UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA

The mechanism of nonsense-mediated mRNA decay

Mestrado em Bioquímica
Especialização em Bioquímica

Nuno Filipe Teixeira Costa

Dissertação orientada por:
Doutora Luísa Romão
Doutora Juliane Menezes

2015

Acknowledgments

Firstly, I would like to express my sincere thanks to Doutora Luísa Romão Loison for giving me the opportunity to work in her research group. Her knowledge, motivation and immense patience helped me in all the time of research.

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

To Juliane Menezes, I want to express my deep sense of gratitude. Always available, even at cost of her own work, she provided excellent technical and theoretical teaching and good advice that not only helped me during this research but will surely help throughout my future career.

I also want to thank Gerson Asper, Paulo Costa, Rafaela Lacerda, Cláudia Onofre and Ana Ramos who provided insight and expertise that greatly assisted this research.

Index

Figures

Tables

Resumo

A expressão genética representa uma das vias mais complexas e importantes para a célula, que culminam na síntese proteica. Como tal, quaisquer erros que possam surgir durante estes eventos podem ser rapidamente amplificados e terem um impacto severo na célula. Portanto surgiram diversos mecanismos de controlo de qualidade para assegurar a fidelidade da expressão genética. Entre estes mecanismos está o decaimento do mRNA mediado por mutações *nonsense* (na sigla inglesa NMD – *nonsense mediated decay*) que permite rapidamente degradar transcritos que contenham codões de terminação de tradução prematuros (na sigla inglesa PTC – *premature translation termination codon*), reduzindo a sua abundância[1]. A eliminação destes transcritos anómalos evita a formação e acumulação de proteínas truncadas no C-terminal potencialmente tóxicas que, caso contrário, danificariam a célula. Assim, o NMD tem um papel protector contra erros genéticos, muitos dos quais originam PTCs. Recentemente, tem-se tornado óbvio que o NMD tem um papel mais abrangente, não só funcionando como mecanismo de controlo de qualidade da expressão genética, mas também apresenta uma importante função na regulação de transcritos fisiológicos[2,3]. De facto foi demonstrado que o NMD controla a abundância de 3 a 10% de todo o transcriptoma em *Saccharomyces cerevisiae*, *Drosophila melanogaster* e em *Homo sapiens*, pois estes transcritos possuem várias características que os tornam sensíveis à acção do NMD[4].

Nas células humanas, este mecanismo envolve a acção dos factores UPF1 (*upframe shift*), UPF2, UPF3 e dos factores SMG1 (*suppressor with morphogenetic effects on genitalia*), SMG5, SMG6 e SMG7[5]. O NMD é activado assim que um PTC é identificado. Os factores UPF1, UPF2, UPF3 e SMG1 são recrutados e interagem com o ribossoma e com os factores de terminação de tradução eRF3 e eRF1 (*eukaryotic release factor*), formando o complexo DECID (*decay inducing complex*)[6,7]. Este complexo estimula a fosforilação do UPF1[8], o que conduz ao recrutamento dos factores SMG5/SMG7 ou SMG6, que activam a degradação do ácido ribonucleico (RNA) pela acção de exossomas[5].

Grande parte da pesquisa sobre o mecanismo do NMD tem-se focado na forma como este mecanismo diferencia um PTC de um codão de terminação normal. De acordo com o modelo prevalente, nas células dos mamíferos a capacidade do NMD reconhecer um PTC depende dos complexos de junção exão-exão (na sigla inglesa EJC). Os EJC são complexos multiproteicos depositados a 20 a 24 nucleótidos (nt) de distância a montante dos locais de excisão de intrões (junções exão-exão) durante o *splicing* do mRNA precursor[9]. Estes complexos mantêm-se associados ao mRNA durante o seu transporte para o citoplasma, servindo como marcadores dos locais de excisão de intrões e também como plataformas de ligação dos factores do NMD, UPF2 e UPF3[10–12]. Durante o primeiro ciclo da tradução, se o transcrito não tiver PTC, o complexo ribossomal dissocia os EJCs para fora da grelha de leitura até chegar ao codão de terminação. No entanto, se o transcrito possuir um PTC e este se localizar a pelo menos 50 a 55 nt de distância a montante da última junção exão-exão, a elongação é terminada prematuramente e pelo menos um EJC fica associado ao mRNA[13]. Nestas condições, o EJC facilita o recrutamento e a interacção dos factores do NMD (UPF1, UPF2 e UPF3) com o complexo ribossomal, activando o NMD. Por outro lado, se o PTC se localizar depois desta fronteira localizada a 55 nts a montante da última junção exão-exão, o ribossoma consegue remover todos os EJCs, que não poderão, assim, desencadear a degradação do transcrito[14].

Contudo esta regra posicional do EJC não explica todos os transcritos susceptíveis de serem degradados por NMD. Existem transcritos que mesmo estando nas condições previstas pelo modelo (i.e.: no qual o PTC está localizado a pelo menos 50-55 nt de distância a montante da última junção exão-exão), conseguem resistir à acção do NMD, como é o caso de mRNAs que contenham PTC próximos do codão de iniciação[15]. De facto, verificou-se que existem outras características do mRNA, para além dos EJC, que influenciam a activação do NMD. Por exemplo, trabalhos feitos pelo grupo laboratorial onde este projecto experimental foi desenvolvido, apontam para a distância física entre um PTC e a proteína citoplasmática de ligação à cauda poli-A do mRNA (na sigla inglesa PABPC1) como sendo um factor que afecta a activação do NMD[16]. Foi demonstrado que a PABPC1 consegue interagir com a proteína eRF3 e pensa-se que a PABPC1 é necessária para estimular um processo de terminação correcto e eficiente[17,18]. Para além disso, a PABPC1 consegue competir com a UPF1

pela ligação ao eRF3, inibindo o NMD, mesmo na presença de um EJC a jusante[16,19]. Propôs-se assim um novo modelo: se o ribossoma ao chegar ao PTC ficar fisicamente perto da PABPC1, ocorre terminação correcta e não há indução do NMD. No caso em que o ribossoma ao chegar ao PTC não possa interagir com a PABPC1, a UPF1 interage com o complexo eRF3/eRF1 e inicia-se o NMD[11]. Portanto, tendo em consideração este modelo e sabendo que o mRNA adquire uma conformação circular devido à interacção do factor eucariótico de iniciação 4G (na sigla inglesa eIF4G) com a estrutura *cap* na extremidade 5', a PABPC1, o complexo eIF3 e o ribossoma, pode-se colocar a hipótese de a PABPC1 ser arrastada para a vizinhança do codão de iniciação pelo ribossoma durante o *scanning*. Nestas condições a PABPC1 encontra-se próxima o suficiente para competir com a UPF1, caso ocorra terminação prematura na vizinhança do codão de iniciação, permitindo assim explicar a resistência ao NMD dos transcritos contendo PTCs próximos do codão de iniciação.

Esta tese teve como objectivo testar esta hipótese. Procedeu-se á construção de plasmídeos necessários para estudar a interacção do PABPC1 com o ribossoma. Os plasmídeos contêm a sequência da β-globina (que origina um pequeno transcrito de três exões cujas mutações *nonsense* já foram extensivamente descritas) e várias repetições da sequência do sítio de ligação da cápside do bacteriófago MS2. Estas sequências são aptámeros de RNA naturais que em conjunto com uma proteína recombinante contendo o N-terminal da proteína da cápside do MS2 e um anticorpo específico, permitem a imunoprecipitação selectiva destes transcritos e de qualquer proteína associada. A construção dos plasmídeos foi iniciada usando um processo de *SOEing PCR*.

O factor de iniciação da tradução eIF3 é um complexo multiproteico que interage com o ribossoma e a proteína eIF4G, funcionando como um ponte entre os dois durante o processo de iniciação da tradução. Trabalhos feitos pelo nosso grupo laboratorial mostraram que as subunidades eIF3h e eIF3f podem estar envolvidas na interacção do ribossoma com a PABPC1 (através da eIF4G)[20] e portanto podem ser necessárias ao deslocamento da PABPC1 para a vizinhança do codão de iniciação pelo ribossoma. Para testar o papel das subunidades eIF3h e eIF3f na ligação entre o complexo ribossomal 40S e a PABPC1 (através do eIF4G), isolou-se por imunoprecipitação, utilizando anticorpos anti-eIF3h ou eIF3f, os complexos

ribonucleoproteícos (na sigla inglesa mRNPs) de células HeLa tratadas com RNAs de interferência (na sigla inglesa siRNA) específicos para as subunidades eIF3h ou eIF3f, respectivamente. Os imunoprecipitados foram testados por *Western Blot* para detectar a presença de PABPC1, eIF4G e proteínas do ribossoma. Nas imunoprecipitações de mRNPs de células HeLa tratadas com siRNAs para a subunidade eIF3f, verificou-se uma ligeira diminuição dos níveis proteicos tanto da PABPC1 como do eIF4G, suportando um modelo no qual a eIF3f é responsável pela aproximação do complexo PABPC1/eIF4G ao ribossoma. Já a imunoprecipitação de mRNPs de células tratadas com siRNA para a subunidade eIF3h, não se observaram variações significativas na expressão da proteína eIF4G e não foi detectada a proteína PABPC1. Contudo a ausência de sinal da PABPC1 pode ter ocorrido devido à perca de proteína durante a remoção dos anticorpos para reutilização da membrana de PVDF. Além disso, a eficiência do silenciamento das subunidades não foi satisfatória, resultando numa diminuição pouco apreciável dos níveis proteicos de eIF3h ou eIF3f. Assim sendo, tanto o processo de imunoprecipitação e o método de silenciamento devem ser optimizados para que se possam tirar conclusões definitivas.

Palavras-chave

mRNA, tradução eucariótica, codão de terminação prematuro, decaimento do mRNA mediado por codões *nonsense*, proteína de ligação à cauda poli-A, factor 3 de iniciação da tradução

Abstract

Nonsense-mediated mRNA decay (NMD) is a quality control pathway that recognizes and rapidly degrades mRNAs containing a premature termination codon (PTC). The prevalent model, proposes that in mammals a translation termination codon is generally interpreted as premature if it is localized more than 50-55 nucleotides upstream of the last exon-exon junction. In these circumstances, during translation, the ribosome is not able to displace the exon junction complexes (EJCs) located downstream of the PTC; thus, when the ribosome reaches the PTC stalls and triggers NMD. However, it has been reported that mRNAs containing PTCs in close proximity to the translation initiation codon (AUG proximal PTCs) can substantially evade NMD, despite the presence of downstream EJCs. Recently, it was demonstrated that the cytoplasmatic poly(A)-binding protein 1 (PABPC1) is capable of inhibiting NMD when placed close to a PTC, even in the presence of EJCs. Taking this into account and that the mRNA acquires a circular conformation due to the interactions of PABPC1 with the scaffold protein eIF4G at the 5'-Cap, eIF3 (eukaryotic initiation factor 3) and the 40S ribosome, we believe that the PABPC1 is relocated into the AUG vicinity during 40S ribosome scanning. In these conditions the PABPC1 would be placed near AUG proximal EJCs and could be able to prevent NMD activation. To prove this hypothesis, we tried to construct by SOEing PCR, the necessary plasmids to demonstrate how PABPC1 is relocated to the AUG vicinity and that the interaction of PABPC1 and eIF4G with the ribosome remains stable during the first steps of elongation. Furthermore, we explored the eIF3h and eIF3f subunits role in bridging the 40S ribosome and the PABPC1/eIF4G complex. For that, HeLa cells were treated with specific short interfering RNAs for these eIF3 subunits and immunoprecipitation assays were performed with antibodies against eIF3h and f, respectively, to test the proteins, such as PABPC1, eIF4G and ribosomal proteins. However, the subunits silencing was not satisfactory and therefore, the knockdown procedure should be optimized before any decisive conclusions can be reached.

Keywords

mRNA, eukaryotic translation, premature termination codon, nonsense-mediated mRNA decay, cytoplasmic poly-A binding protein 1, eukaryotic initiation factor 3

Abbreviations

A	adenine
A-site	aminoacyl-site
ATP	adenosine triphosphate
bp	base pair
BTZ	Barentz
C	cytosine
CBP	cap binding protein
cDNA	complementary DNA
CHX	cycloheximide
COP9	constitutive photomorphogenesis 9
C-terminal	carboxyl-terminal
DEAD	Asp-Glu-Ala-Asp motif
DECID	decay-inducing complex
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
eEF	eukaryotic translation elongation factor
eIF	eukaryotic translation initiation factor
EJC	exon junction complex
eRF	eukaryotic translation release factor
E-site	exit-site
FBS	fetal bovine serum
G	guanine
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GDP	guanosine diphosphate
GFP	green fluorescent protein
GTP	guanosine triphosphate
HRP	horseradish peroxidase
Ig	immunoglobulin
IP	immunoprecipitation
m7G	7-methylguanosine
MAGOH	mago-nashi homolog
Met	methionine
Met-tRNAi	methionine-loaded initiator tRNA
MPN	Mpr1-Pad1-N-terminal
mRNA	messenger ribonucleic acid
mRNP	messenger ribonucleoprotein particle
MS2bs	bacteriophage MS2 binding site
MS2cp	bacteriophage MS2 coat protein
NMD	nonsense-mediated mRNA decay
NP40	nonidet-P40
nt	nucleotide
N-terminal	amino-terminus

ORF	open reading frame
PABP	poly(A)-binding protein
PABPC1	cytoplasmic poly(A)-binding protein 1
PABPN1	nuclear poly(A)-binding protein 1
PAGE	polyacrylamide gel electrophoresis
PAN	poly(A) nuclease
PCR	polymerase chain reaction
PCI	proteasome, COP9 signalosome, eIF3
Pi	inorganic phosphate
PIC	pre-initiation complex
Poly(A)	poly-adenylate
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
RNase	ribonuclease
RPS6	ribosomal protein small subunit 6
RT	reverse-transcription
SDS	sodium dodecyl sulphate
siRNA	short interfering RNA
SMG	suppressor with morphogenetic effects on genitalia
SQ-PCR	semiquantitative PCR
SOEing PCR	gene splicing by overlap extension PCR
ssDNA	single stranded DNA
SURF	SMG1-UPF1-eRFs complex
T	thymine
tRNA	transfer ribonucleic acid
U	uracil
UPF	up-frameshift
UTR	untranslated region
WT	wild type
XRN1	5’-3’ exoribonuclease 1
α-tub	α-tubulin
β15	human β-globin gene with PTC at codon 15
β23	human β-globin gene with PTC at codon 23
β25	human β-globin gene with PTC at codon 25
β26	human β-globin gene with PTC at codon 26
β39	human β-globin gene with PTC at codon 39
βWT	normal human β-globin gene

1. Introduction

Eukaryotic gene expression consists of a multistep pathway that involves a series of highly regulated and interconnected events in which the mRNA (messenger ribonucleic acid) is a central player. Newly synthesized mRNA precursor transcripts (pre-mRNA) are subjected to removal of introns (splicing) and assembling with a set of RNA-binding proteins to form highly compact mRNA ribonucleoprotein particles (mRNPs)[21,22]. The content of the mRNPs evolves dynamically throughout the transcript lifetime, in response to diverse cellular signal networks, mediating mRNA cellular localization, translation and decay. Such plasticity is highly important for cellular homeostasis, because it provides the means to regulate gene expression, allowing the cells to quickly adapt to changes in their environment by altering the patterns of gene expression[23,24].

During eukaryotic transcription, the pre-mRNA is rapidly modified into a complex that contains a methylated cap structure bound to the cap-binding protein (CBP) heterodimer CBP80/CBP20 complex at the 5' end and a poly(A)-tail at the 3' end, bound to the nuclear poly(A)-binding protein 1 (PABPN1). These modifications not only protect the newly synthesized RNA transcript from premature degradation, but also facilitate translation initiation in the cytoplasm[23,25]. The pre-mRNA can also be subjected to splicing to remove introns and ligate the exons together. During this process, a complex of proteins called the exon junction complex (EJC) is deposited at 20–24 nuclotides (nt) upstream of every exon–exon splice site within the mRNP[9]. The EJC is a dynamic structure with a heterogeneous protein composition that evolves throughout the mRNA life cycle. Some of the EJC components dissociate before mRNA export, while others only bind in the cytoplasm. However, it is thought that a core of EJC components comprised of Y14, Mago–Nashi homologue (MAGOH), DEAD-box RNA helicase eukaryotic initiation factor 4A3 (eIF4A3) and Barentsz (BTZ; also known as metastatic lymph node 51: MLN51) remain associated with the mRNA escorting it to the cytoplasm, serving as a marker of the intron excision location, for the translational machinery[10,11].

Once proper 5' and 3' modifications and splicing has been concluded, the mRNA, assembled as an mRNP, migrates through the physical barrier between the nucleus and cytoplasm, to reach the ribosomes and translation factors[26]. At this stage, the 5' cap of the majority of mRNAs are bound by the nuclear CBP80/CBP20 complex, whereas the 3' poly(A) tail carries a mixture of nuclear (PABPN1) and cytoplasmic (PABPC1) poly(A) binding proteins. CBP80/CBP20 interacts with the eukaryotic translation initiation factor 4G (eIF4G), promoting the recruitment of the small ribosomal subunit 40S and initiate 5'→3' scanning along the 5' untranslated region (UTR) for a start codon (AUG). Once the start codon is identified, the large ribosomal subunit 60S is engaged to form an 80S complex competent for protein synthesis, initiating the "pioneering round" of translation.[27] In this first passage, the ribosome scans the mRNA, displacing any remaining nuclear-acquired mRNP proteins, such as EJCs, residing inside the open reading frame (ORF)[14,28,29]. At some point, CBP80/CBP20 and PABPN1 are completely replaced by the major cytoplasmic cap binding protein (eIF4E) and PABPC1, respectively, which directs the steady state of translation, supporting the bulk of cellular protein synthesis[30].

1.1. Eukaryotic mRNA Translation

The translation process can be divided into four distinct stages, initiation, elongation, termination and ribosome recycling[31], each requiring a particular set of conditions and factors.

1.1.1. Translation initiation – the scanning model

Translation initiation is the rate-limiting step of translation and, in eukaryotic cells, requires the participation of several eukaryotic initiation factors (eIF)[32]. The initiation codon recognition generally occurs through a scanning mechanism, wherein every triplet in the mRNA sequence is inspected for complementarity to the anticodon of methionyl initiator transfer RNA (Met-tRNAi). This process begins with the assembly of the Met-tRNAi and the GTP-bound (guanosine triphosphate) form of eIF2 into the

small ribosomal subunit 40S, in a process stimulated by eIF 1, 1A, 5, and the eIF3 complex resulting in the 43S pre-initiation complex (PIC)[33,34]. This structure is then deposited near the 5'-cap with the support of the PABPC1, the eIF3 complex, and the eIF4F complex (which is comprised by the cap-binding protein eIF4E, the eIF4G, and the RNA helicase eIF4A) (fig. 1)[33,35,36].

Mammalian eIF4G is a scaffold protein that harbors binding domains for both PABP and eIF4E in its N-terminus and eIF4A and eIF3 in the middle and C-terminal region, respectively[37]. The interaction between eIF4G and the eIF3 complex establishes a protein bridge between the 43S PIC and the mRNA (fig. 1), while the ATP-dependent helicase activity of eIF4A unwinds the mRNA secondary structure removing secondary structures that impede ribosome attachment, promoting 43S PIC deposition in the mRNA and subsequent scanning[37,38]. Simultaneous, the interaction of eIF4E and PABP with eIF4G brings the 3' UTR in close proximity with 5' UTR, giving the mRNP a circular conformation (fig. 1)[39]. This "closed-loop" structure is thought to protect both ends of the transcript from the mRNA degradation machinery, promotes translational control by regulatory elements in the 3' UTR and stimulates ribosome initiation[17,39].

Once bound near the 5'-cap, the 43S PIC scans the mRNA for an AUG codon in an optimum/consensus context (Kozak context), typically the first AUG triplet with a purine at the -3 and a G at the +4 positions, relative to the A of the AUG codon, which is designated +1[40]. Correct base-pairing between the anticodon of Met-tRNAi and the AUG in the peptidyl-tRNA (P) site of the 40S subunit induces arrest of the scanning PIC. The GTPase-activating factor eIF5 then triggers irreversible hydrolysis of the GTP bound to eIF2, resulting in the release of eIF2-GDP (guanosine diphosphate) and several other eIFs present in the PIC. Finally, loading of the large ribosomal subunit 60S is catalyzed by eIF5B, resulting in an 80S ribosome ready to begin the elongation phase(fig. 1)[31,41,42].

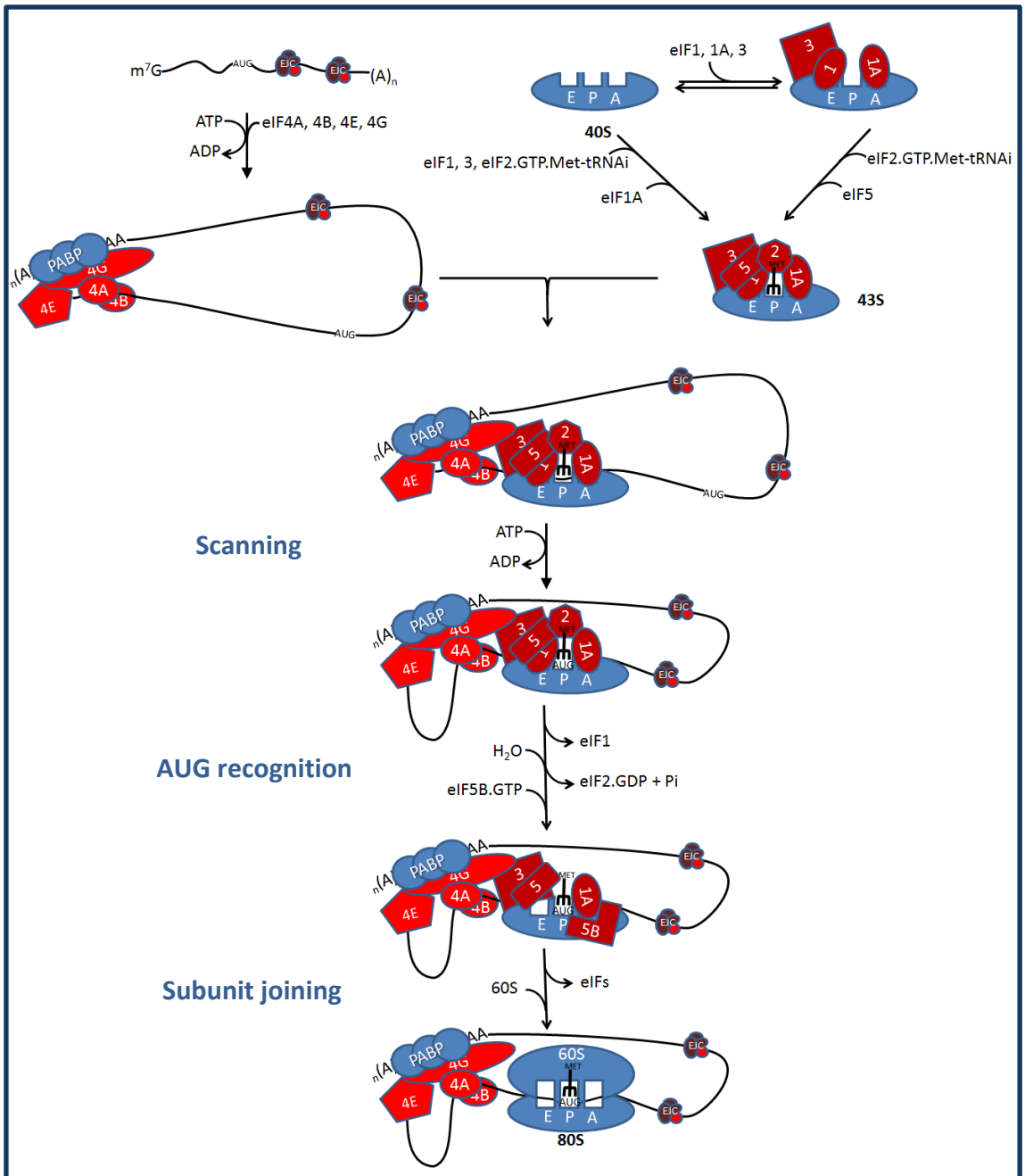

Figure 1: Model of canonical eukaryotic translation initiation pathway (modified from Hinnebusch, 2012)[42]. This pathway begins with loading of the methionine-loaded initiator tRNA bound to the eIF2-guanosine triphosphate (eIF2.GTP.Met-tRNA) into the 40S ribosomal complex, a process that is stimulated by the eukaryotic initiation factors (eIF) 1, 1A, 3 and 5. The resulting 43S preinitiation complex is then deposited near the 5'cap of an activated mRNA, where it scans the mRNA until an AUG codon in a proper context is recognized. Subsequent downstream steps, including eIF2 hydrolysis, induce eIFs release and 60S joining.

1.1.2. Translation elongation

In general, the mechanism of translation elongation is well conserved between eukaryotes and bacteria, and most of our knowledge of the mechanistic details is based on prokaryotic systems[43,44]. The elongation stage, is characterized by the sequentially addition of amino acids to the carboxy-terminal end of the growing polypeptide chain. The string of mRNA codons (read in the 5'→3' direction) determines the order in which amino acids are added to form the primary amino acid sequence of each protein[45].

The ribosome has three tRNA-binding sites: the A- (aminoacyl) site, which receives the incoming aminoacyl-tRNA, the P- (peptidyl) site, which harbors the tRNA with the nascent peptide chain and the E- (exit) site through which the deacylated tRNA leaves the ribosome[46]. Following translation initiation, the 80S ribosome P-site, holds an Met-tRNAi base-paired with the start codon[31,47], while the A-site surrounds the second codon of the ORF, awaiting binding of a cognate (carrying an anticodon with a correct match for the codon) aminoacyl tRNA[48] (fig. 2). Because the codon in the A-site could represent any of the 64 candidates of the genetic code, a process of aminoacyl-tRNA sampling occurs until a cognate tRNA is selected[49]. Various aminoacyl-tRNAs, guided by an accompanying eukaryotic elongation factor 1A (eEF1A) in a GTP-dependent manner, successively enter the A-site (fig. 2). If the tRNA is not a proper match, the tRNA-eEF1a complex dissociates as a unit and aminoacyl-tRNA sampling continues. If correct codon–anticodon interactions are achieved, GTP hydrolysis is activated by eEF1A. This induces a major conformational change in eEF1A, releasing the factor and enabling the aminoacyl-tRNA to be accommodated into the A-site[48,49]. The nascent peptide chain, in P-site, is then transferred to the amino acid of the A-site aminoacyl-tRNA and the translocation of peptidyl-tRNA from A- to P- and of deacylated tRNA from P- to E- sites occurs (fig. 2). This translocation step is promoted by eEF2 in a GTP dependent manner[45]. Once translocation occurs, the ribosomal A-site becomes available for binding of the next aminoacyl-tRNA in a complex with eEF1A. The entire process is then repeated, until a stop codon is encountered.

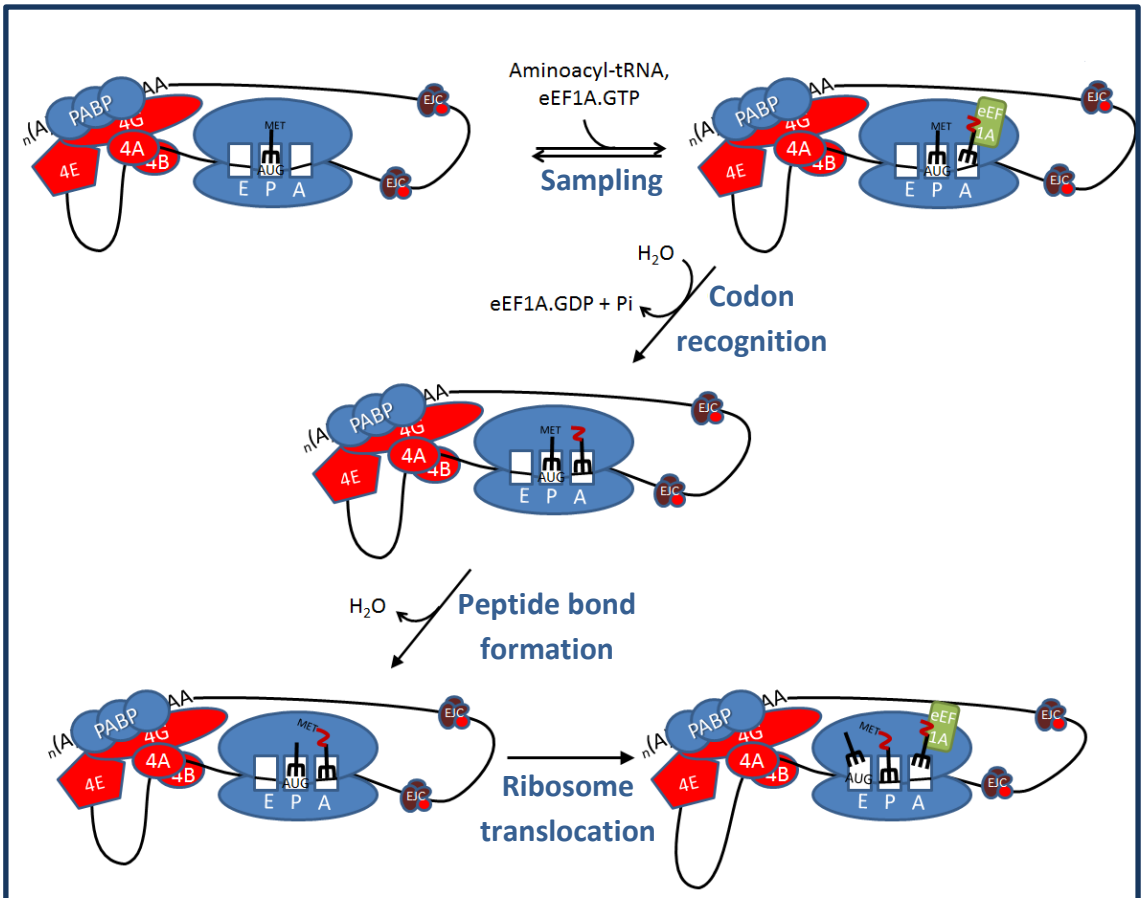

Figure 2: Model of eukaryotic translation elongation pathway (based on Keeling, 2014)[49]. In this pathway an aminoacyl-tRNA is carried by the eukaryotic elongation factor 1A (eEF1A) to the A-site of the ribosome. If the aminoacyl-tRNA possesses a cognate anticodon for the codon enclosed in the ribosomal aminoacyl-site (A-site), eEF1A hydrolysis is triggered allowing aminoacyl-tRNA accommodation within the ribosome. Subsequently, the nascent peptide chain, in the peptidyl-site (P-site), is transferred to the aminoacyl-tRNA in the A-site and ribosome translocation occurs.

1.1.3. Translation termination

Notably, stop codon recognition is the only step where a protein, the eukaryotic release factor 1 (eRF1), instead of a nucleic acid adaptor (tRNA) serves as the recognizing factor. Translation termination begins when a stop codon (UAA, UGA, or UAG) enters the A-site, stalling the ribosome (fig. 3). eRF1 adopts a tRNA-like structure and is able to recognize the termination codon, through its N-terminus, while the C-terminus forms a complex with the C-terminus of the GTPase eRF3. This interaction triggers GTP hydrolysis, inducing a conformational change positioning eRF1 closer to the P-site. Thus, allowing eRF1 to stimulate hydrolysis of the ester bond of

the peptidyl-tRNA, facilitating the release of the completed polypeptide chain (fig. 3)[5,50].

The eRF3 also interacts with the C-terminal domain of PABPC1, through its N-terminal domain[5]. The functional consequence of the PABPC1-eRF3 interaction is not yet fully understood. It has been shown in yeast that when the interaction between the eRF3 orthologue Sup35 and Pab1p is impaired, the terminating ribosome cannot efficiently dissociate from the mRNA[51]. It was also demonstrated that mammalian cells lacking PABPC1 exhibited increased read-through of termination codons[52]. Therefore it is believed that PABPC1 stimulates proper and efficient translation termination[17,18].

When termination is completed and both eRF3 and the polypeptide chain are released, the termination factor ABCE1 is able to interact with eRF1 and the stalled ribosome, triggering dissociation and recycling of the ribosomal subunits[53].

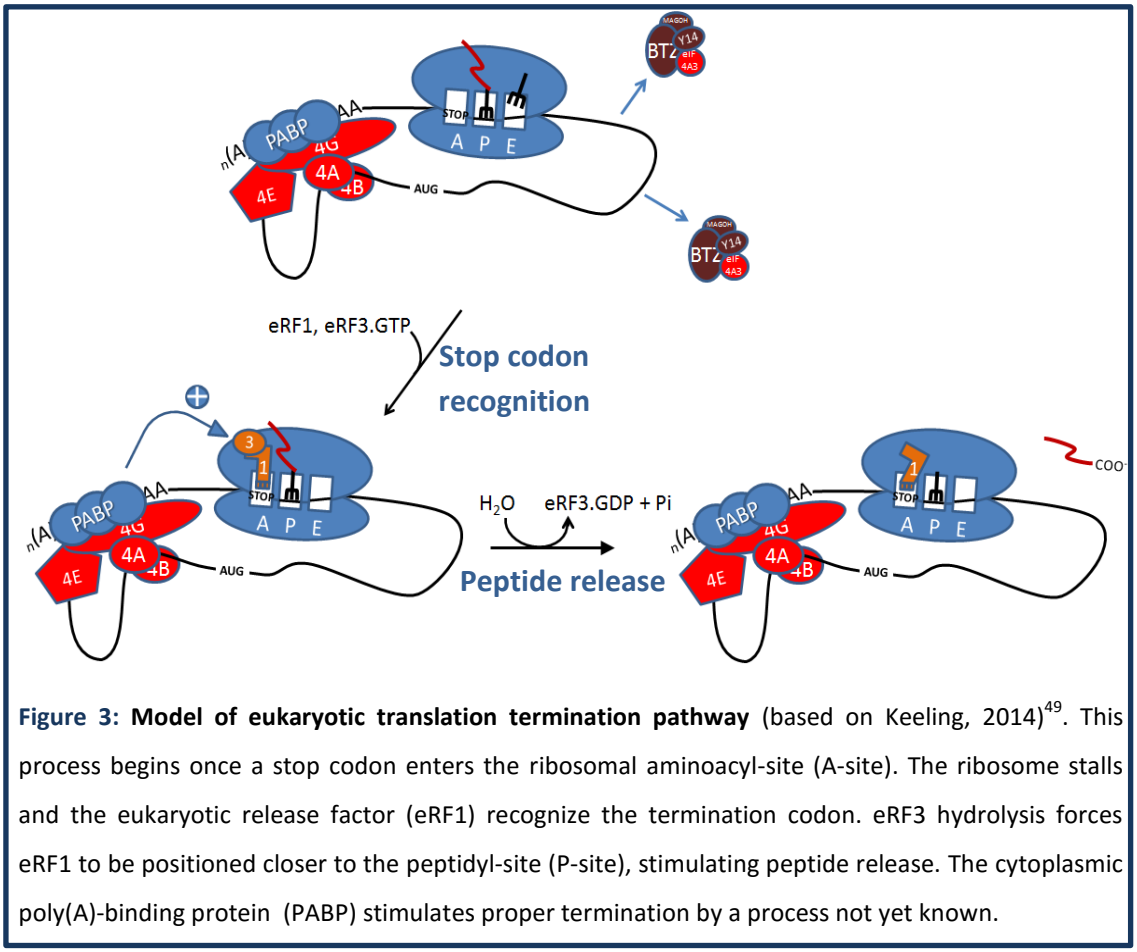

Figure 3: Model of eukaryotic translation termination pathway (based on Keeling, 2014)[49]. This process begins once a stop codon enters the ribosomal aminoacyl-site (A-site). The ribosome stalls and the eukaryotic release factor (eRF1) recognize the termination codon. eRF3 hydrolysis forces eRF1 to be positioned closer to the peptidyl-site (P-site), stimulating peptide release. The cytoplasmic poly(A)-binding protein (PABP) stimulates proper termination by a process not yet known.

1.2. Surveillance mechanisms

Throughout the mRNA biogenesis several errors may occur. Naturally, cells have developed surveillance mechanisms capable of detecting and stimulating degradation of mRNPs that were not properly assembled and/or harboring mutations within the ORF[1]. These surveillance mechanisms operate both in the nucleus and in the cytoplasm, assessing the translatability of the mRNA[54]. So far, three different translation-coupled mRNA surveillance systems have been identified in eukaryotes[50]. The most recently discovered mechanism, the no-go mRNA decay (NGD) induces degradation of mRNAs that have secondary structures that cause a complete block of the translation elongation.[1,55] The non-stop mRNA decay (NSD), appears to have evolved to stimulate degradation of mRNAs that lack a stop codon[56]. In both cases, there is no stop codon and hence no release factors engage with the stalled ribosome[1,50]. The third and the best characterized mechanism in eukaryotes, in which this work will focus, is the nonsense-mediated decay (NMD).

1.2.1. Nonsense-mediated decay

NMD was identified almost forty years ago, when it was observed that nonsense codons truncating the ORF of *ura3* mRNAs in *Saccharomyces cerevisiae* and the β-globin mRNAs in β^0-thalassemic patients dramatically reduced the RNA half-life[57,58]. This resulted in decreased abundance of the affected mRNA transcripts, rather than production of truncated proteins[57,58]. Fast degradation of mRNAs harboring truncated ORFs due to premature translation termination codons (PTCs) was subsequently reported in many other organisms, and it is believed to occur in most if not all eukaryotes[59]. Although important aspects of the NMD mechanism can differ between yeast, worms, flies and humans, the core process and proteins seem to be conserved. Thus, NMD is described as an evolutionary conserved posttranscriptional quality control mechanism that identifies and rapidly induces the degradation of faulty transcripts containing PTCs. Therefore, preventing the synthesis and accumulation of

C-terminally truncated proteins and protecting the cell from potentially deleterious dominant-negative or gain-of-function effects of these truncated proteins[5,60] .

1.2.1.1. NMD factors and mechanism

NMD takes place once a PTC is identified, leading to the recruitment and assembly of the surveillance complex. In human cells, this complex comprises the factors up-frameshift 1 (UPF1), UPF2 and UPF3 (which represents the conserved core of the NMD machinery and is thought to assemble in all organisms) and the suppressor with morphogenetic effects on genitalia 1 (SMG1), SMG5, SMG6 and SMG7[5].

UPF1 is an ATP-dependent RNA helicase and is the most conserved NMD factor and functionally the most important. It is currently unknown how UPF1 is recruited to the terminating ribosome, and if it is present in all termination events. However, once recruited, UPF1 interacts with eRF3 forming a complex with SMG1, eRF1 and eRF3, called the SURF complex (SMG1-UPF1-eRFs complex)[8]. UPF1 also binds to UPF2 and UPF3 forming the decay-inducing complex (DECID), which promotes UPF1 phosphorylation by the protein kinase SMG1 (fig. 4)[6,7]. As a result, UPF1 undergoes a conformational change that increases its affinity for RNA and induces the dissociation of eRF3 from UPF1[8]. Phosphorylated UPF1 can then recruit either the SMG7/SMG5 heterodimer or the endonuclease SMG6.

Interaction of SMG7/SMG5 with the phosphorylated UPF1 promotes deadenylation by the consecutive action of the complexes PAN2/PAN3 and CCR4/CAF1, followed by exonucleolytic degradation of the transcript by the exosome in 3'→5' direction[61,62]. While at the other end of the transcript, the heterodimer DCP1/DCP2 removes the cap structure, leaving the 5' end accessible for rapid degradation by the exonuclease XRN1 in 5'→3' direction (fig. 4)[24,63]. Alternatively, if SMG6 interacts with phosphorylated UPF1, endonucleolytic cleavage is elicited in the vicinity of the PTC and the resulting intermediates are then rapidly degraded in the 3'→5' direction by the exosome and in the 5'→3' direction by XRN1 (fig. 4)[64]. SMG7 and SMG6 also recruit the protein phosphatase 2a (PP2A), which promotes UPF1 dephosphorylation and dissociation from SMG7 or SMG6, enabling the recycle of these proteins for a new round of NMD[54].

In mammalian cells, multiple lines of evidence support the idea that NMD act exclusively on newly synthesized transcripts during the pioneer round of translation, before eIF4E replaces the CBP80/CBP20 complex[28]. It has been shown that short interfering RNA (siRNA) mediated depletion of CBP80 inhibits NMD[65,66], that the core NMD factors UPF1, UPF2, UPF3 and SMG1 co-immunoprecipitate with CBP80-associated mRNPs[28,67] and that levels of CBP80/CBP20-bound PTC-containing mRNA are lower than the levels of corresponding CBP80/CBP20-bound wild-type mRNA[28]. Additionally, CBP80 has also been documented to chaperone UPF1 to the terminating ribosome, located at the PTC, stimulating the SURF complex formation[68]. The first round of translation might even begin before the entire mRNP traverses the nuclear pore[14], with the ribosome acting as a proofreader, triggering NMD and hence eliminating faulty transcripts early on, before they are committed to the eIF4E-associated bulk translation[69]. In contrast, in *Saccharomyces cerevisiae* NMD occurs mainly on eIF4E-bound mRNAs, where the CBP80/CBP20-bound transcripts are largely insensitive to NMD and the yeast homologue of CBP80 is dispensable for NMD[70–72]. In fact, more recently, it has been demonstrated that eIF4E-associated nonsense-transcripts can also be efficiently degraded by NMD in mammalian cells[69,73]. Additionally, it was also observed that UPF1 co-precipitates with eIF4E in an RNA-dependent manner[69], indicating that NMD could be a regulated event that can be activated on substrates already engaged in translation, according to the cell needs. As suggested by Popp and Maquat[74], the apparent mammalian NMD restriction to CBP80/CBP20-bound transcripts may be the result of NMD occurring more rapidly than the replacement of CBP80 by eIF4E[74].

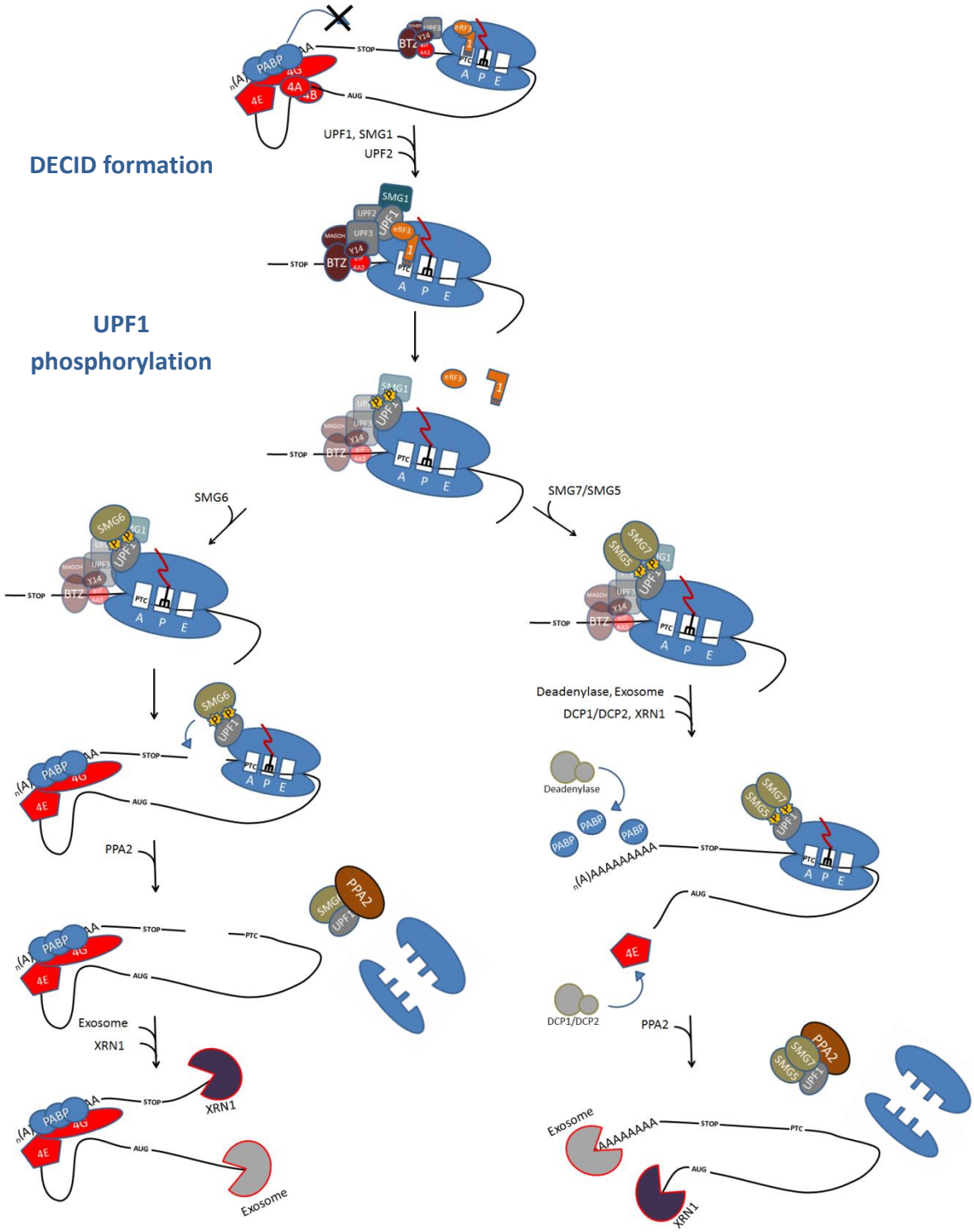

Figure 4: Model of nonsense mediated mRNA decay (NMD) mechanism and degradation stimulation via suppressor with morphogenetic effects on genitalia 6 (SMG6) or SMG7/SMG5 (modified from Nicholson, 2010)[5]. Once a premature translation termination codon (PTC) is identified, the NMD factors up-frameshift 1 (UPF1), UPF2, UPF3 and SMG1 interact with the eukaryotic release factor 3 (eRF3) and eRF1, forming the decay-inducing (DECID) complex. UPF2 and UPF3 recruitment and interaction with UPF1 is facilitated by the presence of a downstream exon junction complex. The DECID complex triggers UPF1 phosphorylation (represented as a P) inducing eRF3 release and recruitment of either SMG7/SMG5 or SMG6. Interaction of SMG7/SMG5 with UPF1 promotes deadenylation and deccaping, followed by mRNA degradation by exonucleases from the 5' and 3' ends. If SMG6 interacts with UPF1, the RNA is cleaved in the vicinity of the PTC and RNA is degraded in the direction of the 5' and 3' ends by exonucleases. Finally, protein phosphatase 2A (PP2A) recruitment enables UPF1, SMG7/5 and SMG6 recycling.

1.2.1.2. PTC definition

Trying to understand how exactly a stop codon is distinguished from a PTC has been the subject of intense ongoing research and discussion. Despite the core NMD factors appear to be conserved amongst species, several models for NMD PTC definition have been proposed from studies in mammalian systems compared with studies in *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster* and plants[75].

According to the prevailing model for mammalian NMD, a stop codon is classified as aberrant if is located more than 50-55 nucleotides upstream of the last exon-exon junction of a transcript. During the pioneer round of translation of a PTC free transcript, the 80S ribosome displaces from the ORF any EJC deposited in the mRNA, until it reaches a stop codon. Conversely, if the transcript possesses a PTC localized at more than 50-55 nucleotides upstream of the last exon-exon junction, the ribosome will stop before being able to remove the EJC[13]. Therefore, at least one EJC will remain bound to the mRNA and trigger recruitment of the NMD factors. The EJC also act as a binding platform for UPF2 and UPF3 (which is loaded onto mRNAs during splicing and represents a genuine EJC component)[12,76], greatly enhancing the interaction between UPF2 and UPF3 with the SURF complex due to their close proximity[5]. In contrast, if the PTC is located at less than 50-55 nucleotides upstream of the last exon-exon junction of a transcript, all EJCs are removed and the NMD is not activated[13].

However, recent data has challenged the generality of this EJC-dependent NMD model[77,78]. For example, the β-globin transcripts possessing a PTC near the AUG fail to trigger NMD, despite the existence of downstream EJCs[15,79]. These exceptions suggest that additional determinants may be involved. Indeed, there is evidence that, similar to what happens in yeast, the decision of whether NMD is to be triggered or not, relies upon competition between UPF1 and PABPC1 for binding to eRF3 on the terminating ribosome[19,80]. If PABPC1 is in close proximity to a stop codon, it interacts with the termination complex, stimulating proper translation termination, and represses NMD[18]. On the contrary, in the event that the spatial distance between the terminating ribosome and the poly(A) tail is too big, the interaction of PABPC1 with the

termination complex is reduced and UPF1 interacts with eRF3 triggering NMD[16,52]. Supporting this model, it has been shown that tethering of PABPC1 in the vicinity of a PTC abolishes NMD[81], even in the presence of a downstream EJC[16] and that native stop codons were found to elicit NMD when 3'UTR length is increased[82].

1.2.1.3. Natural and aberrant targets of NMD

Transcripts harboring PTCs can be generated at several stages during mRNA biogenesis. At the DNA level, PTC containing mRNAs can arise from germline or somatic alterations. Single base pair substitutions can change a sense codon to an in-frame PTC (nonsense mutation) or more frequently frame-shifting deletions and insertions alter the ribosomal reading frame, causing translation ribosomes to encounter a PTC. Mutations at splice sites or splicing regulatory sequences may result in inaccurate intron removal and create an intron-derived PTC or a frameshift[83,84]. In addition, non-faulty regulated processes such as programmed DNA rearrangements occurring in the immunoglobulin and T-cell receptor genes during lymphocyte maturation generate PTCs at a high frequency, due to random deletions and the addition of non-template nucleotides at the recombination sites[5]. NMD, therefore, has an important role eliminating by-products of programmed DNA rearrangements and in the differentiation and maintenance of hematopoietic cells[85].

At the RNA level, alternative splicing constitutes a major source of PTC containing mRNA. A genome wide analysis predicted that ~3,100 of 16,000 human genes examined should produce at least one alternative-splice product, one-third of which would contain PTCs[86]. In some cases, it constitutes a form of auto-regulation of the expression of the canonical protein-encoding isoforms, as the protein products of these genes are responsible for NMD triggering[87].

Other types of faulty mRNAs are transcripts from non-functional pseudogenes, endogenous retroviral and transposon RNAs or mRNA-like non-protein coding RNAs from intergenic regions[2,4]. This variety of targets infers an even broader role of the NMD pathway in dampening the ''transcriptional noise'' of supposedly non-functional

RNAs. Altogether, this advocates that cells produce a large number of faulty PTC mRNAs that are recognized and eliminated by NMD.

It has become clear during recent years that many physiological mRNAs that encode full-length functional proteins are also NMD substrates, indicating a role for NMD not only in mRNA quality control, but also as a translation-dependent post-transcriptional regulator of the steady state level of gene expression[2,3]. Translation regulation of existing mRNAs allows for a spatial and temporal fine-tuning of levels of the encoded proteins, allowing the maintenance of cellular homeostasis. In fact, several microarray studies comparing the mRNA levels of normal cells with NMD-deficient cells in *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Homo sapiens*) revealed that NMD directly and indirectly controls the abundance of 3–10% of the transcriptome in the respective cells[4].

Several features of physiological mRNAs can render them NMD-sensitive. Introns in the 3' UTR, transcripts containing regulatory ORFs that reside upstream of the primary ORF[41]; programmed frameshifts; or long 3' UTRs can activate NMD[88]. Interestingly, mRNAs containing UGA triplets that direct selenocysteine incorporation can also elicit NMD. When selenium is abundant, UGA codes for selenocysteine. But, when the selenium concentration in the cell is low the UGA codon is interpreted as a PTC[89].

The natural NMD targets (identified so far) are involved in a variety of cellular processes such as stress responses, hematopoietic stem cell development, regulation of alternative splice forms[90], genomic stability, cell-cycle, telomere length maintenance[91] and embryonic development[4], potentially allowing NMD to adapt protein expression in accordance with the cellular needs. Nevertheless, these physiological substrates have one feature in common with their pathological counterparts: they possess a translation termination codon that is, by NMD standards, conceived as premature.

1.2.1.4. Biological and medical significance of NMD

NMD plays an important role as a modulator of the severity of the clinical phenotype of many genetic diseases[60]. If translated, the mRNAs containing PTCs give rise to truncated proteins that have either completely lost their function, have acquired dominant-negative function, are still functional, or have gained new functions. As a consequence of these different possibilities, NMD can either mitigate or aggravate the disease outcome. NMD importance as a protective surveillance mechanism is highlighted by the fact that one-third of all genetic disorders, including many cases of cancer, are caused by nonsense mutations or frameshifts, which generate nonsense codons[60]. β-thalassemia demonstrates the protective effects of NMD against the production of faulty proteins. If a mutation causing premature translation termination is localized within exons 1 and 2 of the transcript, the defective β-globin mRNA is degraded by NMD and, therefore, synthesis of truncated β-globin is limited. The resultant excess of free α-globin and any defective β-globin, which are harmful to the cell, are degraded by proteolysis. But in the event, that the nonsense codons resides within the third (final) exon, the transcript evades NMD allowing the generation of truncated protein that overwhelms the cell's proteolytic system and causes toxic precipitation of insoluble globin chains[92,93]. In contrast, there are cases (such as Ullrich disease or cystic fibrosis) in which NMD down-regulates mutant proteins with residual activity that can partially retain normal protein function, resulting in an augmentation of the defects caused by the original mutation. In these cases, selective inhibition of NMD may provide a novel therapeutic method [94–96].

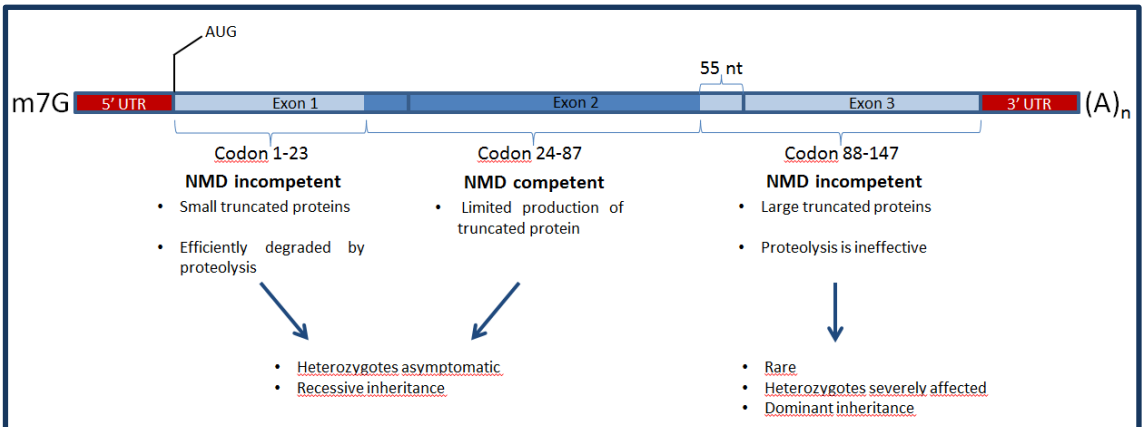

Figure 5: Representation of the position-dependent effects of nonsense mutations in the inheritance pattern and clinical severity of β-thalassemia (based on Holbrook, 2004)[60]. Premature translation termination codons (PTCs) in close proximity to the start codon do not trigger NMD but any translated β-chains are still small enough to be completely hydrolyzed by the red blood cell proteolytic system. If the PTC is located downstream of codon 23 and more than 55 nucleotides (nt) upstream of the last exon-exon junction, the transcript is targeted for NMD limiting protein production. In both cases the organism is protected from the deleterious dominant-negative effects of the truncated peptides. In contrast, transcripts bearing PTCs located less than 55 nt upstream of the last exon-exon junction produce nonfunctional β-globin proteins that are too large to be efficiently degraded, overwhelming the cellular proteolytic system and cause toxic precipitation of insoluble β-globin chains.

1.3. The eukaryotic initiation factor 3

The mammalian eIF3 is the largest (800 kDa molecular mass) and most complex of the eukaryotic translation initiation factors, consisting of 13 different subunits named from eIF3a-eIF3m[97] that form three stable modules; a:b:g:i (module 1); c:d:e:l:k (module 2) and f:h:m (module 3)[98]. Comparative studies of the cDNA sequences of the mammalian eIF3 subunits with the entire genome of the *Saccharomyces cerevisiae* have demonstrated that the subunits eIF3a, b, c, g, i and j are conserved across species[99–101]. The eIF3a, b, c, g and i subunits are essential for translation in vivo[97] and eIF3j despite being nonessential, is capable of enhancing interactions with other eIFs[102], promoting binding of eIF3 to the 40S subunit[103] and has an independent function in 40S ribosome biogenesis[104]. These findings suggest that these essential subunits provide most of the basic functions of eIF3 for translation initiation in vivo, and the remaining nonessencial subunits (of the mammalian eIF3) appear to modulate

its activity[99,105]. This view is supported by the observations that module 1 can promote mRNA binding to 40S ribosomal subunits on its own, but to achieve a maximal efficiency, it requires additional support from module 2 and that downregulation of eIF3c and eIF3a diminishes initiation rates[106].

The mammalian subunits a, c, e, k, l, and m contain a PCI (proteasome, COP9, eIF3) domain, a conserved structural motif shared by the functionally unrelated complex proteasome lid and the COP9 (constitutive photomorphogenesis 9) signalosome[107]. Intriguingly, the subunits f and h possess an MPN (Mpr1–Pad1 N-terminal) domain also found in related proteins in the proteasome lid and COP9[107]. These domains serve as a central structural scaffold that are involved in binding eIF3 to the translation initiation factors eIF1, eIF1A, and eIF2, as well as to the 40S ribosomal subunit[108]. The interactions between eIF3 and other translation initiation factors control the binding of the GTP-eIF2/Met-tRNAi complex to the 40S ribosomal subunit, positioning the mRNA on the 40S subunit, and modulate the stringency of start codon selection[42]. Furthermore, subunits eIF3c, d, and e interact with eIF4G[109], which is known to promote mRNA loading onto the 40S ribosome[37,38]. Therefore, eIF3 plays a central role in assembling the translation initiation complex.

In addition, recent data demonstrate a role of the eIF3 complex in NMD. In fact, the mammalian eIF3e subunit, which is not required for general translation, impairs NMD when silenced[110]. Our group has also demonstrated that human eIF3h and eIF3f subunits are involved in the NMD-resistance of mRNAs of AUG proximal PTCs[20]. These findings further support the notion that some subunits of this multiprotein complex are involved in controlling different aspects of mRNA biogenesis rather than directly participating in translation initiation.

2. Aims

Previous studies from our laboratory group revealed that mRNAs containing PTCs in close proximity to the translation initiation AUG codon (AUG proximal PTCs) escape NMD. This was initially surprising as these mRNAs would be expected to contain residual EJCs and would situate the PTC quite far, in a linear sense, from the poly(A) tail and PABPC1[15], a condition that could induce NMD. The observed NMD resistance was attributed as a direct effect of the translation termination event being located in close proximity to the AUG.

Knowing that the mRNA acquires a circular conformation due to the interaction between eIF4G/PABPC1 and eIF4G/eIF4E and that the interaction between eIF4G and the eIF3 complex establishes a protein bridge between the mRNA and the 43S PIC (fig. 1), we propose a model were the PABPC1 is relocated to the AUG vicinity during 43S scanning. Consequently, if a premature termination event takes place in the vicinity of the AUG, the PABPC1 might be able to compete with UPF1 for binding to eRF3, stimulating proper translation termination, and repressing NMD. This model conjectures that some initiation factors that promote scanning-dependent initiation may remain ribosome associated during translation of first codons[111]. In fact, some data have shown that mRNAs with AUG proximal PTCs become NMD sensitive when translation elongation across the short ORF is slowed down[16], indicating that PABPC1 interactions remain at the AUG vicinity during the first elongation steps, until the ribosome reaches the short ORF stop codon. Furthermore, there are results demonstrating that in HeLa cells treated with eIF3f or eIF3h subunits siRNAs, the AUG proximal nonsense-mutated transcripts become sensitive to NMD[20]. This may imply that these eIF3 subunits are responsible for pulling the PABPC1/eIF4G complex with 43S subunit during ribosomal scanning and translation initiation.

The aim of the work presented in this thesis was to clarify the mechanistic basis for the NMD resistance of mRNAs carrying AUG proximal PTCs and expand the current

models for NMD and translation. In order to accomplish that objective, the following goals were established:

- Proving that mRNA circularization via PABPC1/eIF4G interactions, leads PABPC1 into the AUG codon vicinity, as a consequence of 43S scanning, and that this interaction remains during the first elongation steps.

- Identify and characterize how eIF3 interacts with eIF4G and with the 43S ribosomal subunit.

3. Materials and methods

3.1. Plasmid constructs

A plasmid (pE_MS2-GFP) expressing a fusion protein that contains the N-terminal portion of MS2 coat protein was previously described[16]. The wild-type β-globin gene (βWT), as well as the β-globin variants β15 (UGG→UGA), β23 (GUU→UAG), β25 (GGU→UAG), β26 (GAG→UAG) and β39 (CAG→UAG) subcloned into the pTRE2pur vector (BD Biosciences) were previously described[79]. A β-globin containing three repeats of MS2 coat protein binding site in the 3'UTR (pGEMβNins5'UTR-MS2) was previously constructed by the laboratory group, but not published.

In order to produce a β-globin gene, as well as the β-globin variants β15, β23, β25, β26 and β39 subcloned into a pTRE2pur vector with three repetitions of the MS2 phage coat protein binding site in the 3'UTR (pTre_β#_3xMS2bs(3'UTR); # = WT, 15, 23, 25, 26 or 39) the corresponding pTre_β# and pGEMβNins5'UTR-MS2 plamids where used as templates. All plasmids were sequenced prior to use, to verify the β-globin WT, β-globin variants and MS2 biding site sequence integrity with primers #1 - #5 (table 1). pGEMβNins5'UTR-MS2 (3 µg) was subjected to an enzyme digestion with 1 µL of BstxI (New England Biolabs), 1 µL PciI (Roche), 0.5 µL BSA (100%), 5.0 µL NEBuffer 3 (10X, New England Biolabs) in a total volume of 50 µL, at 37°C during 2h, to cut and isolate a fragment containing the 3xMS2bs (estimated size = 491 bp) to be used as an insert. The vector pTre_β# (# = WT, 15, 23, 25, 26 or 39) (2 µg) was also digested following the same procedure (fragment estimated size = 5109bp). The digestion product was resolved in a 1% agarose gel and the appropriate bands were purified with the GeneJet Gel extraction kit (NZYtech) according to the manufacturer protocol. The vector (pTre_β#) and the insert (3xMS2bs) were used in a ligation reaction (molar ratio of 1:3) with 1 µL T4 DNA ligase reaction buffer (10X, Fermentas), 1 µg T4 DNA ligase enzyme (Fermentas) in a total volume of 10 µL, at 16°C overnight. The ligation product was used to transform NZY5α competent *E.coli* cells according to

the manufacturer protocol. Random colonies were selected and grown in agar medium overnight at 37°C, 220 rpm. The plasmid DNA was purified with the NYZtech miniprep extraction kit according to the manufacturer protocol and sequenced with primers #1, #2 and #5 (table 1). All plasmids with the correct sequence were amplified in NZY5α competent *E.coli* cells and purified with NYZtech miniprep extraction kit.

In order to produce a β-globin gene into a pTRE2pur vector with three repetitions of the MS2 phage coat protein binding site at forty five nucleotides of distance downstream of the codon 39 (pTre_βWT_3xMS2bs(45nt_CD39)); the corresponding pTre_βWT and pTreβWT_3xMS2bs(3'UTR) plamids where used as templates for gene splicing by overlap extension (SOEing) PCR. pTreβWT (0.5 µg) was amplified with 0.5 µL NZYSpeedy DNA polymerase (NZYtech), 5.0 µL reaction buffer (10X, NZYtech), 2.5 µL $MgCl_2$ (50 mM), 2.5 µL dNTPs (10 mM), 5.0 µL DMSO (100%), 1.0 µL BSA (100%) and 2.5 µL of primers (10 µM) #6-#7 (table 1) for Soeing 1 or #8-#9 (table 1) for Soeing 3, in a total volume of 50 µL. Thermocycler conditions were 95°C for 2 min followed by 40 cycles of 95°C for 30 sec, 55°C for 1 min, and 70°C for 45 sec followed by a final extension of 70°C for 5 min. For SOEing 2, pTre_βWT_3xMS2bs(3'UTR) was amplified in with primers #10-#11 (table 1) following the same procedure as above, but with an annealing temperature of 72°C for 1 min. The SOEing 1, 2 and 3 products were analyzed by electrophoresis on a 2% agarose gels. The appropriate bands were purified with the GeneJet Gel extraction kit (NZYtech) according to the manufacturer protocol and quantified in a Nanodrop spectrophotometer (Thermo).

3.2. Cell culture, plasmid and siRNA transfection

HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), at 37°C/5% CO_2. Transfections of cells with siRNAs were carried out when cells had a confluence of 30-40%, using 200 pmol of siRNA oligonucleotides and 4µl of Lipofectamine 2000 Transfection Reagent (Invitrogen), following the manufacturer's instructions. When indicated, a supplemental siRNA transfection was made 24h after the initial siRNA transfection, using 50 pmol of siRNA oligonucleotides and 4µl of Lipofectamine 2000. The siRNA

duplexes (Table 2) were designed as 19-mers with 3'-dTdT overhangs and purchased from Thermo. Transfections of cells with plasmids were carried out when cells had a confluence of 70-80%, using 200 ng of pTRE_β#_3xMS2bs(3'UTR), 800 ng pE_MS2-GFP and 4µl of Lipofectamine 2000 Transfection Reagent (Invitrogen), following the manufacturer's instructions.

Table 1: Sequences of the primers used in the current work

Primer	Sequences (5'-->3')
1	ACATTTGCTTCTGACACAAC
2	GCAATGAAAATAAATGTTTTTTAT
3	GCTCCTGGGCAACGTGCT
4	GTGGATCCTGAGAACTTCAGGCT
5	GTTCATGTCATAGGAAGGGG
6	TCGGTACCCGGGGATCCTCTAGTC
7	CCTCATGTTAACAGCATCAGGAGTGGACAG
8	GATGCTGTTAACATGAGGATCACCCATGTT
9	GTTGCCCATCCCAAACATGGGTGATCCTCA
10	ATGTTTGGGATGGGCAACCCTAAGGTGAAG
11	GAAAGAAAACATCAAGCGTCC
12	CTGCTCATTGCAGGCCAGAT
13	GAGCCTGGGCCATGAAGAG
14	CACCCAGTCATTTTGGCCTC
15	CGACAGTTCCCAACAGGGTC
16	ACCAAGAGAGTTGTCCGCAGTG
17	TCATGGCATTACGGATGGTCC
18	CCATGAGAAGTATGACAACAGCC
19	GGGTGCTAAGCAGTTGGTG

Table 2: Sequences of the siRNAs used in the current work

siRNA	Sequences (5'-->3')
eIF3f1	AUACGCGUACUACGACACU
eIF3f2	GUGAAGGAGAAAUGGGUUU
eIF3h1	GAUCGGCUUGAAAUUACCA
eIF3h2	ACUGCCCAAGGAUCUCUCU
eIF3c	UGACCUAGAGGACUAUCUU
GFP	GGCUACGUCCAGGAGCGCAC

3.3. Immunoprecipitation assay

HeLa cells cultured in 35-mm dishes and treated with siRNAs were collected 48h after transfection. Cells were lysed in 150 µl of NP40 buffer [50 mM Tris-HCl pH=7.5; 10 mM MgCl2; 100 mM NaCl; 10% (v/v) Glicerol; 1% (v/v) Nonidet P-40 and 1% (v/v) protease inhibitor mixture (Sigma). Additionally, 0.4 µL of RNase inhibitor (40 U/µL; NZYtech) were added to the samples not intended to be treated with RNAse A. Total lysates were cleared by centrifugation at 5000 rpm for 10 min at 4 °C and 20 µl were collected for RNA extraction or protein analysis with 4 µl 5x SDS loading buffer (Pre-IP). The remaining lysates were incubated overnight at 4°C with rabbit polyclonal anti-eIF3F (Abcam) or rabbit monoclonal anti-eIF3H (Cell Signaling) at a 1:50 dilution in a vertical rotator (Grant bio). Thirty µl of protein G-agarose beads (Roche) and 100 mg/mL of RNAse A (Quiagen) were then added to the corresponding samples. After an incubation of 1 hour at 4°C, the samples were washed three times with excess NP40 buffer and ressuspended in 25 µl of 2x SDS sample buffer (IP).

3.4. Cycloheximide treatment

HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, in 35-mm plates, at 37°C/5% CO_2. Once they achieved 90-100% confluence, the medium was replaced with a fresh one containing cycloheximide (CHX, Sigma). The cells were in incubated at 37°C/5% CO_2 for 2h and then lysated.

3.5. SDS-PAGE and Western blotting

Protein lysates were resolved, according to standard protocols, in 12% SDS-PAGE electrophoresis and transferred to PVDF membranes (Bio-Rad). The membranes were probed using mouse polyclonal anti-α-tubulin (Sigma) at 1:500 dilution (as a loading control), rabbit polyclonal anti-PABPC1 (Cell Signaling) at 1:500 dilution, rabbit monoclonal anti-eIF3H (Cell Signaling), rabbit polyclonal anti-CBP80 (generous gift from E. Izaurralde), rabbit polyclonal anti-eIF4G (Cell Signaling), mouse monoclonal

anti-eIF4E (Santa Cruz), mouse monoclonal anti-RPS6 (Cell Signaling) at 1:250 dilution, rabbit anti-c-myc (Santa Cruz) at 1:500 dilution and rabbit polyclonal anti-enterobacteriophage MS2 coat protein antibody (Merck Millipore) at 1:250 dilution. Detection was carried out using secondary peroxidase-conjugated anti-mouse IgG (Bio-Rad), anti-rabbit IgG (Bio-Rad) antibodies followed by chemiluminescence.

3.6. RNA isolation

Total RNA from transfected cells was prepared using the Nucleospin RNA extraction (Marcherey-Nagel) following the manufacturer's instructions.

3.7. Reverse transcription

cDNA was prepared by incubating 1 µg of purified total mRNA with 1 µL of dNTPs (10 µM) and 1 µL of random primers (0.25 µg/ µL), in a total volume of 15 µL, during 5 min at 65°C. Then, the mixture was incubated in ice and 2 µL of reaction buffer (10X; NZYtech), 0.1 µL of ribonuclease inhibitor (40 U/µL; NZYtech) and 0.5 µL of reverse transcriptase (200 U/µL; NZYtech) was added, to a final volume of 20 µL. Finally, the mixture was incubated in a thermocycler at 25°C for 10 min, 50°C for 50 min and 85°C for 5 min. cDNA was stored at 4°C.

3.8. Semiquantitative PCR

Three dilutions (1:1; 1:2 and 1:4) were prepared using the RT product. 5 µL of each of the diluted samples were amplified with 0.2 µL of Taq DNA polymerase (5 U/µL; Ambion), 2.5 µL of 10X PCR buffer (with 15 mM $MgCl_2$; Ambion), 0.5 µL of dNTPs (10 mM) and 1 µL of each corresponding forward and reverse primer (table 1; #12-13 for eIF3H; #14-15 for eIF3F and #16-17 for eIF3C), in a total volume of 25 µL. The same procedure was carried out in parallel, but using GADPH specific primers (table 1, #18-19) as an internal standard. Thermocycler conditions were 95°C for 5 min followed by 27 cycles of 94°C for 45 sec, 55°C for 45 sec, and 72°C for 45 sec, followed by a final

extension of 72°C for 5 min. Ten-microliter aliquots from each sample were analyzed by electrophoresis on 2% agarose gels.

4. Results

4.1. ptre_βWT_3xMS2bs(45nt_CD39) construction by SOEing PCR

To test the hypothesis that mRNA circularization via PABPC1/eIF4G interactions, leads PABPC1 into the AUG codon vicinity, as a consequence of 43S scanning, and that this interaction remains during the first elongation steps, we tried to construct by SOEing PCR a plasmid containing the β-globin sequence and three repetitions of MS2 coat protein binding site (3xMS2bs) 45 nt downstream of codon 39. This construct would allow the capture of β-globin mRNPs to study PABPC1 interactions with the ribosome. Using ptre_βWT as a template the 5'UTR of βWT (containing a ClaI restriction site) and codons 1-54 were amplified with primers #6-#7 (fig. 6 and 7; lane S1). Codon 55 until the beginning of exon 2 (containing a BbrPI restriction site) was amplified with primers #8-#9 (fig. 6 and 7; lane S3). Both S1 (expected fragment size = 409 bp) and S3 (expected fragment size = 191 bp) fragments were successfully amplified at annealing temperature of 55°C. The three repetitions of MS2 binding site were amplified with primers #10-#11 using the plasmid ptre_βWT_3xMS2bs(3'UTR) as a template (fig. 6 and 7; lane S2). The 3xMS2bs (expected fragment size = 93 bp) was only successfully amplified at an annealing temperature of 72°C. However, due to the repeated sequence of the MS2 binding site, 1x MS2bs was also amplified (expected fragment size = 43 bp) alongside other unspecified products, resulting in a lower quantity of the desired fragment. This in combination with the 3xMS2bs small size made the purification step very inefficient. Therefore the 3xMS2bs fragment was not obtained with the quantity and quality necessary to continue the plasmid construction.

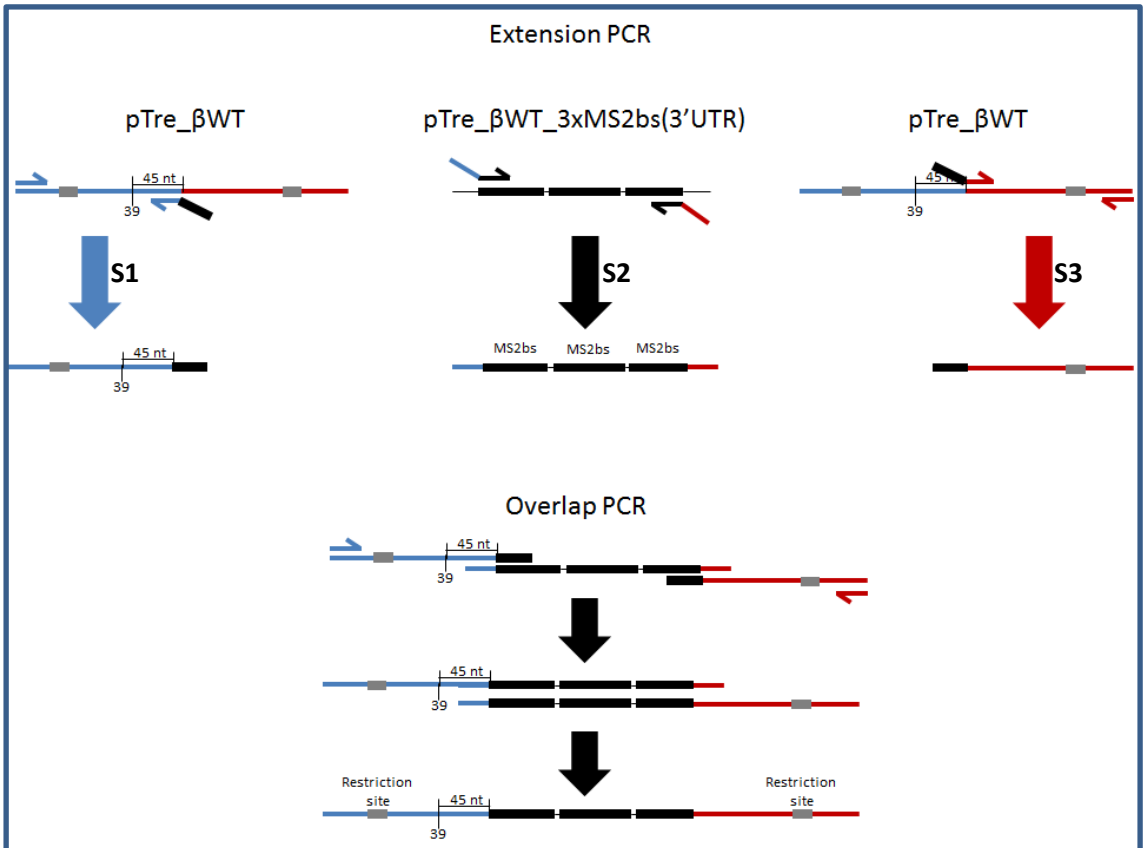

Figure 6: Schematic diagram of SOEing PCR. Using ptre_βWT as a template, the 5' untranslated region (UTR) and codons 1-54 of the normal human β-globin gene (βWT) were amplified with primers #6-#7 (Table 1), while the sequence from codon 55 (localized 45 nucleotides downstream of codon 39) until the beginning of exon 2 were amplified with primers #8-#9 (Table 1). The primers used for S1 and S3 contain overlapping sequences with the MS2 binding site (MS2bs) sequence. The three repetitions of MS2bs were amplified with primers #10-#11 (Table 1) using the plasmid ptre_βWT_3xMS2bs(3'UTR) as a template. The primers used for S2 amplification contained overlapping sequences for the ptre_βWT, upstream and downstream of codon 55. Extension of these overlaps, by DNA polymerase, creates the full-length mutant molecule of S1, S2 and S3 fragments.

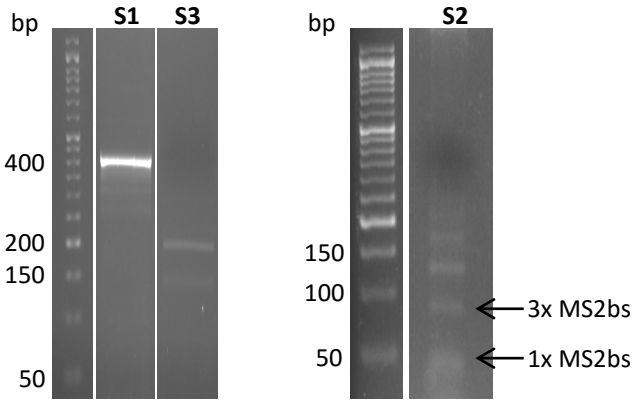

Figure 7: ptre_βWT_3xMS2bs(45nt_CD39) construction by SOEing PCR. Agarose gel photos of the amplification products of the ptre_βWT plasmid with primers #6-#7 (lane S1) or with primers #8-#9 (lane S3) and the ptre_βWT_3xMS2bs(3'UTR) plasmid with primers #10-#11 (lane S2).

4.2. Anti-enterobacteriophage MS2 coat protein antibody test

The coat protein binding site of the MS2 bacteriophage is a naturally occurring RNA aptamer[112] that in conjunction with the MS2 coat protein (MS2cp) and a specific antibody for the MS2cp can selectively immuneprecipitate the corresponding mRNA transcripts and any protein attached, thus providing a robust method to study the RNA–protein interactome[113,114]. Before proceeding with MS2 immunoprecipitations the anti-enterobacteriophage MS2 coat protein antibody specificity and conditions were tested (fig. 8). HeLa cells were transfected with pTre_BWT_3xMS2(3'UTR) plasmid and the pE_MS2-GFP plasmid, which codifies a fusion protein containing MS2cp. Lysates were generated 24 hours after transfection and probed with the anti-MS2cp antibody or anti-α-tubulin as a loading control. According to the manufacturer instructions, the MS2cp signal should appear around 13 kDa, however since the protein expressed is a fusion protein the observed band size corresponding to MS2-GFP was around 40 kDa (fig. 8). The bands at 35 kDa and 25 kDa correspond to uncharacterized interactions. No band was detected in non-transfected HeLa cells (fig. 8; lane 6), confirming the antibody specificity.

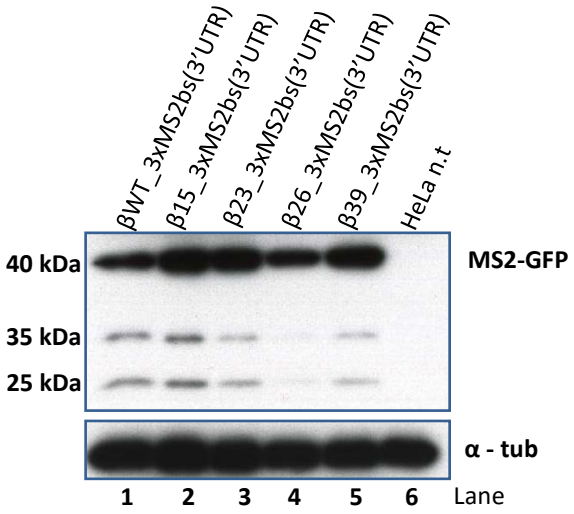

Figure 8: Test of anti-enterobacteriophage MS2 coat protein antibody. Western blot analysis of HeLa cells extracts transfected with a plasmid containing the β-globin sequence with three repetitions of MS2 coat protein binding site at the 3'UTR and a plasmid that codifies for a MS2 coat protein fused with the green fluorescent protein (GFP). Lysates were generated 24h after transfection and analyzed with a specific antibody for MS2 coat protein and with anti-α-tubulin (α-tub) as a loading control.

4.3. Cycloheximide concentration test

Besides using the ptre_βWT_3xMS2bs(45nt_CD39) to capture and study the mRNPs, HeLa cells would also be treated (or untreated) with cycloheximide (CHX). This is a common laboratory reagent that inhibits protein synthesis by binding the ribosome in the E-site and blocking eEF2-mediated translocation. Due to its size, CHX has a bigger affinity for the E-site when it is empty, i.e. when the ribosome is at the AUG. It was shown that CHX can cause about half of the ribosomal population to stop at the very first codon[115]. Taking advantage of this, CHX treatment would facilitate the capture of the mRNPs near the AUG, allowing the analysis of the interaction of PABPC1 and eIF4G with the ribosome (through eIF3) near the start codon. The necessary concentration of CHX to restrict protein synthesis was assessed in HeLa cells treated with different concentration of CHX for 2 hours (fig. 9). Lysates were generated afterwards and CHX effect on translation was evaluated by controlling C-myc protein levels by Western blot. The c-Myc protein is normally degraded very rapidly with a half-life of 20 to 30 min[116]. Therefore if translation elongation was blocked for 2h, it should be expected to see a decrease of the C-my protein levels. In fact, the Western blot shows a decrease of C-myc protein level of about 40% in the samples treated with 150 µg/mL of CHX (fig. 9; lane 6), indicating that at this concentration, the CHX should be able to interfere with translation in HeLa cells.

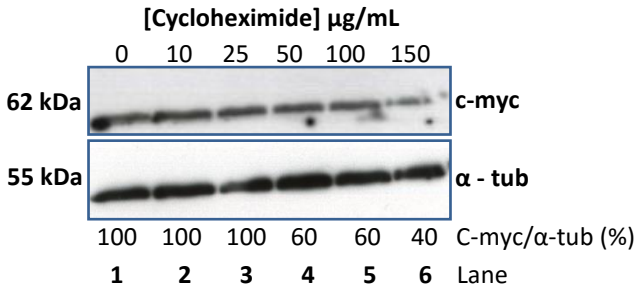

Figure 9: Test of cycloheximide (CHX) concentration necessary to inhibit translation. Western blot analysis of HeLa cells extracts treated with different concentrations of CHX. HeLa cells were incubated with CHX for 2h. Lysates were analyzed with a specific antibody for c-myc and with anti-α-tubulin (α-tub) antibody as a loading control. An estimative of the ratio (in %) of c-myc to α-tub is also represented.

4.4. eIF3h, eIF3f and eIF3c subunits knockdown

Before proceeding to the immunoprecipitation (IP) experiments to assess the eIF3h, eIF3f and eIF3c role in the interaction between PABPC1 and eIF4G with the ribosome, the knockdown efficiency of these subunits was tested at the protein level (fig. 10). HeLa cells were treated with 200 pmol of specific siRNAs for each subunit (table 2 – siRNA eIF3f1, eIF3h1 or eIF3c respectively) or with siRNA GFP as a transfection control. Lysates were generated 48 hours later and analyzed by Western blot with specific antibodies or with α-tubulin as a loading control. Silencing was achieved for both eiF3h (fig. 10A; lane 1) and eiF3c (fig. 10C; lane 1) when compared with the lysates treated with siRNA GFP (fig. 10A and 10C; lane 2), with an estimated knockdown efficiency of 50% and 80%, respectively. For eIF3f silence however, no eiF3f signal was detected (fig. 10B; lane 1 and 2). This may indicate that the quantity or the conditions utilized for the detection of eIF3 were not adequate.

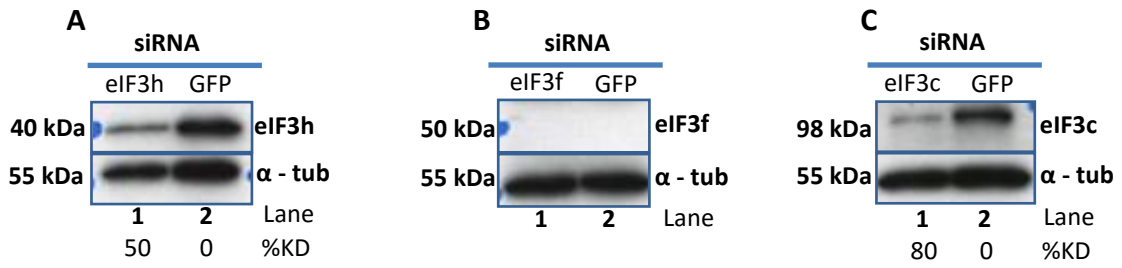

Figure 10: Test of the conditions for the eukaryotic initiation factor 3h (A), eIF3f (B) and eIF3c (C) subunits knockdown. Western blot analysis of HeLa cells extracts transfected with siRNA eIF3h1 (A; lane 1), siRNA eIF3f1 (B; lane 1), siRNA eIF3c (C; lane 1) or with a control green fluorescent protein (GFP) siRNA (lane 2A, 2B and 2C). Cell lysates were obtained 48h post transfection and siRNA silencing efficiency was analyzed with specific antibodies and with anti-α-tubulin (α-tub) as a loading control. An estimative of the percentage of achieved knockdown (% KD) is represented below each lane.

4.5. Immunoprecipitation of mRNP complexes from HeLa cells treated with siRNAs specific for the eIF3f subunit

To test whether the eIF3f subunit mediates the interaction between PABPC1 and eIF4G with the 40S ribosome during initiation scanning, immunoprecipitation (IP) of mRNPs from HeLa cells depleted of eIF3f subunits were carried out, in the conditions previously tested (fig. 10). HeLa cells were transfected with 200 pmol of siRNA GFP or with siRNA eIF3f1 (Table 2). The lysates were obtained 48 hours later and the mRNP complexes were isolated by immunoprecipitation, using an antibody specific for eIF3h. Immunoprecipitation was performed in the presence or absence of RNase A, to test for protein interactions dependent or independent of mRNA. A 20 µL aliquot from each sample, representing the total lysate, was taken before immunoprecipitation (Pre-IP)**.** The presence of PABPC1, RPS6 and other initiation factors was assessed by Western blot (fig. 11). α-Tubulin was used as a loading control and has expected, it was not detected in IP samples (fig. 11; lane 3-5), showing that the washing procedure, to remove any nonspecifically bonded proteins, was effective.

Efficiency eIF3f knockdown was evaluated with an anti-eIF3f antibody. Comparing the eIF3f band of the Pre-IP samples (fig. 11; lane 1 and 2) there is no discernible differences, in terms of eIF3f protein level, between samples treated with control siRNA (fig. 11; lane 1) or with siRNA eIF3f1 (fig. 11; lane 2). Both signals have the same intensity and also there is no difference in the loading control. This indicates that eIF3f silencing did not have any effect, at least at the protein level. In the IP samples no eIF3f protein was detected (fig. 11; lane 3-5). Instead a band (50-55 kDa) corresponding to the heavy chain of the immunoglobulin G (IgG) was observed. This signal is detected because the anti-eIF3h antibody used during IP is also eluted from the beads and size-fractionated by SDS-PAGE. Since the IP antibody and the Western blot primary antibody are from the same species, the secondary antibody will also recognize it. This signal was always detected when a primary antibody of the same species was used in Western Blot and exclusively in the IP samples.

Despite the eIF3f subunit silencing apparently not working as expected, some interesting variations can be pointed out. eIF4G, PABPC1 and CBP80 co-immunoprecipitate with eIF3h and were detected in the IP samples even when treated

with RNase A (fig. 11; lane 4). These results imply that the interaction of these proteins with eIF3h is independent of RNA. Both eIF4G and PABPC1 signals in IP samples treated with control siRNA (fig. 11; lane 3) are stronger than the signals of siRNA eiF3f (fig. 11; lane 4 and 5) treated samples. This suggests that eIF3f silencing interfered with PABPC1 and eIF4G interaction with the eIF3h subunit. On the other hand, samples treated with siRNA eIF3f (fig. 11; lanes 4 and 5) also exhibit a weaker eIF3h protein signal when compared with the control (fig. 11; lane 3). Therefore the decreased interaction of eIF4G and PABPC1 with eIF3h observed in samples treated with siRNA eIF3f, may be due to eIF3f silencing destabilized eIF3h, and not because eIF3f mediates the interaction between eIF4G and PABPC1 with eIF3h and the ribosome. The mRNA levels of eIF3h would have to be assessed to see if the siRNA eIF3f1 is not affecting eIF3h expression. Additionally, this differences seen could also be a case of lane 3 (fig. 11) having more sample quantity than lane 4 and 5 (fig. 11). Since α-tubulin was not detected, has expected, in the IP samples, it is difficult to assess if the protein loading was equal across lanes 3-5 (fig. 11).

RPS6, a component of the 40S ribosomal unit, was not detected in IP samples, in either control (fig. 11; lane 3) or siRNA eIF3f (fig. 11; lane 4 and 5) treated cells. In Pre-IP samples, however, the RPS6 signal of the control sample (fig. 11; lane 1) is more intense than the siRNA eIF3f treated sample (fig. 11; lane 2), suggesting that the silencing treatment interfered with RPS6 expression. RPS6 mRNA levels would have to be analyzed to confirm this difference. The protein eIF4E was not detected in IP samples (fig. 11; lane 3-5) and is barely seen in Pre-IP samples (fig. 11; lane 1 and 2). This was unexpected since eIF4E supports the bulk of translation; therefore, it should be more abundant than CBP80. Nonetheless, the absence of eIF4E and RPS6 in the IP samples suggests that these proteins do not co-immunoprecipitate with eIF3h. Conversely, this could be an experimental artifact. In fact, the PVDF membrane was probed and reprobed several times and part of the protein may have been gradually washed away during removal of primary and secondary antibodies, resulting in a weak protein signal in Pre-IP samples and no signal in the IP samples (which had less protein quantity than the Pre-IP samples to begin with). A newly prepared membrane would have to be probed to assess if eIF4E and RPS6 does or does not co-immunoprecipitate with eIF3h.

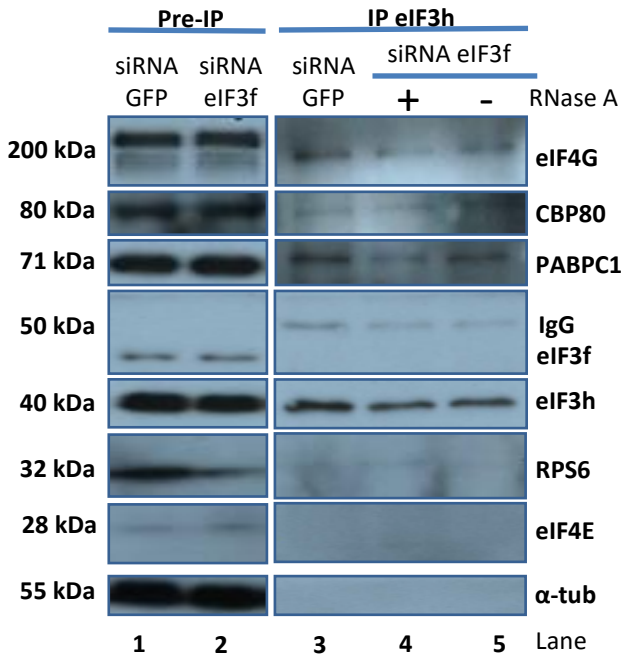

Figure 11: Immunoprecipitation of mRNP complexes from HeLa cells treated with siRNAs specific for the eukaryotic initiation factor 3f (eIF3f). Western blot analysis of HeLa cells extracts transfected with human eIF3f siRNA (lane 2, 4 and 5) or with a control green fluorescent protein (GFP) siRNA (lane 1 and 3). HeLa cells lysates were obtained 48 hours after transfection with GFP siRNA or siRNA targeting the eIF3f subunit. Immunoprecipitation (IP) was accomplished with an antibody specific to eIF3h (lanes 3-5). Immunoprecipitation was carried out in the presence (lane 4) or absence (lane 5) of RNAse A. Lanes 1 and 2 correspond to aliquots of cell lysates prior to IP (Pre-IP). α-tubulin (α-tub) was used as a loading control.

4.6. Immunoprecipitation of mRNP complexes from HeLa cells treated with siRNAs specific for the eIF3h subunit

The eIF3h subunit role in the interaction of PABPC1 and eIF4G with the 40S ribosome during translation initiation was also tested in cells depleted of eIF3h. HeLa cells were either transfected with 200 pmol of siRNA GFP or with siRNA eIF3h1 (Table 2)**.** The lysates were obtained 48 hours later and the mRNP complexes were isolated by immunoprecipitation, using an antibody specific for eIF3f or with protein G-agarose beads alone (mock) to control for nonspecific immunoprecipitation. IP was performed in the presence or absence of RNase A. An aliquot from each sample, representing the total lysate, was taken before immunoprecipitation (Pre-IP) for protein and RNA analysis**.** The presence of PABPC1 and initiation factors was assessed by Western blot and α-tubulin was used as a loading control (fig. 12). eIF3h silencing efficiency was evaluated at the protein level using an anti-eIF3h antibody (fig. 12) and at the RNA level by a semiquantitative PCR (fig. 13).

Western blot analysis of eIF3h knockdown shows no distinct differences between GFP siRNA treated samples (fig. 12; lane 1) and siRNA eIF3h treated samples (fig. 12; lane 2). Contrariwise, the eIF3h RNA endogenous levels in cells treated with siRNA eiF3h (fig. 13; lane 9-11) are clearly lower than those treated with control siRNA (fig. 13; lane 3-5). In fact the lane 9 band (dilution 1:1; fig. 13) intensity is weaker than the band (dilution 1:2) in lane 4, indicating a silencing efficiency of at least 50% at the RNA level.

Both eIF4G and PABPC1 were probed with specific antibodies. eIF4G was again detected in an RNA independent manner in IP samples (fig. 12; lane 4-7). PABPC1, however, was not detected in IP samples, suggesting that it does not co-precipitate with eIF3f. This is surprising, because PABPC1 forms a specific complex with eIF4G, which directly interacts with the eIF3 protein that co-immunoprecipitats with eIF3f. Additionally, PABPC1 is also bound to the poly-A tail, so it was expected that at least the PABPC1 signal could be observed in control IP samples not treated with RNAse A (fig. 12; lane 4). This can be, once again, a case of the protein content loss due membrane washing and reprobing. A newly prepared membrane would have to be probed to confirm these findings.

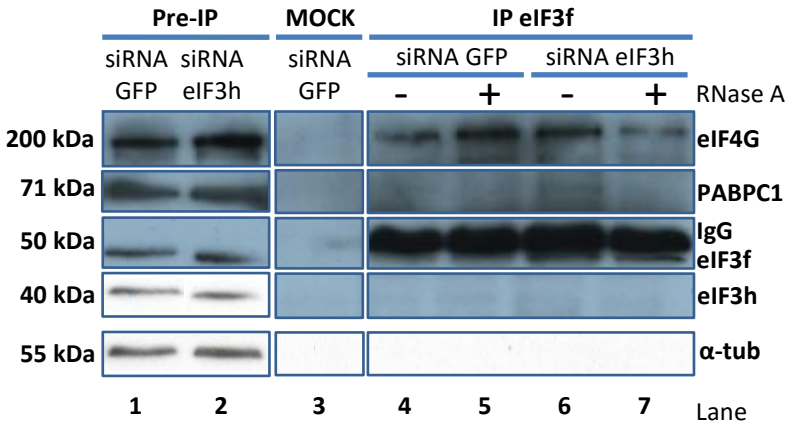

Figure 12: Immunoprecipitation of mRNP complexes from HeLa cells treated with siRNAs specific for the eukaryotic initiation factor 3h (eIF3h). Western blot analysis of HeLa cells extracts transfected with a control green fluorescent protein (GFP) siRNA target (lane 1, 3, 4 and 5) or with human eIF3h siRNA (lane 2, 6 and 7). Cell lysates were obtained 48 hours after transfection with GFP siRNA or siRNA targeting eIF3h and immunoprecipitated with an antibody to eIF3f (lanes 4-7) or with protein G-agarose beads only (lane 3). Immunoprecipitation was carried out in the presence (lane 5 and 7) or absence (lane 4 and 6) of RNAse A. The lysates were analyzed for the presence of PABPC1 and initiation factors by immunoblotting. Lanes 1 and 2 correspond to aliquots of cell lysates prior to IP (pre-IP). α-tubulin (α-tub) was used as a loading control.

Figure 13: Quantification by reverse transcription coupled semiquantitative (SQ) PCR of the eukaryotic initiation factor 3h (eIF3h) knockdown efficiency. Representative SQ-PCR analyses of cDNA obtained from RNAs extracted from green fluorescent protein (GFP; lane 3-8) or eIF3h (lane 9-14) siRNA-treated HeLa cells. SQ-PCR was carried out with eIF3h mRNA specific primers to monitor endogenous eIF3h expression or with glyceraldehyde-3-phosphate dehydrogenase (GADPH) mRNA specific primers as an internal standard. Lane 1 and 2 represent blank samples with eIF3h primers or GAPDH specific primers. Sample dilution is also represented.

4.7. eIF3h, eIF3f and eIF3c subunits knockdown with different siRNAs

Since the previous knockdown procedure was not satisfactory, at least at the protein level, in both immunoprecipitation experiments, eIF3h, eIF3f silencing was repeated using a new set of siRNAs: siRNA eiF3h2 and siRNA eIF3f2, respectively (table 2). Cells were transfected with 200 pmol of the respective siRNAs and 24 hours later, cells were transfected again with 50 pmol of siRNAs. Lysates were generated 48 hours after the initial transfection and analyzed by Western blot with specific antibodies or with α-tubulin as a loading control (fig. 14). Silencing was achieved for both eIF3f (fig. 14; lane 3) and eIF3c (fig. 14; lane 5). However the new set of siRNAs for eiF3h was not able to interfere with eiF3h expression, even with the additional transfection step (fig. 14; lane 1).

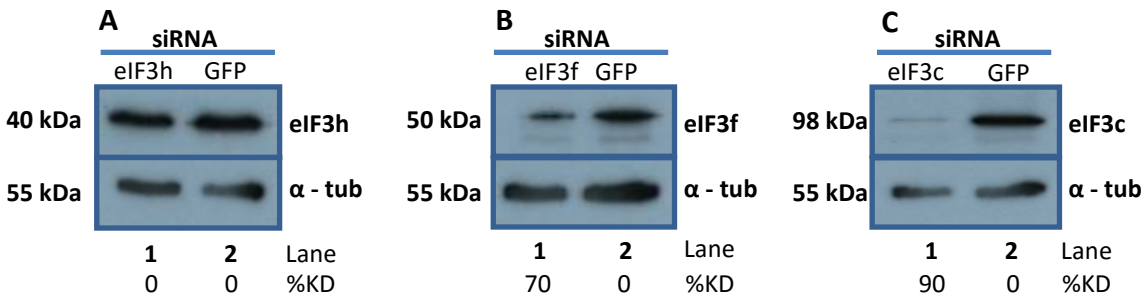

Figure 14: Analysis of the eukaryotic initiation factor 3h (A), eIF3f (B) and eIF3c (C) subunits knockdown efficiency. Western blot analysis of HeLa cells extracts transfected with human eIF3h siRNA (A; lane 1), eIF3f siRNA (B; lane 1), eIF3c (C; lane 1) or with a control green fluorescent protein (GFP) siRNA target (A, B, C; lane 2). Cells were transfected with 200 pmol siRNAs followed by a new transfection of 50 pmol of siRNAs 24h later. After 24h, siRNA silencing efficiency was analyzed with specific antibodies and with anti-α-tubulin (α-tub) as a loading control. An estimative of the percentage of achieved knockdown (% KD) is represented below each lane.

5. Discussion

5.1. Proving that PABPC1/eIF4G remains associated with the ribosome, through eIF3 interaction, during the early stages of elongation

Using β-globin mRNA (a small transcript of three exons, whose NMD profile has been extensively characterized) as the primary model, it was shown that transcripts containing PTCs in the latter half of the first exon, up to the third exon, are targeted for NMD, which can be attributed to the existence of downstream EJCs. However β-globin transcripts bearing PTCs localized up to 23 codons (AUG proximal PTCs) into the open reading frame effectively evade NMD (fig. 5). This was unexpected since there are EJCs still residing downstream. It was demonstrated that this resistance reflects a close proximity of the PTC to the translation initiation codon and are not caused by defects in transcript splicing, impaired translation, or reinitiation of translation 3' to the PTC[15,79,117,118].

Recently, it has been verified that PABPC1 can compete with UPF1 for eRF3 binding, when a ribosome is poised in a stop codon, diminishing NMD response[16,19] . Additionally, it is also known that mRNAs form a closed-loop structure during translation, as a consequence of eIF4G interaction with PABPC1, eIF4E, eIF3 and the ribosome. With this findings, it was proposed that PABPC1 possesses a major role protecting mRNAs harboring AUG proximal PTCs from the NMD surveillance pathway[16,24].

Assuming the retention of ribosome association with eIF3 and PABPC1 (through eIF4G) during scanning and the early stages of elongation, we propose that PABPC1 is repositioned in close proximity to an early PTC, during ribosome scanning, inhibiting NMD and enhancing the efficiency of the translation termination. This model is supported by the fact that tethering of PABPC1 in the vicinity of a PTC inhibited NMD[16] and the observations that eIF3 can remain bound to the translating ribosome during the initial phase of elongation[97,119].

To investigate if PABPC1 is in fact translocated during initiation scanning and that remains in the vicinity of the AUG during the first steps of elongation, the pTre_βWT_3xMS2bs(45nt_CD39) was designed. The introduction of multiple MS2bs into the β-globin sequence would facilitate the isolation of β-globin mRNPs, by immunoprecipitation assays, allowing the analysis of PABPC1 interactions with the ribosome and other proteins of the mRNP. The ptre_βWT_3xMS2bs(45nt_CD39) plasmid assembly was initiated using a SOEing PCR approach (fig. 6), a method of recombining sequences without depending on restriction sites or ligase. This process involves using primers to introduce segments of identical sequence (called the overlap region) in the amplified fragments. Then, during overlap PCR, the overlapping segments anneal and act as primers of one another. Extension of this overlap by DNA polymerase creates the full-length mutant molecule of both fragments. The 3xMS2bs fragment (fig. 7; lane S2), however, proved to be rather difficult to amplify. Owing to the presence of several repetitions of the MS2bs in the pTre_βWT_3xMS2bs(3'UTR), the primer could anneal with any of them, resulting in a PCR product containing fragments with 1 or several repetitions of MS2bs. Several PCR conditions were tested, but none resulted in a satisfactory amount of the fragment of interest (containing 3xMS2bs). Moreover, the fragment is relatively small (93 bp) making purification problematic.

An alternative process of construction could be using site directed mutagenesis to flank 3xMS2bs with a pair of restriction sites, such as SalI and XhoI, at 5' and 3' respectively. These enzymes recognize very similar sequences and produce compatible cohesive ends. The resulting fragment, containing the 3xMS2bs could then be subcloned into the vector containing the β-globin sequence with a XhoI restriction site at 45 nt downstream of the codon 39. The advantage of this method is that the ligation of the insert into the XhoI site of the β-globin will produce an XhoI restriction site at 5', that can be cleaved to insert more 3xMS2bs if needed[114]. The addition of multiple MS2bs admits more coat protein, which could improve pulldown efficiency during immunoprecipitation assays.

5.2. Identification and characterization of eIF3 interactions with eIF4G and the 40S ribosomal subunit

Human eIF3 is an 804 kDa protein complex containing 13 nonidentical subunits, named from 3a to 3m. Six of these subunits (eIF3a, b, c, g, i and j) are conserved in sequence across species, and are thought to provide most of the basic functions of eIF3[120]. eIF3f and h are part of the nonconserved core and with eIF3m constitute a stable module that is located on the periphery of the complex[98]. Both eIF3f and eIF3h contain a MPN domain that is usually found in components of large protein complexes and have been implicated in protein-protein interactions supporting a possible interaction of these proteins with other initiation factor complexes[121].

Our group has shown that AUG proximal nonsense-mutated transcripts are NMD resistant but become sensitive to NMD when HeLa cells are treated with siRNAs specific for the eIF3h, eIF3f[20] and eIF3c (unpublished) subunits, supporting the notion that eIF3 subunits might be involved in the delivery of eIF4G-associated PABPC1 to the vicinity of the AUG proximal PTC.

To study the role of the eIF3h and eIF3f subunits, mRNPs from HeLa cells treated with specific siRNAs for these subunits were isolated and examined. Overall the IP procedure of mRNPs from HeLa cells treated with eIF3f specific siRNAs (fig. 11) was successfully accomplished. The proteins expected to be bound to the mRNP were detected, although some proteins (PABPC1, eIF4G and CBP80) were only visible (in IP samples) with higher (5-10 min) exposure times to the X-ray film, resulting in higher backgrounds. In comparison, IP pulldown efficiency in HeLa cells treated with eIF3h siRNAs was even lower. As a result, the proteins detected in IP samples were barely visible even with a 10 min exposure to the X-ray film. Nonetheless, in both IP experiments, a bead and antibody quantity optimization can be performed to improve precipitation efficiency. Furthermore, during Western blot analysis of the IP samples it was visible the IgG heavy chain signal, around 50-55 kDa. This made eIF3f detection harder. The IgG signal can be omitted simple using a probing antibody raised in a different species than the one used for IP, or using protein G conjugated with horseradish peroxidase (HRP), which binds only to the native IgG structure and not to its reduced forms, instead of conventional HRP-secondary antibodies. Alternatively,

the antibody used for IP can be covalently linked to the agarose beads. Therefore during elution the antibody remains with the beads. This also has the advantage of increasing IP specificity, giving cleaner backgrounds. However, less represented proteins may not be precipitated.

Surprisingly, in both IP experiments, eIF3h and eIF3f protein knockdown efficiency was unexpectedly low. This may have been the result of transfecting an insufficient quantity of siRNAs or using siRNAs that bind weakly to the target mRNA. As such, the silencing procedure was carried out once more, using new siRNAs for eIF3h (siRNA eIF3h2, table 2) and eIF3f (siRNA eIF3f2, table 2) and with an additional transfection, 24h after the initial treatment (fig. 14). This resulted in the depletion of eIF3f protein, but eIF3h knockdown failed, indicating that the siRNA eIF3h2 (fig. 14A) is even more ineffective than siRNA eIF3h1 (fig. 10A versus fig.14A). Interestingly, eIF3c silence using the same siRNA, but with an additional transfection step, had similar results. This is easily explained by the fact that eIF3c protein has a half-life of 15.1h[122] and therefore more time is needed between transfection and lysis to see any difference. Additionally, in IP of mRNPs from HeLa cell lysates treated with eIF3h siRNAs, a decrease of ≥ 50% of the eIF3h mRNA level was observed (fig. 13), while the protein level remained unchanged (fig. 12). This could suggest that eIF3h has a high stability and there was not enough time to decrease the protein population, between the transfection and lysis.

Despite the silence procedure not being satisfactory in both IPs, there are still some results that can be pointed out. Remarkably, in IP samples of HeLa cells treated with eIF3f specific siRNAs (fig. 11), CBP80, eIF4G and PABPC1 co-precipitated with eIF3h and were detected in a RNA independent manner. This supports the idea that CBP80 bound mRNAs also acquire an mRNP circularization conformation analogous to the well-characterized eIF4E-eIF4G-PABPC1 interaction that brings the 5′ and 3′ ends of mRNAs close together[39]. Additionally, this data is in accordance with the results by Lejeune and Chiu[27,123] that suggests that eIF4G and PABPC1 have a function in the pioneer initiation complex rather than being merely a presence during remodeling to the steady-state complex, indicating that there is a degree of mechanistic conservation between the pioneer round of translation and steady-state translation. Our findings also show that treatment with eIF3f siRNAs leads to lower levels of PABPC1 and eIF4G

(fig. 11). This could support a model where the eIF3f subunit is involved in bringing PABPC1 (through eIF4G) into the vicinity of the AUG. However, the difference is small in nature (and arguable), as a result of the low silencing efficiency. If a better efficiency could be obtained, perhaps a greater impact on PABPC1 and eIF4G protein levels would be observed in IP samples treated with eIF3f siRNAs.

6. Conclusion and future directions

NMD is increasingly appreciated as one of the central mechanisms of mRNA surveillance, with an important role both in the physiological control of gene expression and in modulating diseases associated with PTCs. The detailed study of this mechanism will certainly aid in the development of new tools to be implemented in therapies for PTC related pathologies. But, despite having been thoroughly scrutinized for several years, the molecular basis underlying the NMD surveillance pathway has still not been fully elucidated.

PABPC1 has shown a major role in protecting AUG proximal PTCs from NMD. In this thesis we tried to construct the tools necessary to investigate the interaction of PABPC1 with the ribosome, during initiation and the first steps of elongation. Once completed, the pTre_βWT_3xMS2bs(45nt_CD39) could be subjected to direct mutagenesis to create β-globin variants with nonsense mutations in codons 15, 23, 26 and 39. Immunoprecipitation assays of the resultant mRNPs in the presence of RNAse H and specific ssDNA probes, to specifically digest the mRNA in the 5' UTR and downstream of the MS2bs, would permit to capture the mRNA fragment with a ribosome near the AUG or premature stop codon. The presence of PABPC1, ribosomal complex proteins and NMD factors could then be assessed by Western blot. Treatment with cycloheximide would facilitate the capture of the mRNPs near the AUG.

Furthermore, the role of the eIF3f and eIF3h subunits in pulling the PABPC1/eIF4G complex with the 43S subunit during ribosomal scanning and translation initiation was also explored. The data presented here point to a model where eIF3f is indeed involved in this interaction, but both IP and silence procedures should be optimized before solid results can be obtained.

7. References

1. Isken, O. & Maquat, L. E. Quality control of eukaryotic mRNA: Safeguarding cells from abnormal mRNA function. *Genes Dev.* **21,** 1833–1856 (2007).

2. 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,** 1073–8 (2004).

3. 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).

4. Rehwinkel, J., Raes, J. & Izaurralde, E. Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors. *Trends Biochem. Sci.* **31,** 639–46 (2006).

5. 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).

6. 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,** 22709–14 (2001).

7. 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,** 2215–28 (2001).

8. Kashima, I. *et al.* Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.* **20,** 355–67 (2006).

9. Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19,** 6860–9 (2000).

10. Bono, F., Ebert, J., Lorentzen, E. & Conti, E. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* **126,** 713–25 (2006).

11. Mühlemann, O. Recognition of nonsense mRNA: towards a unified model. *Biochem. Soc. Trans.* **36,** 497–501 (2008).

12. Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M. J. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20,** 4987–97 (2001).

13. Nagy, E. & Maquat, L. E. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **23,** 198–9 (1998).

14. Maquat, L. E. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* **5,** 89–99 (2004).

15. Inácio, A. *et al.* Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay. *J. Biol. Chem.* **279,** 32170–80 (2004).

16. Silva, A. L., Ribeiro, P., Inácio, A., Liebhaber, S. a & Romão, L. Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. *RNA* **14,** 563–576 (2008).

17. Mangus, D. A., Evans, M. C. & Jacobson, A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* **4,** 223 (2003).

18. 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,** 50286–92 (2002).

19. Singh, G., Rebbapragada, I. & Lykke-Andersen, J. A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. *PLoS Biol.* **6,** e111 (2008).

20. 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,** 1160–73 (2012).

21. Moore, M. J. & Proudfoot, N. J. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136,** 688–700 (2009).

22. Singh, G. *et al.* The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151,** 750–64 (2012).

23. Moore, M. J. From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309,** 1514–8 (2005).

24. 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).

25. Miller, J. N. & Pearce, D. a. Nonsense-mediated decay in genetic disease: Friend or foe? *Mutat. Res. Rev. Mutat. Res.* **762,** 52–64 (2014).

26. Burns, L. T. & Wente, S. R. Trafficking to uncharted territory of the nuclear envelope. *Curr. Opin. Cell Biol.* **24,** 341–9 (2012).

27. Lejeune, F., Ranganathan, A. C. & Maquat, L. E. eIF4G is required for the pioneer round of translation in mammalian cells. *Nat. Struct. Mol. Biol.* **11,** 992–1000 (2004).

28. Ishigaki, Y., Li, X., Serin, G. & Maquat, L. E. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* **106,** 607–17 (2001).

29. Lejeune, F., Ishigaki, Y., Li, X. & Maquat, L. E. The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J.* **21,** 3536–45 (2002).

30. Maquat, L. E., Tarn, W.-Y. & Isken, O. The pioneer round of translation: features and functions. *Cell* **142,** 368–374 (2010).

31. Sonenberg, N. & Hinnebusch, A. G. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136,** 731–45 (2009).

32. Livingstone, M., Atas, E., Meller, A. & Sonenberg, N. Mechanisms governing the control of mRNA translation. *Phys. Biol.* **7,** 021001 (2010).

33. Asano, K. *et al.* Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBO J.* **20,** 2326–37 (2001).

34. Algire, M. A. *et al.* Development and characterization of a reconstituted yeast translation initiation system. *RNA* **8,** 382–97 (2002).

35. Korneeva, N. L., Lamphear, B. J., Hennigan, F. L. & Rhoads, R. E. Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *J. Biol. Chem.* **275,** 41369–76 (2000).

36. LeFebvre, A. K. *et al.* Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J. Biol. Chem.* **281,** 22917–32 (2006).

37. 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,** 21203–7 (2010).

38. Hinnebusch, A. G. Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol. Mol. Biol. Rev.* **75,** 434–67 (2011).

39. 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).

40. Kozak, M. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266,** 19867–70 (1991).

41. Barbosa, C., Peixeiro, I. & Romão, L. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet.* **9,** e1003529 (2013).

42. Hinnebusch, A. & Lorsch, J. The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb. …* **4,** 1–25 (2012).

43. Rodnina, M. V & Wintermeyer, W. Recent mechanistic insights into eukaryotic ribosomes. *Curr. Opin. Cell Biol.* **21,** 435–43 (2009).

44. Wohlgemuth, I., Pohl, C., Mittelstaet, J., Konevega, A. L. & Rodnina, M. V. Evolutionary optimization of speed and accuracy of decoding on the ribosome. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **366,** 2979–86 (2011).

45. Abbott, C. M. & Proud, C. G. Translation factors: in sickness and in health. *Trends Biochem. Sci.* **29,** 25–31 (2004).

46. Proud, C. G. Peptide-chain elongation in eukaryotes. *Mol. Biol. Rep.* **19,** 161–70 (1994).

47. Gebauer, F. & Hentze, M. W. Molecular mechanisms of translational control. *Nat. Rev. Mol. Cell Biol.* **5,** 827–35 (2004).

48. Dever, T. & Green, R. The Elongation, Termination and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **4,** 1–16 (2012).

49. Keeling, K. M., Xue, X., Gunn, G. & Bedwell, D. M. Therapeutics based on stop codon readthrough. *Annu. Rev. Genomics Hum. Genet.* **13,** 371–94 (2014).

50. 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* **1829,** 612–23 (2013).

51. Amrani, N., Ganesan, R., Kervestin, S. & Mangus, D. A. A faux 3 0 -UTR promotes aberrant termination and triggers nonsense- mediated mRNA decay. *Nature* **432,** 112–118 (2004).

52. 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,** 736–47 (2008).

53. Pisarev, A. V *et al.* The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol. Cell* **37,** 196–210 (2010).

54. 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,** 391–8 (2006).

55. Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440,** 561–4 (2006).

56. Vasudevan, S., Peltz, S. W. & Wilusz, C. J. Non-stop decay--a new mRNA surveillance pathway. *Bioessays* **24,** 785–8 (2002).

57. Losson, R. & Lacroute, F. Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc. Natl. Acad. Sci. U. S. A.* **76,** 5134–7 (1979).

58. Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A. & Ross, J. Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia. *Cell* **27,** 543–53 (1981).

59. Culbertson, M. R. & Leeds, P. F. Looking at mRNA decay pathways through the window of molecular evolution. *Curr. Opin. Genet. Dev.* **13,** 207–14 (2003).

60. Holbrook, J. a, Neu-Yilik, G., Hentze, M. W. & Kulozik, A. E. Nonsense-mediated decay approaches the clinic. *Nat. Genet.* **36,** 801–808 (2004).

61. Yamashita, A. *et al.* Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat. Struct. Mol. Biol.* **12,** 1054–63 (2005).

62. Schmid, M. & Jensen, T. H. The exosome: a multipurpose RNA-decay machine. *Trends Biochem. Sci.* **33,** 501–10 (2008).

63. Unterholzner, L. & Izaurralde, E. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell* **16,** 587–96 (2004).

64. Gatfield, D. & Izaurralde, E. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. *Nature* **429,** 575–8 (2004).

65. Hosoda, N., Kim, Y. K., Lejeune, F. & Maquat, L. E. CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.* **12,** 893–901 (2005).

66. Kim, K. M. *et al.* A new MIF4G domain-containing protein, CTIF, directs nuclear cap-binding protein CBP80/20-dependent translation. *Genes Dev.* **23,** 2033–45 (2009).

67. Lejeune, F., Ishigaki, Y., Li, X. & Maquat, L. The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J.* **21,** 3536–3545 (2002).

68. Hwang, J., Sato, H., Tang, Y., Matsuda, D. & Maquat, L. E. UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. *Mol. Cell* **39,** 396–409 (2010).

69. Rufener, S. C. & Mühlemann, O. eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.* **20,** 710–7 (2013).

70. Gao, Q., Das, B., Sherman, F. & Maquat, L. E. Cap-binding protein 1-mediated and eukaryotic translation initiation factor 4E-mediated pioneer rounds of translation in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 4258–63 (2005).

71. Kuperwasser, N., Brogna, S., Dower, K. & Rosbash, M. Nonsense-mediated decay does not occur within the yeast nucleus. *RNA* **10,** 1907–15 (2004).

72. Maderazo, A. B., Belk, J. P., He, F. & Jacobson, A. Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution. *Mol. Cell. Biol.* **23,** 842–51 (2003).

73. Durand, S. & Lykke-Andersen, J. Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nat. Struct. Mol. Biol.* **20,** 702–9 (2013).

74. Popp, M. W.-L. & Maquat, L. E. The dharma of nonsense-mediated mRNA decay in mammalian cells. *Mol. Cells* **37,** 1–8 (2014).

75. Lejeune, F. & Maquat, L. E. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr. Opin. Cell Biol.* **17,** 309–15 (2005).

76. 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,** 1832–6 (2001).

77. Bühler, M., Steiner, S., Mohn, F., Paillusson, A. & Mühlemann, O. EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. *Nat. Struct. Mol. Biol.* **13,** 462–4 (2006).

78. Eberle, A. B., Stalder, L., Mathys, H., Orozco, R. Z. & Mühlemann, O. Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS Biol.* **6,** e92 (2008).

79. 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).

80. Amrani, N., Sachs, M. S. & Jacobson, A. Early nonsense: mRNA decay solves a translational problem. *Nat. Rev. Mol. Cell Biol.* **7,** 415–25 (2006).

81. Amrani, N. *et al.* A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432,** 112–8 (2004).

82. Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V. & Izaurralde, E. A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J.* **26,** 1591–601 (2007).

83. Kuzmiak, H. A. & Maquat, L. E. Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges. *Trends Mol. Med.* **12,** 306–16 (2006).

84. Stalder, L. & Mühlemann, O. The meaning of nonsense. *Trends Cell Biol.* **18,** 315–21 (2008).

85. Weischenfeldt, J. *et al.* NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. *Genes Dev.* **22,** 1381–96 (2008).

86. Lewis, B. P., Green, R. E. & Brenner, S. E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 189–92 (2003).

87. Neu-Yilik, G. & Kulozik, A. E. NMD: multitasking between mRNA surveillance and modulation of gene expression. *Adv. Genet.* **62,** 185–243 (2008).

88. Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M. & Mühlemann, O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* **17,** 2108–2118 (2011).

89. 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).

90. Martins, R. *et al.* Alternative polyadenylation and nonsense-mediated decay coordinately regulate the human HFE mRNA levels. *PLoS One* **7,** e35461 (2012).

91. 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,** 1259–69 (2014).

92. 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).

93. Thein, S. L. *et al.* Molecular basis for dominantly inherited inclusion body beta-thalassemia. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 3924–8 (1990).

94. Usuki, F. *et al.* Specific inhibition of nonsense-mediated mRNA decay components, SMG-1 or Upf1, rescues the phenotype of Ullrich disease fibroblasts. *Mol. Ther.* **14,** 351–60 (2006).

95. Usuki, F. *et al.* Inhibition of nonsense-mediated mRNA decay rescues the phenotype in Ullrich's disease. *Ann. Neurol.* **55,** 740–4 (2004).

96. Linde, L. & Kerem, B. Nonsense-mediated mRNA decay and cystic fibrosis. *Methods Mol. Biol.* **741,** 137–54 (2011).

97. Hinnebusch, A. G. eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem. Sci.* **31,** 553–62 (2006).

98. Zhou, M. *et al.* Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 18139–44 (2008).

99. Asano, K. *et al.* Structure of cDNAs encoding human eukaryotic initiation factor 3 subunits. Possible roles in RNA binding and macromolecular assembly. *J. Biol. Chem.* **272,** 27042–52 (1997).

100. Phan, L. *et al.* Identification of a translation initiation factor 3 (eIF3) core complex, conserved in yeast and mammals, that interacts with eIF5. *Mol. Cell. Biol.* **18,** 4935–46 (1998).

101. Block, K. L., Vornlocher, H. P. & Hershey, J. W. Characterization of cDNAs encoding the p44 and p35 subunits of human translation initiation factor eIF3. *J. Biol. Chem.* **273,** 31901–8 (1998).

102. Valásek, L., Phan, L., Schoenfeld, L. W., Valásková, V. & Hinnebusch, A. G. Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. *EMBO J.* **20,** 891–904 (2001).

103. Nielsen, K. H., Valásek, L., Sykes, C., Jivotovskaya, A. & Hinnebusch, A. G. Interaction of the RNP1 motif in PRT1 with HCR1 promotes 40S binding of eukaryotic initiation factor 3 in yeast. *Mol. Cell. Biol.* **26,** 2984–98 (2006).

104. Valásek, L., Hasek, J., Nielsen, K. H. & Hinnebusch, A. G. Dual function of eIF3j/Hcr1p in processing 20 S pre-rRNA and translation initiation. *J. Biol. Chem.* **276,** 43351–60 (2001).

105. Asano, K., Phan, L., Anderson, J. & Hinnebusch, A. G. Complex formation by all five homologues of mammalian translation initiation factor 3 subunits from yeast Saccharomyces cerevisiae. *J. Biol. Chem.* **273,** 18573–85 (1998).

106. Wagner, S., Herrmannová, A., Malík, R., Peclinovská, L. & Valášek, L. S. Functional and biochemical characterization of human eukaryotic translation initiation factor 3 in living cells. *Mol. Cell. Biol.* **34,** 3041–52 (2014).

107. Querol-Audi, J. *et al.* Architecture of human translation initiation factor 3. *Structure* **21,** 920–8 (2013).

108. Sun, C. *et al.* Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). *Proc. Natl. Acad. Sci. U. S. A.* **108,** 20473–8 (2011).

109. Villa, N., Do, A., Hershey, J. W. B. & Fraser, C. S. Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J. Biol. Chem.* **288,** 32932–40 (2013).

110. Morris, C., Wittmann, J., Jäck, H.-M. & Jalinot, P. Human INT6/eIF3e is required for nonsense-mediated mRNA decay. *EMBO Rep.* **8,** 596–602 (2007).

111. Pestova, T. V *et al.* in *Translational Control in Biology and Medicine* (eds. Mathews, M. B., Sonenberg, N. & Hershey, J. W. B.) **48,** 87–128 (Cold Spring Harbor Laboratory Press, 2007).

112. Peabody, D. S. The RNA binding site of bacteriophage MS2 coat protein. *EMBO J.* **12,** 595–600 (1993).

113. Iioka, H., Loiselle, D., Haystead, T. A. & Macara, I. G. Efficient detection of RNA-protein interactions using tethered RNAs. *Nucleic Acids Res.* **39,** e53 (2011).

114. Gehring, N. H., Hentze, M. W. & Kulozik, A. E. Tethering assays to investigate nonsense-mediated mRNA decay activating proteins. *Methods Enzymol.* **448,** 467–82 (2008).

115. Schneider-Poetsch, T. *et al.* Inhibition of Eukaryotic Translation Elongation by Cycloheximide and Lactimidomycin. *Nat. Chem. Biol.* **6,** 209–217 (2010).

116. Gregory, M. A. & Hann, S. R. c-Myc proteolysis by the ubiquitin-proteasome pathway: stabilization of c-Myc in Burkitt's lymphoma cells. *Mol. Cell. Biol.* **20,** 2423–35 (2000).

117. Romão, L. *et al.* Nonsense mutations in the human B-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation. *Blood* **96,** 2895–2901 (2000).

118. Silva, A. L. *et al.* The canonical UPF1-dependent nonsense-mediated mRNA decay is inhibited in transcripts carrying a short open reading frame independent of sequence context. *RNA* **12,** 2160–2170 (2006).

119. Park, H. S., Himmelbach, A., Browning, K. S., Hohn, T. & Ryabova, L. A. A plant viral 'reinitiation' factor interacts with the host translational machinery. *Cell* **106,** 723–33 (2001).

120. Hershey, J. W. B. The role of eIF3 and its individual subunits in cancer. *Biochim. Biophys. Acta* (2014). doi:10.1016/j.bbagrm.2014.10.005

121. Pick, E., Hofmann, K. & Glickman, M. H. PCI complexes: Beyond the proteasome, CSN, and eIF3 Troika. *Mol. Cell* **35,** 260–4 (2009).

122. Roobol, A. *et al.* The chaperonin CCT interacts with and mediates the correct folding and activity of three subunits of translation initiation factor eIF3: b, i and h. *Biochem. J.* **458,** 213–24 (2014).

123. Chiu, S.-Y., Lejeune, F., Ranganathan, A. C. & Maquat, L. E. The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex. *Genes Dev.* **18,** 745–54 (2004).