independent 3'-5' ribonuclease, with special affinity for uridylylated transcripts. Nonetheless, not much else is known for certain about its activity, especially in humans, including if it is coupled to NMD or NSD mRNA degradation. As a consequence, we intended to evaluate hDIS3L2's possible involvement in mRNA degradation pathways, by performing knockdown of hDIS3L2, with siRNAs (small-interfering RNAs) in HeLa cells transfected with plasmids containing the human β -globin wild type gene (β WT), the variants carrying a nonsense mutation at codon 15 (β 15), 26 (β 26) or 39 (β 39), and also a variant lacking an in-frame stop codon (nonstop) (β NS). We then evaluated the human β -globin mRNA levels, as well as *HFE*'s, which is an NMD natural target. Our results show that human DIS3L2 is involved in NMD, NSD and possibly normal transcript degradation (mRNA turnover).

Key-words: Nonsense-mediated mRNA decay (NMD); Nonstop mRNA Decay (NSD); G418; PTC124; hDIS3L2.

RESUMO

A expressão génica nos eucariotas envolve uma série de passos interligados e acoplados entre si, tendo a molécula de RNA (*ribonucleic acid*) como mensageiro entre os grandes passos. Resumidamente, a mensagem codificada pelas bases nucleotídicas do ácido desoxirribonucleico (DNA) (*deoxyribonucleic acid*) é transferida para uma molécula de RNA (transcrição), que, após processamento no núcleo, é transferida para o citoplasma onde é lida e transformada numa cadeia polipeptídica (tradução). Por vezes, contudo, podem ocorrer erros, em qualquer uma das fases da expressão, erros esses que podem resultar em mRNAs aberrantes que, se forem traduzidos, podem dar origem a proteínas truncadas com possíveis efeito deletérios. Para contornar este problema, as células eucarióticas desenvolveram mecanismos de controlo de qualidade do mRNA de modo a assegurarem a fidelidade da expressão génica através da detecção e degradação de transcritos aberrantes. O decaimento do mRNA mediado por mutações nonsense (NMD; *Nonsense-mediated mRNA decay*) é o mais conhecido, detecta e degrada transcritos que contêm codões de terminação prematuros (CTPs). O decaimento *nonstop* (NSD; *Nonstop mRNA decay*) detecta e degrada transcritos que não possuem codões de terminação em fase na grelha de leitura, mas existem outros, como o NGD (*No-go decay*) que evoluiu para lidar com os transcritos que possuem uma qualquer mutação que impeça a normal elongação da tradução. O NMD em particular, é um mecanismo de vigilância conservado em todas as células eucarióticas e é também o mais estudado. Os mRNAs que contêm CTPs poderiam dar origem, sem o NMD, a proteínas truncadas na extremidade C-terminal tóxicas para a célula, que podem adquirir um ganho de função prejudicial ou um efeito dominante-negativo. A importância fisiológica do NMD é ainda adicionalmente demonstrada pelo facto de que cerca de um terço das doenças genéticas associadas a doenças gerarem CTPs. Recentemente, alguns estudos têm vindo a demonstrar que compostos de baixo peso molecular, aminoglicósidos e não-aminoglicósidos podem suprimir CTPs em contexto de fibrose cística, distrofia muscular de Duchenne e deficiência em carnitina palmitoltransferase 1A como uma nova abordagem terapêutica, a terapia de supressão, a qual usa estes compostos para induzir a recodificação de um codão *nonsense* num codão *sense*. Uma doença também associada a mutações *nonsense* é a β -talassémia. A β -talassémia é uma das doenças genéticas mais comuns no mundo e cujo tratamento para

os fenótipos mais agressivos requer inevitavelmente transfusões regulares sanguíneas, com os riscos que isso acarreta como a acumulação excessiva de ferro no organismo. Uma cura que se possa chamar definitiva ainda não existe e, portanto, qualquer nova abordagem terapêutica constitui uma mais valia. O NMD é um modelador do fenótipo da β -talassémia, podendo contribuir para o melhoramento das manifestações da doença. Em relação à terapia de supressão, permanece ainda por esclarecer se a β -talassémia responderia também à mesma. Alguns estudos recentes mostram resultados positivos para o composto PTC124 (ou Ataluren) na supressão de mutações *nonsense* no gene da CFTR (associadas à Fibrose Cística) bem como noutros genes, associados à Distorfia Muscular de Duchenne, Síndrome de Usher e Deficiência em Carnitina Palmitoltransferase 1C; para além disso, resultados obtidos previamente pelo nosso laboratório demonstraram que o aminoglicósido G418 pode suprimir uma mutação *nonsense* no codão 39 do mRNA do gene da β -globina humana, embora a baixos níveis em células eritróides em cultura. Tendo em conta esta informação, como uma primeira parte deste trabalho, decidimos investigar se a terapia de supressão pode restaurar β -globina suficiente para conseguir corrigir as manifestações da β -talassémia. Propusémo-nos a testar se o G418 e/ou o PTC124 seriam capazes de induzir níveis suficientes de supressão, duma maneira dependente da dose, em células HeLa transfectadas com plasmídeos que contêm o gene da β -globina humana, variante selvagem, *wild type* (β WT), ou as outras variantes contendo uma mutação *nonsense* no codão 15 (β 15) ou 39 (β 39). Contudo, não fomos bem sucedidos nesta primeira parte do projecto devido a dificuldades intransponíveis associadas à clonagem de genes.

O próximo passo para os RNAs sinalizados pelas maquinarias do NMD ou do NSD, bem como os RNAs normais, que não se acumulam indefinidamente, é a degradação. A degradação é, na verdade, uma parte essencial do metabolismo do mRNA, constituindo até um mecanismo pelo qual a expressão génica é regulada. Na degradação do mRNA, o exossoma é um componente essencial. O exossoma eucariótico é um complexo ribonucleolítico multi-subunidade, e é responsável pela degradação na direcção 3'-5' de todos os tipos de RNA na célula, entre outras coisas. As proteínas humanas Dis3 ou Dis3L1 são os seus elementos catalíticos. Contudo, outra ribonuclease foi também identificada, a Dis3L2, da qual pouco ainda se sabe, especialmente em humanos. Pensa-se que esta proteína seja uma ribonuclease citoplasmática na direcção 3'-5', independente do exossoma, que parece ter especial afinidade para transcritos uridilados. Tendo em conta

as lacunas no conhecimento acerca da Dis3L2 humana (hDis3L2) e da importância dos mecanismos de vigilância do mRNA e da sua degradação, decidimos avaliar a possibilidade da Dis3L2 humana estar efectivamente envolvida nas vias de degradação do mRNA NMD e/ou NSD. Para esse efeito, foram efectuadas experiências de silenciamento do gene da Dis3L2 humana recorrendo à tecnologia de RNA de interferência, mais especificamente siRNAs (*small interfering RNAs*) em células HeLa transfectadas com plasmídeos contendo a variante selvagem do gene da β -globina humana (β WT), com as variantes contendo uma mutação *nonsense* no codão 15 (β 15), 26 (β 26) ou 39 (β 39), e também a variante que não contém um codão de terminação na grelha de leitura, *nonstop*, (β NS). Seguidamente avaliámos os níveis celulares de mRNA da β -globina humana, bem como os do *HFE*, que é um alvo natural do NMD. Os nossos resultados sugerem que a hDis3L2 está envolvida no NMD, NSD e possivelmente também na degradação de transcritos normais e consequentemente no *turnover* do RNA.

Palavras-chave: Decaimento do mRNA mediado por mutações *nonsense* (NMD); Decaimento do mRNA *nonstop* (NSD); G418; PTC124; hDIS3L2.

1. INTRODUCTION

1.1. Gene Expression in Eukaryotes

The central dogma of molecular biology states that protein synthesis starts with transcription of a DNA template into a complementary messenger RNA molecule (mRNA) in the nucleus of the eukaryotic cell, which is then transported to the cytoplasm where it will be translated into protein by the ribosomes¹. This seemingly straightforward mechanism involves, in reality, a series of interconnected steps and a complex coupled network of machineries that guarantees the efficiency and specificity of gene expression, with messenger RNAs being the key intermediaries throughout the whole process^{2,3}.

Briefly, the DNA is transcribed into a precursor mRNA (pre-mRNA) molecule by the ribosome, to which a 5'cap and a 3'poly(A) tail will be added mainly for protection against degradation. Also this pre-mRNA molecule will be subject to splicing, which consists in removing introns and joining together the exons. By the end of this stage we have a mature mRNA molecule. This mRNA is transported to the cytoplasm where it will finally be translated into protein and eventually degraded^{3,4}.

However, sometimes errors occur, at any step of gene expression, and aberrant mRNAs are produced⁴. If translated, these result in aberrant proteins that can be detrimental to the cell⁵. To deal with this problem, evolution came up with mRNA surveillance mechanisms. These mechanisms target and degrade aberrant transcripts before they can be translated thus avoiding the possible deleterious effects of aberrant proteins. There are three main mRNA surveillance pathways, although they are not the only ones, which are: Nonsense-Mediated Decay (NMD), Nonstop Decay (NSD) and No-Go Decay (NGD); they target and degrade, respectively, transcripts containing a premature termination codon (PTC), mRNAs lacking a termination codon and mRNAs containing any mutation that stalls the ribosome and impedes normal translation elongation⁶. Of these three pathways, NMD and NSD are the ones relevant for this study.

1.2. mRNA Translation

One of the most important steps in mRNA life and gene expression is the translation of messenger RNA into protein. Like virtually all the steps in gene expression, mRNA

translation is a tightly regulated process, aiding the fine tuning of eukaryotic gene regulation⁷. In order to better understand the translation-dependent mechanisms of NMD and NSD, as well as gene therapy itself, it is useful to get familiar with the basic concepts of mRNA translation.

Translation is conceptually divided into three stages (initiation, elongation and termination) where the codon sequence in the mRNA directs the synthesis of a polypeptide chain, with ribosomes, which are two-subunit ribonucleoproteins, being the key players in the reaction^{8,9,10}. The three stages are explained below.

1.2.1. Translation Initiation

In the canonical, cap-dependent, mechanism of translation, the interaction of the ternary complex with the 40S ribosomal subunit marks the beginning. This complex is formed by the eukaryotic initiation factor eIF2 bound to both the methionyl-initiator tRNA (Met-tRNA_i^{Met}) and GTP¹¹. When the ternary complex, along with eIF1, 1A, 3 and 5, binds to the 40S ribosomal subunit, the 43S pre-initiation complex (PIC) is formed^{7,11,12,13}. This complex, according to Kozak's scanning model has the task of scanning the 5'UTR (5' untranslated region) until it finds the AUG initiation codon, where translation starts¹⁴. But before that happens, another complex must be present: eIF4F. The eukaryotic initiation factor 4F recognizes the 5'-cap proximal region of mRNA and its functions are illustrated by each of its three comprising elements: the ATP-dependent DEAD-box RNA-helicase eIF4A, which resolves secondary structures in the 5'UTR, making the scanning process possible; eIF4E which is the cap-binding element; and eIF4G that acts as a scaffold protein for the interaction of the other two and also interacts with the poly(A)-binding protein (PABP) and by doing so, circularizes the mRNP (messenger ribonucleoprotein) molecule, because PABP is bound to the 3'-poly(A)-tail^{15,16,17}.

When both complexes, eIF4F and PIC, are present and bound to their respective targets, scanning, by PIC, of the 5'UTR in the 5'-3' direction takes place until it finds an AUG codon in an optimum context, which is the start codon. When this happens, complementary base pairing takes place between the start codon, which by now is at the peptidyl-tRNA (P) site of the 40S subunit, and the anticodon present in the Met-tRNA_i^{Met}. Then, eIF5 activates GTPases, which hydrolyze the ternary complexes' GTP bound to the eIF2 and this triggers the displacement of the initiation factors, eventually leading to the joining of the 60S

ribosomal subunit, catalyzed by eIF5B. The 80S eukaryotic ribosome is formed and the translation initiation phase has ended^{18,19,20}. The entire process is pictured in figure 1.

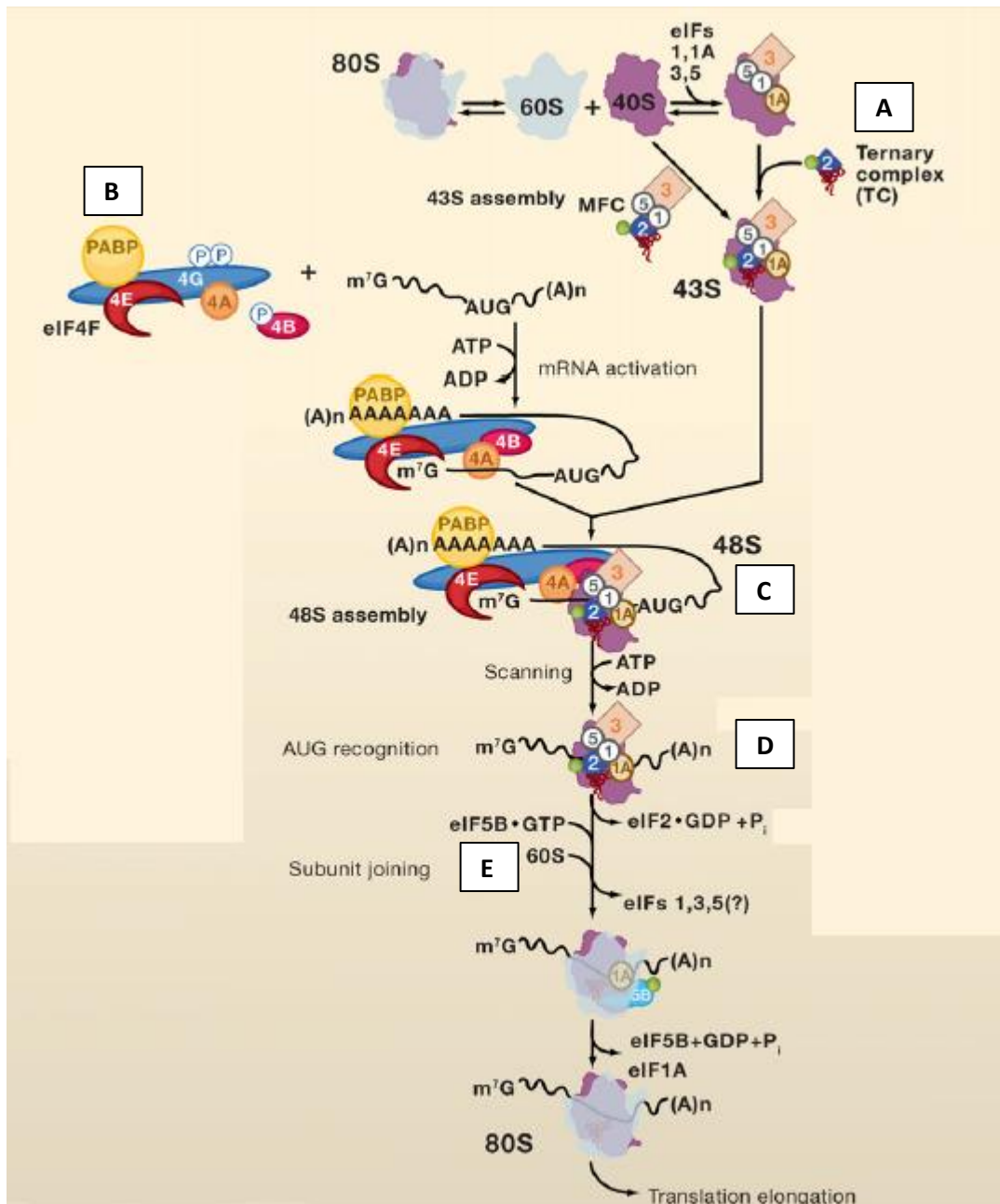


Figure 1 – Canonical model of cap-dependent translation initiation. **A:** The ternary Complex (TC) binds to the 40S ribosomal subunit, along with eIF1, 1A, 3 and 5 forming the 43S pre-initiation complex (PIC). **B:** Also, eIF4F recognizes the 5'-cap proximal region of mRNA and interacts with PABP via the eIF4G element. **C:** When both eIF4F and PIC are bound to the mRNP, PIC scans the 5'UTR in the 5'-3' direction until it finds the AUG initiation codon. **D:** The start codon is recognized, complementary base pairing takes place between the nucleotide elements of the Met-tRNA^{Met} and the mRNP. Then, eIF5 activates GTPases, which hydrolyze the ternary complexes' GTP bound to the eIF2, triggering the displacement of the initiation factors. **E:** Joining of the 60S ribosomal subunit to the 40S takes place, catalyzed by eIF5B. The 80S eukaryotic ribosome is formed and the translation initiation phase is complete. Adapted from Sonenberg & Hinnebusch (2009).

1.2.2. Translation Elongation

This stage consists of the sequential decoding of the mRNA nucleotide sequence, in the 5' - 3' direction, into a polypeptide chain^{21,22}; the order of the mRNA codons specifies the order in which amino acids are added²³.

By the end of initiation, we are left with the initiator aminoacyl-tRNA located at the P-site of the ribosome, base-paired with the start codon, and the adjacent A-site empty and ready to receive a cognate aminoacyl-tRNA, thereby continuing the decoding process of mRNA into protein^{10,11,18,24}. In this sequential process, the eukaryotic elongation factor eEF1A has the task of bringing the aminoacyl-tRNAs to the ribosome A-site: if codon and anticodon are complementary, then eEF1A activates GTP hydrolysis, which in turn induces eEF1A release, thereby permitting the entry of aminoacyl-tRNA to the A site²⁴. Then, the two tRNAs present at the moment at the ribosome are translocated, thanks to the aid of the GTP-dependent eEF2: the deacylated tRNA is translocated from the P to the E-site (E from "exit") and the peptidyl-tRNA from the A to the P-site (figure 2) and the ribosome moves 3 nucleotides in the 5'-3' direction²³. Also eEF1A interacts with eEF1B, which promotes the exchange of GDP for GTP, recycling the GTP-bound eEF1A²³. After translocation, the A-site is once again empty and ready for the next aminoacyl-tRNA. This process goes on until the ribosome reaches a stop codon.

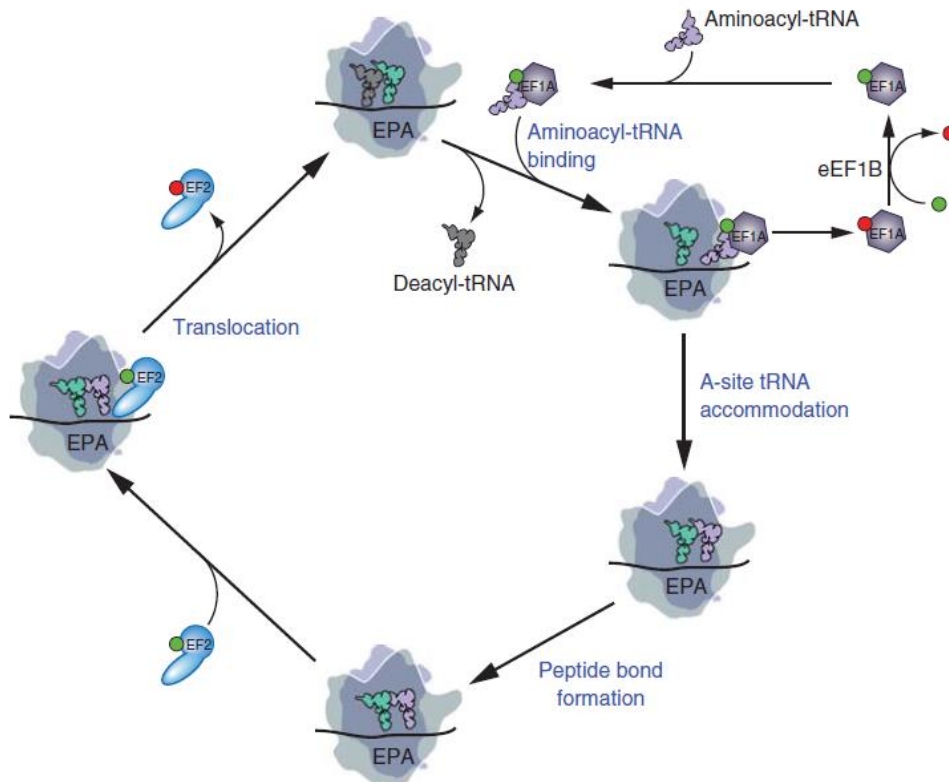


Figure 2 – Translation Elongation. Starting at the top: the elongation factor eEF1A brings the aminoacyl-tRNA to the ribosomal A-site, where complementary base pairing occurs. Then eEF1A activates GTP hydrolysis, inducing eEF1A release, which allows the entry of the aminoacyl-tRNA to the A-site. GDP is recycled to eEF1A-GTP by the exchange factor eEF1B. After the tRNA accommodation into the A-site, peptide bond formation occurs. Then, the two tRNAs present at the ribosome are translocated, by binding of the GTP-dependent eEF2. The ribosome is now ready for the next cycle of elongation with release of the deacylated tRNA from the E-site and binding of the appropriate eEF1A-GTP-aminoacyl-tRNA to the A-site. In this figure GTP is depicted as a green circle and GDP as a red one. *In Dever & Green (2012).*

1.2.3. Translation Termination

Translation termination takes place when the ribosome recognizes a stop codon, making the whole translation process come to an end and the nascent protein is released²⁵. It all starts when a stop codon, either UAG, UGA or UAA, is present at the A-site. Because there is no cognate aminoacyl-tRNA for a stop codon, the ribosome stalls and eRF1 recognizes the stop codon instead and also forms a complex with the GTPase eRF3, which causes GTP hydrolysis. This hydrolysis, in turn, induces a conformational change in eRF1 that allows it to move closer to the P-site and promote cleavage of the peptidyl-tRNA ester bond, finally leading to the release of the newly synthesized polypeptide chain^{23,26,27} and both ribosome subunits dissociate, ready to be recycled^{23,24,28}.

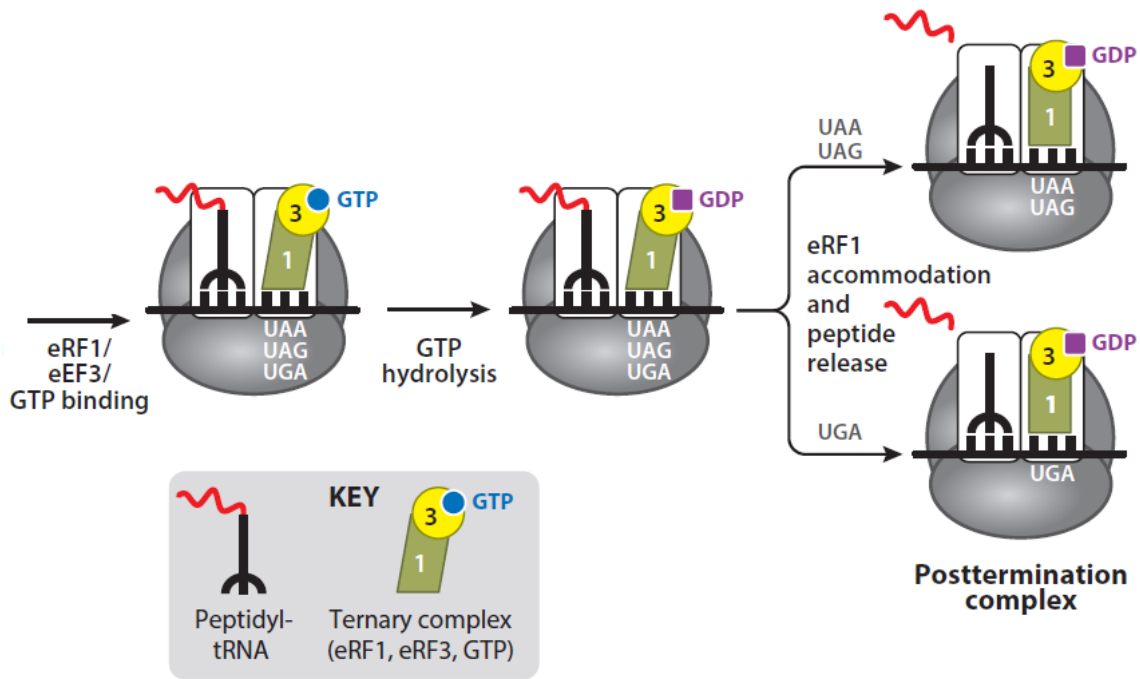


Figure 3 – Translation Termination. From left to right, when a stop codon, UAG, UGA or UAA, is present at the A-site, the ribosome stalls and eRF1, complexed with the GTPase eRF3, recognizes the stop codon, which causes GTP hydrolysis. This hydrolysis induces a conformational change in eRF1 that allows it to move closer to the P-site and promote cleavage of the peptidyl-tRNA ester bond, leading to the release of the newly synthesized polypeptide chain. Adapted from Keelint *et al.* (2014).

1.3. Nonsense Mediated mRNA Decay

NMD is a post-translational surveillance mechanism that is evolutionary conserved in eukaryotic cells. It is also the most studied quality control mechanism in eukaryotes⁶. Briefly, NMD function is to identify and eliminate mRNAs that contain premature translation termination codons (PTCs), also known as nonsense codons, thus preventing the synthesis and accumulation of C-terminally truncated proteins that can have a dominant-negative or gain-of-function effect potentially deleterious for the cell^{4,26,29,30,31}. It is important to note, however, that in mammals there are other quality control mechanisms, such as NSD (nonstop mRNA decay), which we will also approach in this work, and NGD (no-go mRNA decay)^{8,6}. In order to understand the type of gene therapy that we tried to achieve in this work, which was one where we used the knowledge of the NMD mechanism to our advantage, one must first comprehend NMD itself.

1.3.1. PTC Recognition and NMD Eliciting

Generally, PTCs can arise from nonsense or frameshift mutations in the DNA sequence, transcription errors, anomalous splicing events, among others^{31,32,33}. When these PTCs are

located at more than 50-55 (or 50-54 depending on the references; for convenience purposes, we will refer to this boundary as 50-55) nucleotides (nts) upstream the last exon-exon junction, they are indeed recognized as what they are, aberrant stop codons, and are thus able to elicit NMD, but if they are located downstream of this boundary, NMD is not elicited^{4,27,32}. This happens as such because, when pre-mRNA splicing occurs in the nucleus of the eukaryotic cell, multiprotein complexes appropriately called exon junction complexes (EJCs) are deposited approximately 20-24 nts upstream of each exon-exon junction. During the pioneer round of translation, the 80S ribosome displaces the EJCs assembled in the open reading frame (ORF), but, if the transcript has a PTC positioned at \geq 50-55 nts upstream the last exon-exon junction, the ribosome will reach a stop codon (which is, obviously, premature in this case) before it actually can be in a close enough position to displace the next EJC (figure 4) and therefore, the EJC will not be removed as it normally would^{34,35}. The EJC will trigger recruitment the NMD factors UPF3 and UPF2, which may be attributable to the MAGOH/Y14 heterodimer (a component of the EJC)^{31,36}. It has been shown, however, that AUG-proximal PTCs fail to elicit NMD, in spite of fulfilling the rule of being positioned at \geq 50-55 nts upstream the last exon-exon junction; this implies that the EJC-dependent model is not the only mechanism by which NMD is triggered^{32,37}. How is NMD elicited then? Work previously done in our lab, as well as other studies, have suggested that the mechanism behind this is a competitive relationship between PABP and UPF1 for binding eRF3. When PABP is in close proximity to a PTC (which happens when the PTC is AUG-proximal; this is made possible by the circular conformation of the mRNA), it interacts with the release factor eRF3 and promotes translation termination; however when the PTC is not physically close enough to PABP, then it won't interact with eRF3, instead, the NMD factor UPF1 will, and this in turn will trigger NMD^{38,39,40}.

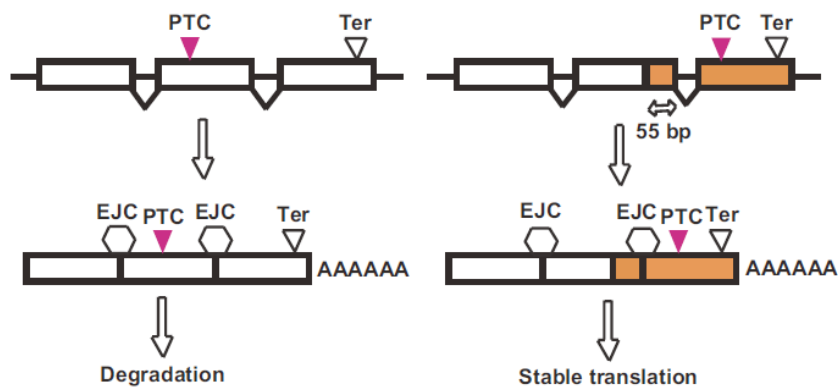


Figure 4 - Simplified model of mammalian EJC-dependent NMD eliciting. On the left, a PTC is located more than 50-55 nts upstream the last exon-exon junction, therefore it will elicit NMD and hence, degradation of the transcript will occur. On the right, there's a PTC located within the 50-55 nts boundary and as such it won't be recognized by the NMD machinery and, as a consequence, that transcript won't be rapidly degraded and will result in a truncated protein. In the figure, "Ter" represents the canonical termination codon. *In Khajavi et al. (2006).*

1.3.2. mRNA Degradation via NMD

Following PTC recognition, the next step for the NMD machinery is the degradation of the faulty mRNA molecule. However, unlike the previous step, mRNA degradation via NMD is not a very well understood process to date; still, some studies did manage to provide insights and models.

It is known that, in NMD, phosphorylated UPF1 has better affinity for RNA and marks the mRNA for degradation^{27,41}. UPF1 phosphorylation occurs as follows: once UPF1 has been recruited to the terminating ribosome, as it has already been mentioned, it will interact with eRF3 and form the SURF complex (SMG1-UPF1-eRF1-eRF3)⁴². The SURF complex, via UPF1, binds to UPF2 and UPF3, the former, in turn, binds to the EJC; all these elements together constitute the DECID complex (Decay-Inducing Complex). Finally, the DECID complex promotes UPF1 phosphorylation by SMG1^{43,44,45,46}. This phosphorylated UPF1 recruits either SMG6, which has endonucleolytic activity, or the SMG7/SMG5 heterodimer⁴⁶.

If SMG6 is the one recruited and interacts with phosphorylated UPF1, endonucleolytic cleavage of the transcript in the vicinity of the PTC occurs by the action of SMG6 itself. This cleavage produces two cleavage products: one 3' product, which contains the EJC and NMD components, and another 5' product that includes the PTC. The 5' cleavage product undergoes 3'-5' degradation, as it seems, by the exosome, while the 3' product is first stripped off of its proteins by the action of UPF1 (because when it binds UPF2, as is the

case in the DECID complex, UPF1's helicase activity is enabled) and then is cleaved in the 5'-3' direction by XRN1 following decapping^{47,48,49}.

If, however, SMG7/SMG5 heterodimer is recruited, then its interaction with the phosphorylated UPF1 induces deadenylation, carried out by the PAN2/PAN3 and CCR4-NOT complex; also, decapping by the action of DCP1/DCP2 takes place. This allows the transcript to be degraded in the 3'-5' direction by the exosome and in the 5'-3' direction by the exonuclease XRN1^{32,50,51,52}.

Additionally, SMG7 and SMG6 recruit protein phosphatase 2a (PP2A), which induces UPF1 dephosphorylation and dissociation, thereby allowing recycling for the next round of NMD⁵³. The mRNA degradation via NMD pathway is simplified bellow on figure 5.

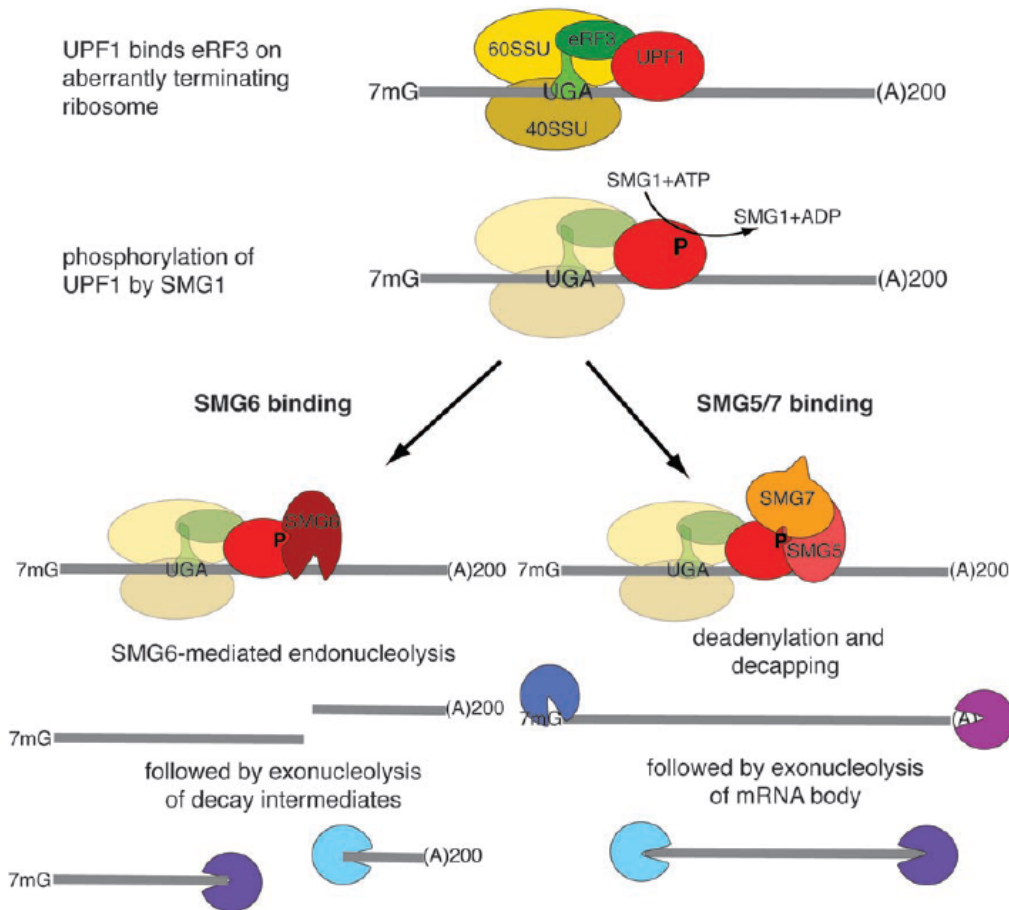


Figure 5 – Simplified decay pathway of human NMD. UPF1 binds to the stalled ribosome through interaction with eRF3 and is subsequently phosphorylated. Phosphorylated UPF1 binds either SMG6 or the SMG5/SMG7 heterodimer. Binding of the endonuclease SMG6 results in RNA cleavage in the vicinity of the PTC, producing decay intermediates that will be rapidly degraded in both 5'-3' (light blue PacMan) and 3'-5' (violet PacMan) directions by exonucleases. On the other hand, binding of SMG5/SMG7 results in the recruitment of deadenylases (purple PacMan) and decapping enzymes (dark blue PacMan). After being both decapped and deadenylated, the RNA is degraded by exonucleases. *In Mühlemann & Lykke-Andersen (2010).*

1.3.3. NMD Targets

During recent years, it has been increasingly clear that NMD isn't simply an mRNA quality control mechanism targeting nonsense mRNAs, as wide variety of physiological transcripts that don't contain any PTC whatsoever are regulated by it^{54,55}. In fact, studies have revealed that NMD participates in the control of steady-state levels of 3-10% of the transcriptome⁵⁶. So the question arises: what renders these non-faulty mRNAs NMD-sensitive? Some characteristics possessed by these NMD-sensitive transcripts include upstream ORFs (uORFs), long 3'UTRs or signals for programmed frameshift; mRNAs encoding proteins that contain selenocysteine, where the UGA codon is interpreted either as a stop codon or a PTC depending on selenium concentration; bicistronic mRNAs, among others^{33,57,58}. Many of these aforementioned features make the stop codons of these mRNAs, called natural NMD targets, to be in a position where they are recognized as *de facto* PTCs by the NMD machinery, in spite of them being canonical ones. NMD natural targets perform a variety of functions in the cell, ranging from regulation of alternative splicing to stress response and cell-cycle progression^{59,60}. One of those natural NMD targets, which is important in this work, is *HFE* mRNA, whose gene, when mutated, might be involved in hereditary hemochromatosis, a disease of iron metabolism where an excessive intestinal iron absorption is present, leading to organ damage⁵⁹.

1.3.4. NMD and Human Disease

One can easily infer the clinical relevance of NMD from the fact that approximately one third of all human genetic disorders are attributable to PTCs, many of which are NMD targets^{29,61}. There are several disease phenotypes that result from nonsense or frameshift mutations that are indeed modulated by NMD^{29,34}. Examples of disease clinical phenotypes modulated by NMD include cystic fibrosis, Duchenne muscular dystrophy and β -thalassemia, among others.

In diseases where NMD is a clinical manifestation modulator, the surveillance mechanism impedes the nonsense mRNA to be translated into truncated proteins. These truncated proteins can have lost their function, have acquired a dominant-negative effect or gain of function, but they can also be still functional or residually functional and in those cases, it's obvious that NMD, instead of protecting the cell against the deleterious effect of proteins resulting from faulty mRNAs, is in fact having a detrimental effect instead. So, NMD can either be beneficial or detrimental to the clinical manifestation of a given disease^{26,30}.

Cystic fibrosis constitutes an example where NMD worsens the clinical outcome of a disease. Some truncated proteins encoded by the mutated CFTR gene partially retain normal function, however, NMD targets and degrades their PTC-containing mRNAs, thereby downregulating those proteins, which of course aggravates the clinical manifestations of the disease^{26,62}.

In β -thalassemia, on the other hand, NMD has the opposite effect; it has a well-documented beneficial role. PTC-containing β -globin mRNAs are targeted and degraded by NMD and this prevents the synthesis of C-terminally truncated proteins toxic for the cell^{26,63}.

1.3.5. β -thalassemia as a Model

The β -thalassemias comprise a heterogeneous group of inherited human anaemias, and are among the most common genetic diseases worldwide arising most frequently as a result of nonsense or frameshift mutations in the β -globin gene. It generally has a pattern of autosomal recessive inheritance, except when nonsense mutations occur on the last exon, being dominant in that case. The disease is characterized by reduced or absent β -globin chains, which causes low levels of haemoglobin and reduced production of red blood cells^{64,65}.

Several nonsense mutations have been characterized for the β -globin gene, of which CD15, CD26 and CD39 are of particular importance for this study. In CD15-thalassemia the PTC, being on codon 15, is near the 3'-end of the transcript and this proximity to PABP will inhibit NMD, making this transcript produce a C-terminally truncated protein. In CD39-thalassemia, however, there is no AUG-proximal effect and this PTC will elicit NMD^{66,67}.

The profusion of this disease in the world, the small size of the β -globin gene and the wide range of nonsense mutations that it can harbour, "reacting" in different ways to NMD, are all reasons that turn this genetic disorder a good model for studying this surveillance mechanism^{64,66}.

1.3.6. Nonsense Suppression Therapy

In the most aggressive manifestation of β -thalassemia (Thalassemia Major) and in some phenotypes of Thalassemia Intermedia, patients require regular red blood cell transfusions and appropriate chelation therapy to counterbalance the iron overload often present due to those necessary transfusions. Bone marrow transplantation remains the only cure, but it doesn't constitute by any means a true solid hope for patients, given the compatibility

issues and other risk factors like the extent of iron accumulation present at the time of transplantation. Given the fact that NMD is a modulator of the clinical manifestation of β -thalassemia, it constitutes a promising and needed therapeutic strategy for this disease and many others caused by PTC-generating mutations that don't have effective treatments^{22,33,64,68}. At the current point in time, these therapies have already been analysed in clinical trials for Cystic Fibrosis and Duchenne's muscular dystrophy⁶⁹.

In nonsense suppression therapy, one takes advantage of the fact that some truncated proteins arising from PTC-containing mRNAs are still functional but, however, they do elicit NMD and are therefore down-regulated. If this NMD eliciting could be hampered, then the protein would still be produced and hence at least phenotype amelioration of the disease would take place^{33,70}. The way this is done is through the PTC "readthrough" strategy, which is the process that recodes a stop codon, in this case premature, into a sense codon, thereby allowing translation elongation to continue until the canonical stop codon is reached and a full-length protein is produced, which would otherwise be impossible⁷⁰. Mechanistically, suppression therapy works by facilitating the ability of near-cognate aminoacyl tRNAs (tRNAs that are complementary for two of the three codon nucleotides) to be incorporated into the ribosomal A-site (figure 6), thereby effectively increasing the number of times a PTC will be read as a sense codon and a full-length protein is produced, restoring enough normal protein function in the cell⁷⁰.

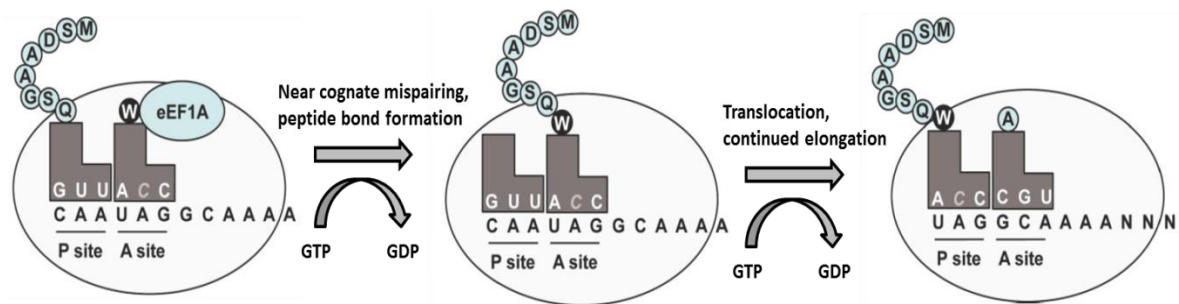


Figure 6 – Suppression of translation termination. From left to right, a near-cognate aminoacyl-tRNA complementary base pairs (at two of the three bases, of course) with the premature termination codon with the help of eEF1A. The aminoacid carried by this near-cognate tRNA is added to the growing polypeptide chain, translocation occurs and premature translation termination is aborted, taking place normal translation elongation instead, until the canonical stop codon is eventually reached. Adapted from Keeling & Bedwell (2011).

The ribosomal readthrough is accomplished by the use of small molecular weight compounds, namely aminoglycoside antibiotics such as gentamicin and geneticin (G418), and also non-aminoglycoside compounds such as ataluren (PTC124). These drugs bind to the ribosomal decoding centre, decreasing the accuracy of complementary base pairing between codon and anticodon and, in this way, facilitates near-cognate aminoacyl-tRNA incorporation. Aminoglycosides have antibacterial activity, may cause cell toxicity and, at an organismal level, nephrotoxicity; PTC124, a non-aminoglycoside, on the other hand, is free from the previously mentioned disadvantages, however its effectiveness is controversial, with one study⁷¹ reporting that there is no evidence for PTC inducing readthrough, however, the majority of studies point to successful PTC readthrough in mammalian cells and it is currently undergoing clinical trials for several diseases^{33,70,72,73,74,75,76}.

Suppression therapy efficiency, besides being different for different compounds and depending on each individual, also depends on the PTC nature, the ranking order for aminoglycoside readthrough efficiency being UGA≥UAG>UAA; and PTC context: in this regard, the nucleotide immediately downstream of the PTC is the most important one, being the ranking order of readthrough efficiency according to the it, C>U>A>G, but up to 6 nucleotides downstream and 2 nucleotides upstream of the PTC can have an effect on readthrough efficiency^{77,70,69}.

1.4. Nonstop mRNA Decay

Nonstop decay is by far much less well understood than NMD and this lack of knowledge is even more apparent for human cells; a great amount of what is known for the process relates to yeasts. Nonstop decay has evolved as a surveillance mechanism that targets and degrades transcripts that lack an in-frame stop codon. These transcripts don't allow normal translation termination to take place and instead the ribosome reads the ORF and into the 3'UTR, where it eventually stalls, as translation of poly(A) results in positively charged poly-lysine that interacts with the very negative exit channel of the ribosome; as little as 6 incorporated lysine residues are enough to incite ribosome stalling^{8,78,79,80}.

Nonstop mRNAs can arise from nonstop or frameshift mutations in the DNA that disrupt the stop codon, instances when transcription aborts or even from cryptic poly(A) sites, which cause premature polyadenylation^{80,81,82}. In particular, when cryptic poly(A) sites are present, upstream of the stop codon, these mRNAs can give rise to truncated proteins that, in similarity with NMD, can acquire deleterious dominant negative or gain of function

effects^{79,80,83}. It is estimated that nonstop mRNAs can arise from premature polyadenylation 5-10% of the time^{83,84}. However, there are many questions regarding the initial recognition of NSD-targeted ribosome complexes on a mechanistic level⁸.

In yeasts, nonstop mRNA degradation probably starts when the Ski7 GTPase recognizes the stalled ribosome through its C-terminal domain, promoting transcript degradation by the exosome^{83,85}. Ski7 belongs to a family of proteins that also includes the eukaryotic release factor 3 (eRF3) and the eukaryotic elongation factor eEF1 α , however, Ski7 is different in two ways: it doesn't act as a heterodimer and may not be dependent on GTP, which is an integral part of the specificity of both eRF3 and eEF1 α ^{86,87,85}. Nonstop mRNA degradation, as it has already been previously mentioned, is accomplished by the 3'-5' exonucleolytic action of the exosome, without the need for prior deadenylation⁸².

1.5. Eukaryotic Exosome

The exosome, as well as its cytoplasmic factors are conserved in all eukaryotes^{88,89}. It is a 3'-5' multi-subunit ribonuclease complex of approximately 400 kDa, present in nuclear, cytoplasmic and nucleolar forms, whose main function is degrading almost every type of RNA in the cell, thereby guaranteeing the following: the degradation of RNA processing by-products, elimination of aberrant transcripts in various mRNA surveillance pathways (including NMD and NSD), removal of incorrectly spliced mRNAs and it also prevents defective rRNA incorporation into the ribosome structure. Furthermore the exosome plays a role in RNA molecules maturation in the nucleus, such as ribosomal RNAs, for instance^{90,91,92,93}.

The eukaryotic exosome has a conserved structure constituted by a 9-subunit core, in similarity with bacterial PNPase and the archaeal exosome, which is made up of a hexameric ring and a "cap" consisting of three RNA-binding subunits (figure 7). Depending on the subcellular localization, to this 9-subunit core can bind Rrp44 (Ribosomal RNA Processing Protein) also known as Dis3, Rrp6, or both Rrp44 (Dis3) and Rrp6; it is possible that the exosome co-factors influence how a "binding path" is chosen (Rrp6 or Dis3) by competing for interaction with the exosome core or by influencing activities of both subunits^{94,95}.

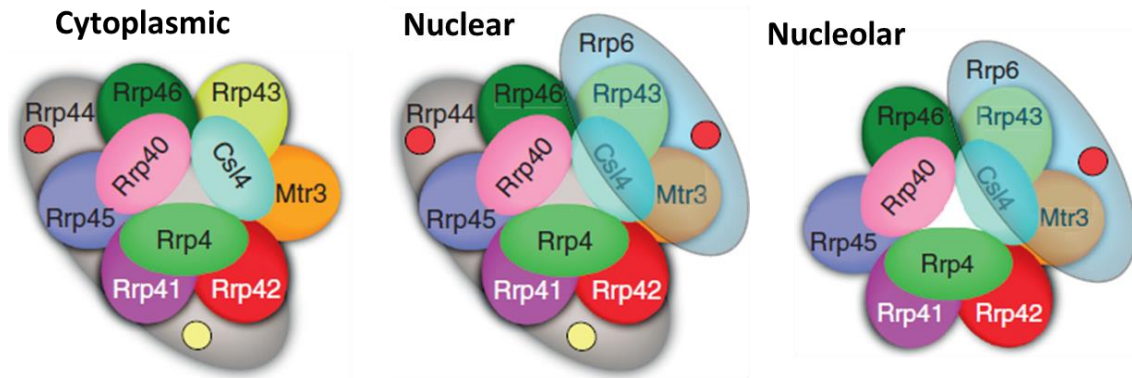


Figure 7 – Eukaryotic exosome structures. On the left, we can see the cytoplasmic 10-subunit exosome, at the centre it is pictured the canonical nuclear 11-subunit exosome and on the right, we have the nucleolar 10-subunit exosome. The 9-subunit core is present in all three forms. In the figure, the red circles represent phosphorolytic and hydrolytic exoribonuclease active sites and the pale yellow circles represent endoribonuclease active sites. Adapted from Januszyk & Lima (2014).

Dis3 is the largest subunit of the exosome (approximately 110 kDa) and is responsible for the major exoribonucleolytic activity of the complex. Dis3 is indispensable for cellular viability and is predominantly present in the nucleus, but also in the cytoplasm. Its structure consists of: an N-terminal pilus-forming N-terminus domain (PIN) with three aspartic acid residues; the PIN domain mediates the majority of the contacts with the N-terminal of the Rrp41 subunit of the exosome core, so its main function is really to bind Dis3 to the core; two cold shock domains (CSD1 and CSD2) responsible for RNA-binding; central ribonuclease domain (RNB), which contains an active site for exoribonuclease activity; and a C-terminal S1 domain also responsible for RNA binding (figure 8). Dis3 binds both to Rrp41 and Rrp45^{94,95}.



Figure 8 – Model for Dis3 domains' structure. On the left one can see the N-terminal exosome-binding domain, PIN, followed by the RNA-binding CSD domains. Next are the catalytic RNB site and the C-terminal RNA-binding S1 domain. *In Tomecki et al.* (2010).

1.6. Human Dis3L2

Recent studies indicate that human DisL2 (DIS3 Like 3'-5' Exoribonuclease 2) protein is a cytoplasmic processive 3'-5' ribonuclease that contributes to RNA homeostasis and whose activity is independent to that of the exosome. Regarding its structure, it is known that it lacks a PIN domain, hence it not binding to the exosome. Furthermore, Dis3L2 doesn't co-

precipitate or co-localize with the exosome components. It also has an extended RNA-binding CSD1 domain, which can explain Dis3L2's stronger activity for structured substrates with short two nucleotide overhangs or blunt-ended double stranded RNA when compared to hDis3^{96,97}. Dis3L2 main targets constitute poly-uridyated transcripts (figure 9), being inhibited by poly(A) sequences, and it has recently been linked to apoptosis-triggered global RNA degradation, which involves 3'-uridyated intermediaries^{96,98}. Transcript uridylation can trigger decapping and 5'-3' degradation and protect 3' ends from exoribonucleases; however, it can also promote 3'-5' degradation. This only makes sense because of the spatial separation between 5'-3' degradation and decapping, as well as Dis3L2 localization on docked structures to processing bodies (P-bodies), which is where 5'-3' degradation takes place in the eukaryotic cell; even more, under stress conditions, Dis3L2 completely co-localizes with P-bodies, which can be a way of accelerating the degradation of uridyated RNAs^{96,97}. Recently, human Dis3L2's role in mRNA metabolism has been highlighted and it has been seen that it may well be essential for cell growth regulation and division, as its depletion causes mitotic abnormalities due to deregulation of mitotic control proteins, being associated with Perlman syndrome and susceptibility to Wilms tumor⁹⁹; furthermore, it has a role in target directed miRNA degradation¹⁰⁰. Whether Dis3L2 is coupled to NMD or NSD is still unclear.

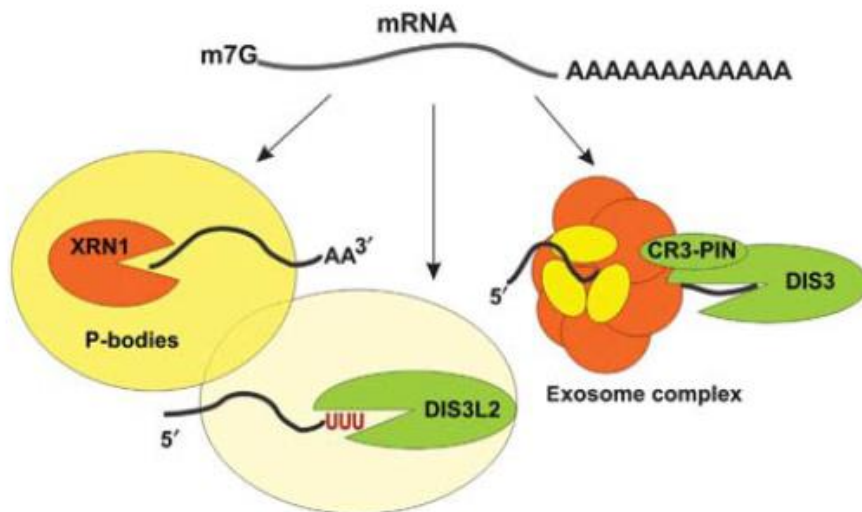


Figure 9 – mRNA degradation pathways in the cytoplasm of *S. pombe*. From left to right, mRNAs can be degraded in the 5'-3' direction by Xrn1 exonuclease, in the P-bodies; in the 3'-5' direction by Dis3L2 that can be localized on P-bodies or not; and again in the 3'-5' direction by the exosome in the cytoplasm. What applies to *S. pombe* may also apply to humans. *In* Malecki *et al.* (2013).

2. AIMS

The aims for this work are divided in two parts, corresponding to the two chapters of this dissertation.

2.1. Chapter I

β -thalassemia is one of the most common genetic diseases worldwide and its treatment can be bothersome and inconvenient, with no risk-free and easy to attain cure^{64,65,68}. Furthermore, β -thalassemia happens mainly as a result of nonsense mRNAs committed to degradation by NMD; with NMD being a clinical modulator of the manifestations of this disorder, it constitutes a hope for new treatments^{26,63,101}. Under that light, if one can use suppression therapy to induce PTC readthrough on the faulty mRNAs that trigger this disease, one would, in theory, be able to induce synthesis of full-length protein that could maintain normal physiological function. Already preliminary results in our lab have shown that G418 can suppress a PTC at codon 39 of the human β -globin mRNA, although at low levels in cultured erythroid cells; also, some recent studies have shown high readthrough efficiencies for a non-aminoglycoside compound, PTC124^{74,75,76}.

The aim for chapter I of this work was to test whether suppression therapy can restore enough β -globin to correct the manifestations of the disease and, under that light, we decided to test whether β 15 (nonsense mutation CD15) or β 39 (nonsense CD39) mutated versions of the *HBB* gene allow their suppression by either G418 or PTC124 to occur, in a dose-dependent manner, using for this effect HeLa cells transiently transfected with plasmids containing either the wild-type (β WT), β 15 or β 39 versions of the human β -globin gene.

2.2. Chapter II

DIS3L2 is a novel 3'-5' exonuclease independent of the exosome of which little is yet known, especially in humans^{96,97}. Currently it is known that this protein has a preference for poly-uridylated transcripts, but either if this protein is linked to the degradation coupled to surveillance mechanisms or just to global uridylated mRNA is yet to be known^{96,97,98,99}. Also, recently, mutations in the germline *DIS3L2* gene have been linked to mitotic abnormalities and human diseases⁹⁹. Given the fact that surveillance mechanisms are vital to the cell and are involved in many human diseases, and taking into account the

possible important role in human disease and the current gap in the knowledge regarding hDis3L2, it would be of interest investigate if this protein is involved in the degradation coupled to surveillance mechanisms, augmenting the overall knowledge of mRNA metabolism.

The aim for the chapter II of this work was to investigate if hDis3L2 is involved in the NMD and/or NSD mRNA degradation, as well as normal transcript degradation. To this effect, we utilized a knockdown approach using siRNAs specific to hDis3L2 in HeLa cells transiently transfected with plasmids containing either the β WT, β 15, β 26, β 39 or β NS (a variant which lacks a stop codon) mutated version of the human β -globin gene.

3. MATERIALS AND METHODS

3.1. Plasmid Constructs

The plasmids containing β WT (wild type version of the β -globin gene), β 15 (with nonsense mutation at codon 15) [CD 15 (TGG→TGA)], β 26 (with nonsense mutation at codon 26) [CD 26 (GAG→TAG)], β 39 (with nonsense mutation at codon 39) [CD 39 (CAG→TAG)], or β NS (without the canonical stop codon) of the human β -globin gene were obtained as previously described in Silva *et al.* (2006). All variants were created within the 428-bp NcoI-BamHI fragment of the β -globin gene template by overlap-extension PCR. Competent *Escherichia coli* were transformed with the plasmid DNA, and transformants were selected on luria-bertani (LB) agar/ampicillin plates. The corresponding plasmid DNAs were purified from overnight cultures of single colonies with the NZYMini prep kit (NZYTech, Portugal) following the manufacturer's instructions. Confirmation of the correct cloned sequences containing the relevant mutation was carried out by automatic sequencing.

3.2. Cell Culture and Plasmid Transfections

HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM 1x + GlutaMAX™-I; Gibco® by Life Technologies™, USA) supplemented with 10% (v/v) foetal bovine serum (FBS; Gibco® by Life Technologies™, USA), incubated at 37°C in a humidified atmosphere of 5% CO₂.

Transient transfections were performed using Lipofectamine 2000 Transfection Reagent (Invitrogen® by Life Technologies™, USA), following the manufacturer's instructions, in 35-mm plates containing HeLa cells plated 24h prior to transfection, using 200 ng, of plasmid DNA of each variant (β WT, β 15 or β 39). Cells were lysed, 24h after transfection, via solubilisation in Passive Lysis Buffer (PLB; Promega, USA) for posterior protein analysis.

3.3. Knockdown Experiments

HeLa cells were plated into P-60mm cell plates to a concentration of approximately 100×10^3 cells/mL and 24h later, the cells were transfected with the siRNAs to a final concentration of 20nM using Lipofectamine 2000 Transfection Reagent (Invitrogen® by Life Technologies™, USA), following the manufacturer's instructions. Approximately 36h later, a second siRNA transfection was performed in the same way as the first one and, at the same time, the cells were transiently transfected with 400ng of the plasmid containing the respective variant (β WT, β 15, β 26, β 39 and β NS) using Lipofectamine 2000

Transfection Reagent (Invitrogen® by Life Technologies™, USA), following the manufacturer's instructions. Cells were lysed, 24h after this last transfection, by solubilisation in RA1 Buffer (NZYTech, Portugal) for posterior RNA isolation.

3.4. RNA isolation

Total RNA from cultured HeLa cells was isolated using the RNA extraction kit NucleoSpin RNA II (Macherey-Nagel, Germany) according to the manufacturer's instructions. RNA samples were treated with RNase-free DNase I (Ambion® by Life Technologies™, USA).

3.5. Reverse transcription-coupled quantitative PCR (RT-qPCR)

cDNA synthesis was carried out using 2 µg of total RNA and Reverse Transcriptase (NZYTech, Portugal), according to the manufacturer's instructions. Real-Time quantitative PCR (RT-qPCR) was performed in ABI Prism 7000 Sequence Detection System, using SybrGreen Master Mix (Applied Biosystems® by Life Technologies™, USA). Primers specific for the gene of interest, β -globin (primer forward 5'-CTCAAGGGCACCTTTGCCAC-3' and primer reverse 5'-CAGCACACAGACCAGCACGT-3') and for the control, luciferase reporter gene were used. Quantification was performed using the relative standard curve method ($\Delta\Delta C_t$, Applied Biosystems® by Life Technologies™, USA). The following cycling parameters were used for the cDNA synthesis: 10 min at 25°C, then 50°C for 40 min and 5 min at 85°C. Technical triplicates from each experiment were assessed in all cases.

3.6. SDS-PAGE and Western Blotting

Cells lysates were denatured for 10 minutes at 95°C. Five µl of SDS sample buffer 5x [Bromophenol blue (0.25%), DTT (dithiothreitol; 0.5 M), Glycerol (50%), SDS (sodium dodecyl sulfate; 10%), Tris-Cl (0.25 M, pH 6.8)] was added to 20 µl of purified lysates and these were loaded into a 12% polyacrylamide gel and resolved for 1 hour. Afterwards, they were transferred to a PVDF membrane (Bio-Rad, USA) at 30V for 30minutes. The membranes were then either fixated with a 0,4% PFA (v/v) solution in PBS 1x (Phosphate Buffer Saline) for 30min at room temperature and blocked in TBS 5% (w/v) non-fat dry milk for 1 hour, or immediately blocked in SuperBlock™ Blocking Buffer (ThermoFisher Scientific, USA). Membranes were probed using mouse anti- α -tubulin antibody (loading control; Roche, Switzerland) at 1:10 000 dilution and mouse monoclonal anti- β -globin (Santa Cruz Biotechnology, USA) at 1:200 overnight. After incubation with the primary antibody, membranes were washed 3 times in TBS 1x. Detection was carried out by

incubating the membranes for 1 hour with the secondary antibodies, peroxidase-conjugated anti-mouse IgG (Bio-Rad, USA), anti-rabbit IgG (Bio-Rad, USA) antibodies, followed by enhanced chemiluminescence reaction.

3.7. SOEing PCR

SOEing (Gene Splicing by Overlap Extension) PCR: four SOEing PCRs were required in order to extract the FLAG sequence (GACTACAAGGATGACGATGACAAAGACTACAAGGATGACGATGACAAA) from pTRE+mTOR plasmid to pTRE2pur+(β -globin variant). The first three SOEings were carried out using the following conditions: 5 μ L Pfu Buffer 10x, 5 μ L DMSO 100%, 1 μ L MgSO₄, 5 μ L dNTPs (2mM), 0,85 μ L BSA, 1 μ L Pfu enzyme, 2 μ L of each primer (forward and reverse) and 1 μ L template DNA and water to the final volume of 50 μ L. The cycling program was: 95°C during 4min30sec.; 40 cycles of 95°C 1min, 54°C 30sec., 72°C 1min.; and 72°C 10min. The final SOEing4 was carried out using a final reaction volume of 100 μ L, seeing that 15 μ L of that volume was the template DNA (4 μ L of SOEing1 PCR reaction, 8 μ L of SOEing2 and 3 μ L of SOEing 3, respecting equimolar concentrations). The sequence of these primers is listed in Annex I.

3.8. Digestions and Ligation Reactions

Following purification from the agarose gel using the NZYGelpure kit (NZY-Tech, Portugal), following the manufacturers' instructions, 6,5 μ L of SOEing4 PCR product reaction was submitted to an A-tailing reaction following the manufacturer's instructions in NZY-A PCR cloning kit (NZYTech Potugal). After A-tailing, ligation was performed using T4 DNA Ligase and a proportion of 1 part vector to 3 parts fragment in a 10 μ L final volume at 4°C overweekend following the manufacturer's instructions in the NZY-A PCR cloning kit.

Following that, digestion of the ligation product was performed with the enzymes BstXI (10 U/ μ L) (Thermo Scientific) and BspLU11I (10 U/ μ L) for 2 hours at 37°C. We used the volume corresponding to 2 μ g of the ligation product, 1 μ L of each enzyme, 5 μ L of Buffer3 (100mM NaCl; 50mM Tris-HCl; 10mM MgCl₂; 100 μ g/ml BSA) in a final volume of 50 μ L.

3.9. Statistical Analysis

Results are expressed as mean \pm standard deviation of 3 experiments in which the mRNA levels expressed from β 15, β 26, β 39 and β NS-containing plasmids are normalized to the mRNA levels of the β WT transfected with siRNA for Luciferase, which are arbitrarily set to 1. Student's t test was used for estimation of statistical significance (unpaired, two-tailed). Significance for statistical analysis was defined as a $p < 0.05$.

CHAPTER I

Nonsense-mediated mRNA Decay as a Tool for Gene Therapy

4. RESULTS & DISCUSSION

4.1. Human β -globin protein wasn't detectable by western blot analysis

SDS-PAGE Western blot analysis of protein samples extracted from HeLa cells transfected with each one of the pTRE+(β -globin variant) plasmids, or not transfected at all (NT), was performed in order to evaluate protein levels produced by each transcript (fig. 10). Also, one of the membranes was blocked using SuperBlock solution, while the other was blocked with TBS-milk 5% as usual and incubated with a protein fixation solution (PBS-PFA 0,4%). We wanted to test this protocol with a fixation solution because as β -globin is a small protein, it could be that it got washed away easily from the membrane.

As one can observe the figure bellow, no significant difference was produced by either using a fixation solution or not. Furthermore, one can see that α -tubulin was detected in both membranes but not β -globin. As the antibody used for β -globin had already been previously tested in the lab with the same negative result even with posterior optimization of the western blot protocol, we decided to make changes that hadn't yet been tested.

We hypothesized that β -globin was being lost in the transfer process. Hence, we repeated western blot analysis of the cell lysates adjusting the polyacrylamide gel percentage to 14%, transferring just for 30 minutes at 30V in a cold room using a transfer buffer without SDS, because small proteins are more likely to be hindered by this detergent from binding to PVDF membranes; we also performed every washing and blocking step using TBS without Tween 20 or any other detergent. However, the results were no different from those on fig. 1 and we concluded that most likely the problem lies with the antibody itself.

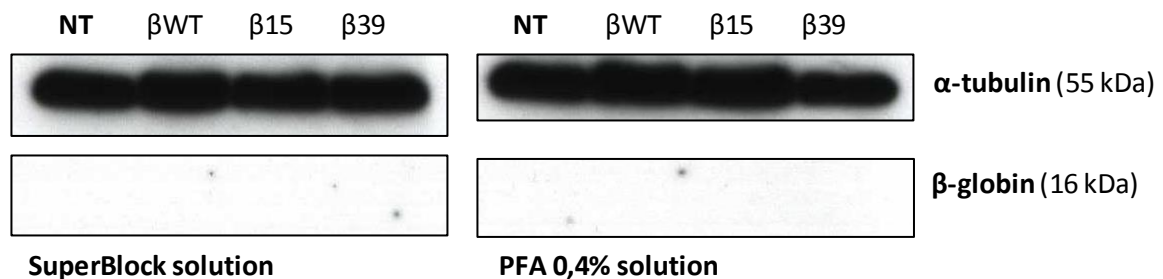


Figure 10 – Western blot analysis of cell lysates. The molecular weights of the α -tubulin (55 kDa) and β -globin (16 kDa) proteins are indicated on the right hand side of the autoradiographs. Below each autoradiograph, there is the indication of whether that membrane went through a fixation procedure (PFA 0,4% solution) or if it was just blocked using a commercial solution (SuperBlock solution). NT indicates the non-transfected cell lysate.

In order to still be able to detect β -globin, we then approached the problem in another way: we tried to clone a tag at the C-terminal region of the β -globin sequence for which we had a working antibody in the laboratory.

4.2. SOEing PCR was not successful enough as an approach to clone a FLAG-tag into the C-terminal region of Human β -globin

We used the Gene Splicing by Overlap extension PCR technique in order to clone the FLAG-tag sequence at the C-terminal end of the β -globin sequence as we already had a functioning antibody for FLAG and also we had FLAG already cloned into another plasmid in the laboratory (pTRE+mTOR), which made this the fastest option. FLAG consists of two repeats of 5'-GACTACAAGGATGACGATGACAAA-3' sequence.

Our strategy consisted of taking FLAG from the plasmid it was located in and place it in the C-terminal region of the β -globin sequence previously cloned in pTRE2pur vector. For this we designed and utilized specific primers whose sequences are present in table 1 in Annex I. In our so-called SOEing 1, we used a specific reverse primer for the C-terminal region of the β -globin sequence containing a linker sequence to the FLAG (that is located, as previously mentioned, in other plasmid), as well as a forward primer for β -globin; the PCR product contains the BstXI restriction site (figure 11A and 13A) for posterior digestion. In SOEing 2, we used a specific forward primer for the 5' end region of the FLAG sequence with a linker sequence to β -globin (that is located in other plasmid, pTRE2pur), as well as a reverse primer for the 3' end of the FLAG sequence containing a linker sequence for the 3'UTR of the β -globin sequence with a termination codon (TAA) immediately 3' of the FLAG sequence (figure 11B and 13A). In SOEing 3, we used a specific forward primer for the 3'UTR of the β -globin sequence containing a linker sequence for the FLAG along with a TAA termination codon just 3' of the linker, as well as a reverse primer for the pTRE plasmid sequence that contains the BspLU11I enzyme restriction site (figure 11C and 13A). Our objective with these three SOEing reactions was to get our insert (FLAG-tag), which will be present in the SOEing 2 PCR reaction and create overlap regions (with the linker sequences) in the other two PCR products (which contain the 3' end region of the β -globin gene in SOEing 1 and the 3'UTR region in SOEing 2) to the FLAG sequence thereby tethering the FLAG-tag to the C-terminal region of the β -globin gene with a final SOEing PCR 4 where we would add together in equimolar proportions the products from previous

SOEings and amplify them with the two most external primers (the forward primer of SOEing 1 and the reverse primer of the SOEing 3) (figure 12 and 13B and C). This would allow us, by the use of a FLAG-specific primary antibody, to effectively verify if PTC readthrough had occurred in transfected HeLa cells with the β -globin variants treated with G418 or PTC124 drugs.

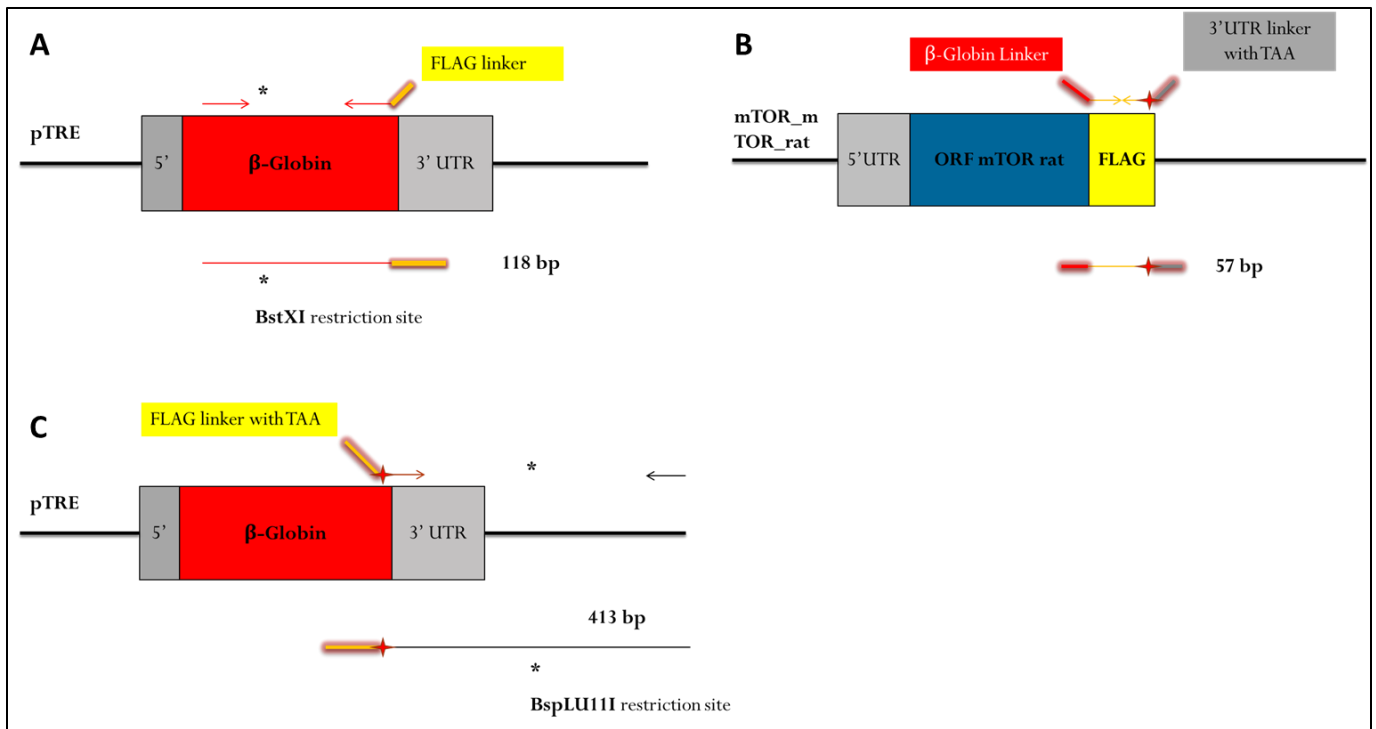


Figure 11 – Illustrations for SOEing PCR 1, 2 and 3. On panel **A** one can see the model for SOEing 1. In this model, the pTRE plasmid sequences are represented as a bold black line, the 5' and 3' UTRs are shown in gray and β -globin in red. On panel **B** there's the model for SOEing 2, where the 5'UTR of the mTOR sequence is represented in gray, the mTOR sequence itself in blue and the FLAG sequence in yellow. On panel **C** SOEing3 is schematized. In the figure, the asterisks represent enzyme restriction sites and the arrows represent the primers and their amplification direction; the stronger ends of some primers represent the linker sequences, they are elevated as a way to show that they are not complementary to the sequence below them. The red 4-pointed stars represent the canonical stop codon of the β -globin gene.

Following SOEing 1, 2 and 3 reactions, the PCR products were resolved in an agarose gel (figure 13A) and subjected to column purification to be then utilized in the SOEing 4 reaction (figure 13B). However, as one can see from figure 13B, the expected SOEing 4 band didn't appear, which goes to show that indeed much is lost during the purification process. We therefore tried performing the SOEing 4 reaction with the direct un-resolved and un-purified products of the previous SOEings (figure 13C). As one can clearly see, this time, our SOEing 4 band did appear, although, predictably, with many other unspecific products.

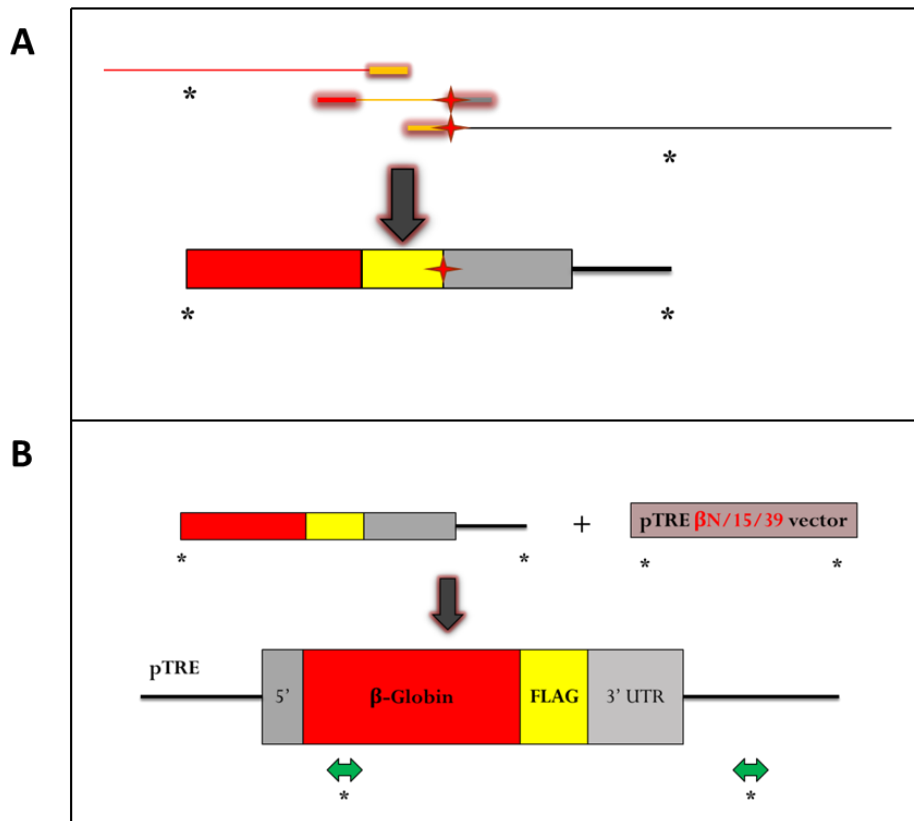


Figure 12 – Model for SOEing 4 PCR and Ligation to pTRE vectors. On panel **A** one can see the products from previous SOEings 1, 2 and 3; after the PCR reaction takes place and digestion is performed on the product, which is represented by the grey arrow, we are left with our final insert, which contains part of the 3' region of the β-globin gene (in red), followed by the FLAG-tag sequence (in yellow), the canonical stop codon (that is represented as a red 4-pointed star) and finally the 3'UTR of β-globin as well as some plasmid sequences (bold black line). The asterisks represent enzyme restriction sites. On panel **B**, ligation of the digested SOEing 4 product to the digested pTRE vector containing a β-globin gene variant is schematized. After ligation takes place, represented by the grey arrow, we have our final product. The green arrows on the bottom show the areas where ligation by T4 DNA ligase took place. The primers are not represented here because they could only be the most exterior ones.

We tried to optimize the SOEing 4 PCR reaction by changing the concentrations of its components and even having performed the PCR at various annealing temperatures so that the better one could be chosen. However, no significant differences in efficiency were obtained. The intended SOEing 4 product band was purified from the gel, digested using BstXI and BspLU11I enzymes, and this digestion was run on a gel and the intended product purified, which was then ligated to a commercial plasmid pNZY28-A using A tailing and T4 DNA Ligase and amplified through bacterial transformation, this in order to have an acceptable quantity of our product, despite the successive purifications. Grown colonies were selected and its DNA digested with BstXI and BspLU11I and run on an agarose gel. As one can see from figure 14, the enzyme BspLU11I cuts our template DNA at more than one

site, but still that doesn't invalidate our result. Only one colony selected presented a band that had the possibility of containing our product, the last one loaded on the gel.

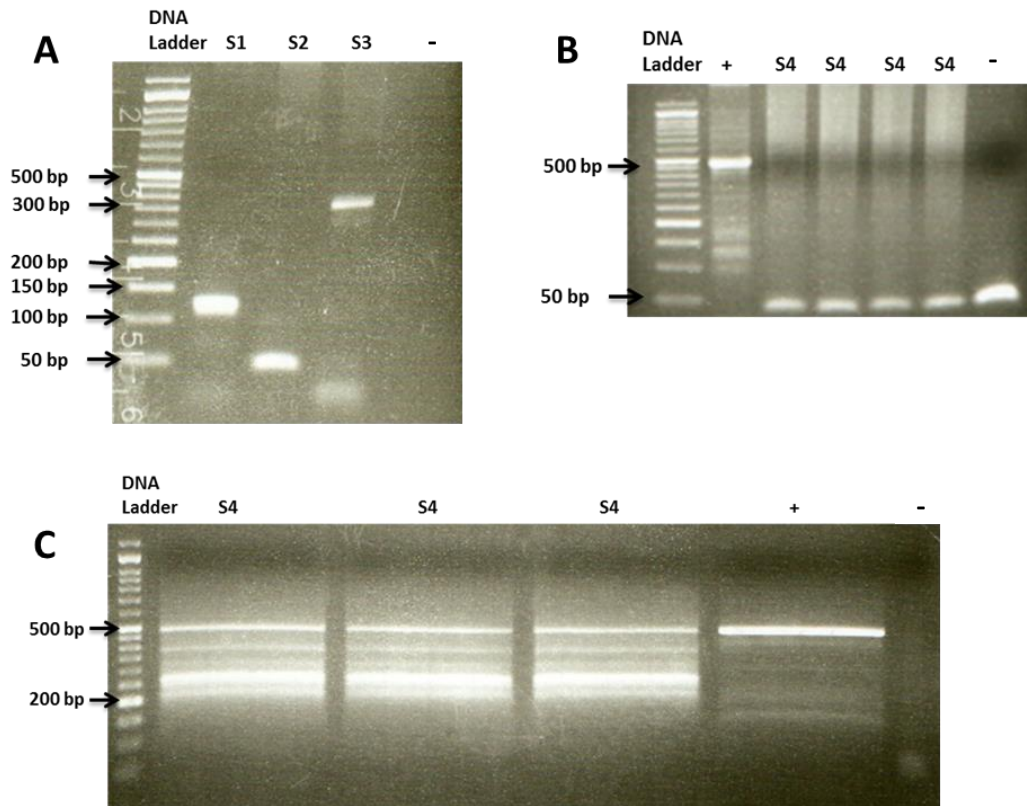


Figure 13 – Gel Electrophoresis of the SOEing PCR reactions' products. On panel **A** one can see SOEing 1, 2 and 3, 10 μ L of each reaction product was loaded and resolved in an agarose gel. S1 stands for SOEing 1 PCR reaction product, S2 for SOEing 2 PCR and so on. SOEing 1 PCR product is 118bp long, SOEing 2's is 57bp and SOEing 3's is 413 bp. "-" stands for negative control, which consisted of the PCR mix reaction only, without any template DNA. On panel **B** one can see the electrophoresis of the SOEing 4 reaction product that was performed using the purified products of SOEings 1, 2 and 3 from panel A. The + symbol represents the positive control, which was the product of a PCR reaction using the same primers as in SOEing 4 but with the pTRE2pur carrying human β -globin as a template. Samples were resolved in an agarose gel. The positive control (+) consists of a PCR using pTRE β WT as a template for the SOEing 4 primers. SOEing 4 PCR product length is 571 bp. Panel **C** depicts the same procedure as panel B, with the exception that, this time, SOEing4 PCR was carried out with the direct un-purified products of SOEings 1, 2 and 3. All gels depicted had 1,5% agarose concentration.

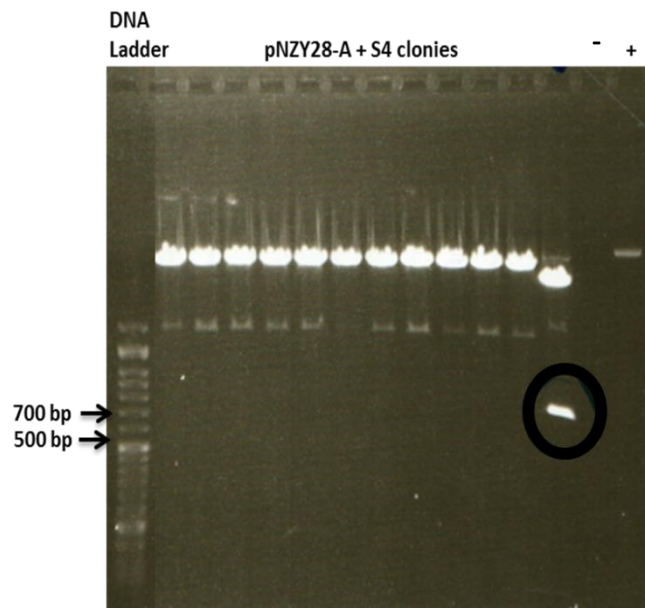


Figure 14 – Gel electrophoresis of the colony PCR products. Digested (with BstXI and BspLU11I) DNA from the selected bacterial colonies transformed with pNZY28-A + SOEing 4 ligation product was resolved in a 0,8% agarose gel. “-“ stands for negative control, where the PCR mix without any template was loaded and the + sign designates the positive control, which was the same PCR as the others but with pTRE+βWT as template. The black circumference indicates the band of interest.

We then extracted the DNA band of interest, purified it and sequenced it. It did contain the FLAG sequence in the expected location, but unfortunately, something wrong must’ve occurred in the PCR reactions because what we obtained was the two FLAG repetitions with, however, other small sequences of the β-globin gene inserted between them. For this reason, we couldn’t use it for our experiments.

We retried this approach several times but we never succeeded. We think that the main problem lies in the SOEing 4 PCR reaction, for it has too many nonspecific products and also because the sizes of the different sequences are very dissimilar among them, which makes it difficult to have an equilibrium that makes our intended product the most energetically favourable one.

5. FINAL CONSIDERATIONS & FUTURE DIRECTIONS

Given the challenges we faced for Chapter I of this work, we couldn't test whether suppression therapy can restore enough β -globin to correct the manifestations of β -thalassemia. It would be easier as a future approach to carry out this study using a different primary antibody for human β -globin, since the other strategies involving cloning were demonstrated to be of difficult achievement and to take too much time.

It could also be interesting, as an augmentation to the suppression therapy assays using G418 or PTC124, to test the same readthrough efficiency under NMD inhibiting conditions. These two treatments taken together could yield a stronger efficiency and therefore constitute a promising therapeutic strategy to treat β -thalassemia and other diseases caused by NMD eliciting nonsense mutations. This could be done taking advantage of the previous knowledge our lab and others^{4,38} that tethering PABPC1 in close proximity of a NMD-competent PTC inhibits NMD, transfecting cells with antisense oligonucleotides specific for the β -globin mRNA sequence, but with a degenerated poly(A)-tail, thus inhibiting UPF1 interaction with eRF3 and inhibiting NMD.

Also, it would be interesting, following confirmation of our assumptions in HeLa cells, to make the same experiments in Murine Erythroleukemia (MEL) cells as expression models, as they differentiate into erythroid cells and thus provide a model closer to reality for the human β -globin gene, further strengthening the results¹⁰². Also, western blot analysis wouldn't be a problem since an antibody specific for human β -globin would be utilized.

CHAPTER II

The Role of Human Dis3L2 in Transcript Degradation

4. RESULTS & DISCUSSION

4.1. DIS3L2 knockdown correlates with increased levels of β -globin mRNAs independently from their decay mechanism

We transfected HeLa cells with plasmids expressing β WT, β 15, β 26, β 39 or β NS gene variants of the human β -globin gene. Twenty-four hours later, cell cultures were either treated with siRNA for hDIS3L2 or Luciferase (LUC) as negative control. The β -globin mRNA levels were quantified by RT-qPCR analysis (figure 15).

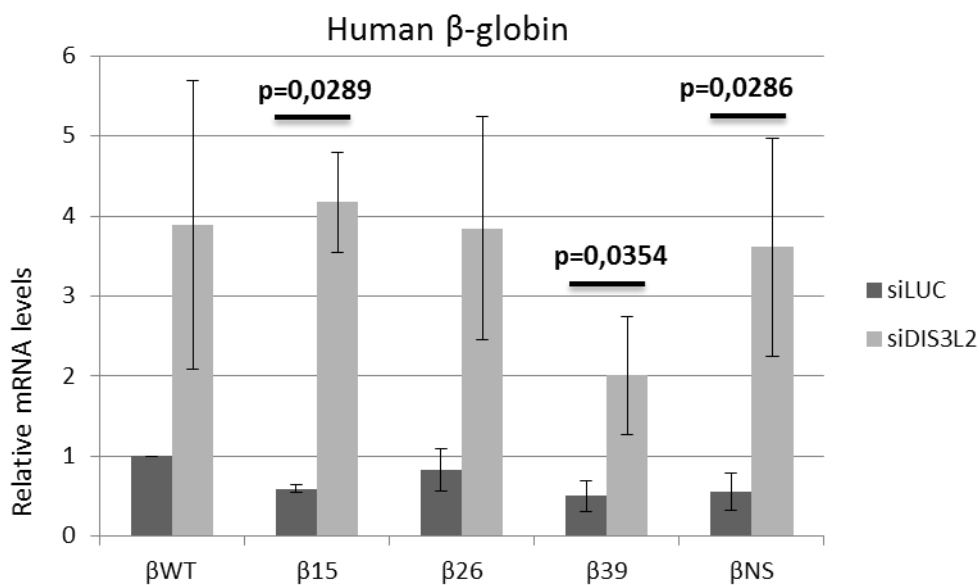


Figure 15 – RT-qPCR analysis of Dis3L2 knockdown experiments performed for the Human β -globin gene. The analysis was done on RNA isolated from HeLa cells containing the β -globin variants' transcripts treated with either siRNA specific for Dis3L2 (siDIS3L2) or for the control, Luciferase reporter gene (siLUC). The bars represent the relative mRNA levels. Standard deviations are represented as vertical error bars. The p-values of the statistically significant differences are given above each pair of data.

These results suggest that when Dis3L2 is present at low levels in the cell, mRNA decay is not as efficient, as evidenced by the fact that β 39 mRNA is up-regulated with Dis3L2 knockdown compared to normal Dis3L2 levels, this being statistically significant ($p=0,0354$). Furthermore, this happens independently of the decay mechanism, for the same happens, with a statistically significant difference, among variants that undergo NSD (β NS) ($p=0,0286$) or just normal transcript degradation (β 15) ($p=0,0289$). However we didn't obtain a statistically significant difference for β WT and β 26. Those could be explained by the high standard deviations, but even in these conditions, we see a tendency

for mRNA levels to rise with Dis3L2 knockdown. hDis3L2's role in mRNA degradation has already been established^{96,97}, our results further contribute to assert this proteins' effective role in mRNA metabolism and, furthermore, point to the direction that it may be an ubiquitous degrading component, present in normal transcript degradation but also in surveillance mechanism directed decay.

4.2. hDIS3L2 knockdown correlates with increased *HFE* mRNA levels

In order to have further confirmation that Dis3L2 is involved in the NMD directed decay pathway, we decided to evaluate the expression of a known NMD target, the *HFE* (human hemochromatosis protein) mRNA. We did this experiment for two groups, the first one where we performed knockdown of Dis3L2 and the other where the knockdown target was UPF1 mRNA. Knowing that UPF1 is a key factor in the NMD pathway, this group constituted in fact a positive control for the experiment, where we expected NMD inhibition and consequent rise in *HFE* mRNA levels.

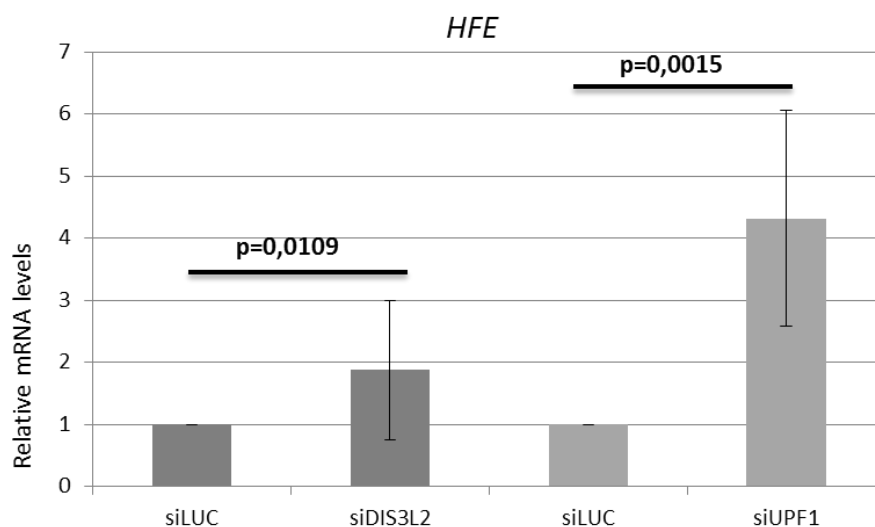


Figure 16 - RT-qPCR analysis of Dis3L2 or UPF1 knockdown experiments performed for the *HFE* gene. The analysis was done on RNA isolated from HeLa cells containing the β -globin variants' transcripts divided in two groups. In the first group, represented as darker in shade, cells were either treated with siRNA specific for Dis3L2 (siDIS3L2) or for the control Luciferase reporter gene (siLUC); in the second group, cells were treated with siRNA specific for UPF1 (siUPF1) or for the control Luciferase reporter gene once again. Standard deviations are represented as vertical error bars. The p-values of the statistically significant differences are given above each pair of data.

As one can see in figure 16, the expected rise in *HFE* mRNA levels did occur when UPF1 was silenced, and that difference is statistically significant. This is in accordance with previous results from our lab that indicate that *HFE* mRNA is an NMD natural target⁵⁹.

That's because, if NMD is inhibited, *HFE* mRNA levels will rise as they are targeted by the NMD machinery; hence *HFE* mRNA higher levels are an indicator of NMD inhibition. More importantly, though, Dis3L2 knockdown correlated with increased *HFE* mRNA levels with a statistically significant difference. As higher *HFE* mRNA levels indicate NMD inhibition, this means that when human Dis3L2 was silenced, NMD was in some way inhibited. These results further point to the assumption that human Dis3L2 is indeed involved in the NMD pathway, possibly being one of the recruited components for degradation of detected nonsense transcripts. Furthermore, the fact that the rise in *HFE* mRNA levels was significantly less than that observed for *UPF1* mRNA, suggests that UPF1 is more vital to NMD function than Dis3L2. In fact, UPF1 is required for NMD triggering^{61,31,29,42}. So one can deduce that hDis3L2 takes part in NMD but it isn't, by any means, a vital component such as UPF1, it is more likely an accessory one.

5. FINAL CONSIDERATIONS & FUTURE DIRECTIONS

Our results allow us to conclude that hDis3L2 protein is involved in surveillance mechanisms directed mRNA decay, namely NMD and NSD, but also possibly normal transcript degradation overall. However, further confirmation is needed. With this, our work contributes to the narrowing of the gap in the knowledge of the functions of this protein in human cells.

As a practical measure, it would be of interest to further repeat the real time PCRs for the hDis3L2 knockdown experiments, as a means to tackle the big standard deviations present in the wild-type and β 26 samples and by doing so, effectively asserting this work's results. Also, other NMD natural targets could be used in place of HFE, such as SMG5, SLC7A11 and GADD45A as a way to diversify and give a stronger certainty for the Dis3L2 involvement in the NMD pathway.

It would also be interesting to enrich the results, in a later phase, by looking for and using other genes aside from β -globin to ensure that these results aren't specific for this gene family.

6. REFERENCES

1. Crick, F. Central Dogma of Molecular Biology. *Nature* **227**, 561–563 (1970).
2. Maniatis, T. & Reed, R. An extensive network of coupling among gene expression machines. *Nature* **416**, 499–506 (2002).
3. Shandilya, J. & Roberts, S. G. E. The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1819**, 391–400 (2012).
4. Behm-ansmant, I., Wittkopp, N., Kashima, I., Rehwinkel, J. & Izaurralde, E. mRNA quality control : An ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Lett.* **581**, 2845–2853 (2007).
5. Le Hir, H., Nott, A. & Moore, M. J. How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* **28**, 215–220 (2003).
6. Schoenberg, D. R. & Maquat, L. E. Regulation of cytoplasmic mRNA decay. *Nat. Rev. Genet.* **13**, 448–448 (2012).
7. Jackson, R. J., Hellen, C. U. T. & Pestova, T. V. The Mechanisms of Eukaryotic Translation and Principles of its Regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 113–127 (2010).
8. Shoemaker, C. J. & Green, R. Translation drives mRNA quality control. *Nat. Struct. Mol. Biol.* **19**, 594–601 (2012).
9. Preiss, T. & Hentze, M. W. Starting the protein synthesis machine: Eukaryotic translation initiation. *BioEssays* **25**, 1201–1211 (2003).
10. Ramakrishnan, V. Ribosome structure and the mechanism of translation. *Cell* **108**, 557–572 (2002).
11. Gebauer, F. & Hentze, M. W. Molecular mechanisms of translational control. *Nat.Rev.Mol.Cell Biol.* **5**, 827–835 (2004).
12. Asano, K. *et al.* Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBOJ.* **20**, 2326–37 (2001).
13. Algire, M. a *et al.* Development and characterization of a reconstituted yeast translation initiation system. *RNA* **8**, 382–97 (2002).

14. Kozak, M. The scanning model for translation:an update. *J. Cell Biol.* **108**, 229–241 (1989).
15. Lorsch, J. R. & Dever, T. E. Molecular view of 43 S complex formation and start site selection in eukaryotic translation initiation. *J. Biol. Chem.* **285**, 21203–21207 (2010).
16. Hinnebusch, A. G. Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes. *Microbiol. Mol. Biol. Rev.* **75**, 434–467 (2011).
17. Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Mol. Cell* **2**, 135–140 (1998).
18. Sonenberg, N. & Hinnebusch, A. G. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **136**, 731–745 (2009).
19. Barbosa, C., Peixeiro, I. & Romão, L. Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. *PLoS Genet.* **9**, 1–12 (2013).
20. Hinnebusch, A. G. & Lorsch, J. R. The mechanism of eukaryotic translation initiation - new insights and challenges.pdf. *Cold Spring Harb. Perspect. Biol.* **4**, 1–25 (2012).
21. Proud, C. G. Peptide-chain elongation in eukaryotes. *Mol. Biol. Rep.* **19**, 161–170 (1994).
22. Keeling, K. M., Xue, X., Gunn, G. & Bedwell, D. M. Therapeutics Based on Stop Codon Readthrough. *Annu. Rev. Genomics Hum. Genet.* **13**, 371–394 (2014).
23. Abbott, C. M. & Proud, C. G. Translation factors: In sickness and in health. *Trends Biochem. Sci.* **29**, 25–31 (2004).
24. Dever, T. E. & Green, R. The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harb Perspect Biol* **4**, (2012).
25. Clancy, S. & Brown, W. Translation: DNA to mRNA to protein. *Nat. Educ.* **1**, 101 (2008).
26. Nicholson, P. *et al.* Nonsense-mediated mRNA decay in human cells: Mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cell. Mol. Life Sci.* **67**, 677–700 (2010).
27. Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A. & Mühlemann, O. Nonsense-mediated mRNA decay - Mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1829**, 612–623 (2013).

28. Karijolic, J. & Yu, Y. T. Therapeutic suppression of premature termination codons: Mechanisms and clinical considerations (Review). *Int. J. Mol. Med.* **34**, 355–362 (2014).
29. Holbrook, J. A., Neu-Yilik, G., Hentze, M. W. & Kulozik, A. E. Nonsense-mediated decay approaches the clinic. *Nat. Genet.* **36**, 801–8 (2004).
30. Nicholson, P. & Mühlemann, O. Cutting the nonsense: the degradation of PTC-containing mRNAs. *Biochem. Soc. Trans.* **38**, 1615–1620 (2010).
31. Chang, Y.-F., Imam, J. S. & Wilkinson, M. F. The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **76**, 51–74 (2007).
32. Silva, A. L. & Romão, L. The mammalian nonsense-mediated mRNA decay pathway: To decay or not to decay! Which players make the decision? *FEBS Lett.* **583**, 499–505 (2009).
33. Mühlemann, O., Eberle, A. B., Stalder, L. & Zamudio Orozco, R. Recognition and elimination of nonsense mRNA. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1779**, 538–549 (2008).
34. Khajavi, M., Inoue, K. & Lupski, J. R. Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *Eur. J. Hum. Genet.* **14**, 1074–1081 (2006).
35. Nagy, E. & Maquat, L. E. A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. *Trends Biochem. Sci.* **23**, 198–199 (1998).
36. Kim, V. N., Kataoka, N. & Dreyfuss, G. Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. *Science* **293**, 1832–1836 (2001).
37. Pereira, F. J. C. *et al.* Resistance of mRNAs with AUG-proximal nonsense mutations to nonsense-mediated decay reflects variables of mRNA structure and translational activity. *Nucleic Acids Res.* **43**, 6528–6544 (2015).
38. Silva, A. L., Ribeiro, P., Inácio, Â., Liebhaber, S. A. & Romão, L. Proximity of the poly (A) -binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay Proximity of the poly (A) -binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. *RNA* **14**, 563–576 (2008).
39. Uchida, N., Hoshino, S. I., Imataka, H., Sonenberg, N. & Katada, T. A novel role of the mammalian GSPT/eRF3 associating with poly(A)-binding protein in cap/poly(A)-dependent translation. *J. Biol. Chem.* **277**, 50286–50292 (2002).
40. Ivanov, P. V, Gehring, N. H., Kunz, J. B., Hentze, M. W. & Kulozik, A. E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest

- an integrated model for mammalian NMD pathways. *EMBO J.* **27**, 736–747 (2008).
41. Mühlemann, O. & Lykke-andersen, J. How and where are nonsense mRNAs degraded in mammalian cells ? *RNA Biol.* **7**, 28–32 (2010).
 42. Kashima, I. *et al.* Nonsense-mediated mRNA decay (NMD) is a surveillance mechanism that degrades mRNA containing premature termination codons (PTCs). In mammalian cells, recognition of PTCs requires translation and depends on the presence on the mRNA with the splicing-depend. *Genes Dev.* **20**, 355–367 (2006).
 43. Denning, G., Jamieson, L., Maquat, L. E., Thompson, E. A. & Fields, A. P. Cloning of a novel phosphatidylinositol kinase-related kinase. Characterization of the human SMG-1 RNA surveillance protein. *J. Biol. Chem.* **276**, 22709–22714 (2001).
 44. Palacios, I. M. Nonsense-mediated mRNA decay: From mechanistic insights to impacts on human health. *Brief. Funct. Genomics* **12**, 25–36 (2013).
 45. Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y. & Ohno, S. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes Dev.* **15**, 2215–2228 (2001).
 46. Eberle, A. B., Lykke-Andersen, S., Mühlemann, O. & Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* **16**, 49–55 (2009).
 47. Gatfield, D. & Izaurralde, E. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature* **429**, 575–578 (2004).
 48. Popp, M. W.-L. & Maquat, L. E. Organizing Principles of Mammalian Nonsense-Mediated mRNA Decay. *Annu Rev Genet.* **47**, 139–165 (2013).
 49. Nagarajan, V. K., Jones, C. I., Newbury, S. F. & Green, P. J. XRN 5'→3' exoribonucleases: Structure, mechanisms and functions. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1829**, 590–603 (2013).
 50. Unterholzner, L. & Izaurralde, E. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell* **16**, 587–596 (2004).
 51. Yamashita, A. *et al.* Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat. Struct. Mol. Biol.* **12**, 1054–1063 (2005).

52. Schmid, M. & Jensen, T. H. The exosome: a multipurpose RNA-decay machine. *Trends Biochem. Sci.* **33**, 501–510 (2008).
53. Behm-Ansmant, I. & Izaurralde, E. Quality control of gene expression: A stepwise assembly pathway for the surveillance complex that triggers nonsense-mediated mRNA decay. *Genes Dev.* **20**, 391–398 (2006).
54. Mendell, J. T., Sharifi, N. a, Meyers, J. L., Martinez-Murillo, F. & Dietz, H. C. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat. Genet.* **36**, 1073–1078 (2004).
55. Wittmann, J., Hol, E. M. & Jäck, H.-M. hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. *Mol. Cell. Biol.* **26**, 1272–87 (2006).
56. Rehwinkel, J., Raes, J. & Izaurralde, E. Nonsense-mediated mRNA decay: target genes and functional diversification of effectors. *Trends Biochem. Sci.* **31**, 639–646 (2006).
57. Moriarty, P. M., Reddy, C. C. & Maquat, L. E. Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA. *Mol. Cell. Biol.* **18**, 2932–9 (1998).
58. Chapin, A. *et al.* In vivo determination of direct targets of the nonsense-mediated decay pathway in *Drosophila*. *G3 (Bethesda)*. **4**, 485–96 (2014).
59. Martins, R. *et al.* Alternative polyadenylation and nonsense-mediated decay coordinately regulate the human HFE mRNA levels. *PLoS One* **7**, 1–12 (2012).
60. Holstein, E. M., Clark, K. R. M. & Lydall, D. Interplay between nonsense-mediated mRNA decay and DNA damage response pathways reveals that Stn1 and Ten1 are the key CST telomere-cap components. *Cell Rep.* **7**, 1259–1269 (2014).
61. Bhuvanagiri, M., Schlitter, A. M., Hentze, M. W. & Kulozik, A. E. NMD: RNA biology meets human genetic medicine. *Biochem. J.* **430**, 365–77 (2010).
62. Linde, L. & Kerem, B. Nonsense-Mediated mRNA Decay and Cystic Fibrosis. *Methods Mol. Biol.* **741**, 137–154 (2001).
63. Hall, G. W. & Thein, S. Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia. *Blood* **83**, 2031–7 (1994).
64. Higgs, D. R., Engel, J. D. & Stamatoyannopoulos, G. Thalassaemia. *Lancet* **379**, 373–383 (2012).

65. Galanello, R. & Origa, R. Beta-thalassemia. *J. Rare Dis.* **5**, (2010).
66. Romão, L. *et al.* Nonsense mutations in the human beta-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation. *Blood* **96**, 2895–2901 (2000).
67. Salvatori, F. *et al.* Production of Production of β -globin and adult haemoglobin following G418 treatment of erythroid precursor cells from homozygous β 039 thalassemia patients-globin and adult hemoglobin following G418 treatment of erythroid precursor cells from homozygous ?? *Am. J. Hematol.* **84**, 720–728 (2009).
68. Gaziev, J., Sodani, P., Polchi, P., Andreani, M. & Lucarelli, G. Bone marrow transplantation in adults with thalassemia: Treatment and long-term follow-up. *Ann. N. Y. Acad. Sci.* **1054**, 196–205 (2005).
69. Welch, E. M. *et al.* PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **447**, 87–91 (2007).
70. Keeling, K. M. & Bedwell, D. M. Suppression of nonsense mutations as a therapeutic approach to treat genetic diseases. *Wiley Interdiscip. Rev. RNA* **2**, 837–852 (2011).
71. McElroy, S. P. *et al.* A Lack of Premature Termination Codon Read-Through Efficacy of PTC124 (Ataluren) in a Diverse Array of Reporter Assays. *PLoS Biol.* **11**, (2013).
72. Hirawat, S. *et al.* Safety, tolerability, and pharmacokinetics of PTC124, a nonaminoglycoside nonsense mutation suppressor, following single- and multiple-dose administration to healthy male and female adult volunteers. *J. Clin. Pharmacol.* **47**, 430–444 (2007).
73. Salvatori, F. *et al.* Production of β -globin and adult hemoglobin following G418 treatment of erythroid precursor cells from homozygous β 039 thalassemia patients. *Am. J. Hematol.* **84**, 720–728 (2009).
74. Goldmann, T., Overlack, N., Wolfrum, U. & Nagel-Wolfrum, K. PTC124-mediated translational readthrough of a nonsense mutation causing Usher syndrome type 1C. *Hum. Gene Ther.* **22**, 537–547 (2011).
75. Tan, L., Narayan, S. B., Chen, J., Meyers, G. D. & Bennett, M. J. PTC124 improves readthrough and increases enzymatic activity of the CPT1A R160X nonsense mutation. *J. Inherit. Metab. Dis.* **34**, 443–447 (2011).
76. Du, M. *et al.* PTC124 is an orally bioavailable compound that promotes suppression of the human CFTR-G542X nonsense allele in a CF mouse model. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 2064–9 (2008).

77. Lai, C.-H. *et al.* Correction of ATM Gene Function by Aminoglycoside-Induced Read-Through of Premature Termination Codons. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15676–81 (2004).
78. Lu, J. & Deutsch, C. Electrostatics in the Ribosomal Tunnel Modulate Chain Elongation Rates. *J. Mol. Biol.* **384**, 73–86 (2008).
79. Ito-Harashima, S., Kuroha, K., Tatematsu, T. & Inada, T. Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes Dev.* **21**, 519–524 (2007).
80. Akimitsu, N., Tanaka, J. & Pelletier, J. Translation of nonSTOP mRNA is repressed post-initiation in mammalian cells. *EMBO J.* **26**, 2327–2338 (2007).
81. Isken, O. & Maquat, L. E. Quality control of eukaryotic mRNA: Safeguarding cells from abnormal mRNA function. *Genes Dev.* **21**, 1833–1856 (2007).
82. Klauer, A. A. & van Hoof, A. Degradation of mRNAs that lack a stop codon: A decade of nonstop progress. *Wiley Interdiscip. Rev. RNA* **3**, 649–660 (2012).
83. Frischmeyer, P. a, Hoof, A. Van & Dietz, H. C. An mRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination. **295**, 2258–2262 (2002).
84. Sparks, K. a. & Dieckmann, C. L. Regulation of poly(A) site choice of several yeast mRNAs. *Nucleic Acids Res.* **26**, 4676–4687 (1998).
85. Van Hoof, A. Exosome-Mediated Recognition and Degradation of mRNAs Lacking a Termination Codon. *Science* **295**, 2262–2264 (2002).
86. Carr-Schmid, A., Durko, N., Cavallius, J., Merrick, W. C. & Kinzy, T. G. Mutations in a GTP-binding motif of eukaryotic elongation factor 1A reduce both translational fidelity and the requirement for nucleotide exchange. *J. Biol. Chem.* **274**, 30297–30302 (1999).
87. Shoemaker, C. J., Eyler, D. E. & Green, R. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop off to initiate no-go decay. *Science* **330**, 369–372 (2010).
88. Schilders, G., van Dijk, E., Raijmakers, R. & Pruijn, G. J. M. Cell and Molecular Biology of the Exosome: How to Make or Break an RNA. *Int. Rev. Cytol.* **251**, 159–208 (2006).
89. Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A. & Conti, E. Structure of the Active Subunit of the Yeast Exosome Core, Rrp44: Diverse Modes of Substrate Recruitment in the RNase II Nuclease Family. *Mol. Cell* **29**, 717–728 (2008).

90. Allmang, C. *et al.* Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **18**, 5399–5410 (1999).
91. Allmang, C., Mitchell, P., Petfalski, E. & Tollervey, D. Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res.* **28**, 1684–1691 (2000).
92. Bousquet-Antonelli, C., Presutti, C. & Tollervey, D. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102**, 765–775 (2000).
93. Kadaba, S. *et al.* Nuclear surveillance and degradation of hypomodified initiator tRNA Met in *S. cerevisiae*. *Genes Dev.* **18**, 1227–1240 (2004).
94. Januszyk, K. & Lima, C. D. The eukaryotic RNA exosome. *Curr. Opin. Struct. Biol.* **24**, 132–140 (2014).
95. Tomecki, R., Drazkowska, K. & Dziembowski, A. Mechanisms of RNA degradation by the eukaryotic exosome. *ChemBioChem* **11**, 938–945 (2010).
96. Lubas, M. *et al.* Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. *EMBO J.* **32**, 1855–68 (2013).
97. Malecki, M. *et al.* The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *TL - 32. EMBO J.* **32**, 1842–1854 (2013).
98. Thomas, M. P. *et al.* Apoptosis Triggers Specific, Rapid, and Global mRNA Decay with 3' Uridylated Intermediates Degraded by DIS3L2. *Cell Rep.* **11**, 1079–1089 (2015).
99. Astuti, D. *et al.* Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. *Nat. Genet.* **44**, 277–284 (2012).
100. Haas, G. *et al.* Identification of factors involved in target RNA-directed microRNA degradation. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw040
101. Peixeiro, I. *et al.* Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. *Nucleic Acids Res.* **40**, 1160–1173 (2012).
102. Needham, M. *et al.* LCR/MEL: a versatile system for high-level expression of heterologous proteins in erythroid cells. *Nucleic Acids Res* **20**, 997–1003 (1992).

ANNEXES

ANNEX I

Annex I – Primer sequences for the FLAG-tag cloning SOEing PCRs and the respective length of the reactions products.

SOEing PCR number	Primer forward (5' – 3')	Primer reverse (5' – 3')	Product length (base pairs)
1	CAACGTGCTGGTCTGTG TG	CTTGTAGTCGTGATACTT GTGGGCCAGGGC	118 bp
2	AAGTATCACGACTACAA GGATGACGATGAC	GCGAGCTTATTTGTCAT CGTCATCCTTGTA	57 bp
3	GATGACAAATAAGCTCG CTTTCTTGCTGTC	TATGGAAAAACGCCAGC AAC	413 bp
4	CAACGTGCTGGTCTGTG TG	TATGGAAAAACGCCAGC AAC	571 bp