



INVESTIGATING THE LINK BETWEEN TRANSLATION AND NONSENSE-MEDIATED MRNA DECAY

by

MARIJA PETRIĆ HOWE

Supervised by Dr Saverio Brogna

February 2019

School of Biosciences
College of Life and Environmental Sciences
University of Birmingham

A thesis submitted to the University of Birmingham
for the degree of DOCTOR OF PHILOSOPHY

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Summary

Nonsense mediated mRNA decay (NMD) is widely regarded as an active cellular surveillance mechanism that eliminates mRNAs that contain a premature translation termination codon (PTC). In this study, I present an alternative hypothesis, formulated as the ribosome release model, which proposes that NMD may occur as a passive consequence of general roles that NMD factors play in ribosome release upon translation termination. I tested this directly, by looking at whether NMD factors affect NMD reporter mRNA association with the ribosomes. I also analysed the role of NMD factors in general translation and their association with translating mRNAs. My data indicates that NMD factors do bind mRNAs undergoing translation, regulating translation of many transcripts, and indeed possibly triggering ribosome release from mRNAs upon termination. Additionally, I found that UPF1 associates with ribosome-free RNA granules in the cell, into which in stress conditions, when global translation is impaired, it almost entirely re-distributes, possibly sequestering translationally repressed mRNAs from the translational pool. Finally, whilst investigating the role of NMD factors in translation by means of a puromycylation assay, I made unexpected observations that indicate that many complete proteins can be released from the ribosome with a tRNA still attached to the last C-terminal amino acid.

Acknowledgements

It is with immense gratitude that I acknowledge the support and help of my supervisor, Dr Saverio Brogna, for welcoming me to the world of research, constantly providing encouragement, patience, advice and a truly inspiring enthusiasm for science.

I would like to thank all the members of the lab, past and present, Tina, Jianming, Anand, Vibha, Andrea, Roy, Preethi, Akilu, with whom I've shared my PhD journey and who have created a brilliant supportive environment that has made my PhD experience a real joy. A special mention to Tina, Jianming, Joe and Tom, who have been such amazing friends, I will always remember the nights spent in the lab, our outings, many travels (all the way to China, even) that created wonderful memories I will always treasure. I would like to thank the students that I have supervised; all of you have been lovely and have made me into a more patient and enthusiastic teacher. I also want to thank all the members of the 6th floor, past and present, for being helpful, supportive and a true inspiration.

I would like to thank the Darwin Trust of Edinburgh for kindly sponsoring my PhD and the University of Birmingham for excellent research and teaching facilities. Thank you to all the lecturers who have given me the opportunity to do lab demonstrating, as well as public engagement. This has been rewarding in so many ways. I want to express my gratitude to Dr Robin May and Dr Klaus Futterer for their support, advice and understanding whilst monitoring my progress, as well as to Dr Yun Fan, Dr Alicia Hidalgo and Dr Matthias Soller, for sharing equipment and reagents.

A very special gratitude goes to the University of Belgrade and the Institute for Molecular Genetics and Genetic Engineering and their amazing professors and researchers. Had it not been for them, their enthusiasm for science and excellence in teaching, I would not have embarked on my PhD journey.

Special thanks to the wonderful people in my life outside of academia, to all my friends in Birmingham with whom I share the most amazing memories, here's to Sue, Marina, Frits, Sylvia, Bogdan, Giacomo, and many more. Huge thanks to my little Serbian-

Balkan group of amazing friends, Marta, Kristina, Anja, Doris, that have really helped me battle nostalgia and with whom I've spent late nights playing board games, gone travelling, eaten enormous amounts of Serbian food, lived with and more!

I want to thank all my friends in Serbia, for being so supportive of me, so understanding about my desire to go abroad to study, and who have always been by my side, making our friendship even stronger. I love you all. To Svetlana, Katarina, Dusica, Igor, Svetlana, Snezana, Marija, Dragana, Mihajlo, Jelena, Branislava and many more...

I want to thank my love, my wonderful husband, Nicholas, for his unconditional love, support, care and understanding, for being someone I can always confide in and rely on. I love you so much.

Last but by no means least, I want to thank my beautiful family. To my late granddad Petar who has always been so supportive of me pursuing my degree and always so enthusiastic and interested about my studies. I really miss you. To my grandma, Vida, who has brought me up and has always known what to say when I struggled and known how to make me feel better. Finally, I am truly indebted to my mom, Milena. Words cannot express how much I appreciate your sacrifice, unconditional love, care, generosity, positive attitude to live, support and strength. I love you more than anything.

I would love to dedicate this work to my mom, Milena Petric.

You are an inspiration to everything I aspire to be, and you are truly my hero.

Волела бих да посветим овај рад својој мами, Милени Петрић.

Ти си неисцрпна инспирација свему што тежим и за шта се борим,

ти си заиста мој херој.

Abbreviations used in this thesis

40S	Eukaryotic small ribosomal subunit
60S	Eukaryotic large ribosomal subunit
80S	Eukaryotic ribosome
CBC	Cap binding complex
CDS	Coding sequence
Cryo-EM	Cryo-electron microscopy
CTD	C terminal domain
ChIP	Chromatin immunoprecipitation
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DSE	Downstream sequence element
EDTA	Ethylenediaminetetraacetic acid
eEF	Eukaryotic translation elongation factor
eIF	Eukaryotic translation initiation factor
EJC	Exon junction complex
EM	Electron microscopy
eRF	Eukaryotic release factor
ES	Expansion segments
mRNP	Messenger ribonucleoprotein particle
GFP	Green fluorescent protein
NGD	No go decay
NMD	Nonsense mediated mRNA decay

NSD	Nonstop decay
ORF	Open reading frame
PABP	Poly(A) binding protein
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PTC	Premature translation termination codon
PMSF	Phenylmethanesulfonyl fluoride
Pol (I, II and III)	RNA polymerases (I, II and III)
RNA	Ribonucleic acid
RP	Ribosomal protein
RPM	Ribopuromycylation
RRM	RNA recognition motif
PCR	Polymerase chain reaction
SDS	Sodium dodecyl sulphate
SMG	Suppressor with morphogenic effects on genitalia
SS	Splice site
SURF	SMG1-UPF1-RFs complex
UPF	Up frameshift
UTR	Untranslated region

Contents

Summary	ii
Acknowledgements	iii
Abbreviations used in this thesis	vi
CHAPTER 1: INTRODUCTION.....	1
1.1. Eukaryotic gene expression	1
1.2. Transcription	4
1.3. Pre-mRNA processing.....	5
1.4. Translation	7
1.4.1. Ribosome structure and functional implications	7
1.4.2. The mechanism of translation	10
1.4.3. Methods of studying translation.....	18
1.5. Quality control of eukaryotic gene expression.....	23
1.6. Nonsense mediated mRNA decay (NMD)	25
1.7. Factors involved in NMD	27
1.7.1. Identification of the NMD factors.....	27
1.7.2. Structure and interactions of NMD factors.....	37
1.7.2.1. UPF1 structure and function	37
1.7.2.2. UPF2 structure and function	43
1.7.2.3. UPF3 structure and function	44
1.7.3. Role of NMD factors beyond NMD.....	45
1.8. Standard NMD models.....	48
1.8.1. DSE Model.....	48
1.8.2. The EJC Model.....	52
1.8.3. The Faux 3'UTR Model.....	57
1.9. Ribosome release model	60
1.10. CRISPR/Cas9 gene editing	65
Chapter 2. Materials and methods.....	68
2.1. <i>Schizosaccharomyces pombe</i> strains.....	68
2.2. <i>S. pombe</i> media.....	68
2.3. <i>S. pombe</i> culturing and maintenance	69
2.4. <i>S. pombe</i> cell counting.....	70
2.5. Genomic DNA extraction.....	71

2.6. RNA extraction from yeast cell cultures	72
2.7. Protein purification from <i>S. pombe</i> cell cultures	73
2.8. PCR based gene targeting – two-step PCR approach.....	74
2.9. CRISPR/Cas9 gene editing in <i>S. pombe</i>	78
2.10. pREP41 expression vector.....	79
2.11. Agarose gel electrophoresis and DNA gel extraction.....	81
2.12. Transformation of <i>S. pombe</i> strains	82
2.13. Colony PCR.....	84
2.14. Polysome profiling.....	85
2.14.1. Preparation of sucrose gradients	85
2.14.2. Cell lysis	86
2.14.3. Sedimentation of translation complexes	88
2.14.4. RNA and protein extraction from polysomal fractions.....	88
2.15. SDS-PAGE and Western blotting	90
2.16. Northern blotting	92
2.16.1. Gel electrophoresis and blot set-up.....	92
2.16.2. RNA UV crosslinking and hybridisation with the probe	93
2.16.3. Random priming – probe synthesis and purification.....	95
2.16.4. Membrane washes and exposure	96
2.17. Protein affinity purification from polysomal fractions	97
2.18. Silver staining and mass spectrometry analysis of stained proteins.....	98
2.19. GFP fluorescence imaging.....	100
2.20. <i>Escherichia coli</i> methods.....	101
2.20.1. <i>E. coli</i> strains.....	101
2.20.2. <i>E. coli</i> media and culturing	101
2.20.3. Gibson Assembly Cloning.....	102
2.20.4. <i>E. coli</i> transformation.....	103
2.20.5. Large scale plasmid preparation	103
2.20.6. Restriction enzyme digestion	104
RESULTS.....	105
Chapter 3: Re-evaluating methods of studying translation	105
3.1 Introduction	105
3.2 Results and discussion.....	109
3.2.1 Emetine does not stabilise polysomes in <i>S. pombe</i> , in contrast to cycloheximide	109

3.2.3 Puromycylation can be driven independently of ribosomes.....	115
3.2.5 Puromycylation potentially occurs in light fractions on peptidyl-tRNAs	124
3.2.6 Ribosome-independent puromycylation is not impaired upon EDTA treatment	133
3.2.7 Ribosome-independent puromycylation is impaired upon hydroxylamine treatment	136
3.2.8 Proteins puromycylated in a ribosome-independent manner represent highly expressed proteins.....	142
3.3 Conclusions	145

Chapter 4: Beyond NMD: Role of UPF1 in global translation and stress

response148

4.1 Introduction	148
4.2 Results	150
4.2.1 UPF1 associates with translating mRNAs.....	150
4.2.2 UPF3 associates with translating mRNAs.....	157
4.2.3 UPF1 and UPF3 have a role in translation efficiency.....	161
4.2.4 UPF1 is directed into heavy ribosome-free cytoplasmic compartments upon treatment with high salt	166
4.2.5 Over-expressed UPF1-GFP shows the same distribution pattern as that of endogenous UPF1-FLAG.....	169
4.2.6 UPF1-GFP localises in cytoplasmic granules upon heat-shock treatment	172
4.2.7 UPF1-GFP localises in cytoplasmic granules under higher concentrations of KCl.....	175
4.2.8 UPF1-GFP localisation in cytoplasmic granules is reversible.....	177
4.2.9 The composition of heavy cytoplasmic ribosome-free compartments suggests that UPF1 associates with stress granules	180
4.3 Discussion	183

Chapter 5: What is the mechanism of NMD?187

5.1. Introduction.....	187
5.2. Results	189
5.2.1. NMD is suppressed in upf1 Δ and upf2 Δ strains	189
5.2.2. NMD factors seem to affect ribosome release from mRNAs	191
5.2.3. NMD is position-dependent and it is fully suppressed in newly transformed upf1 Δ strains.....	196

5.2.4. NMD reporter mRNA was detectable only at an unexpectedly low level in polysome fractionation samples of upf1Δ strain	198
5.2.5. NMD reporter mRNA does not pellet during polysomal clearing stages.....	203
5.2.6. Modifications to the lysis buffer lead to an improved RNA yield and demonstrate that UPF1 and UPF3 may have a minor effect on ribosome release	205
5.3. Discussion.....	216
Chapter 6: CRISPR/Cas9 approach as a novel and efficient method to manipulate <i>S. pombe</i> genes.....	219
6.1 Introduction	219
6.2 Results	225
6.2.1 CRISPR/Cas9 gRNA design	225
6.2.2 Creating HR templates for CRISPR/Cas9 induced gene mutagenesis	234
6.2.3 Creating HR templates for CRISPR/Cas9 induced protein tagging.....	238
6.2.4 CRISPR/Cas9 gene editing approach efficiency	245
6.3 Discussion	252
Chapter 7: Conclusions	256
7.1. Ribosome release model	256
7.2. NMD factors partially associate with polysomes and affect translation	257
7.3. UPF1 associates with stress granules.....	259
7.5. Monosomes translate full length proteins.....	262
7.7. Ribosome-independent puromycylation uncovers that complete proteins can be released from the ribosome as peptidyl-tRNAs.....	262
References	265
8. APPENDIX.....	286
8.1. <i>S. pombe</i> strains.....	286
8.2. Media	288
8.3. Plasmids	291
8.4. Plasmid maps.....	292
8.5. Primers used for C-terminal protein tagging	295
8.6. CRISPR/Cas9 gene editing	296

CHAPTER 1: INTRODUCTION

1.1. Eukaryotic gene expression

Gene expression is a fundamental cellular process, which converts genetic information deposited in DNA into either RNA molecules that exert specific functions in the cell or messenger RNAs (mRNAs) that are translated into proteins, the main structural and functional cellular components.

Whilst there are fundamental similarities between prokaryotes and eukaryotes in terms of gene expression, this process is far more complex in eukaryotic organisms. Eukaryotic protein-coding gene expression comprises of transcription, pre-mRNA processing, formation of messenger ribonucleoprotein particles (mRNPs) and their transport into the cytoplasm, translation as well as posttranslational modifications of resulting proteins that regulate their localisation and functionality (Watson *et al.*, 2014) (Figure 1.1).

In contrast to prokaryotic cells, which do not contain nuclei so transcription and translation are coupled as ribosomes can bind to nascent transcripts, eukaryotic cells are compartmentalised. Transcription and pre-mRNA processing occur in the nucleus, the resulting mRNA is then transported through the nuclear pore complex into the cytoplasm where translation and further modifications take place (Moore and Proudfoot, 2009) (Figure 1.1).

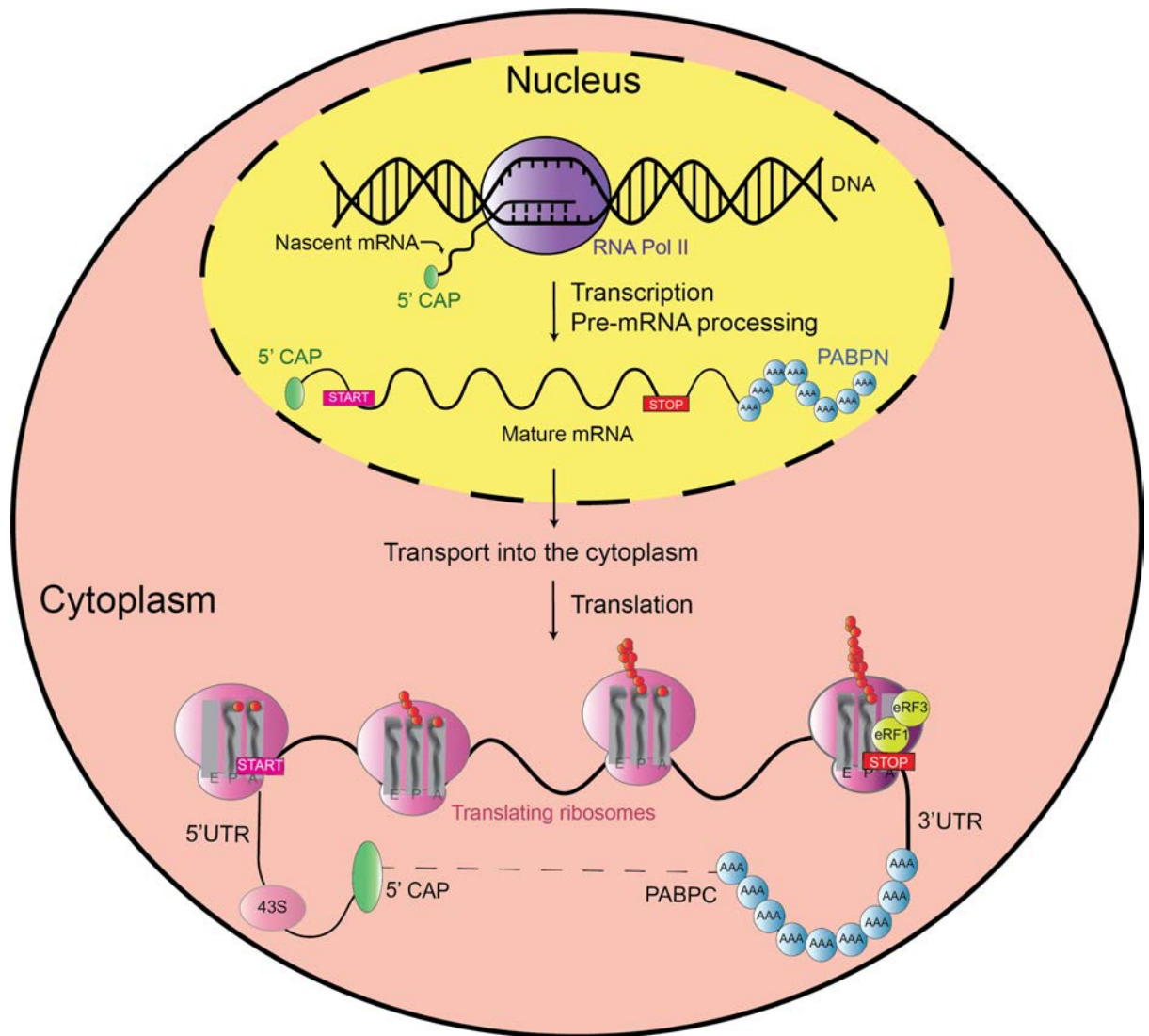


Figure 1.1 – **Overview of eukaryotic protein-coding gene expression.** Schematic displaying the key steps in eukaryotic gene expression of protein coding genes, detailed in the text. In brief, transcription uses a DNA molecule as a template to generate an mRNA molecule, a process catalysed by RNA Polymerase II (in purple). During transcription, mRNA is processed, which involves 5' capping, splicing, and cleavage and polyadenylation, creating a mature mRNA molecule that is complexed with a number of proteins, generating an mRNP. Mature mRNA is then transported through the nuclear pore complex into the cytoplasm, where it engages with ribosomes (in light purple) that translate its nucleotide sequence into the sequence of a polypeptide chain (amino acids presented as red circles). The newly synthesised polypeptide is released from the ribosome upon termination (terminating ribosome indicated in dark purple), due to a concerted action of eukaryotic release factors. 5' CAP – 5' capping complex; START – start codon; STOP – stop codon; PABCN – poly(A) binding complex, nuclear; 43S – translation pre-initiation complex; UTR – untranslated region; eRF1 – eukaryotic release factor 1; eRF3 – eukaryotic release factor 3; PABPC - poly(A) binding complex, cytoplasmic.

1.2. Transcription

A protein-coding gene is copied into a primary mRNA (pre-mRNA) molecule by means of transcription, a process catalysed by a multi-unit enzyme, RNA polymerase II. Association of transcription factors to promoter elements as well as chromatin remodelling allow RNA polymerase II to bind the DNA, the double helix to unwind locally, which enables RNA polymerase II to commence generating an mRNA molecule using DNA as a template (Shandilya and Roberts, 2012; Venters and Pugh, 2009).

Rpb1, the RNA polymerase II largest subunit, contains C-terminal domain (CTD) consisting of multiple amino-acid heptad repeats of the consensus motif Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (Harlen and Churchman, 2017). CTD is subjected to phosphorylation/dephosphorylation cycles of specific amino-acid residues during transcription. Phosphorylation patterns of Rpb1 CTD domain facilitate dynamic interactions of RNA Polymerase II with transcription factors, required for its progression through transcription stages (Buratowski, 2009). Key modifications, demonstrated by chromatin immunoprecipitation studies in budding yeast as well as mass spectrometry studies, include phosphorylated serine residue at position 5 affiliated with RNA polymerase II present at or near the promoter of a gene and serine 2 phosphorylation, which is a hallmark of elongating RNA polymerase II (KomarnitskyCho and Buratowski, 2000; Suh *et al.*, 2016; Mayfield *et al.*, 2017). Hence, transition from the initial low-processivity stage into productive transcription elongation, pre-mRNA processing steps as well as transition into termination and RNA cleavage and polyadenylation, is regulated by CTD phosphorylation/dephosphorylation

cycles, and appropriate factor association and activity (Buratowski, 2009). As a result, mRNA transcript is generated and processed, as detailed in section 1.3.

1.3. Pre-mRNA processing

Pre-mRNA molecules undergo different modifications during transcription, termed pre-mRNA processing, which include 5' cap formation (5' capping), splicing and 3' polyadenylation.

Once the nascent transcript is about 25-30 nucleotides long, it undergoes 5' capping, which involves a multi-step addition of a 7-methylguanosine to the transcript's 5' end via 5'-5' triphosphate bond (Topisirovic *et al.*, 2011). This is followed by the association of the cap binding complex (CBC) proteins, CBC20 and CBC80, with CBC20 directly binding to the cap (Shatkin, 1976).

Eukaryotic genes typically consist of exons, protein-coding sequences, and introns which, being non-protein-coding sequences, are eliminated from pre-mRNAs during splicing when the exons are joined together, generating an open reading frame. Splicing is catalysed by the spliceosome, a ribonucleoprotein particle, consisted of five small nuclear uridine-rich RNAs (U1, U2, U4, U5, U6 snRNAs) and a number of associated proteins (MooreQuery and Sharp, 1993). Intron defining features within a gene are mainly the 5' and 3' splice sites (ss), the branch point and the polypyrimidine tract (MooreQuery and Sharp, 1993; BurgeTuschl and Sharp, 1999). Spliceosome assembly and catalysis of the two consecutive transesterification splicing reactions are carried out via complex and dynamic sets of interactions with the nascent mRNA, mainly with the intron-defining features (WahlWill and Lührmann, 2009). Additionally, it was shown in

mammalian cells that during splicing, exon junction complexes (EJCs) are being deposited 22-24 nucleotides upstream of each exon-exon junction, which, among other functions, coat and protect the mRNA (Le Hir *et al.*, 2000; Le Hir *et al.*, 2001).

At the 3' end of the mRNA, sequences termed as signals for polyadenylation lead to the cleavage of nascent mRNA followed by a sequential addition of multiple adenine residues to form a poly(A) tail, to which poly(A) binding proteins (PABPs) bind (Colgan and Manley, 1997).

The 5' cap and the poly (A) tail protect the transcript from degradation by exonucleases, enable nuclear export, interact with a number of factors and have a role in translation initiation and mRNA circularisation, which is thought to allow stabilisation of the translation circuit and its rapid re-initiation (Watson *et al.*, 2014).

Therefore, an mRNA undergoes a series of modifications from its synthesis up to the final degradation step. It is associated with a number of specific proteins to form an mRNP, which is being remodelled as mRNA progresses through its life cycle (McCarthy and Kollmus, 1995; Mühlemann *et al.*, 2008). One of the key changes in mRNP composition occurs during its export into the cytoplasm, when the nuclear mRNP changes into a cytoplasmic mRNP (Kuersten and Goodwin, 2005).

1.4. Translation

Processed mature mRNAs are transported through the nuclear pore complex into the cytoplasm where they are translated into proteins (MehlinDaneholt and Skoglund, 1992; MehlinDaneholt and Skoglund, 1995). Translation is catalysed by the ribosome; a cellular ribonucleoprotein complex consisted of ribosomal RNAs (rRNAs) and proteins (Yusupova and Yusupov, 2014). The mRNA sequence determines the order of amino acids in the resulting protein; where each triplet of nucleotides, termed as codon, encodes an amino acid. Nucleotide triplets, anticodons, present in amino acid-carrying transport RNAs (tRNAs) recognise the codons via base pairing and consequently bring appropriate amino acids into the growing polypeptide chain. Therefore, tRNAs serve as adapter molecules, key for decoding the mRNA nucleotide sequence into amino acids of the resulting polypeptide (Kapp and Lorsch, 2004; Dever and Green, 2012).

1.4.1. Ribosome structure and functional implications

Eukaryotic ribosome (80S) is consisted of a large 60S and a small 40S ribosomal subunit. S indicates the Svedberg unit that measures sedimentation rate of particles upon centrifugation. The small ribosomal subunit is comprised of the 18S rRNA and 33 ribosomal proteins while the large ribosomal subunit consists of 28S (25S in yeast), 5.8S and 5S rRNAs forming a complex with 46 ribosomal proteins (Ben-Shem *et al.*, 2011). Bacterial ribosome (70S), in turn, has a similar structure with its large 50S subunit consisting of 23S and 5S rRNAs and 33 proteins, while its small 30S subunit consists of 16S rRNA and 21 associated proteins (Armache *et al.*, 2010; Yusupov *et al.*, 2001).

When the two subunits come together during translation, three major sites are formed on the interphase – E, P and A sites with specific functions during translation, detailed in section 1.4.2. In brief, A site is where aminoacyl-tRNAs form the codon-anticodon interactions bringing amino-acids, protein building blocks, to the nascent chain. P site is the location where the growing polypeptide chain resides, while the deacetylated tRNAs exit the ribosome within the E site (Melnikov *et al.*, 2012).

Structural and biochemical studies have resolved the structure of the two subunits and provided significant insight into the ribosome functioning. These ranged from studies resolving the structure of the bacterial ribosome to some of the key eukaryotic studies, including X-ray crystallography of yeast ribosome structure at 3.0Å (Ben-Shem *et al.*, 2011), and a cryo-EM study resolving the structure of the human ribosome in different phases of translation (Behrmann *et al.*, 2015).

One of the key findings regarding ribosome activity was that the active site of the ribosome, which catalyses peptide bond formation as well as peptidyl-tRNA hydrolysis, is comprised solely of rRNA, it is in fact a ribozyme (Cech, 2000). This was hypothesised early on, in a study that identified a region of 23S rRNA that associates with the peptidyl transferase site, indicating an important role of rRNA in the ribosome enzymatic activity (Barta *et al.*, 1984). Moreover, peptidyl transferase reaction remains largely unaffected even after approximately 95% of proteins from the large ribosomal subunit are removed while the reaction is in turn abolished upon treatments that destroy the RNA component of the ribosome (NollerHoffarth and Zimniak, 1992). The definitive confirmation that

the catalytic activity resides in the rRNA was the finding that a specific, conserved 23S rRNA region forms the catalytic site to which substrates bind, with the ribosomal protein component being as far as 18.0Å from the site of the reaction (Nissen *et al.*, 2000). Testing the catalytic activity of purified ribosomes containing an array of mutations in the 23S rRNA led to identification of conserved nucleotides responsible for peptidyl transferase and hydrolysis reactions (Youngman *et al.*, 2004).

The catalytic core of the ribosome, regarding both its rRNA and protein component, is highly evolutionarily conserved, with the structure of the peptidyl transferase centre between bacterial and eukaryotic ribosomes being nearly identical (Ben-Shem *et al.*, 2011; Dever and Green, 2012). The key tendency during evolution of the eukaryotic ribosomes was increasing the ribosome size, particularly with the appearance of outer rRNA expansion segments that do not affect the structure of the conserved core but provide a platform for new interactions with different cellular components (Gerbi, 1996; Spahn *et al.*, 2001). Interestingly, a recent study establishing the ribosome affinity purification method of mammalian ribosomes, uncovered an array of ribosome associated proteins (RAPs) that regulate cell cycle and cell metabolism-related translation as well as RNA- and protein-modifying enzymes that affect ribosome functionality (Simsek *et al.*, 2017). Additionally, it was found that ribosomes could be heterogeneous in terms of ribosomal proteins they consist of. Thereby, ribosomes of a certain RP composition direct translation of a subset of transcripts, adding another layer to translational regulation (Shi *et al.*, 2017).

1.4.2. The mechanism of translation

In eukaryotes translation consists of four stages: initiation, elongation, termination and ribosome recycling, as detailed below (Figure 1.2). In short, during initiation stages, the 40S subunit, together with the initiating tRNA carrying the first amino acid, associates with the 5' cap of the mRNA, scans the 5'UTR until it encounters the start codon, upon which the large ribosomal unit joins and forms the complete ribosome. During elongation, ribosome propagates across the mRNA, catalysing polypeptide chain formation based on the information in the mRNA. Once the stop codon is encountered, the ribosome catalyses hydrolysis of the peptidyl-tRNA and release of the polypeptide. Finally, post-termination ribosomes dissociate from mRNA into ribosomal subunits becoming available to engage in a new cycle of translation (Hinnebusch and Lorsch, 2012; JacksonHellen and Pestova, 2010).

Figure 1.2 – **The mechanism of translation.** (A) Schematic of a translation circuit composed of translating ribosomes trailing on an mRNA, displaying the key stages of translation. Detailed schematics of individual translation stages are shown below (B-D), as follows. (B) Translation initiation. A detailed explanation of this stage is presented in section 1.4.2.1. In brief, nuclear cap binding complex is replaced in the cytoplasm by the eIF4F factors: eIF4E, eIF4G and eIF4A, as indicated. A pre-initiation 43S complex (PIC) assembles, composed of the small ribosomal subunit, the ternary complex eIF2-GFP-Met-tRNA^{Met} and associated factors eIF1 and eIF3. Once the PIC binds to the 5' cap, it generates a 48S complex. eIF4A is a 5'-3' RNA helicase that translocates across the 5' UTR, unwinding secondary structures and thus enabling small ribosomal subunit with associated factors to move across the 5' UTR, scanning for the start codon. Once the initiating tRNA successfully base pairs with the start codon, eIF5 stimulates GTP hydrolysis of the ternary complex and eIF5B-mediated association of the large ribosomal subunit. (C) Translation elongation, detailed in section 1.4.2.2. eEF1-α-GTP-aminoacyl-tRNA complementary base pairs with the second codon. If the base pairing is perfect, GTP hydrolysis ensues, releasing eEF1-α and allowing peptidyl-transferase centre to properly position the two amino acids and to catalyse peptide bond formation. The tRNAs are then found in the P/E and A/P hybrid states, as indicated. eEF2 stimulates ribosome translocation three nucleotides across the mRNA which resolves the hybrid state. Deacetylated tRNA exits the ribosome via the E site. tRNA carrying the growing polypeptide chain is present in the P site and the A site is free, allowing the next aminoacyl-tRNA to bind. (D) Translation termination and ribosome recycling, detailed in section 1.4.2.3 and 1.4.2.4. Termination occurs once the ribosome encounters a stop codon (UAA, UAG or UGA). The stop codon is recognised by the eRF1-eRF3-GTP complex, where eRF1 recognises the stop codon and eRF3 then catalyses GTP hydrolysis that stimulates eRF1-mediated peptidyl-tRNA hydrolysis and release of the polypeptide. Ribosomes are then dissociated from the

mRNA via an unclear mechanism. It is proposed that eIF3 and ABCE1 play a key role, stimulating dissociation of the large ribosomal subunit. eIF1 then stimulates tRNA dissociation while eIF3j releases the mRNA, allowing ribosomal subunits to engage in a new translation cycle. 5' CAP – 5' capping complex; UTR – untranslated region; START – start codon; STOP – stop codon; 40S – small ribosomal subunit; 60S – large ribosomal subunit; PABPC – poly(A)-binding protein, cytosolic; eIF – eukaryotic initiation factor; eEF – eukaryotic elongation factor; eRF – eukaryotic release factor; GTP- guanosine-triphosphate.

1.4.2.1. Translation initiation

Upon mature mRNA transport into the cytoplasm, the cytoplasmic eIF4F complex consisting of eIF4E, eIF4G and eIF4A replaces the nuclear CBC20/CBC80 cap-binding complex (Pestova *et al.*, 2001). eIF4E directly binds to the cap while eIF4G interacts with the PABPC bound to the 3' end poly(A) tail, causing the mRNA to form a closed loop, thought to be responsible for stabilisation of the translation circuit and translation re-initiation (Wells *et al.*, 1998). eIF4A is a helicase that unwinds any secondary mRNA structures, allowing the 48S pre-initiation complex (PIC) to scan the 5' UTR of the mRNA (Figure 1.2B) (Pestova *et al.*, 2001; PrévôtDarlix and Ohlmann, 2003).

The initiation stage commences when the PIC, consisted of the small ribosomal subunit with associated eukaryotic initiation factors eIF1, eIF1A, eIF3 and the ternary complex comprised of eIF2 complexed with initiating aminoacyl-tRNA carrying methionine and a GTP molecule (eIF2-GTP-Met-tRNA_i^{Met}), assembles on the mRNA (Pestova *et al.*, 2001; JacksonHellen and Pestova, 2010). The assembly is mediated via an interaction between eIF4G, bound to the mRNA, and eIF3, associating with the PIC, generating a 48S complex (Korneeva *et al.*, 2000; Villa *et al.*, 2013; Eliseev *et al.*, 2018). The complex then scans the mRNA until it reaches the AUG initiation codon, detected by the initiator tRNA anticodon-start codon base pairing (JacksonHellen and Pestova, 2010; Hinnebusch and Lorsch, 2012). Once the initiation codon is identified, eIF5, via interaction with eIF2 and eIF3, triggers hydrolysis of the GTP bound to eIF2, which further leads to dissociation of the initiation factors and eIF5B-mediated binding of the large ribosomal subunit (Das and Maitra, 2001; Nag *et al.*, 2016).

Initiating Met-tRNA_i^{Met} remains hybridised with the mRNA, residing within the ribosome P site (Figure 1.2B) (Hinnebusch and Lorsch, 2012).

1.4.2.2. Translation elongation

As previously outlined, the ribosome contains E, P and A sites. Initiating tRNA^{Met} resides within the P site upon the completion of the initiation step. Elongation step commences when the eukaryotic elongation factor eEF-1 α complexed with GTP binds an aminoacyl-tRNA and directs it to the ribosome A site, allowing it to base pair with the second mRNA codon (Figure 1.2C) (Kapp and Lorsch, 2004; Dever and Green, 2012). Upon successful anticodon-codon base pairing, GTP is hydrolysed and the elongation factor released from the ribosome. The GTP hydrolysis is a rate-limiting step as the time required for it allows tRNAs whose anticodons do not bind perfectly to the codon to dissociate from the ribosome, serving as a proofreading mechanism (Dever and Green, 2012). Once eEF-1 α is released, the peptidyl transferase centre catalyses peptide bond formation between the methionine and the second aminoacyl-tRNA (Polacek and Mankin, 2005). When the peptide bond is formed, the two aminoacyl-tRNAs are found in the P/E and A/P hybrid states, where the peptide is bound to the tRNA with its acceptor arm present in the P site and its anticodon located in the A site. The deacetylated tRNA in turn has its acceptor arm in the E site while its anticodon is within the P site (Schuller and Green, 2018). The subsequent translocation of the ribosome three nucleotides downstream within the mRNA sequence, mediated by the elongation factor eEF-2 and utilising energy of GTP hydrolysis, leads to deacetylated tRNA transferring into the E site and exiting the ribosome, while the peptidyl-tRNA

now fully residing in the P site (Dever and Green, 2012). The A site is free for a new aminoacyl-tRNA to bind and introduce the next amino acid into the peptide chain (Figure 1.2C). For a new aminoacyl-tRNA to bind to the ribosome, GDP now bound to eEF-1 α has to be converted back to GTP, a reaction catalysed by eEF-1 β (Gromadski *et al.*, 2007). The ribosome progresses through elongation via repeated cycles of aminoacyl-tRNA binding to the ribosome, peptide bond formation and translocation, until it reaches the stop codon (UAA, UAG or UGA triplet).

1.4.2.3. Translation termination

When the ribosome encounters one of the three stop codons, located in its A site, the stop codon cannot be recognised by an aminoacyl-tRNA, as there are no anticodons compatible to it; it is recognised by release factors instead (Dever and Green, 2012). Release factor eRF1 can recognise each of the three stop codons (Chavatte *et al.*, 2002). eRF3 factor complexed with GTP, via its interaction with eRF1, also binds the terminating ribosome. The energy produced by GTP hydrolysis stimulates a conformational shift of eRF1, triggering hydrolysis of the ester bond in the peptidyl-tRNA and generating a newly synthesised polypeptide chain and an empty deacetylated-tRNA (Figure 1.2D) (Frolova *et al.*, 1996). Polypeptide release is followed by post-termination ribosome dissociation.

1.4.2.4. Ribosome recycling

Upon successful release of the synthesised polypeptide chain, post-termination ribosome contains a deacetylated tRNA in its P site and the release factors in its A site. A ribosome-recycling factor (RRF) was identified in bacteria, playing a crucial role in the dissociation of the ribosomes upon translation termination (Hirokawa *et al.*, 2005). A homologue has not been identified in eukaryotes and the exact mechanism of post-termination ribosome disassembly is yet not completely clear. Two proteins, eukaryotic initiation factor complex eIF3, as well as an ABC-family ATPase ABCE1 are thought to play a role in these processes (Figure 1.2D) (PisarevHellen and Pestova, 2007; Pisarev *et al.*, 2010).

The only factor found to be capable of stimulating post-termination ribosome release alone is eIF3, with its activity being stimulated by eIF3j, eIF1 and eIF1A (PisarevHellen and Pestova, 2007). Once incubated with *in vitro* assembled translation termination complexes, eIF3 stimulated dissociation of approximately 45% of ribosomes, with eIF1 and eIF1A stimulating disassembly of an extra 15%, while eIF3j had a more substantial effect, with eIF3 and eIF3j together stimulating dissociation of 70-75% of terminating ribosomes into subunits (PisarevHellen and Pestova, 2007). Concerted actions of eIF3 and eIF1A generate 40S/mRNA/tRNA complexes where eIF1 then stimulates release of deacetylated tRNA and eIF3j triggers mRNA dissociation from the 40S subunit (Figure 1.2D). eIF3 was also shown to remain with the dissociated 40S subunits, preventing them from re-associating with 60S. This mechanism of ribosome recycling is functional only in low Mg^{2+} conditions (PisarevHellen and Pestova, 2007).

ABC-family ATPase ABCE1, on the other hand, was shown to significantly stimulate the disassembly of *in vitro* assembled termination complexes into 60S subunits and mRNA/tRNA/40S complexes over a wide range of Mg^{2+} concentrations (Figure 1.2D) (Pisarev *et al.*, 2010). It was postulated that ABCE1 promotes the rate of peptide release via association to the release factors, playing a role in post-termination ribosome disassembly by transforming ATP hydrolysis energy into movements that can separate the two subunits (Pisarev *et al.*, 2010; Barthelme *et al.*, 2011). The joint action of all the above-discussed factors allows the ribosome recycling to take place; however, further studies are necessary to acquire full understanding of this process.

1.4.3. Methods of studying translation

A wide range of structural and biochemical methods have been developed to study the process of translation, *in vitro* translation assays being among the most important. Another significant method is polysome fractionation, which enables separation of translation complexes using sucrose gradients, thereby allowing studies into the cellular translation profiles, translation of individual mRNAs, as well as into the roles of different translation factors in the process (Chassé *et al.*, 2017). Ribosome profiling, in turn, is a more recently developed genome-wide method that enables tracking ribosome occupancy on mRNAs by analysing ribosome footprints, at a nucleotide resolution (Ingolia *et al.*, 2009).

One of traditional ways to investigate translation-related processes is by utilising a range of translation inhibitors, typically antibiotics isolated from various bacterial sources that inhibit translation at specific stages. Cycloheximide and emetine are among the standard translation elongation

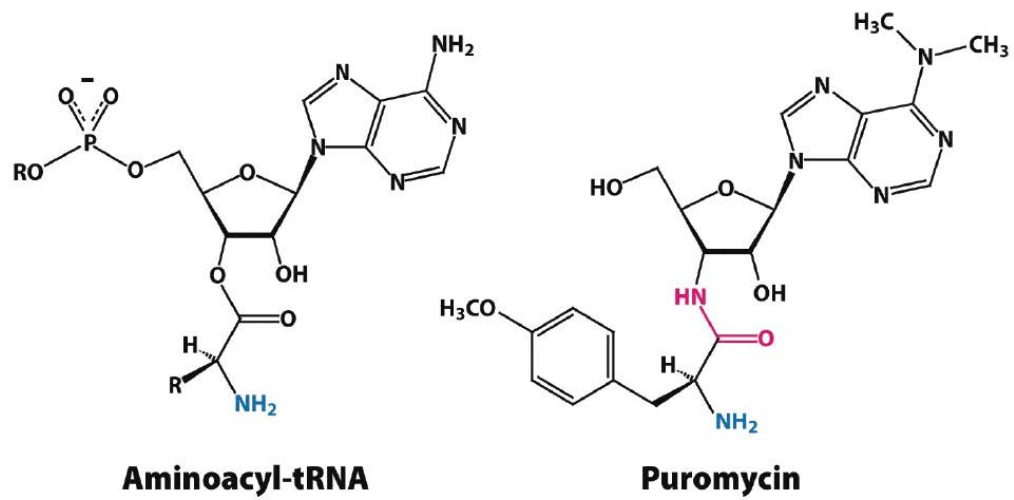
inhibitors, typically used in polysome and ribosome profiling studies as they trap the ribosomes on mRNAs and allow creating a snapshot of cellular translation profile at a given time (Pestka, 1971). Cycloheximide acts by binding to the large ribosomal subunit, within its E site, impairing eEF2-mediated mRNA-tRNA translocation, thereby inhibiting translational elongation (Grollman, 1966; Schneider-Poetsch *et al.*, 2010). Cycloheximide can bind to the E site regardless of whether the site is occupied with a deacetylated tRNA or not, therefore, it can inhibit translation elongation at any given time (Schneider-Poetsch *et al.*, 2010). Emetine was shown to inhibit translation by binding to the small ribosomal subunit, within its E site, and again, by inhibiting mRNA/tRNA translocation, similarly to cycloheximide (Grollman, 1966; Wong *et al.*, 2014). On the other hands, anisomycin, another potent translation inhibitor, binds to the ribosome A site, blocking new amino-acid incorporation and, hence, stimulating translation arrest (Grollman, 1967; IoannouCoutsogeorgopoulos and Synetos, 1998).

Puromycin is an additional important translation inhibitor used from the very classic studies until the present day. Structurally, it mimics amino-acyl-tRNA and as such, it binds to the ribosome A site (Figure 1.3A) (Pestka, 1971). The ribosome catalyses its incorporation into the nascent peptide chain which leads to an immediate polypeptide release and dissociation of translation complexes (Figure 1.3B) (Azzam and Algranati, 1973). Puromycin was used in structural studies, as well as in studies investigating the mechanism of translation. Since it gets incorporated into nascent peptides, it can be used to analyse proteins being synthesised in the cells at a given time. Also, a ribopuromycylation

technique has recently been developed where puromycin treatment was combined with pre-treatment with cycloheximide or emetine (DavidBennink and Yewdell, 2013). The authors reported that immobilising the ribosomes using translation elongation inhibitors will prevent the release of the puromycylated peptides, therefore, this technique, paired with immunostaining, could be used to visualise translation loci in the cells (DavidBennink and Yewdell, 2013).

It is widely acknowledged that puromycylation is catalysed exclusively by the ribosome. Interestingly, in this study, I demonstrated that puromycylation could potentially be driven independently of the ribosome (detailed in Chapter 3).

A



B

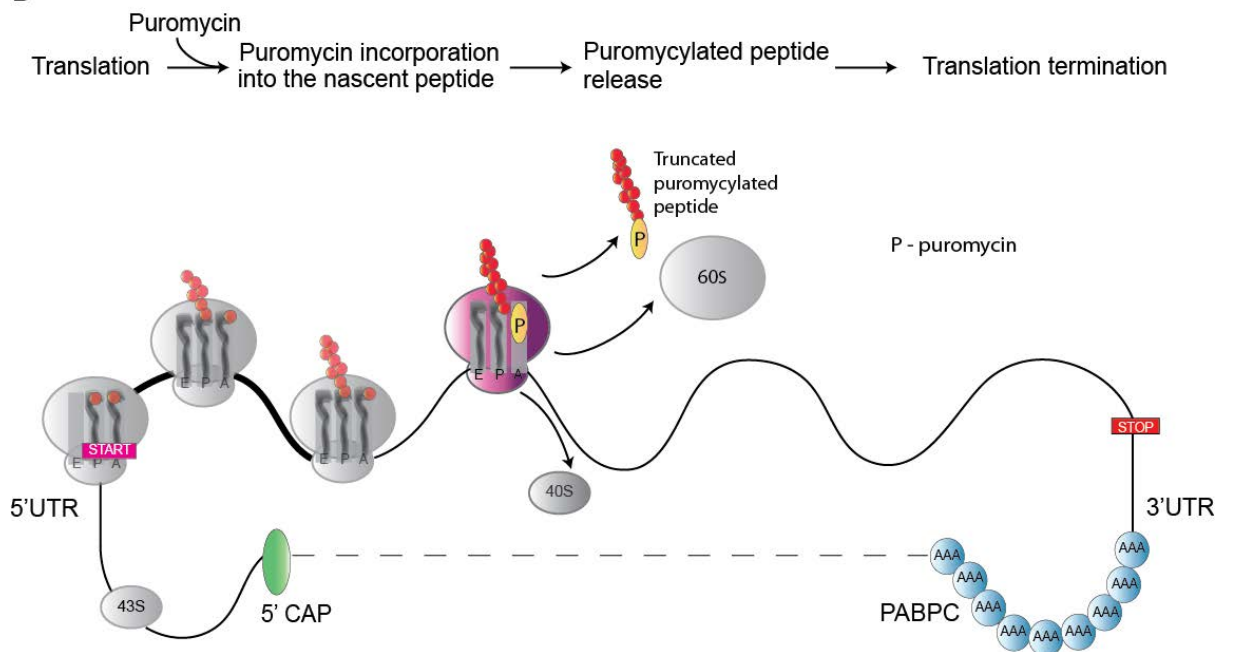


Figure 1.3 – Puromycin gets incorporated into the nascent polypeptide via reaction catalysed by translating ribosomes. (A) Comparison between the structure of an aminoacyl-tRNA and puromycin. **(B)** Schematic of the puromylation reaction. In brief, puromycin, upon entering an empty ribosomal A site, gets incorporated into the nascent peptide chain. This leads to its release and, consequently, release of ribosomal subunits and translation inhibition. 40S – small ribosomal subunit; 60S – large ribosomal subunit; 80S – monosome; UTR – untranslated region; PABPC – poly(A)-binding protein, cytosolic.

1.5. Quality control of eukaryotic gene expression

Many mechanisms have evolved to regulate gene expression on every step of the complex mRNA life cycle, ensuring that it occurs in a correct temporal and spatial manner. On the other hand, as gene expression steps are prone to mistakes, which can lead to production of aberrant transcripts and proteins, various quality control mechanisms are thought to have evolved to ensure gene expression fidelity, by detection and elimination of such aberrant products, together with their potential toxic effects (Mühlemann and Jensen, 2012).

Mechanisms of eukaryotic mRNA surveillance coupled with translation are no go decay (NGD), nonstop decay (NSD) and nonsense mediated mRNA decay (NMD) (AtkinsonBaldauf and Hauryliuk, 2008). NGD happens when a ribosome stalls on an mRNA undergoing translation due to an occurrence of an RNA secondary structure or sequence elements (Doma and Parker, 2006; Becker *et al.*, 2011). NSD eliminates mRNAs that lack a stop codon or occurs when ribosomes have read through the normal stop codon (Klauer and van Hoof, 2012; Sundermeier *et al.*, 2008). An additional mechanism has been discovered in mammalian cells – Staufen mediated decay (SMD) (Kim *et al.*, 2005). In this process a double stranded mRNA binding protein, STAU1, triggers degradation by recognising its binding site in the 3'UTR of target mRNAs. This binding site can either be an intramolecular secondary structure or an intermolecular base pairing of a 3'-UTR sequence with a long-noncoding RNA (lncRNA) (Kim *et al.*, 2005; Park and Maquat, 2013).

Moreover, as a consequence of translation errors, nascent polypeptide might be released from the ribosome with a tRNA still attached to it, as a peptidyl-tRNA,

instead of being hydrolysed and released as a polypeptide (Harigaya and Parker, 2010; Cruz-Vera *et al.*, 2003). This process has mainly been described in bacteria, where the peptidyl-tRNAs released prematurely from ribosomes are typically very short (Cruz-Vera *et al.*, 2003; Cruz-Vera *et al.*, 2004). Peptidyl-tRNA accumulation in prokaryotes can impede translation as it can significantly reduce the cellular pool of aminoacyl-tRNAs, protein synthesis building blocks. These aberrant translation products are destroyed by the action of peptidyl-tRNA hydrolase (Pth) in bacteria (Kössel, 1970). It is yet unclear how frequently the described process occurs in eukaryotic organisms as these translation by-products have largely been overlooked. Peptidyl-tRNAs were mainly reported as a consequence of ribosome stalling that triggers the NGD process (Harigaya and Parker, 2010). Two homologues of the bacterial enzyme, Pth and Pth2, were identified in eukaryotes, one of which was shown to indeed possess peptidyl-tRNA hydrolase activity (Pth2), while the second, Pth, has shown no such activity and its function remains unknown (BurksMcFeeters and McFeeters, 2016). Pth2 purified from rabbit reticulocytes catalyses peptidyl-tRNA hydrolysis only when they are released from the ribosomes, and has no effect on ribosome-associated peptidyl-tRNAs (Gross *et al.*, 1992; GrossCrow and White, 1992). Human Pth2 crystal structure has been resolved and the protein was found to catalyse peptidyl-tRNA hydrolysis at a much higher rate than the bacterial Pth, but at a rate comparable to that of Pth2 in archaea (De Pereda *et al.*, 2004; Fromant *et al.*, 2003). These observations suggest that peptidyl-tRNAs do arise during translation, being released from the ribosome and eliminated from the cell by peptidyl hydrolase activity. Results presented in

this study suggest that peptidyl-tRNAs dissociation from ribosomes could have a higher occurrence in eukaryotic cells than originally thought; how much these translation by-products contribute to overall translation should be explored further.

1.6. Nonsense mediated mRNA decay (NMD)

NMD describes a universal translation-coupled phenomenon of nonsense mutations causing a reduction in mRNA levels. It is recognised as an evolutionarily conserved mechanism that initiates rapid degradation of transcripts that possess a premature translation termination codon (PTC) (Brognia and Wen, 2009; BrogniaMcLeod and Petric, 2016). A PTC can occur as a consequence of frameshift or nonsense mutations or as a secondary consequence of errors during transcription or pre-mRNA processing (Mühlemann *et al.*, 2008; Schweingruber *et al.*, 2013). It typically occurs as a result of inefficient splicing, in alternately spliced isoforms, in presence of upstream reading frames (uORFs), atypically long 3'UTRs or in genes particularly prone to mutations such as T-cell-receptor, immunoglobulin and other antigen-receptor genes (CelikKervestin and Jacobson, 2015; Mühlemann *et al.*, 2008).

NMD was initially regarded exclusively as a surveillance mechanism that eliminates PTC-containing transcripts, preventing synthesis of truncated proteins that would possibly have toxic effects on normal cell functioning (Czaplinski *et al.*, 1998). However, it was shown in a variety of organisms, ranging from *S. cerevisiae* to human cells, that upon deletion of NMD factors, expression of not only NMD targets but also normal transcripts becomes

upregulated (Guan *et al.*, 2006; Mendell *et al.*, 2004). This suggested that NMD could have a more widespread role in gene expression regulation as expression of 3-10% of all physiological transcripts is altered in NMD mutants, depending on the organism (Stalder and Mühlemann, 2008). Many of these transcripts are involved in cellular processes such as metabolism, cell cycle progression, DNA repair and telomere maintenance (Stalder and Mühlemann, 2008).

The study of NMD is medically important, as approximately 30% of all human disease-inducing mutations are nonsense mutations. Modulation of NMD can therefore affect the clinical outcome of such diseases (Stalder and Mühlemann, 2008; Mühlemann *et al.*, 2008). Due to its significance, NMD has been studied to a large extent, various factors involved in it have been identified and several models have been proposed to explain how this process is carried out. However, some of the key aspects of its functioning in the cell still remain unclear.

1.7. Factors involved in NMD

1.7.1. Identification of the NMD factors

NMD factors were initially identified in a genetic screen of randomly induced mutants (CulbertsonUnderbrink and Fink, 1980). A *his4-38* mutant contains a G/C insertion, which causes a frameshift in the *HIS4* gene, introducing a premature translation termination codon (PTC). Consequently, the enzyme encoded by this gene, important for later phases of histidine biosynthesis, is not expressed and the cells, therefore, cannot produce histidine. In a genetic background of a tRNA frameshift suppressor in which this four-base codon is also recognized and translated, the open reading frame is restored and the enzyme is expressed, allowing cells to grow in the absence of histidine (His^+ phenotype). The frameshift suppressor is efficient only at 30°C in normal cells, while at 37°C the suppression resulting in *HIS4* protein synthesis occurs only in the presence of mutations in three *S. cerevisiae* genes. Proteins encoded by these three genes were consequently named Up-frameshift 1, 2, and 3 (UPF1, UPF2 and UPF3) (CulbertsonUnderbrink and Fink, 1980) since mutations in the genes encoding these proteins would cause up-frameshift suppression. Their homologues, which comprise the core NMD machinery, were identified in multicellular organisms, together with additional factors.

Six NMD factors were identified in a mutagenesis screen in *Caenorhabditis elegans*, and due to the phenotype of these mutants, they were named suppressors with morphogenic effects on genitalia 1-6 (SMG1-6) (Hodgkin *et al.*, 1989; Pulak and Anderson, 1993). A seventh was later discovered and named SMG7 (Cali *et al.*, 1999). Two additional factors named SMG lethal 1

and 2 (SMGL-1 and SMGL-2) were identified in a genome wide RNAi screen in *C. elegans* (Longman *et al.*, 2007b; Longman *et al.*, 2007a). Finally, the last two proteins involved in the NMD process were discovered in the most recent study in mammalian cells, named SMG8 and SMG9 (Yamashita *et al.*, 2009).

SMG2, SMG3 and SMG4 are homologous to the yeast UPF1, UPF2 and UPF3, respectively. Orthologues of the remaining SMG proteins were not identified in yeast; all are present in mammalian cells, while *Drosophila melanogaster* lacks only the orthologue of SMG7 (Mühlemann *et al.*, 2008).

UPF1, UPF2 and UPF3 are considered to be the core of the NMD complex and are highly evolutionarily conserved, as demonstrated by the protein sequence alignment between a number of eukaryotic organisms shown in Figures 1.4-1.6 (Kerényi *et al.*, 2008). The results demonstrate that UPF1 is most evolutionarily conserved protein of the three, as UPF1 of *S. pombe* shares 54.6% of amino acid identity with human UPF1 and 51.3% with *S. cerevisiae* (Figure 1.4) (calculated as a percentage of identity between the two sequences using the pairwise alignment function of the Jalview software). As for UPF2, there is 25.1% amino acid identity between *S. pombe* and human, and 22.3% between *S. cerevisiae* and human (Figure 1.5). UPF3 is the least conserved of the three, with *S. pombe* sharing 20.5% amino acid identity with human UPF3 and 22.7% with budding yeast UPF3 (Figure 1.6).

Sc_UPF1
Sp_UPF1
At_UPF1
Dm_UPF1
Ce_UPF1
Xl_UPF1
Mm_UPF1
Hs_UPF1

10 20 30 40 50 60 70

MDSQSDSLFDTASQPDIV--ADEYTFLEFNTQGDSEFYQDFGSPATAWPTPSDSISIAADVADRGEAGAAADHH
---MSVDTYAPSSA-LSLDMDDNELLPAGADTQPTQYDYRDTMPSTSSQSTQND-----QLE-----
MDDSDDEYSRSHGETLTFVDPEDDGV-SIGNTCDQSQAAYEQFSVPTQSSQATDLL-----PGGTDGTTN---
---MSVEAYGPSSQTLTFLDTEEAEL-LGADTCGSEFEFTDFTLPSQTQGGSQ-----
---MSVEAYGPSSQTLTFLDTEEAEL-LGADTCGSEFEFTDFTLPSQTQTPPGGP-----GG-----AGGP
---MSVEAYGPSSQTLTFLDTEEAEL-LGADTCGSEFEFTDFTLPSQTQTPPGGP-----GGPGGGGAGGP

80 90 100 110 120 130 140

-----MVGSG-SHTPYDINSNSPDVNVQPATQLNSTLVEDDDVDNQLFEE-AQVTEIGFRS
-----MSLGLQPNNDISLVSSKNMT-----SENGLEHQFEE-----LLVE
SEASSPSSLSAGAGNGAKVGRGVGSGGVSSSSQVD-AL-----AAGVGNLFEEETGDDGFDYK
-----IAQ-----RCSAGD--SHPLRA-SI-----TNDLADLFEEEDDEPGSSY-V
-----DLFFHDVEDDES-----SEKSLTE-----EQHE
-----SQLDGQVNGPDMLNQGT--VDDGVG-KT-----SQLGLGELNFE-EDEEDTYY-T
GGAG--AGGAAGQLD-AQVGPEGLONGA--VDDSVA-KT-----SOLLAEELNFE-EDEEDTYY-T
GGAG--AGAAAGQLD-AQVGPEGLONGA--VDDSVA-KT-----SOLLAEELNFE-EDEEDTYY-T

150 160 170 180 190 200 210

PSASDNSCAYCGIDSAKCVIKCN--SCKKWFNCNTKNGTSSSHIVNHLVLSHNVVSLHPDSDLGDTVLECYNC
KQYSEEHCAVCHIKNFNSILKCL--HCNKWFCNVRGKSGASHIISHLVRRARHQVALHSHSSLSDTVLECYNC
NDFTEHACKYCGISNPACVVRCNVASCRKWFCNVRGNTSGSHIVNHLVRAKHKEVCLHRDSPLGETILECYNC
KELPFAHACKYCGIHDPATVVMCN--NCRKWFCNVRGNTSGSHIINHLVRAKHKEVCLHGEGLGETILECYNC
QKLPFAHACKYCGISDPLCAVAKCT--VCRKWFCNVRGNTSGSHIIVHVMVRSQHEAYTHKSPCGDTQLECYNC
KDLPLVHACS YCGIHDPACVVCYN--TSKKWFCNVRGNTSGSHIVNHLVRAKHKEVCLHKGDLGETVLECYNC
KDLPLVHACS YCGIHDPACVVCYN--TSKKWFCNVRGNTSGSHIVNHLVRAKHKEVCLHKGDLGETVLECYNC
KDLPLVHACS YCGIHDPACVVCYN--TSKKWFCNVRGNTSGSHIVNHLVRAKHKEVCLHKGDLGETVLECYNC

220 230 240 250 260 270 280 290

GRKNVFLVLFVSIAKSEA VVVLVLCRI PCAQT---KNA NWDTDOWQPLIEDRQLSWVAEQPTEEKKARLITP
GTNRVFLVLFIPAKAKTIVVVLVLCRQPCARASIAKDMNWDLTOWQPLISDRQFLPWLITPSEEEQKLAIPITS
GCRNVFLVLFISAKTDSVVVLVLCRQPCALVNALKDMNWDLTOWQPLIDDRCFPLWLKVVPSEEEQRLARQISA
GVNRVFLVLFIPAKADSVVVLVLCRQPCAAQNSLKDMNWDLTOWQPLIADRCFLAWLVKQPSQEQQLRARQISA
GSKNVFLVLFIPGKKDQVVVLCRTPCASIAFQNDNDWSPEDWKSVAEKQLLSWIVNVPSEEEQVARKIITA
GCRNVFLVLFIPAKADSVVVLVLCRQPCASQSSLKDI NWDSQWQPLIQDRCFLSWL VKIPSEEEQRLARQITA
GCRNVFLVLFIPAKADSVVVLVLCRQPCASQSSLKDI NWDSQWQPLIQDRCFLSWL VKIPSEEEQRLARQITA
GCRNVFLVLFIPAKADSVVVLVLCRQPCASQSSLKDI NWDSQWQPLIQDRCFLSWL VKIPSEEEQRLARQITA

CH domain

300 310 320 330 340 350 360

SDISKLEAKWRSNKDATINDIDAEEQEAIPPLLLRYQDAYEYORSYGPLIKLEADYDQQLKESQALEHISVS
QCMVRLEELWRKDPNANLELDKPIEDDSLPSVRLRYKDAHAYQAVLSPLIQAEADYDKRLKESOTQKDVVVR
QQINKIEELWKNTPDATLEDLEKPGVDEEPQVQPKYEDAYQYQNVFAPLIKLEADYDKMMKESQSKENLTVR
AQINKLEELWKENIEATFQDLEKPGIDSEPAHLLRYEDGYQYKTFGLPLVLEAEYDQKLKESATQENIEVR
TQAVRMEELVRDHPATVDDLNKPCLDREPDHQLRYVDAAHYSKVFRPLVATEAEYDRRVKESASAVGTVR
QQINKLEELWKENPSATLEDLEKPGVDEEPQVHLLRYEDAYQYQNI FGPLVKLEADYDKKLKESQTDNITVR
QQINKLEELWKENPSATLEDLEKPGVDEEPQVHLLRYEDAYQYQNI FGPLVKLEADYDKKLKESQTDNITVR
QQINKLEELWKENPSATLEDLEKPGVDEEPQVHLLRYEDAYQYQNI FGPLVKLEADYDKKLKESQTDNITVR

370 380 390 400 410 420 430

WSLALNNRHLASFTLSTFESN-----ELKVAIGDEMILWYSGMQ-HEDWEGRGYIVRLPNSFQDTFTL
WQAINKRYTAWFLPKLESG-----ELRLAIGDEMILTYEGEL-RAPWSSTGYVIKIPNVSEVGL
WDIGLNKKRVAYFVFPKEEN-----ELRLVPGDELRLRYSGDAVHSGWSQSVGHVILKTA--QEEVAL
WDVGLNKKTIAYFTLAKTDS-----DMKLMHGDELRLHYVGEI-YNPSEIGHVIVKVPDNFGDDVGL
WEOGLRQSVLAFFHLQFADG-----VMKLAKGDELRLKHSQTVDSGSEWTKIGSVFKIPDNHGEVGI
WDIGLNKKRIAYFTLPKTDS-----DMRLMQGDEICLRVKGDL-AFLWKGIGHVIVKVPDNYGDEIAI
WDIGLNKKRIAYFTLPKTDSGNEDLVIWLRDMRLMQGDEICLRVKGDL-AFLWKGIGHVIVKVPDNYGDEIAI
WDIGLNKKRIAYFTLPKTDSGNEDLVIWLRDMRLMQGDEICLRVKGDL-AFLWKGIGHVIVKVPDNYGDEIAI

440 450 460 470 480 490 500 510

ELKPSKTPPTHTLTGTAEFIWKGTSDYRMDALKKFAIDKKSISGLYLYILGHQV--DISFDVPLPK
ELKRSKDP-VPIECTHNSVDYVVKSTSFDRMOTALRLFATDGSRLSSFLYHKLGHGDI--PSFLPKPLPS
ELRANQG-VPIDVNHGFSVDFVVKSTSFDRMQGAMKNFAVDETSVSGYTYHOLLGHEVE--AQMVNRNTPR
ELKSSSTN-APVKCTSNFTVDVFIWKGTSDYRMDALCKFAIDRNSVSNFIYSRLLGHRADSNDEVLFRGPQK
ELRGAVDKSVMSRIMFTVDVFNATTFERQYKALAAALNDSKATISPLYQLKLLGHPAE--EMMLKFDLPR
ELRSSVG-APVEVTHNQVDFVVKSTSFDRMQSALKTFVAVDETSVSGYIYHKLGHGHEVE--DVIIKQCLPK
ELRSSVG-APVEVTHNQVDFVVKSTSFDRMQSALKTFVAVDETSVSGYIYHKLGHGHEVE--DVIIKQCLPK
ELRSSVG-APVEVTHNQVDFVVKSTSFDRMQSALKTFVAVDETSVSGYIYHKLGHGHEVE--DVIIKQCLPK

520 530 540 550 560 570 580

EFISIPNFAQLNSQSNNAVSHVLRPLSLIQGPPGTGKTVTSATIVYHLISKI-----HKDRILVCAPSNAVAVD
DLSSVPLPKLNASQSEAVRAVLSKPLSLIQGPPGTGKTVTSASVYHLATMQSRKRKSHSPVLVCAPSNAVAVD
RFGVPLPELNASQVNAVSVLQKPIISLIQGPPGTGKTVTSAAIVYHMAKQ-----GQGVLCAPSNAVAVD
LFSAPHLPOLNRSOVYAVKHALORPLSLIQGPPGTGKTVTSATIVYQLVKL-----HGTVLCAPSNAVAVD
RLSVAGLPOLNSSQMAVKKVLTPLSLIQGPPGTGKTVTSATIVYHLVQK-----TEENVLCAPSNAVAVD
RFTAQGLPOLNHSQVYAVKTVLQRLSLIQGPPGTGKTVTSATIVYHLARQ-----GNCPVLVCAPSNAVAVD
RFTAQGLPOLNHSQVYAVKTVLQRLSLIQGPPGTGKTVTSATIVYHLARQ-----GNCPVLVCAPSNAVAVD
RFTAQGLPOLNHSQVYAVKTVLQRLSLIQGPPGTGKTVTSATIVYHLARQ-----GNCPVLVCAPSNAVAVD

ATP-binding

590 600 610 620 630 640 650

HLAAKLRDLGLKVVRLAKSREDVESVSNLALHNLVGRG--AKGELKNLKLKDEVGELSSADTKRFVKLV
QLAEKIHRTGLRVVRAAKSREDVSSVSFLSLHEQIKNYKF--NPQLQRLKLKRSENNELSIQDEKKLRILV
QLAEKISATGLKVVRLCAKSREAVSPVEYLLHYQVIRHLDTSEKSELHKLQQLKDEGELSSADEKKYKRLK
QLTEKIHRTGLKVVRLCAKSREADSPVSFLALHNQIRNMET--NSELKKLQQLKDEGELSSADEKKYKRLK
HLAEKIHRTGLKVVRLCAKSREHSETTVPYLLQHLKVMG--GAELQKLILQKDEAGELFKDQLRYMQLK
QLTEKIHRTGLKVVRLCAKSREADSPVSFLALHNQIRNMES--MPLEQLKQQLKDEGELSSADEKKYKRLK
QLTEKIHRTGLKVVRLCAKSREADSPVSFLALHNQIRNMDS--MPLEQLKQQLKDEGELSSADEKKYKRLK
QLTEKIHRTGLKVVRLCAKSREADSPVSFLALHNQIRNMDS--MPLEQLKQQLKDEGELSSADEKKYKRLK

660 670 680 690 700 710 720 730

Sc_UPF1 R K T E A E I L N K A D V V C C T C V G A G D K R L D - T K F R T V L I D E S T O A S E P E C L I P I V K G A K Q V L V G D H Q Q L G P V I L E
Sp_UPF1 A A E K E L L R A A H V I C C T C V G A G D R L S K Y K F R S V L I D E A T O A S E P E C M I P L V L G A K Q V L V G D H Q Q L G P V V M N
At_UPF1 R A T E R E I T Q S A D V I C C T C V G A A D L R L S N F R F R O V L I D E S T O A T E P E C L I P L V L G V K Q V L V G D H C Q L G P V I M C
Dm_UPF1 R A A E N Q L L E A A D V I C C T C V G A G D G R L S R V K F T S I L I D E S M O S T E P E C M V P V V L G A K Q L I L V G D H C Q L G P V V M C
Ce_UPF1 R V K E H E L L A A A D V I C C T C S S A A D A R L S K I R T R T V L I D E S T O A T E P E T L V S I M R C V R Q L V L V G D H C Q L G P V V I C
Xl_UPF1 R T A E R E L L M N A D V I C C T C V G A G D P R L A K M Q F R S I L I D E S T O A T E P E C M V P V V L G A K Q L I L V G D H C Q L G P V V M C
Mm_UPF1 R T A E R E L L M N A D V I C C T C V G A G D P R L A K M Q F R S I L I D E S T O A T E P E C M V P V V L G A K Q L I L V G D H C Q L G P V V M C
Hs_UPF1 R T A E R E L L M N A D V I C C T C V G A G D P R L A K M Q F R S I L I D E S T O A T E P E C M V P V V L G A K Q L I L V G D H C Q L G P V V M C

30 740 750 760 770 780 790 800

R K A A D A G L K Q S L F E R L I S L G H V P I R L E V Q Y R M N P Y L S E F P S N M F Y E G S L O N G V T I E Q R T V P N S K F P W P I R G I P
K K V A L A S L S Q S L F E R L I L G N S P F R L V V Q Y R M H P C L S E F P S N T F Y E G T L O N G V T T S E R I A R H V D F P M I Q P D S P
K K A A R A G L A Q S L F E R L V T L G I K P I R L Q V Q Y R M H P A L S E F P S N S F Y E G T L O N G V T I E R Q T T G I D F P W P V P N R P
K K A A R A G L S Q S L F E R L V V L G I R P F R L E V Q Y R M H P E L S Q F P S N F F Y E G S L O N G V C A E D R R - L K L D F P W P O P E R P
K K A A I A G L S Q S L F E R L V L G I R P F R L Q V Q Y R M H P V L S E F P S N V F Y D G S L O N G V T E N D R H M T G V D W H W P K P N K P
K K A A K A G L S Q S L F E R L V V L G I R P I R L Q V Q Y R M H P A L S A F P S N I F Y E G S L O N G V T A A D R V K K G F D F Q W P O P D K P
K K A A K A G L S Q S L F E R L V V L G I R P I R L Q V Q Y R M H P A L S A F P S N I F Y E G S L O N G V T A A D R V K K G F D F Q W P O P D K P
K K A A K A G L S Q S L F E R L V V L G I R P I R L Q V Q Y R M H P A L S A F P S N I F Y E G S L O N G V T A A D R V K K G F D F Q W P O P D K P

Helicase domain

810 820 830 840 850 860 870

M M W A N Y G R E E I S A N G T S F L N R I E A M N C E R I T K L F R D E V K P E Q I G V I T P Y E G Q R A Y I L Q Y M O M N G S L D K D L Y
L M F Y A N F G Q E E L S A S G T S F L N R T E A S T C E K L V T F L R S N V L P E O I G I V T P Y D G O R S Y I V Q Y M O N N G S M Q K D L Y
M F F Y V Q L G Q E E I S A S G T S Y L N R T E A A N V E K L V T A F L K S G V V P S Q I G V I T P Y E G Q R A Y I V N Y M A R N G S L R Q Q L Y
M F F L V T Q G Q E E I A S G T S F L N R T E A A N V E K I T T R F L K A G I K P E Q I G I I T P Y E G Q R A Y I V Q Y M O Y Q G S L H S R L Y
A F F W H C S G S E E L S A S G T S F L N R T E A A N V E K L V S K L I K A G V Q P H Q I G V I T S Y E G O R S F I V N Y M T Q G L N S K L Y
M F F Y V T Q G Q E E I A S S G T S Y L N R T E A A N V E K I T T K L K A G A K P D Q I G I I T P Y E G O R S Y L V Q Y M O F S G S L H T K L Y
M F F Y V T Q G Q E E I A S S G T S Y L N R T E A A N V E K I T T K L K A G A K P D Q I G I I T P Y E G O R S Y L V Q Y M O F S G S L H T K L Y
M F F Y V T Q G Q E E I A S S G T S Y L N R T E A A N V E K I T T K L K A G A K P D Q I G I I T P Y E G O R S Y L V Q Y M O F S G S L H T K L Y

880 890 900 910 920 930 940

I K V E V A S V D A F Q G R E K D Y I L S C V R A N E Q Q A I G F L R D P R R L N V G L T R A K Y G L V I L G N P R S L A R N T L W N H L L I H
K A V E V A S V D A F Q G R E K D F I L S C V R S S E H Q G I G F V N D P R R L N V A L T R A K Y G V I V L G N P K V L A K H A L W Y H F V L H
K E I E V A S V D S F Q G R E K D Y I L S C V R S N E H Q G I G F L N D P R R L N V A L T R A R Y G I V I L G N P K V L S K O P L W N G L L T H
Q E I E I A S V D A F Q G R E K D I I M S C V R S N E R Q G I G F L N D P R R L N V A L T R A K F G I I V G N P K V L A K Q Q L W N H L L N F
E N V E I A S V D A F Q G R E K D Y I V T C V R S N D I L G I G F L S D P R R L N V A I T R A K Y G L V L V G N A K V L A R H D L W H E L I N H
Q E V E I A S V D A F Q G R E K D F I L S C V R A N E H Q G I G F L N D P R R L N V A L T R A R Y G V I V G N P K A L S K O P L W N H L L N Y
Q E V E I A S V D A F Q G R E K D F I L S C V R A N E H Q G I G F L N D P R R L N V A L T R A R Y G V I V G N P K A L S K O P L W N H L L N Y
Q E V E I A S V D A F Q G R E K D F I L S C V R A N E H Q G I G F L N D P R R L N V A L T R A R Y G V I V G N P K A L S K O P L W N H L L N Y

950 960 970 980 990 1000 1010 1020

F R E K G C L V E C T L D N L Q L C T V Q L V R P Q - P R K T E R P M N A Q F N V E S E M G D F P K F Q D F D A Q S M V S F S G Q I G D F G N A F
C K E K G Y L V E G T L N S L Q K F S L T L T P Q K P Q K F K R D L N V Q R S L - - S - - - - P - I - - Q N A G S - A M L P S F S N L - - - P N
Y K E H E C L V E G P L N N L K O S M V O F Q - - - K P R K I Y N D R R L - - - - - F - - Y G G G A - G M I G N D N F G S G N P N
Y K D R K V L V E G S L N N L K E S L I H F Q - - - K P K K L V N S M N I G A H F M S T - - - - I - I - - A D A K E - V M V P G S I Y D - - - -
Y K S K E M L Y E G P I N A L K P L N L A L P K A T - - - - I R T K N I A - - - - - G N A N R F G - - - -
Y K Q K V L V E G P L N N L R E S L M Q F S - - - K P R K L V N T I N P G A R F M T T - - - - A - M - - Y D A R E - A I I P G S V Y D - - - -
Y K E Q K A L V E G P L N N L R E S L M Q F S - - - K P R K L V N T V N P G A R F M T T - - - - A - M - - Y D A R E - A I I P G S V Y D - - - -
Y K E Q K V L V E G P L N N L R E S L M Q F S - - - K P R K L V N T I N P G A R F M T T - - - - A - M - - Y D A R E - A I I P G S V Y D - - - -

1030 1040 1050 1060 1070 1080 1090

V D N T E L S S Y I N N E Y W N F E N F K S A F S Q K Q N R N E I D D R - - - - - N L Y Q E E A S H L - - - - - N S
L - - - - Y S S Y L E E W N V - - F A Q Y K R R E S N A T D F E D F - R S Q V G D D E S K F D E P T R F - - - - -
A D R - - - - R G S R G R A G - - - - G S Y L P S G - - - - - P P N G A R P G L H P A G Y P - - - - -
- - - - - R S G G Y G Q G - - - - - R Q M V G Q S M N G G Q Y G G S G G G P Y G N S P L G Y G T P S S N S M V G F L G N G G N G A A G G N N
- - - - - I K R M Q Y T F N E Y K S N D P - - - - - S Q P R L P - P T Y S N S Q N L L - - - - -
- - - - - R S S Q C R P S - - - - - N M Y F - - - - -
- - - - - R S S Q C R P S - - - - - N M Y F - - - - -
- - - - - R S S Q C R P S - - - - - S M Y F - - - - -

1100 1110 1120 1130 1140 1150 1160

N F A - - - - - R E L Q R E E Q K H E L S K D F S - - - - - N L G I - - - - -
- - - - - I P R V P L S P F P G G P S S Q P Y A I P T R G P V G A V - - - - -
N F G G A G P S W A A A H L H D S I G Y I S N E H G A A A L G N M P V P V G M F M N M S N I P P - R F Y N Q - H Q Q A I M A V K Q N R A I Q Q Q
- - - - - S M S - - - - - K L A Q T F N K V P I P A H M M D P - - - - - N V Y A A A R N Q K D R R - - - - -
- - - - - Q T H D Q I G M I S G S P S H V T A M I P I P N L V - - M P P M P P P G Y F G Q - A N G P A - - - - -
- - - - - Q T H D Q I S M I S A G P S H V A A M I P I P N L V - - M P P M P P P G Y F G Q - A N G P A - - - - -
- - - - - Q T H D Q I G M I S A G P S H V A A M I P I P N L V - - M P P M P P P G Y F G Q - A N G P A - - - - -

1170 1180 1190 1200 1210 1220 1230 124

- - - - - P H A P Q P G N H G F A G R G T S V G G H L P H Q A T Q H N V G T I G P S - L N F P L - - - - D S P N S Q P S P - G G P L S Q P G Y
T G N F S P G N S G P G V T V G V G R S A T P G G N K K T N K L G K S R V - - T G G G T G G A P L T G G S S V C N A A P Y S - Q H P M P L - S -
R G D - - - - - Q R R P P P Q - - - - - A E A A M D L S G G M S S Q Q S Q Y P P Q G A S S Q S Q Y
- - - - - P G R - - - - - M P K G K - S T R G G R P K Q R G I G L - - - - - Q G M S Q G N M P N S C A S Q D V V S O P F S - Q C P L T Q - G Y
- - - - - A G R - - - - - G T P K T K - T G R G G R Q K N - R F G L - - - - - P G S Q T T L P N S C A S Q D V A S O P F S - Q C A L T Q - G Y
- - - - - A G R - - - - - G T P K G K - T G R G G R Q K N - R F G L - - - - - P G S Q T N L P N S C A S Q D V A S O P F S - Q C A L T Q - G Y

1250 1260 1270 1280 1290 1300 1310

G S Q A - - - - - F R D G F S M - - - - - G G I S Q D F L A D I I K S Q G S H D P Y N M A D F A T Q A S P G G F A V D Y A T Q G A H
L Q M T - - - - - Q P S G F A L - - - - - S Q C P E L S Q D F - - - - - G Q - - - - -
L L D G A S S L G W S Q S Q T T T T T R H H H H R Q N R N S Q Q M S Q D M D D I Q - - - - - Q K M D D - - - - -
I S M S - - - - - Q P S Q M S Q - - - - - P G L S Q C P E L S Q D S - - - - - Y L G D E F - - - - -
V S M S - - - - - Q P S Q M S Q - - - - - P G L S Q C P E L S Q D S - - - - - Y L G D E F - - - - -
I S M S - - - - - Q P S Q M S Q - - - - - P G L S Q C P E L S Q D S - - - - - Y L G D E F - - - - -

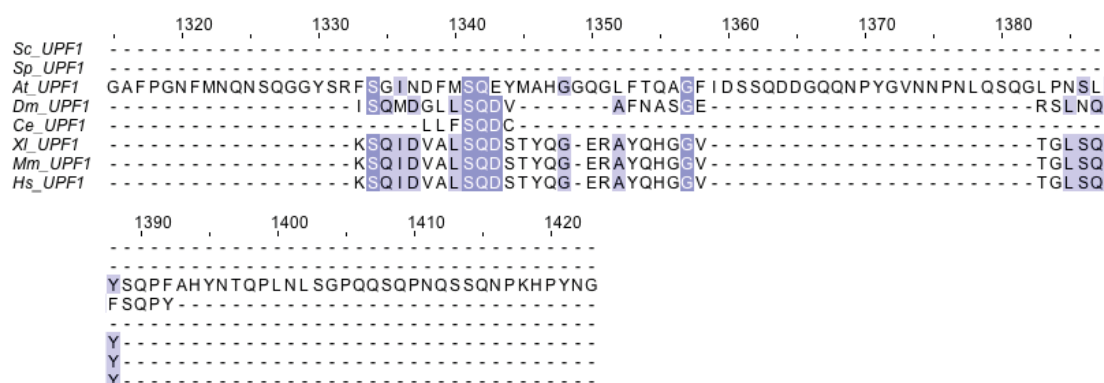


Figure 1.4 – **UPF1 sequence alignment.** Protein sequence alignment of UPF1 originating from *S. cerevisiae* (*Sc*), *S. pombe* (*Sp*), *Arabidopsis thaliana* (*At*), *D. melanogaster* (*Dm*), *C. elegans* (*Ce*), *Xenopus laevis* (*Xl*), *Mus musculus* and *Homo sapiens* (*Hs*). The intensity of the blue colour indicates how much identity between the amino acid residues do the different organisms share, with stronger blue indicating higher correlation between the organisms. The sequence alignment was generated using an integrated Clustal omega plug-in of Jalview software, version 2.10.5. Protein sequences corresponding to the key UPF1 domains are indicated below the alignment.

Sc_UPF2
Sp_UPF2
At_UPF2
Dm_UPF2
Ce_UPF2
Xl_UPF2
Mm_UPF2
Hs_UPF2

10 20 30 40 50 60 70

-----MLA-----NDSATTDSSEVSTPPSSRKDLADG-----

MPAERKKPAGMEEKESPSSPRDKECTEKKATSSKDRVKDEPKACTKRDQSKASEEKKQTDEDKQRKEDKECK
MPAERKKKSASMEEKESLLNNKEKDCSERRPVSSKEKPRDDLKVTAKKEVSKVPEDKKKRL EEDKRRKKEDKERK
MPAERKKPASMEEKDSLNNKEKDCSERRTVSSKERPKDDIKLTAKKEVSKAPEDKKKRL EEDKRRKKEDKERK

80 90 100 110 120 130 140

-----MDDGRKKELH-----DNLNT
-----MSRFEQIKKLNQYLD-----NRELAF
-----VAEGDNDNDIDTGDTDANAD-----EDVDAAAALAEAEERELQQFISLREKIESKRQL-----
-----MRRKVETSWLELKKRVKYQETSRL
TK EEDQVKADDEQMNTDG-----KNQDDEMCKKQEEEEANQREAAALQIKEKDEANQLHQAWERHHTRKELRS
KKEEEKVKAEFEELKKKEEEEKKKQEEEEKKQEEQAKRQOE-EAAAQLKEKEESLQLHQAWERHQLRKELRS
KKDEEEKVKAEFEESKKKEEEEKKKQEEEEKKQEEQAKRQOEAAAQMKKEESLQLHQAWERHHLRKELRS

150 160 170 180 190 200 210

RAWI--GEEVFPLKSKKLDSSIKRNTGFIKKLKGKGVKGSSESLKLDLSEASEIKYLSIIVTVTECLNVL
RAKD--GDKNIFHTESQLDSSLKKNTAFMKRCSSLTSENYSFIFEIKTSLKKFIPEITAAIVEGMMCKCK
RQN--LNPERDSAYLRTLDSSIKRNTAVIKKLKQ-INEEOREGLMDLRLGVNLSKFVSEAVTAICEAKLKSS
RLQN-STIELGGEYFAKLDNLKKNTAFVKKLKL-FTATOLDGLLREISALNLSKYISLCAALVEAKLKMT
QMVEIRETLLDNKEKLSLDSTLKKTTTFMKKVKL-LSAATVPOLIEELSKNLNLSKFVEMASGIIETKLKLS
KNQN-APENRDEESFFSRLDSSLKKNTAFVKKLKT-ITEQQRDSLSDFNLSLSKYIAEAVASIVEAKLKLS
KNQN-APDNRDEENFFSRLDSSLKKNTAFVKKLKT-ITEQQRDSLSDFNGLNLSKYIAEAVASIVEAKLKLS
KNQN-APDSRDEENFFSRLDSSLKKNTAFVKKLKT-ITEQQRDSLSDFNGLNLSKYIAEAVASIVEAKLKLS

220 230 240 250 260 270 280 290

KNKDDVIAVEIISGLHGRFNGRFTSPILGAFLO-----AFENPSVDIESERDELORIT
AT-KDILSSKIVWALNLR-----FSTAFTGPMLANLYCALYPNPGYSLCHESYFELKQNEVESEKDRSSHL
DI-QA--AVQICSLHGRYK-EFSASLTQGLL-----VFPFGK-SAEDEADKNSKAM
DV-PA--VVTLASRLCTYA-DFDVHFL EAWQK-----ALNIKA-----TEKIGNPS
DI-PK--VTCLLAISAKYS-NFSEQMAGEFK-----VLPVKK-----SDKIANVA
DV-NC--AVHLCSLFHGRYA-DFASQLLQVWKK-----HFEARK-----EEKTPNIT
DV-NC--AHLCSLFHGRYS-DFAPSLQVWKK-----HFEARK-----EEKTPNIT
DV-NC--AVHLCSLFHGRYA-DFAPSLQVWKK-----HFEARK-----EEKTPNIT

mIF4G domain

300 310 320 330 340 350 360

RVKGNLRVFTELYLVGVHRTLDDIESKDAIPNF-----LQKKTGRKDPPLLSIL
KVRPLRLFLIEFWLNGVGTPEDFVSYPSTDSNDKKFRKPWFEEQNKKPLVLLFNDLMDTRFGFLLLPVL
KKRSTLKLLELYYVGVIEDSNIFINIKDLTS-----V-EQLKDRDTHQNLTL
KL RVDLRLFAELVSSGVLMKPLAQGLGVVLVH-----LAL-DKDDHSNFSII
KL RVDITFLAELCLCGVNEKEGLQVLGAVLSY-----LQT-DKTDVFNVGLL
KLRTDLRFIAELTIIIGITDKEGLSLNYEQLKN-----I-NA-DRETHTHSVSV
KLRTDLRFIAELTIVGIFTDKEGLSLIYEQLKS-----I-NA-DRESHTHSVSV
KLRTDLRFIAELTIVGIFTDKEGLSLIYEQLKN-----I-NA-DRESHTHSVSV

370 380 390 400 410 420 430

REILNYKFKLGFTTTIATAFIKK-----FAPLFRD-DDNSWDDL IYDSKLKGALQSIFKNFIDATFARA
TSLV--RT--FSCELFT-----TEDFEDKETELVNRLNPVWRTYLKKSINSYVDKLEVYC
LSFA--RQ--GRIFLGLPI SG-----QDEDFFKGLDV-----TADQKKSFKAFNTYYDAADLL
LSFC--RH--CGEEYAGLVPRKMQQL--ATKYGVGVPKSDFL-----TADQKLNLRMTLKGFKALCKHV
ATVS--RT--VGWQIANI VPI PMDDAETVEIQDEELPESSAL-----SVEQKKSIRELFSKYDLSYKST
ISFC--RH--CGDDIAGLVPRKVKCA--AEKYNLNFPPSEII-----NPEKQPPFQNLKEYFTSLTKHL
ISFC--RH--CGDDIAGLVPRKVKSA--AEKFNLSFPPSEII-----SPEKQPPFQNLKEYFTSLTKHL
ISFC--RH--CGDDIAGLVPRKVKSA--AEKFNLSFPPSEII-----SPEKQPPFQNLKEYFTSLTKHL

440 450 460 470 480 490 500 510

TELHKVNVKLQREHQKQCI RTGKLRDEYVEEYDKLLPIFIRFKTSAITLGEFFKLEIPEL EGASNDOLKETAS
QKRKSLFEELNKQYQEQSIRADPNNEKFQRLANFSKSI ESEFSSYASLSEVNRKASEDLLELNFMEKASSG
QSEFKLLLMQEKENAKLVNAKGELSEDSASSYELKRKSYDHL YRNISSLAEALDMQPPVMPEDGTTTLT---
LAEQAEALMMMTKNIRRTMECKGETSTEKREKCLMQAGFDKLASAQSLSELIGVEMKPEKEASEG---C--N
EKTCSARNKAMKRVKQERSRQGDAGDEEKTTLNDQSELDL RKMAIEISSAIGVEMKPEKEASEDD--E--E
KRDHRELQNIERQNRRI LHSKGELSEDRHKQYEEFATSYQKLASCQSLADLDENMPDLPODKPAP--EH--G
KRDHRELQNTERRQNRRI LHSKGELSEDRHKQYEEFAMSYQKLANSQSLADLDENMPDLPODKPAP--EH--G
KRDHRELQNTERRQNRRI LHSKGELSEDRHKQYEEFAMSYQKLANSQSLADLDENMPDLPODKPAP--EH--G

520 530 540 550 560 570 580

P-----MITN--QILPPNQRLLWENEDTRKFYEILPDI SKTVEESQSSKTEKDSN-----
TNSVFNASGERSANVETAQVWDDREQYFFYEVFPNFNESGIAEMKSSI-----YESSQ--
ADEASPSGTV--KDTSVPEPIWDDDEDTKTFYECPLDRAFPVAVLLGEAPKSNEQSAKAKEKLSSESESVV
PGTVIDNMLDS--ASFG-VLPDWPGEETRAFYTDPLDRLQFLNFSAPKVDLLETLEPSE---LTEEAI D--
A--ANLEMGRK--LA-EGAIKLSWDEETKAFYEDLIDL RQMVPKDL YKESQRTLSKAKMAER---IE--
PGIDIFTGPKP--GDYDLEGGIWEDEDARNFYENIDLKAFVPAILFKDNEKSSQSKETGKD--NVKDTK--
PGIDIFTGPKP--GEYDLEGGIWEDEDARNFYENIDLKAFVPAILFKDNEKSSQSKETGKD--DSKEAKEPK
PGIDIFTGPKP--GEYDLEGGIWEDEDARNFYENIDLKAFVPAILFKDNEKSCQNKESNKD--DTKEAKESK

590 600 610 620 630 640 650

-----VN-----EGIRS--SENN-----KKEDDLK-----DS-----T
ENQ-QTTEDTTVSADS--ASMDR--SNAEOPKE--KFEVEKEKA-KDTKKEKGEKDKSEKKMEHE
-----ANLDAEMDLDPP--STTSOTTENPIEEQPTTPVAAAEDLKPKQKM-GNALMELGRQQQQSQLNQN
-----DVDMENINESGAMDAKR--SSLQ--RGES-EKETTPEDSQLQALLKE-----A
-EKDVPTTEELLELETL D--IND--DGIEM--EGAEDAEDLT-----KKLLDE--Q-----EQEDE
DNKEASSPDDLLELENLE--IND--DTLEL--EGAEDAEDLT-----KKLLDE--Q-----EQEDE
ENKEVSSPDDLLELENLE--IND--DTLEL--EGGDEAEDLT-----KKLLDE--Q-----EQEDE

660 670 680 690 700 710 720 730

Sc_UPF2 -----SKNINLFFTDLEMAADCKIIDL SNRYWSSYLDNKATNRNRLKFFME--TQDWSKLPVYSRFIAT
 Sp_UPF2 GDLNTTQVSSRVNDFLLKPSMVSLEITNEMA-LEFY-DLNTKASNRRIKALCTIPRTSSFVPIYVRLARI
 KEKGSGLDVANFERLQRLPGCVSRDLIDQLT-VEYCY-LNSKTNRRKKLVKALFNVPRTSLLELLAYSSRMVAT
 At_UPF2 PIQIQNQMRQOQDGLFVNLFNCVNKELIDSA-IEFLLNFNTKHQRKKLTRTITFSQRTRLDILPYLSRFVAI
 Dm_UPF2 VDASEERGVSQWQSFVLDLDHLVSKYSTDQAA-IYFVSNLNNKGCRRKLVKLMIDPPPSRIDVVPFYARLVAT
 Ce_UPF2 EASTGSHLKLIVDAFLQQLPNCVNROLDKAA-MDFCMNMNTKANRRKLVRLFIIVPRORLDLLPFYARLVAT
 XI_UPF2 EASTGSHLKLIVDAFLQQLPNCVNROLDKAA-MDFCMNMNTKANRRKLVRLFIIVPRORLDLLPFYARLVAT
 Mm_UPF2 EASTGSHLKLIVDAFLQQLPNCVNROLDKAA-MDFCMNMNTKANRRKLVRLFIIVPRORLDLLPFYARLVAT
 Hs_UPF2 EASTGSHLKLIVDAFLQQLPNCVNROLDKAA-MDFCMNMNTKANRRKLVRLFIIVPRORLDLLPFYARLVAT

740 750 760 770 780 790 800

NSKYIPEIVSEFINYDNGFESQLHSNK-----INVKNIIFFSEMIKFLQIPSFMIHFHKIRTIMYMQ
 LSQLSSEFSTSVDHARHSKRMIRHMA-----KHEYDTRLILVRYSELTKPOLMPFHMVFQYKLCIN-EF
 LASCKIKIPSMVQMTEDEFNSLVHKKD-----QNIETKIRNIRFIELCKPKIVPAGLVFSCKACID-EF
 VHMNTDVAATDLAELRKEFKWHIHKKN-----QLNIETKIRNIRFIELCKPKIVPAGLVFSCKACID-EF
 THHNIDVACNLETCGRFLYSPETTLRMTNMDILMRLKNVKNLDPQSTLVENAYLYCKPPERSARIS-
 QHHQLEMACAFVEVTGVYLYNCRDARLLNNVFLDQMLRLKTATAMDSHAAQIESVYLYVKPPRESSREP-
 RGHVSVENTATFMESSGLYLYHTESHAKMKRLLVVKA-KRERMKDOQAMLDIDNAFTCLPPEDSKERLRL
 SHHILEMACTLLETTCGRFLRSSDSHLRTSVLLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-
 SHHILEMACTLLETTCGRFLRSSDSHLRTSVLLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-
 SHHILEMACTLLETTCGRFLRSSDSHLRTSVLLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-

810 820 830 840 850 860 870

VPNNVILTVEHSCKFLLNKPEYKELMEKMVQLIKDKKNDROINMNMKSALENITLLYPPSVKSLNV-
 TPFDLVLLALLESQGRFLRYPETKIQMQSFLTAIQKKKLSALASQDQLVLENALHFVNPKRGIIYS-
 THHNIDVACNLETCGRFLYSPETTLRMTNMDILMRLKNVKNLDPQSTLVENAYLYCKPPERSARIS-
 QHHQLEMACAFVEVTGVYLYNCRDARLLNNVFLDQMLRLKTATAMDSHAAQIESVYLYVKPPRESSREP-
 RGHVSVENTATFMESSGLYLYHTESHAKMKRLLVVKA-KRERMKDOQAMLDIDNAFTCLPPEDSKERLRL
 SHHILEMACTLLETTCGRFLRSSDSHLRTSVLLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-
 SHHILEMACTLLETTCGRFLRSSDSHLRTSVLLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-
 SHHILEMACTLLETTCGRFLRSSDSHLRTSVLLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-

880 890 900 910 920 930 940

-TVTITITPEQFYRIIRSESSLDKFHIVKLVKAHWDVAIQKVLFSLSKPKHISQNIPLLTKVIGGLY
 ---KKSLKEEFLYDLIQIRUKDDNVFPTLLLRKFDWKDD--YQILYNTIMEVWNIKYNSNALARLISALY
 ---KVRPPLHQVVRKLLFSQDKDSIANVLKOLRKPWSEC--EQYILKCFMKVHKGYGQTHLIASISGLS
 ---MVRPAMHEIYRILFEELCKQNVDRCLKMLRRIDWQDPEITNCYAKGLSKAYLLRFQLRCLADLVSGLS
 KLDEEDTPMKRFIRHIL-ILNESTVDVFKCLRRLEWYDPEVADYARYLSSTLLPIENLOHVASAIAGLC
 ---KRRPPOEYIRKLLYKOLSKVTTEKVLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-
 ---KRRPPOEYIRKLLYKOLSKVTTEKVLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-
 ---KRRPPOEYIRKLLYKOLSKVTTEKVLQMMRKKQAMHLDAYVTMVENAYYCNPPPAETVK-

950 960 970 980 990 1000 1010 1020

S--VRRDFVRCIDQVLENIERGLEINDYGOIMHRI SNVRYITEIFNEMIKSDVLLDTIYHIFRSHINNQP
 K--FHPEFCFHVDDTLESLSAVNNSDHVEKQKRLAQAIFSELCVIHMLDVRAITNFIHLLPLEKFESFL
 R--HHDEFVVAVDVELEETRVGLEMEYGAAQKRLAHMRFLGELYNYEHVDSVIFETLYLTLLYCHDTS--
 S--YQPRAVTIVIDNVEDIRAGLEIHSPRMAORRIMAKYLGEMNYKLVESTNILLIYSLISLGSND--
 NLVHLQWIGMAVIDSTIETIRISLENP-GVYVQWAHSAVYLAELYSFELCDEDLVLLIYQLISYPE----
 L--YQEDVGHHVVDGVLEDIRLGMEVNOPKFNORRISAKFLGELYNYRMVESAVIFRTLYSFTSFGVNPDP--
 L--YQEDVGHHVVDGVLEDIRLGMEVNOPKFNORRISAKFLGELYNYRMVESAVIFRTLYSFTSFGVNPDP--
 L--YQEDVGHHVVDGVLEDIRLGMEVNOPKFNORRISAKFLGELYNYRMVESAVIFRTLYSFTSFGVNPDP--

1030 1040 1050 1060 1070 1080 1090

NPFYLNYSDDPDNYFRIQLVTTILLNINRTP--AAFTKCKKLLLRFEYTYFIEKEQ-----
 TMKASTLTNINNDMFRLRLIVVVLQTCGPSIRSKTKTMTLYLLAYCYFLIQP-----
 ---EQEVLDPPEDFFRVVMVILLETGCHYFDGSSKKRLDQFLIHFORIILSKG-----
 -QNVVSPLDPPDSLFRLKLAAMLLDTQAPYTSQATRKLDYFLVFFCHYYWFKKSHFPVSKTEN-----
 ---PENSWRDLHRIRMIAMLEILREFMKGSGIKIMRYFLSYFHRYYYIKKD--AWDQVAEQQPNSSG
 -G-SPSPLDPPEHLFRRLVCTILDTCGQYFDGSSKKRLDCLVYFQRYVFWKKSLDVWTKD-----
 -G-SPSSLDPPEHLFRRLVCTILDTCGQYFDGSSKKRLDCLVYFQRYVFWKKSLDVWTKD-----
 -G-SPSSLDPPEHLFRRLVCTILDTCGQYFDGSSKKRLDCLVYFQRYVFWKKSLDVWTKD-----

UPF3-binding

1100 1110 1120 1130 1140 1150 1160

-----PLPKETEFVSSTFKKYENIFGNTKFERSENLVESAASRLSLLKSLNAIKSKDDRVRKSS
 -----EMPLDMLYEFEDVIGYVFPMSK-----VYMVHARNALTE---LQAIS
 -----HLPLDTEFDLQDLFANLRPNMT-----RYSTIDVNAAILQLLEEREPHASS
 -----TSDLFPIILMDHTYRCLLNVRPKLK-----IYKSLQAKAAIDHLQERLYPQ
 ETGENQNPLDVQTSFPEVELEYTELCRQFRQKSN-----SLRWPKNLKAAQDITVKIEKKFKGEL
 -----YFPFIDIDYMI SDTLELLRPKMK-----LCVSLNESVRQVODLEREFLTKL
 -----HPFPIIDIDYMI SDTLELLRPKIK-----LCNLSLESIHQVODLEREFLIKL
 -----HPFPIIDIDYMI SDTLELLRPKIK-----LCNLSLESIHQVODLEREFLIKL

1170 1180 1190 1200 1210 1220 1230 124

ASIHNGKE--SAVPIESITEDDEDEDDEDDGVLLDEDEDAEIST-----PNTESAPQKHQAKQDESEDED
 DDWEED--DTRPV-----FQANDGDISNEESVY--LP---
 GDKVSIERHSDTKPSNK-----SSSDVISNKKSTAKDIRENGEAHEESDSSGSGSVVRDGO---
 KTTNNA--Q-----DPSLGTIS-----EISEI-DEGGTDEDSGSSNDQRRQVS---
 KEFAGE-----HKDI-EGDGLN--ELN---
 GLMNDK--E-----SKHSMTEOLVYEDDDDEE-EEGGAETEEQSGNES---EMN---
 GLVNDK--E-----SKDSMTEENLE-EDEEE-EEGGAETEEQSGNES---EVN---
 GLVNDK--D-----SKDSMTEENLE-EDEEE-EEGGAETEEQSGNES---EVN---

1250 1260 1270 1280 1290 1300 1310

DEDDDDDDDDDDDDGEGDEDDDEDDDDDDDEEEDSDSDLEYEGGDLADRDIEMKRMYYEYERKLK
 ---EDISDESETDESSGLE---ESLLDSED---EIDNEM---
 ---NELLDGNH--FRGS-SGDGDYDDGDG-P--GSD---DDKFRVRQKV---
 ---GQDQSDNWTN---AEPPLPPPPP---SEEDROKIEENFS---
 ---VIEDDEDEDEN---RDSEDEDD---SEEDROKIEENFS---
 ---EQEED-ASENE-DREEEEENTYMTDSN-KENETDENNEVMIKGGGI---
 ---EP-EEEGGSEE-EEGEEEEEENTYLTDSN-KENETDENNEVMIKGGGI---
 ---EP-EEEGGSDNDDDEGEEEEEENTYLTDSN-KENETDENNEVMIKGGGI---

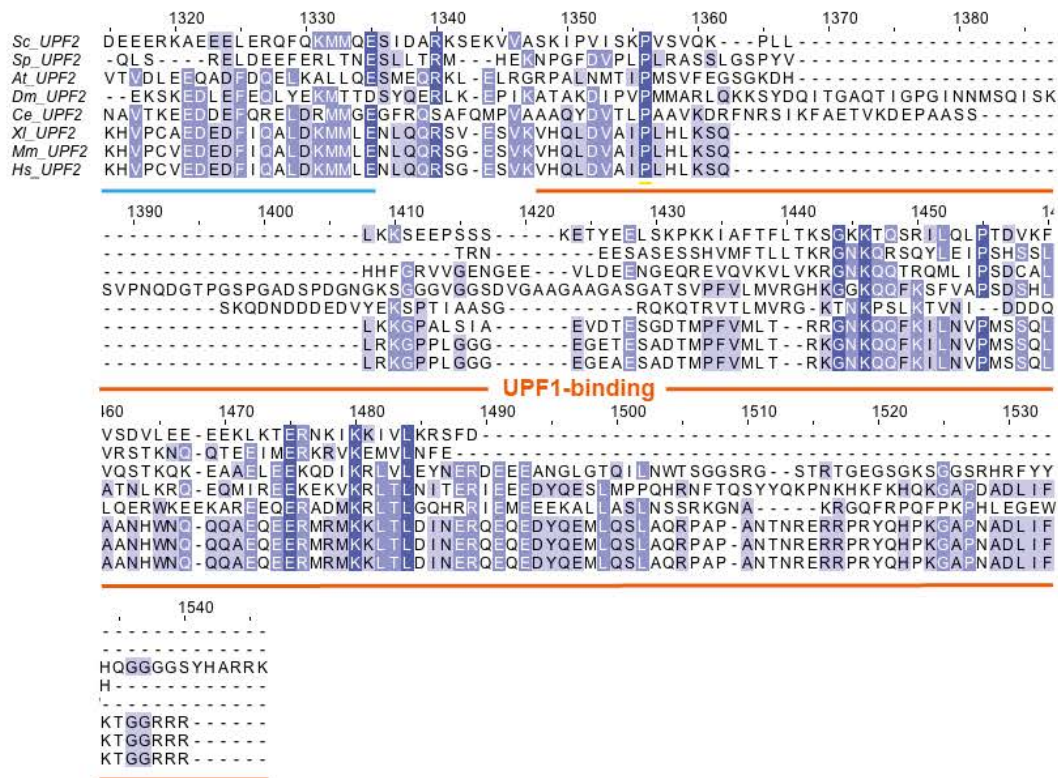


Figure 1.5 – **UPF2 sequence alignment.** Protein sequence alignment of UPF2 originating from *S. cerevisiae* (*Sc*), *S. pombe* (*Sp*), *Arabidopsis thaliana* (*At*), *D. melanogaster* (*Dm*), *C. elegans* (*Ce*), *Xenopus laevis* (*Xi*), *Mus musculus* and *Homo sapiens* (*Hs*). The intensity of the blue colour indicates how much identity between the amino acid residues do the different organisms share, with stronger blue indicating higher correlation between the organisms. The sequence alignment was generated using an integrated Clustal omega plug-in of Jalview software, version 2.10.5. Protein sequences corresponding to the key UPF2 domains are indicated below the alignment.

10 20 30 40 50 60 70
Sc_UPF3 MSNVAGELKNSEGKKKGRGNRYHNKNRGKSKNETVDPKKNNKVNATNATHNNSKG---RRNNKKRNREYYN
Sp_UPF3 -----
At_UPF3 -----
Dm_UPF3 -----MAETFA---EEKPKEKS-KAS-----
Ce_UPF3 -----
Xl_UPF3 -----MRNEREPAPAVLRANREN-RTFDKQCRELETENT-PSN-----
Mm_UPF3 -----MKEEKDHRPKKEKRVTLTTPQATGGCIGATAEGAKGDD-----
Hs_UPF3 -----MKEEKEHRPKKEKRVTLTTPAGATGSGGGTSGDSSKGED-----

80 90 100 110 120 130 140
 YKRKARLGKSTENEGFKLVIIRLLPPLNTADEFFAILRDNNNDGDGKQDIQGLKYSOWCFEFGHYSKVFKNK
 ---MAPDISKKRLPCVVLVFNLPPTLPEQVFLQSIN-SFLP-----HVEVHRFSK GKATVGRSE
 ---MKEPLQKKKVVVRHLPPSLSQSDLLSQIDPRFAD-----RYNWSFRPGKSSYK--NQ
 ---RRKDKDKTNQIVKIVMRHLPPMTIEAQFLDQVGPLPEN-----DSY-YYCKADWLSG--QE
 ---MTDSSKDGHVKVVLRRLPKYMTIEHEVLEQISPLPEE-----VIGTYFHPANFSFD--RC
 ---GLRHRERKATLSKVVIIRLLPPLNFKKEQLLEQLRPLPAH-----DYF-EFCTADPSLY--PH
 ---RQDRNRDKKEALS KVVIRRLPPLTKEQLQEHLPMPHEH-----DYF-EFSDNTSLY--PH
 ---KQDRNKEKKEALS KVVIRRLPPLTKEQLQEHLPMPHEH-----DYF-EFSDNTSLY--PH

UPF2-binding

150 160 170 180 190 200 210
 TYSRCNFLFDN--LSLEKCANFIKTKCFIDNKNITIPDMKLSFYVKKFTQT--SKKDAALVGTIEDEIFK
 LLSFAYLKFS--ATAVQEFRRVYQCHTIDKKNNITYRAIVTIAPYQKIPPSK---VKADSLGLEQDPKFK
 KYSRAYVSFKA--PEDVYFAAFNGHVFNNEKGAQFKAIVEYAPSQRVPKPS--DKKDPREGSTSKDPDYL
 ATCRAYIDMSKSDIGEVVQFRDRFDGYVFDHKKVEYMAIVEYAPFQCFLLNK--ARNDDSKVNTIESEPHYK
 AATLTVNFSEY-CDSMMEFERRFDGYIFVDSRDNDAAVVEAASNQNFACDRNRMKEDTRVGAILTDQKYYL
 LFSRAYINFRN--PEDILLFRDRFDGYIFIDNKKQEYPAVVEFAPFQKISKKK--LKKKDAKAGSIADLEYK
 MYARAYINFRN--QEDILLFRDRFDGYVFLDNKKQEYHAIVEFAPFQKAAKKK--IKKRTTKVGTIEDDPEYR
 MYARAYINFRN--QEDILLFRDRFDGYVFLDNKKQEYHAIVEFAPFQKAAKKK--TKKRTTKVGTIEDDPEYR

220 230 240 250 260 270 280 290
 TMNSMKQLNENDEYSFQDFSVLKSLEKEFSKSIENENKIAERTERVLTETVGTGDKVKN-----K-NKKK
 EKVKQRESYSQTAS-----NDD-----VIEKLQTSPTLLQYLAEKKNNAVVEKGSKSPSKSVK-
 EFLKVIQAPVENLPSA-E--IQLEERRAEQSG-----ASKAPIVTPLMEFIRKRAVVMGPOGLSDIRRGGR-
 EFKRLAQERAEASRMGD--VKI-DFN--FER---RTEKVKSTPLQYLANKKEK-----RREEAR
 DCKKLEEEAIPILTL--QQIRKLNQPDPA--RTQIDKMETPLVKYFFEKETGK-----
 KLENYCAEEKVYANPE--TLLEGEI--AKT---KELIARRTTPLLEFIKNRKLKLEK---QRIRREEKREERR
 KLESYATDNKMTSTPE--TLLEEIE--AKN---RELIAKRTTPLLSTLKNK-----QRMREEKREERR
 KLESYATDNKMTSTPE--TLLEEIE--AKN---RELIAKRTTPLLSTLKNK-----QRMREEKREERR

300 310 320 330 340 350 360
 KNKNAKKKFKEEASAKIPKKKRNRGKKRENREIST-----ISKTKNSNVV
 ---AKKKLALAEKPASNNS-KAGKS---SQESKKS-----SKAPAESAASAAVI
 ---RTVVVSANKPSRPS--KRNSEKKYVEKESKKNVPRKTTADVSSSKPDYRQSNSSGKELPGNETAAII
 RRNEEKRRKQEEQKL-----LRLAAQSDA-----SKL
 RRQDYDARRQRDEKRAEKR-----DKVDIFEKHKNVLD--ILMKPSVPMAS--T--SATTSKKDL
 RRELEKKRIREEKKRRREE-ER---RKRKEADQKKI SEKEIR-----IKLL
 RREIERKRRQEEERRKWKEE-EK---RKRKDIENLKKIERIPEREKIRDEPKIK-----LL
 RREIERKRRQEEERRKWKEE-EK---RKRKDIENLKKIDRIPERDKLKDEPKIKVHR--FLLQAVNQKNLL

370 380 390 400 410 420 430
 IIEEAGKEV-LKQRKKMLLQEKLLKISNSSQPSQSSAQ--TOPSFQPK-----NLFVPRV
 HEDK-----VSDRKSKKKPKKTPVSN--TASQASENASDKKT-----
 DSSPPGIALTMDSGKKILLLRSKDRDNPDNPPQPEQHI DNL SRNSTDSRQ-----
 KEAEGGGG-GDTKKPAK-----KDVKDQSSGSQ--ANDAKA-SRSKRRTERDQRRREEHEQRKLVKRD
 KKEKPMTE-----KEKERWEK-QDAKRKE-----RNMIRK--QKFLDEKKKKHEEREQDGPVPRD
 KPEKGDAASETQKEKIEEAED-EDTKLEKPPVSGSI--KVRPFENSVKELKEKNPNCEDKQRDKERRFD
 RKPEKGDKELDKRDKTK-RLDK---ENLNEERASGHS--YTLPRR-SDVELKDEKPKRLDDEGV-RDYRDRD
 KPEKGDKELDKREKAK-KLDK---ENLSDERASGQS--CTLPKR-SDSELKDEKPKRPEDESG-RDYRERE

440 450 460 470 480 490 500 510
 KILHRDDTKK-----SSGKKKIASKKKQDLT--TDNV-----
 KEKK-----DVGGLIKGI--LLR--NDSRP-----SQSSTFV-----Q
 NQKS-----SKQKNKNSKSMIDVILKKANKPESNESLDVPSTSNAAEASKEKPA
 KKDDGGQKQDKQDK--QDKGKPS--SKQKNKNSKSMIDVILKKANKPESNESLDVPSTSNAAEASKEKPA
 K-----KERLAKPPRPSPAKTVEGQGGEDWIKKLTDPKSAPM-KKKHDL SVTTKLNKE-----
 KDLERQFRMDD-----GRKHRNHYEY-----
 RQYERDQERMIRERERLKRQEEER--RRQQKERYEK-----
 REYERDQERILRERERLKRQEEER--RR--QKERYEK-----

520 530 540 550 560 570 580
 SEQRVEPSEAEENYKRP-----SRPANTR-----AGKDY-HTS-----GTISEKQERTRNKDRPDR
 TENTETPEAAVKVFTPAARGGRKEKRTGADKQ-----PAGSAHETAEEAIIKAAQFRASEERRIRNKDRPSI
 TENTETPEAAVKVFTPAARGGRKEKRTGADKQ-----PAGSAHETAEEAIIKAAQFRASEERRIRNKDRPSI
 -----DKFMRRNEDDQKWGKGYNQDRCKKGSIISSFASDAVDKLGKEDKCDLASKKDRIRNKDRPAM
 -----EKAFKRKEEEMKREKEALRDGKKSSENTESIC-----SLEKIE-KKEEVVKRDRIRNKDRPAM
 -----EKTFRKEEEMKKEKDTLRDKGKKAESTESIG-----SSEKTE-KKEEVVKRDRIRNKDRPAM

590 600 610 620 630 640 650
 VMWAPRRDGSSEDQ-PLS-SAGNNGEVKDRMFQSRSGE--VVNSSGGHTLENCSARHSSRRVGGRRNRKEEVVIG
 AIIYQPKARIRASDEL PQGAGGKDGSD-----GEASVVEEKT SKRNKRPNRRNPKP-----PKEVKEC
 TKLPTRAT--TSNHPKRPSTAPNH-----
 QLYQPGARNRITTGFTS--KYPEENF-----YRKNNTNSVTRRGTEKSEDAG-----
 QLYQPGARSRLCPADDSTKPGDSP-----VDKKQE-----SGISHRKEGEE-----
 QLYQPGARSRLCPADDSTKPGDSA-----AERKQE-----SGISHRKEGEE-----

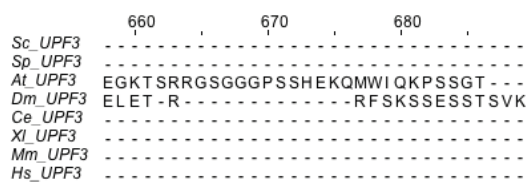


Figure 1.6 – **UPF3 sequence alignment.** Protein sequence alignment of UPF3 originating from *S. cerevisiae* (Sc), *S. pombe* (Sp), *Arabidopsis thaliana* (At), *D. melanogaster* (Dm), *C. elegans* (Ce), *Xenopus laevis* (XI), *Mus musculus* and *Homo sapiens* (Hs). The intensity of the blue colour indicates how much identity between the amino acid residues do the different organisms share, with stronger blue indicating higher correlation between the organisms. The sequence alignment was generated using an integrated Clustal omega plug-in of Jalview software, version 2.10.5. Protein sequences corresponding to the key UPF3 domains are indicated below the alignment.

1.7.2. Structure and interactions of NMD factors

Deletion or silencing of genes encoding the three proteins that comprise the core NMD machinery, UPF1, UPF2 and UPF3, leads to stabilisation of NMD substrates also in *S. pombe* (Wen and Brogna, 2008; Wen and Brogna, 2010). The three proteins were shown to interact by yeast two-hybrid assays and co-immunoprecipitation studies, with the interacting domains characterised (detailed below) (He and Jacobson, 1995; HeBrown and Jacobson, 1997). UPF2 was shown to mediate the interaction, as UPF1-UPF3 association is abolished upon UPF2 deletion while overexpression of the UPF2 regions identified to interact with UPF1 and UPF3 drastically stimulates UPF1-UPF2-UPF3 complex formation (HeBrown and Jacobson, 1997).

1.7.2.1. UPF1 structure and function

UPF1 is the most evolutionarily conserved protein involved in NMD, considered to be the crucial NMD regulator, being essential for NMD practically in all eukaryotes. UPF1 belongs to the RNA helicase superfamily 1 and it contains several protein domains: a cysteine- and histidine-rich (CH) domain at its N terminus, a helicase domain towards its C terminus and an ATPase domain. CH domain comprises of two zinc knuckle structures whilst the RNA helicase domain contains two core helicase segments (RecA1 and RecA2) and two regulatory segments (1B and 1C) (CelikKervestin and Jacobson, 2015; Cheng *et al.*, 2007) (Figure 1.7).

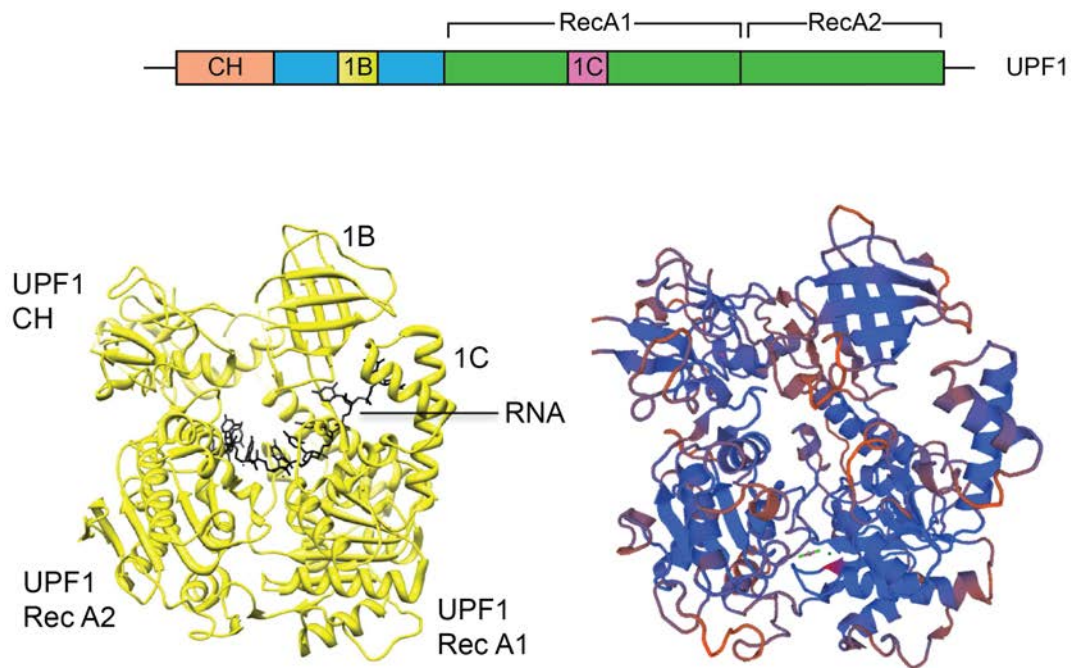


Figure 1.7 – **UPF1 domain organisation and structure.** (A) Schematic representation of UPF1 protein domains. (B) Crystal structure of human UPF1 in the closed conformation, with indicated domains, as well as the RNA binding region (Llorca, 2013). (C) A predicted *S. pombe* UPF1 structure based on its amino acid sequence, using ExPASy SWISS-MODEL, a protein structure-modelling server. The blue indicates regions whose structure has been modelled to a high confidence, while the red determines regions modelled with a low confidence. The two protein structures bear significant similarities.

The protein exhibits ATPase as well as helicase activity. UPF1 ATPase activity was demonstrated in an *in vitro* assay, by incubating the yeast UPF1, in the presence or absence of RNA or DNA, with radiolabelled [γ - ^{32}P]-ATP molecule and by measuring the quantity of the released phosphate. UPF1 was shown to exert nucleic acid-dependent ATPase activity, as the reaction was greatly stimulated only when nucleic acids were present in the reaction (Czaplinski *et al.*, 1995). The helicase activity of UPF1, on the other hand, was assessed via a strand displacement assay, in which a radiolabelled oligonucleotide was hybridised with a long synthetic DNA or RNA molecule, generating double stranded regions used to test UPF1 unwinding properties. UPF1 displayed 5'-3' helicase activity only in the presence of ATP, suggesting that this enzymatic activity is dependent on energy generated by ATP hydrolysis (Czaplinski *et al.*, 1995). UPF1 remodelling properties were again tested in a recent *in vitro* study that used magnetic tweezers, tethering an RNA molecule containing a hairpin between a glass coverslip and a paramagnetic bead (where the 3' end of the RNA was tethered) (Fiorini *et al.*, 2015). Unwinding of the RNA secondary structure was assessed by the extension of the molecule, demonstrated via bead movements. Using this assay, in the presence of supplied ATP, human UPF1 was shown to translocate over single stranded RNA stretches with a processivity of > 10 kb whilst unwinding the secondary, double-stranded structures as well as translocating across protein-bound sequences, indicating its role as a potent remodeller of ribonucleoprotein complexes (Fiorini *et al.*, 2015).

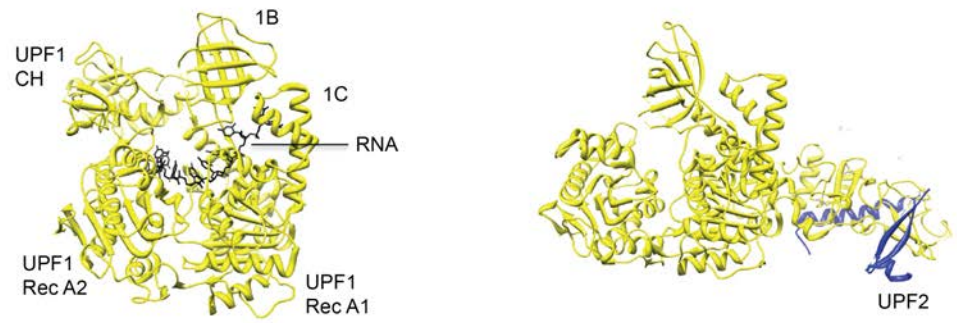
In vitro, UPF1 was shown to bind and remodel both RNA and DNA molecules, however *in vivo*, UPF1 binds solely mRNAs, best demonstrated via CLIP experiments, indicating that UPF1 functions specifically as an RNA helicase in cells (Zünd *et al.*, 2013). UPF1 primarily associates with 3' UTR of mRNAs, being mainly present in the cytoplasm where it is thought to remodel the 3'UTR and recruit the decay machinery during NMD (Zünd *et al.*, 2013). Indeed, the ATPase dependant mRNase activity of UPF1, which remodels the mRNP in a 5' to 3' direction seems to be crucial in the NMD pathway (FranksSingh and Lykke-Andersen, 2010). This activity triggers the dissociation of NMD factors from PTC containing mRNAs, which is obligatory for efficient and complete degradation of NMD substrates and recycling of NMD factors (FranksSingh and Lykke-Andersen, 2010).

UPF1 enzymatic activities are regulated via interactions between the UPF1 CH-domain and the C-terminal region of UPF2. UPF1 is in its closed conformation when its CH domain interacts with the helicase domain. After UPF2 binds to it, a conformational change occurs, CH domain distances from the helicase domain which brings UPF1 into its unwound, active conformation where its ATPase and helicase enzymatic activities are at an increased rate (Figure 1.8) (Chakrabarti *et al.*, 2011).

A



B



C

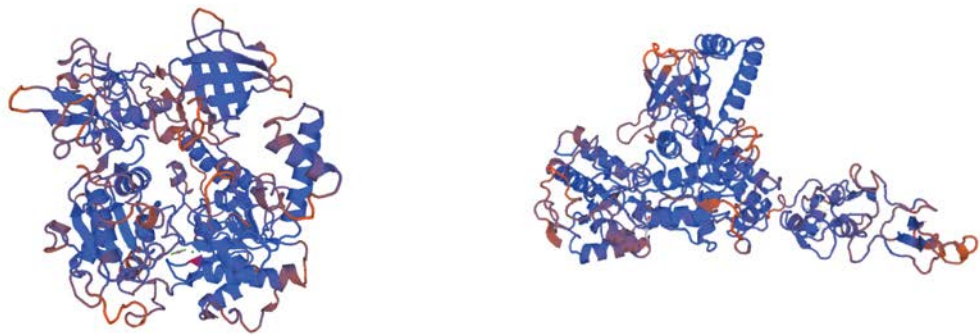


Figure 1.8 – **UPF1-UPF2 interaction.** (A) Crystal structure of the human UPF1-UPF2 complex showing a bipartite mode of interaction (Structure obtained from Protein Data Bank in Europe – <http://www.ebi.ac.uk/pdbe> – PDB code: 2wjv). UPF1 is presented in green and blue, while the domain of UPF2 it is interacting with is shown in purple and yellow. (B) Conformational change that ensues when UPF2 binds to UPF1 – crystal structure of the closed UPF1 conformation is shown on the left, while the active form is presented on the right (UPF2 is shown in blue). In closed conformation CH domain interacts with the helicase domain. UPF2 binding induces a conformational change – CH domain moves apart and UPF1 transits into its active conformation (Llorca, 2013). (C) I used ExPASy SWISS-MODEL server to obtain the predicted *S. pombe* structures of both the closed (left) and the active (right) UPF1 conformations. The blue indicates regions whose structure has been modelled to a high confidence, while the red determines regions modelled with a low confidence.

So far, no evidence has been found that the phosphorylation of NMD factors plays a role in NMD regulation in yeast. However, in metazoans, UPF1 activity is regulated by phosphorylation and dephosphorylation cycles catalysed by other NMD factors (Ohnishi *et al.*, 2003; Chiu *et al.*, 2003). In fact phosphorylation of UPF1 is thought to activate the protein and consequently initiate the NMD process. The sites of phosphorylation are the numerous serine/threonine-glutamine rich (SQ-rich) motifs in the C-terminal region of UPF1 (Mühlemann *et al.*, 2008).

UPF1 was shown to selectively co-purify with eukaryotic release factors 1 (eRF1) and 3 (eRF3) recruited to the terminating ribosome (Czapinski *et al.*, 1998). In human cells UPF1 is considered to interact with eRF1, eRF3 and SMG1, leading to formation of the Smg-1-Upf1-Release Factors (SURF) complex. SMG8 and SMG9 inhibit SMG1 activity which keeps UPF1 in a dephosphorylated state in the SURF complex until it interacts with UPF2 and UPF3 (CelikKervestin and Jacobson, 2015). Only then SMG1 phosphorylates UPF1 after which UPF1 dissociates from eRF1/eRF3 and triggers NMD (Kashima *et al.*, 2006).

1.7.2.2. UPF2 structure and function

UPF2 contains three mIF4G domains (structures which are comparable to the middle domain of eukaryotic initiation factor eIF4G) (Clerici *et al.*, 2014). Mutations in a conserved motif of two of the mIF4G domains were shown to suppress NMD in *S. pombe*, suggesting the importance of the domains in this process (Mendell *et al.*, 2000). The mIF4G closest to the UPF2 C-terminus mediates its interaction with UPF3 (Kadleclzaurralde and Cusack, 2004) while

its N- and C-terminal regions were mapped to interact with UPF1, regulating its conformational change from a closed, catalytically inactive, into an open conformation that enables UPF1 helicase activity, as discussed above (Kadleclzaurralde and Cusack, 2004; Clerici *et al.*, 2014). Therefore, UPF2 binds both UPF1 and UPF3, representing a bridge that interconnects these two proteins, allowing the concerted functions of the three proteins in NMD. Human UPF2 was also shown to interact directly with eRF3, in a UPF3 independent manner, only being able to bind one of the two proteins at a time, functional implications of which remain unclear (López-Perrote *et al.*, 2016). Additionally, a highly conserved region of the N-terminal mIF4G that does not mediate the interactions with UPF1 and UPF3 was shown to be essential for NMD. This region might mediate interactions of UPF2 with other factors, which would stimulate NMD (Fourati *et al.*, 2014).

1.7.2.3. UPF3 structure and function

UPF3 is the third core NMD factor, absence of which results in NMD suppression. A single gene encoding this protein is present in lower eukaryotes while human cells contain two UPF3 homologues, UPF3A and UPF3B, encoding two UPF3 isoforms. The isoforms display antagonistic roles, UPF3A is considered to have an inhibitory, while UPF3B has a stimulatory effect on NMD (Shum *et al.*, 2016).

UPF3 contains a ribonucleoprotein-type RNA-binding domain (RBD), otherwise termed as an RNA recognition motif (RRM); however, this protein has no RNA-binding properties even though the region indeed folds as a canonical RBD (Kadleclzaurralde and Cusack, 2004). UPF3 was found to interact with UPF2,

which in turn, mediates formation of the UPF1-UPF2-UPF3 complex. The interaction occurs between the C-terminal mIF4G domain of UPF2 and the region of UPF3 that includes the RNP domain, involving highly conserved residues of both proteins. Once the interaction is established, the complex, interestingly, obtains RNA-binding ability (Kadleciszaurralde and Cusack, 2004).

UPF3 was also found to associate with Y14, a protein that forms the EJC complex deposited on exon-exon junctions during splicing and it is this observation that led to the postulation of the EJC model of NMD, detailed below (Gehring *et al.*, 2003).

1.7.3. Role of NMD factors beyond NMD

Most observations obtained by investigating the structure, enzymatic activity and interactions of NMD factors have been interpreted in the context of the roles that these factors play in NMD. However, NMD factors may play roles in other cellular processes, which may or may not have a consequence on NMD.

UPF1 binds mRNAs globally, as demonstrated by CLIP studies, with the protein not discriminating between NMD targets and normal transcripts (Zünd *et al.*, 2013). Primarily it binds the 3'UTR with its distribution being shifted towards the coding sequence upon translation inhibition suggesting that it binds mRNA across its coding sequence and gets displaced from it into the 3' UTR by the translating ribosome (Zünd *et al.*, 2013). As, being an ATP-dependent mRNA helicase, UPF1 has mRNA remodelling properties, it could potentially translocate along the mRNA molecule, unwinding secondary structures, displacing RNA-bound proteins which could help ribosome translocation. UPF1, therefore, may have a more general role in translation efficiency. Additionally,

UPF1 might stimulate translation termination events as mutations in the CH domain lead to inhibition of translation termination (Czaplinski *et al.*, 1998). Its association with eIF3, demonstrated by co-immunoprecipitation studies, may also indicate a role of UPF1 in post-termination ribosome release (Isken *et al.*, 2008). Therefore, the role of UPF1 in NMD may be a consequence of its wider roles in mRNP remodelling and translation.

UPF1 is most abundant in the cytoplasm where it exerts its role in NMD and potentially in other processes discussed above. However, UPF1 is shown to localise in the nucleus, too, co-purifying with DNA polymerase δ as well as playing a role in DNA damage response and telomere maintenance (Azzalin and Lingner, 2006a; Azzalin and Lingner, 2006b). Some of the key recent findings on the nuclear role of UPF1 were made in the Brogna lab, where it was shown that UPF1 globally associates with transcription active sites in *D. melanogaster*, preferentially binding exons (unpublished). The authors have shown that UPF1 plays a key role in the release of polyadenylated mRNAs from the transcription sites and in their export into the cytoplasm. The data suggests that UPF1 may play a global role not only in NMD, but also in all stages of mRNA biogenesis, from its processing in the nucleus to the cytoplasm.

UPF3 is an NMD factor that shuttles between the nucleus and the cytoplasm. It was shown to interact with the EJC complex and to elicit NMD via its interaction with the remaining core NMD factors (KadlecZaurralde and Cusack, 2004; Gehring *et al.*, 2003). Recent findings have demonstrated that this protein, however, plays a more general role in translation termination. By using an *in vitro* re-constituted translation termination assay, UPF3B isoform was shown to

interact with the release factors and the ribosome directly and to play a critical role in two phases of translation termination, firstly delaying termination and inhibiting peptide release and then stimulating post-termination ribosome release upon peptide dissociation (Neu-Yilik *et al.*, 2017; Gao and Wilkinson, 2017; Mühlemann and Karousis, 2017). These observations suggest that a key role of UPF3B in NMD could be a consequence of its role in translation termination events.

NMD factors were shown to associate with cellular granules, namely P bodies as well as stress granules (Brognan and Ramanathan and Wen, 2008). The association with P bodies was mainly interpreted as a consequence of NMD substrates being targeted to these cytoplasmic bodies for degradation (Sheth and Parker, 2006). Normally, the NMD substrates go through the P bodies transiently, however, when the NMD process is impaired, the mRNAs accumulate in these granules instead. Indeed, it was demonstrated that PTC-containing mRNAs accumulate in P bodies in UPF deletion strains, as their degradation is impaired (Sheth and Parker, 2006). On the other hand, the UPF factors might be sequestered into the stress granules together with the pool of mRNAs that are silenced during the course of stress conditions. Whether these factors play a role in the assembly and functionality of the cellular granules remains to be investigated.

1.8. Standard NMD models

NMD is a translation dependent process, which is thought to be the joint function of NMD factors (discussed above), required for PTC recognition and mRNA degradation. It is not yet understood whether and how NMD distinguishes between natural and premature stop codons and how degradation of PTC-containing mRNAs is initiated. Current NMD models offer different explanations as to how NMD occurs in cells. However, what all these models have in common is the presumption that NMD occurs as an active surveillance mechanism, a result of coordinated actions of UPF proteins that recognise PTC-containing transcripts and eliminate them (BrognnaMcLeod and Petric, 2016).

Therefore, in the following, I will discuss the models proposed so far which explain a number of characteristics of this phenomenon. Finally, I will discuss an alternative model that we propose, assessed in this study.

1.8.1. DSE Model

One of the first observations regarding NMD was the position-dependent effect of a PTC, meaning that the transcripts containing PTCs early within their sequences are strongly targeted by NMD. In a study carried out in budding yeast *S. cerevisiae*, where nonsense mutations were introduced at different positions within the *URA3* gene, mutants that contained nonsense mutations closer to the 5' end produced lower levels of *URA3* mRNA compared to PTCs introduced towards the 3' end (Losson and Lacroute, 1979). This phenomenon has since been observed in other organisms, such as *D. melanogaster* and mammals (PeltzBrown and Jacobson, 1993; Maquat and Li, 2001).

The presence of one or more sequence motifs, which elicit NMD, termed downstream sequence elements (DSEs), was first proposed as an explanation of the described phenomenon. The DSEs were predicted to occur more often in the coding regions compared to the 3' UTRs (Czaplinski *et al.*, 1998; Wang *et al.*, 2001). Consequently, if a PTC is encountered early in the coding sequence, the higher is the probability for a DSE motif to occur in the downstream region and to trigger NMD. The DSE model was, therefore, postulated based on this hypothesis (Figure 1.9).

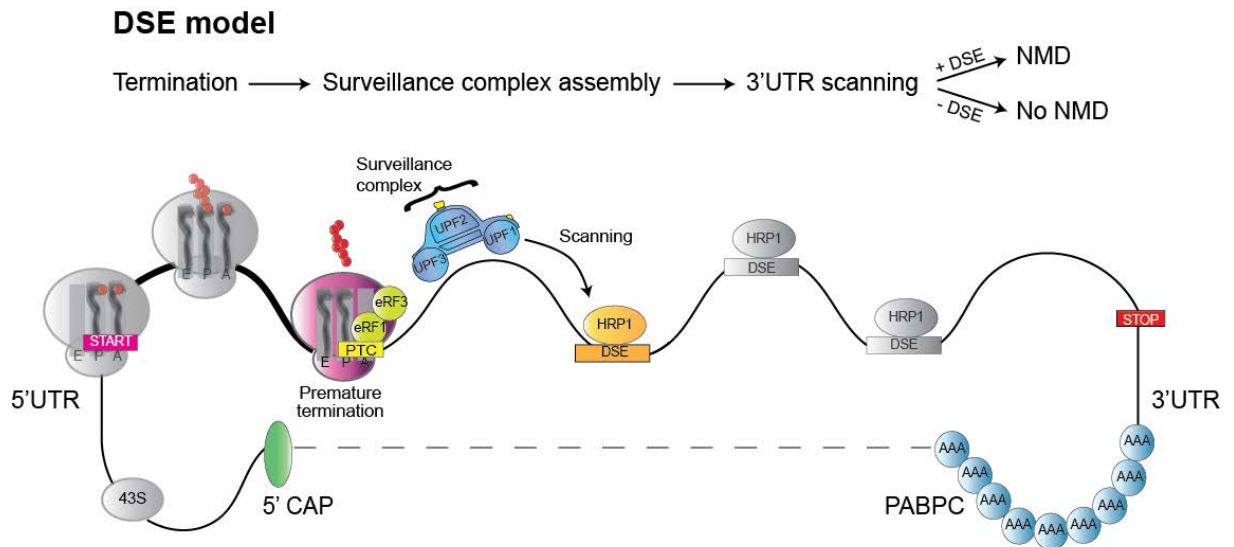


Figure 1.9 – **The DSE model.** Translation termination is labelled as premature due to the presence of a downstream sequence element (DSE). HRP1, an RNA binding protein, binds to the putative DSE in the nucleus and remains associated during mRNA transport into the cytoplasm. Upon translation termination, a surveillance complex, composed of UPF1, UPF2 and UPF3 proteins, assembles on the terminating ribosome and scans the downstream sequence. The interaction between UPF1 of the surveillance complex and the HRP1 bound to the DSE indicates that translation termination is premature, which triggers rapid mRNA degradation (the figure was published in (BrognamcLeod and Petric, 2016). 5' CAP – 5' capping complex; UTR – untranslated region; START – start codon; PTC – premature translation termination codon; STOP – stop codon; PABPC – poly(A)-binding protein, cytosolic; eRF – eukaryotic release factor; DSE – downstream sequence element.

This model predicted that HRP1, an RNA binding protein, plays a crucial role in NMD. This protein would bind a DSE in the nucleus and remain bound to the mRNA during its export to the cytoplasm (Kessler *et al.*, 1997). When translation termination occurs, a surveillance complex comprised of UPF1, UPF2 and UPF3 assembles on the terminating ribosome and scans the downstream sequence. Only if it encounters a DSE, the interaction between HRP1 bound to the DSE and UPF1 of the surveillance complex determines that translation termination is premature, recruiting the decay factors and initiating mRNA rapid degradation (BrognnaMcLeod and Petric, 2016). HRP1 was found to interact *in vitro* with both UPF1 and a putative DSE, which was in favour of the model (González *et al.*, 2000). Additionally, NMD is suppressed if sequences downstream of a nonsense mutation are deleted. However, as a consensus has not been established between the DSE sequences of different mRNAs and as artificial extension of 3'UTRs also strongly stimulates NMD, the hypothesis of specific sequence motifs that induce NMD has been largely abandoned (Muhlrad and Parker, 1999; Ruiz-EchevarríaGonzález and Peltz, 1998).

1.8.2. The EJC Model

Similarly to the previous hypothesis, the EJC model postulates that a downstream signal, which in this case is a splice junction, distinguishes premature translation termination from a normal termination event.

NMD was first linked to splicing in mammalian cells where it was shown that a PTC located upstream of an intron strongly stimulated NMD (Zhang *et al.*, 1998) (Figure 1.10). Moreover, artificial insertion of an intron downstream of a natural stop codon leads to an NMD-induced degradation of the transcript (CarterLi and Wilkinson, 1996; Wen and Brogna, 2010). These observations suggested that an intron downstream of a PTC might serve as the signal that distinguishes a PTC from a natural termination codon.

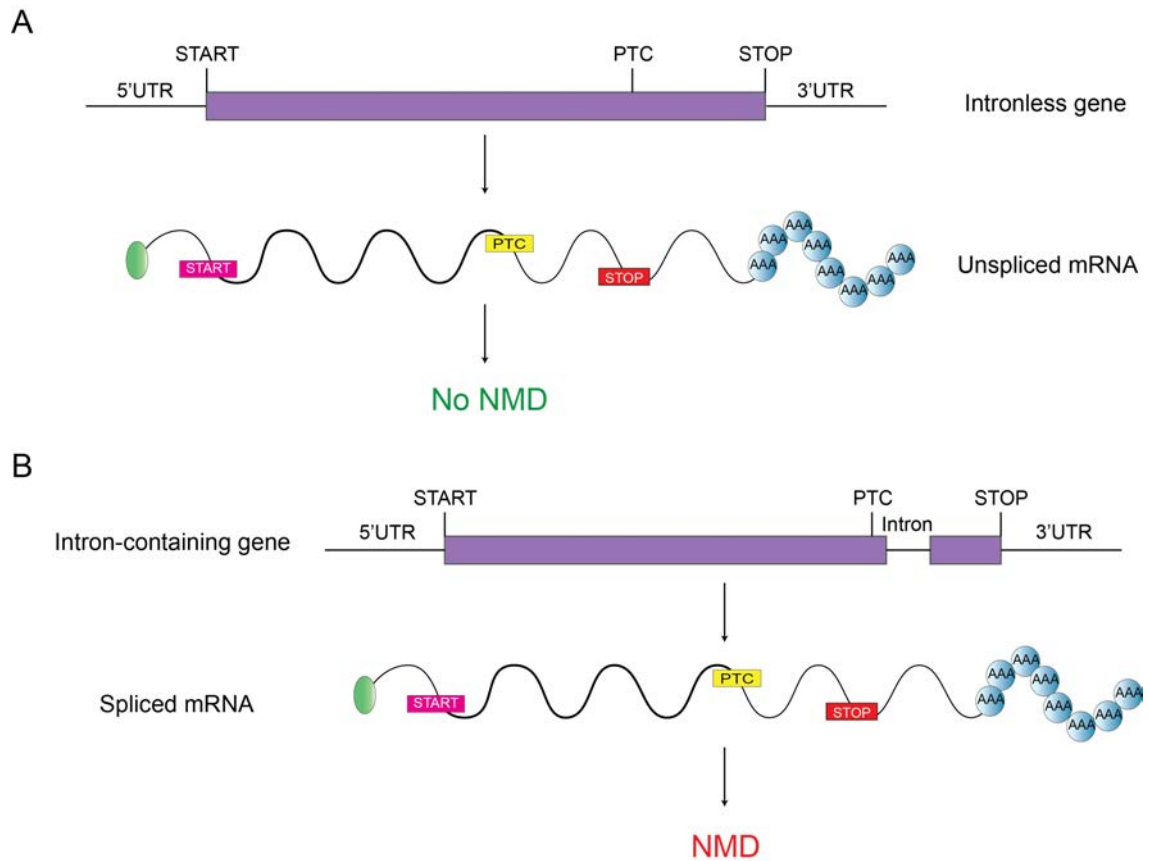


Figure 1.10 – **Splicing enhances NMD.** (A) An intronless construct that contains a PTC late in its sequence escapes NMD. (B) The same construct containing an intron introduced downstream of the PTC undergoes strong NMD, with undetectable mRNA levels. UTR – untranslated region; START – start codon; PTC – premature translation termination codon; STOP – stop codon.

However, how splicing, being a nuclear process, can affect the cytoplasmic process of NMD has long been unclear, until the discovery of the Exon Junction Complex (EJC), which led to the postulation of the EJC model of NMD (Figure 1.11).

As outlined in section 1.3, the EJC is a multiprotein complex deposited 22-24 nucleotides upstream of exon-exon junctions during splicing, which stays bound to mRNAs during their transport to the cytoplasm (Brognia and Wen, 2009; Le Hir *et al.*, 2000). Proteins eIF4AIII, Y14 and MAGO constitute the core of the EJC to which additional proteins associate, including UPF2 and UPF3 (Le Hir *et al.*, 2001; BrogniaMcLeod and Petric, 2016). The current EJC model predicts that, as discussed briefly in the chapter 1.7.2, SMG1 and UPF1 bind to all terminating ribosomes and, together with the eukaryotic release factors (eRF1, eRF3), form the SURF complex on the ribosome (Kashima *et al.*, 2006). UPF2 can associate with both UPF3 bound to the EJC and UPF1 bound to the ribosome. Therefore, if the termination occurs upstream of an exon-exon junction, the model postulates that UPF2, by binding to UPF1 and UPF3, forms a bridge between the SURF complex on the terminating ribosome and the downstream EJC (Kashima *et al.*, 2006; Chamieh *et al.*, 2008). This interaction induces UPF1 helicase activity and subsequent mRNA degradation (Figure 1.11) (Brognia and Wen, 2009; Mühlemann *et al.*, 2008).

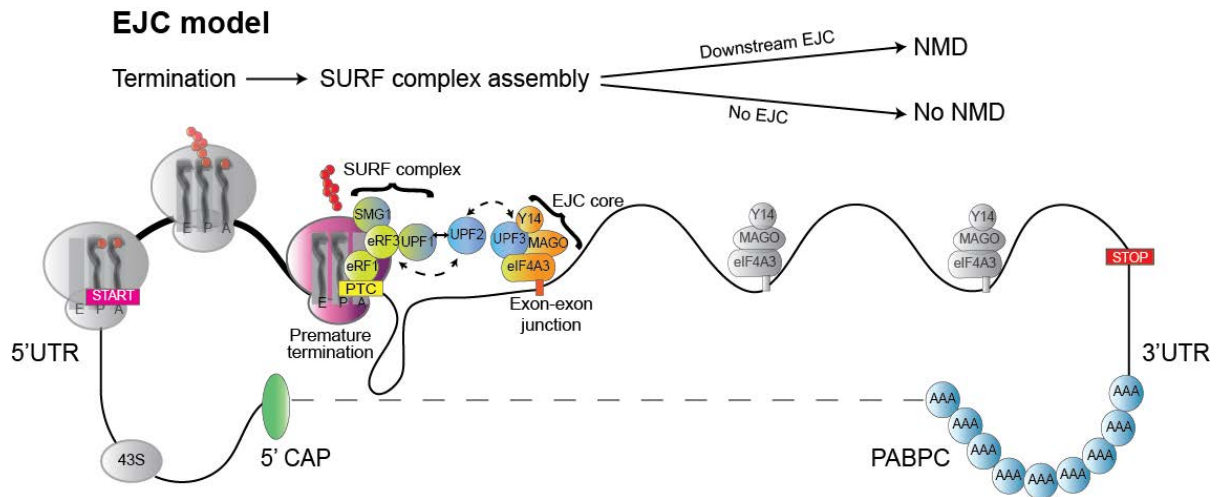


Figure 1.11 – **The EJC model**. Translation termination at a PTC is abnormal. NMD is again triggered due to a presence of a downstream signal, in this case, an EJC complex that binds close to mRNA exon-exon junctions. Translation termination leads to the formation of the SURF complex, composed of SMG-1, UPF1 and release factors 1 and 3 (eRF1 and eRF3). UPF2 and UPF3, on the other hand, were found to associate with the EJC complex consisting of eIF4AIII, Y14 and MAGO. Therefore, interaction between UPF1 associated with the SURF complex and UPF3 associated with the downstream EJC, mediated via UPF2, labels translation termination as premature and triggers NMD (the figure was published in (BrognMcLeod and Petric, 2016). 5' CAP – 5' capping complex; UTR – untranslated region; START – start codon; PTC – premature translation termination codon; EJC – exon junction complex; STOP – stop codon; PABPC – poly(A)-binding protein, cytosolic; eRF – eukaryotic release factor.

The EJC model cannot explain a number of observations regarding NMD. For example, taking the EJC model into account, it was puzzling to find that splicing enhances NMD in *S. pombe* regardless of whether an intron is positioned upstream or downstream of a PTC (Wen and Brogna, 2010). If an intron were upstream of a PTC, the EJC deposited on it would be removed by the translating ribosome before the PTC is reached. The EJC would no longer be present to trigger NMD, as a consequence. Moreover, introns induced strong NMD even in the absence of the core EJC proteins, suggesting that the EJC is not required for NMD in *S. pombe* (Wen and Brogna, 2010), *D. melanogaster* nor in *C. elegans* (Stalder and Mühlemann, 2008). Additionally, even though *S. cerevisiae* does not contain the EJC proteins, splicing-dependent NMD can be efficiently induced in this organism (Brogna lab, unpublished). It is possible, however, that *S. cerevisiae* could contain some EJC-like structures.

The link between splicing and NMD, therefore, is still not completely understood, but, to a large extent, it does not seem to rely on the EJC.

1.8.3. The Faux 3'UTR Model

The faux 3'UTR model postulates that premature translation termination is distinctly different from that at a normal stop codon and it triggers NMD when it occurs far from the 3' end of the mRNA (BroгнаMcLeod and Petric, 2016).

Mature mRNA contains multiple adenine residues at the 3' end (poly(A) tail), to which cytoplasmic poly(A) binding protein (PABPC) binds (Cosson *et al.*, 2002b). During translation termination, eukaryotic release factor eRF3 forms a complex with eRF1 which recognises all three stop codons in the A site of the terminating ribosome, as detailed in the section 1.4.2.3. PABPC was found to interact with the eRF3, via a yeast two-hybrid approach and *in vitro* binding studies (Hoshino *et al.*, 1999; Cosson *et al.*, 2002a). Therefore, if translation termination occurs at a normal stop codon, eRF3 associated with the terminating ribosome interacts with the PABPCs, which has a stimulatory effect on polypeptide release and recycling of the ribosomal subunits. However, if a termination event occurs at a PTC which is far from the 3' end (early in the sequence), release factors cannot interact with the PABPC, which alters the termination event and most probably causes the delay in ribosome release (BroгнаMcLeod and Petric, 2016). UPF1 and other NMD factors then associate to the terminating ribosome and initiate degradation of such mRNAs (Broгна and Wen, 2009). Therefore, this model suggests that the distance between a PTC and the 3' end, more specifically the poly (A) tail, is the main determinant of how translation termination will occur (Figure 1.12).

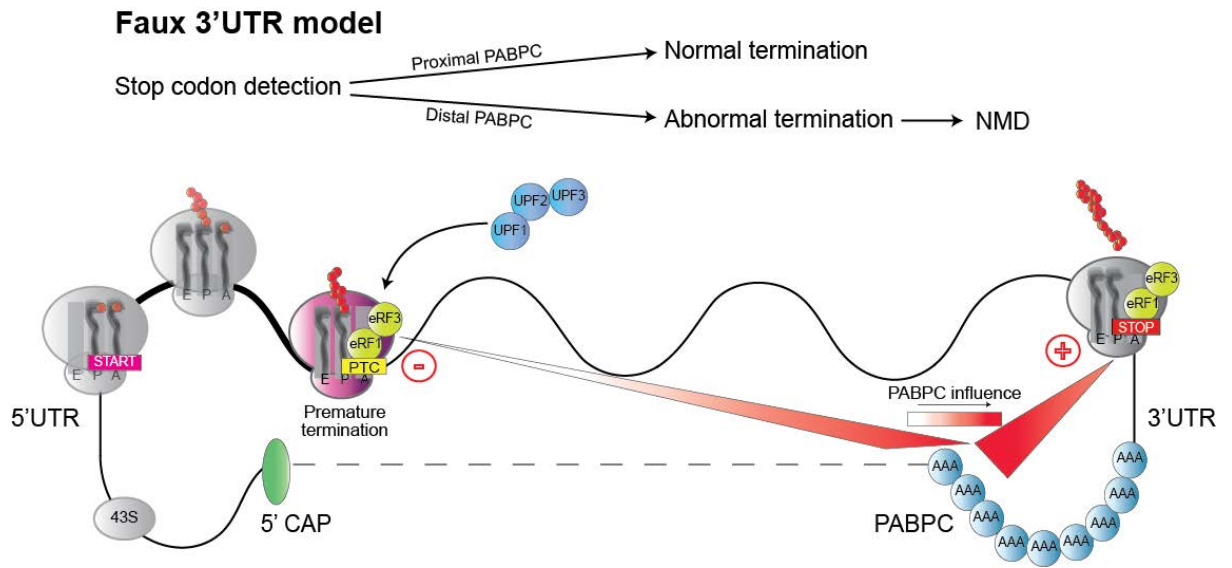


Figure 1.12 – **The faux 3'UTR model.** Translation termination at a normal stop codon is stimulated via an interaction between eRF3 bound to the terminating ribosome and PABPC associating with the poly(A) tail. This interaction stimulates polypeptide release and ribosome disassembly. Termination at a PTC is instead abnormal, as it occurs far from the poly(A) tail. The eRF3 is therefore incapable of interacting with the PABPCs, which delays normal termination and leads to association of NMD factors that in turn initiate rapid mRNA degradation (the figure was published in (BroгнаMcLeod and Petric, 2016). 5' CAP – 5' capping complex; UTR – untranslated region; START – start codon; PTC – premature translation termination codon; STOP – stop codon; PABPC – poly(A)-binding protein, cytosolic; eRF – eukaryotic release factor.

One of the key observations consistent with this model is that artificial tethering of PABPCs towards a PTC leads to mRNA stabilisation (Ivanov *et al.*, 2008). The main supporting evidence, in turn, comes from experiments which showed that ribosomes undergoing premature and normal termination leave different mRNA footprints (toeprints), suggesting an altered translation termination at a PTC (Amrani *et al.*, 2004).

On the other hand, mRNAs generally form closed loop structures whilst translated, so 5' proximal PTCs would still be close to PABPC, meaning that eRF3 could potentially still interact with the PABPC and should trigger normal translation. However, this does not happen and these PTCs are inducing strong NMD instead. Moreover, inconsistent with the faux 3'UTR model, mRNAs that lack a poly(A) tail can still be targeted by NMD (Meauxvan Hoof and Baker, 2008). In addition, deletion of the yeast PABPC does not affect NMD in *S. cerevisiae* (Meauxvan Hoof and Baker, 2008; Brogna and Wen, 2009), nor does deletion of the C-terminal region of PABPC that interacts with the terminating ribosome (Simón and Séraphin, 2007).

Therefore, current NMD models cannot fully explain the process of NMD and further studies need to be carried out to provide a general, unified explanation of how NMD is indeed carried out.

1.9. Ribosome release model

NMD models to date presume that an active mechanism or a downstream signal discriminate PTCs from normal stop codons. However, it is possible that, due to more generalised roles of NMD factors in translation, NMD occurs as a passive consequence of ribosome release from an mRNA upon premature translation termination, which renders the mRNA exposed and thereby susceptible to degradation by nucleases.

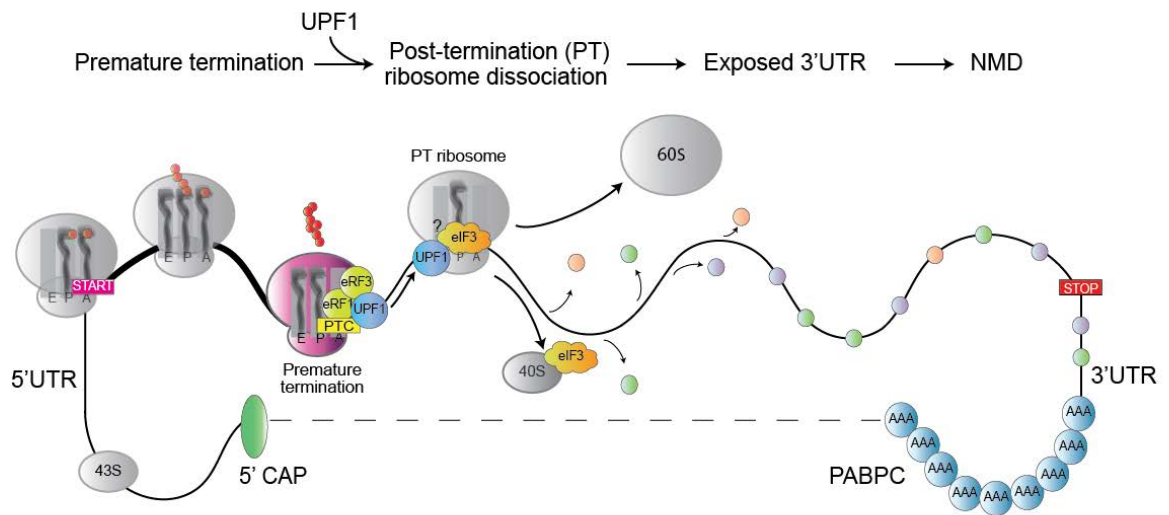
After translation termination occurs, post-termination ribosomes dissociate into subunits to be recycled for another round of translation. Ribosome recycling is a multi-step process that requires activity of a number of proteins, as discussed in section 1.4.2.4, with eIF3 considered to be one of the main factors mediating the dissociation of ribosome subunits (PisarevHellen and Pestova, 2007). Several studies indicated that this protein has an important role in NMD in yeast and mammalian cells, as mutations in eIF3 subunits can cause NMD suppression (Welch and Jacobson, 1999). Moreover, UPF1 was shown to interact directly with eIF3, as demonstrated by a co-immunoprecipitation assay (Isken *et al.*, 2008).

In light of these observations, which indicate that UPF1 might help in dissociating ribosomal subunits from an mRNA after translation termination, the underlying hypothesis is that NMD might occur as a consequence of UPF1-mediated release of post-termination ribosomes stalled at a PTC (CelikKervestin and Jacobson, 2015; Brogna and Wen, 2009). UPF1 also, being a highly processive helicase shown to possess mRNP-remodelling

properties, might stimulate dissociation of other mRNA-associated proteins and potentially of ribosomal subunits, as well. As a result, mRNP would become unstable as mRNA is now unprotected and, as such, more susceptible to the effect of nucleases, which would initiate its rapid degradation (Figure 1.13, wild type). On the other hand, in UPF1 null mutant strain, ribosomes and mRNA-associated proteins would not be efficiently released from an mRNA. The transcript would remain covered with mRNA-bound proteins, which would protect it from degradation by nucleases in a passive manner, inhibiting decay (Figure 1.13, UPF1 null mutant) (BrogaMcLeod and Petric, 2016).

Release model

Wild type: *efficient ribosome release*



UPF1 null mutant: *inefficient ribosome release*

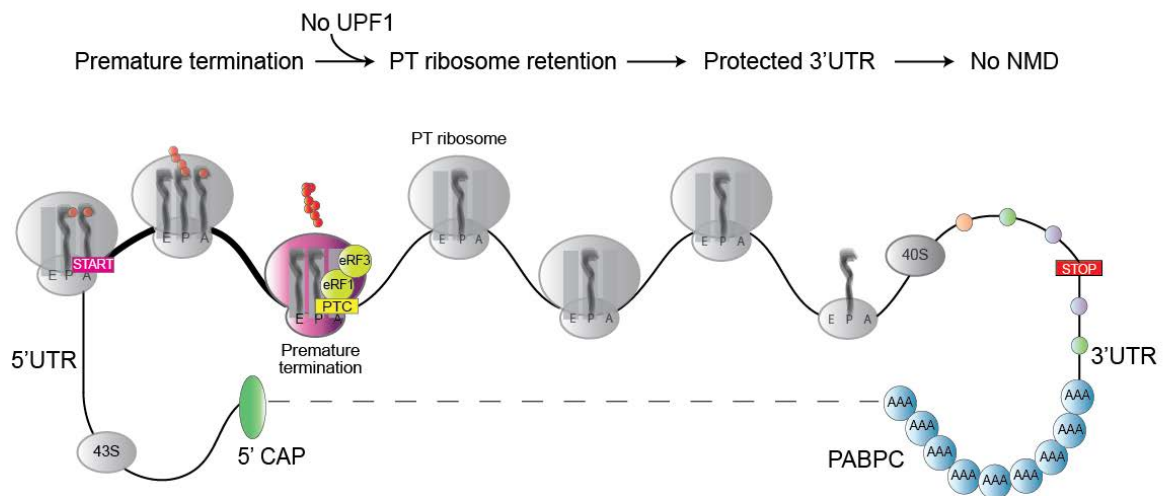


Figure 1.13 – **The ribosome release model.** In normal conditions (top diagram), upon translation termination, the post-termination ribosome remains associated with the mRNA until UPF1 is recruited. UPF1, via its interaction with eIF3 and/or via its ATP 5'-3' helicase activity, stimulates release of the terminating ribosome as well as other mRNA-associated proteins. In the presence of a PTC, the resulting unprotected mRNA becomes susceptible to degradation by the decay machinery (top diagram). In the absence of UPF1 (bottom diagram), post-termination ribosomes remain associated with the mRNA, shielding (together with RNA-binding proteins) the mRNA from degradation (published in (BrogaMcLeod and Petric, 2016). 40S – small ribosomal subunit; 60S – large ribosomal subunit; UTR – untranslated region; eRF1 – eukaryotic release factor 1; eRF3 – eukaryotic release factor 3; eIF3 – eukaryotic initiation factor 3; UPF1 – up-frameshift protein 1; PTC – premature translation termination codon; PT – post-termination; PABPC – poly(A)-binding protein, cytosolic. Colourful circles represent mRNA-associated proteins.

Regardless of the PTC position, ribosomal subunits should be released from an mRNP equally efficiently, making position-dependent NMD difficult to explain using this model in the first instance. However, if a premature termination occurs early, a long region of an mRNA downstream of the PTC would become unprotected, which would destabilise the mRNP and make it a target for destruction either by exosome or endonucleases. In case a PTC occurs closer to the 3' end, NMD would not be induced as efficiently, since only a short region of mRNA would be depleted of ribosomes (Broгна and Wen, 2009). Also, as mRNAs are circularised during translation, dissociated ribosomal subunits would find themselves close to the 5' end, which could lead to efficient re-initiation of translation and stabilisation of the translation circuit.

It was reported that more ribosomes associate with a PTC containing mRNA in the UPF1 null mutant background in *S. cerevisiae* (Hu *et al.*, 2010), consistent with the model. Additionally, as previously discussed, the disassembly and the turnover of NMD-targeted mRNPs were shown to depend on UPF1 ATPase activity. In cells expressing UPF1 which lacks ATPase activity, NMD factors stay bound to the partially degraded NMD substrates which concentrate in processing bodies (FranksSingh and Lykke-Andersen, 2010). Therefore, the mRNP remodelling catalysed by UPF1 seems to be crucial for NMD to occur. Factors which help either UPF1 recruitment or its activation should enhance NMD, such as the EJC proteins (Broгна and Wen, 2009), which could explain their stimulatory effect on NMD.

In light of recent observations, however, that suggest a key role of UPF3B for ribosome release and recycling, where UPF1 was found to mainly play a role in

triggering assembly of the decay machinery, the model needs to be revisited (Neu-Yilik *et al.*, 2017). This model, nevertheless, provides a fresh outlook on the mechanism of NMD, even though it requires further experimental validation. In this study, I made attempts to assess this model, as detailed in Chapter 5.

1.10. CRISPR/Cas9 gene editing

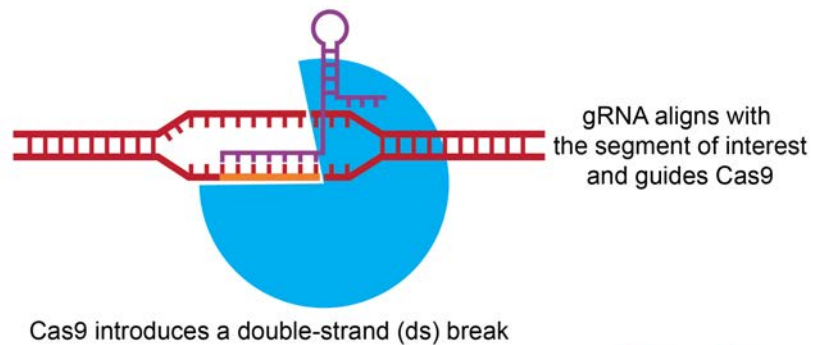
The final aim of my project is to optimise CRISPR/Cas9 gene editing approach to generate reagents to study NMD and translation in the cells – endogenous NMD reporter genes and ribosomal protein-tagged strains. CRISPR/Cas9 (clustered, regularly interspaced short palindromic repeats- associated proteins system) is a means of adaptive prokaryotic immunity that prevents recurring invasion of mobile genetic elements derived from plasmids and bacteriophages, by destroying the invading DNA whilst retaining a memory of previous infections (Mojica *et al.*, 2005; Bolotin *et al.*, 2005). The foreign DNA gets incorporated into the CRISPR locus and upon transcription of CRISPR pre-mRNA and its processing into short RNAs, Cas9-bound RNA molecules hybridise with the invading pathogen's DNA, which is followed by its rapid degradation catalysed by Cas9 (PourcelSalvignol and Vergnaud, 2005; Makarova *et al.*, 2006; Barrangou *et al.*, 2007).

This system has had a revolutionary application in eukaryotic organisms where it is being used for gene editing – creating targeted deletions, insertions, point mutations, and epitope tags (Mali *et al.*, 2013). The system relies on the expression of a guide RNA (gRNA) molecule that aligns to the target gene, and Cas9 that binds the DNA in a complex with the gRNA and introduces a double strand break. The double strand break can be repaired either by non-

homologous end joining (NHEJ) that introduces random mutations, ideal to inactivate the gene of interest, or, in case a homologous template bearing the desired change is supplied to the cells, the repair occurs via homologous recombination, introducing alteration of interest (Figure 1.14) (Ran *et al.*, 2013; WestraBuckling and Fineran, 2014; MaZhang and Huang, 2014).

As previously outlined, and detailed in Chapter 6, I describe the implementation of CRISPR/Cas9 approach *in S. pombe*, adapting it in our laboratory conditions based on published data (Jacobs *et al.*, 2014), to generate strains expressing endogenous NMD reporter genes and strains in which desired proteins contain a small epitope tag.

A



B



C

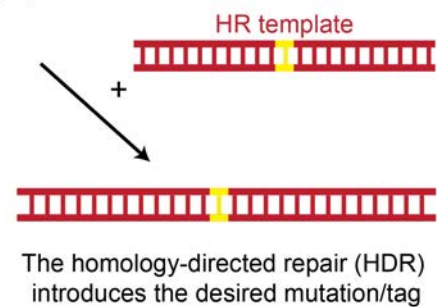


Figure 1.14 – **CRISPR/Cas9 gene editing overview.** (A) gRNA aligns to the gene of interest, bringing Cas9 to the target locus. Cas9 introduces a double strand break, which is then repaired either via (B) non-homologous end joining that introduces random mutations or, in case a homologous recombination template containing desired alterations is supplied to the cells, via (C) homology-directed repair.

Chapter 2. Materials and methods

2.1. *Schizosaccharomyces pombe* strains

S. pombe strains used in this study are listed in the Appendix 8.1.

2.2. *S. pombe* media

S. pombe was grown in the following types of media: rich medium, otherwise referred to as yeast extract with supplements (YES) and minimal medium which is based on Edinburgh minimal medium (EMM). Typically, rich medium was used for nonselective or antibiotic resistance-selective growth, growth before long-term storage at -80°C as well as for strain recovery from glycerol stocks. A variant of EMM, EMM containing glutamic acid as a nitrogen source (EMMG), was used to select prototrophic strains after transformation and, since, contrary to YES, it does not contain thiamine, it was used in cases when the gene of interest is expressed from a no-message in thiamine (*nmt*) promoter, for which YES is not suitable. Additionally, prior to carrying out a transformation, cells were grown in MB medium, which is extremely stringent; it weakens the cell wall and therefore increases transformation efficiency. All media components are listed in the Appendix 8.2 (Petersen and Russell, 2016). Media was sterilised by autoclaving at 120°C for 30 minutes except for MB medium, which was filter sterilised.

2.3. *S. pombe* culturing and maintenance

All strain stocks were stored at -80°C in YES medium containing 25% of glycerol. The strains either designed by previous lab members or obtained from other groups were retrieved from Brogna lab glycerol stocks where individual strains were streaked on YES agar plates and grown at 32°C to obtain single colonies.

Upon transformation, a successfully transformed strain was selected on an appropriate selection plate (in case of antibiotic resistance on a YES plate containing 100 µg/mL of the antibiotic; in case of leucine or uracil prototrophy on an EMM plate deprived of either leucine or uracil (EMM leucine or uracil drop-out, respectively)). Single colonies were re-streaked onto another selection plate to confirm selection.

Liquid cultures were obtained by a single colony resuspension in 2 mL of appropriate medium and overnight growth at 30°C in a shaking incubator. The resulting start-up culture was scaled up to a desired volume the following day and subsequently grown overnight to a desired optical density/growth phase.

2.4. *S. pombe* cell counting

Cells were counted using a haemocytometer to ensure that they are in an appropriate phase of cell growth (typically mid-log phase) for a required experiment. A clean coverslip was first placed onto the counting chamber. Cells were then diluted 10-fold (or more if the culture appeared very dense) and 10 μL of the dilution was carefully loaded onto the chamber, with cells being evenly dispensed beneath the coverslip by capillary action. Cells from squares indicated in red (Figure 2.1) were then counted under the microscope using a hand tally counter. The resulting cell count multiplied by 2×10^5 would represent the number of cells per mL of culture, prior to the 10-fold dilution.

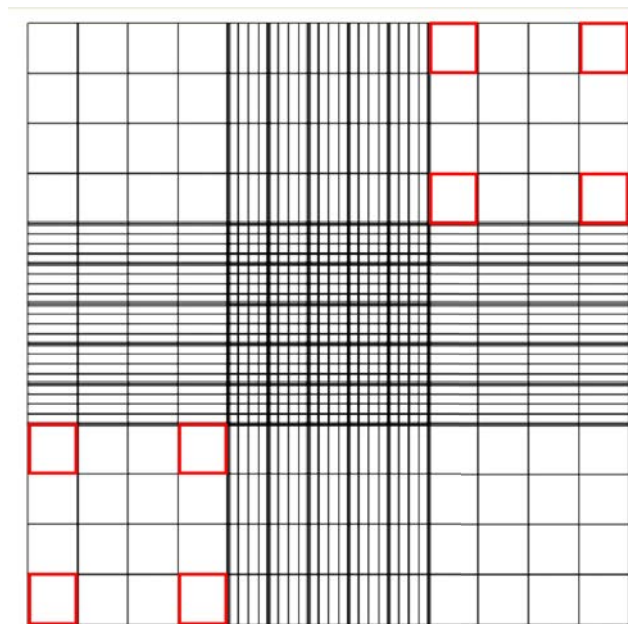


Figure 2.1 – **Haemocytometer grid.** The cells are counted from squares highlighted in red (details in the text).

2.5. Genomic DNA extraction

S. pombe genomic DNA was extracted using GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific), according to manufacturer's instructions. Up to 1×10^8 cells (10 mL culture, OD₆₀₀ of 0.5) were harvested by centrifugation at 3000 rpm for 5 min at 4°C. The pellet was resuspended in 1 mL of yeast lysis buffer (5 mg/mL Zymolyase 20T (AMS Biotechnology), 1 M sorbitol, 0.1 M EDTA), the suspension transferred to a 1.5 mL microcentrifuge tube and incubated for 1 hour at 37°C. Cells were then centrifuged at 5500 rpm for 10 min and the pellet was resuspended in 180 µL of Digestion solution, to which 20 µL of Proteinase K was added. The suspension was mixed vigorously until uniform and incubated for approximately 45 minutes at 56°C in a dry bath with occasional vortexing, until the cells were completely lysed. To the lysate 20 µL of RNase A was then added, followed by a 10-minute incubation at room temperature. With the addition of 200 µL of Lysis solution, the sample was mixed by vortexing for 15 seconds until a homogenous mixture was obtained, to which 400 µL of 50% ethanol was added and resuspended by pipetting. The lysate was transferred into a GeneJET Genomic DNA Purification column inserted into a collection tube and spun at 8000 rpm for 1 min. The column was transferred into a new collection tube, washed with 500 µL of Wash buffer I with added ethanol and centrifuged at 9000 rpm for 1 min. The flow-through was discarded, the column placed back into the collection tube and washed with 500 µL of Wash buffer II, followed by a 3-minute centrifugation at 13000 rpm. The column was transferred to a clean 1.5 mL microcentrifuge tube to which 200 µL of Elution buffer was carefully added. The sample was incubated for 5 min at

room temperature and then centrifuged at 9000 rpm for 2 minutes to recover genomic DNA. The DNA quality was verified by running the sample on a 1% agarose gel.

2.6. RNA extraction from yeast cell cultures

S. pombe RNA was isolated by extraction using hot acidic phenol (Collart and Oliviero, 2001). All steps following initial acid phenol incubation were carried out at 4°C to minimise RNA degradation. A 10 mL *S. pombe* cell culture was grown until its OD₆₀₀ reached 0.5 units (10⁷ cells/mL, mid-log phase), since, as a cell culture approaches stationary phase, the RNA yield varies, and the results are inconsistent. The cell culture was transferred into a 50 mL centrifuge tube and spun at 3000 rpm for 5 minutes at 4°C. After discarding the supernatant, the pellet was resuspended in 1 mL of sterile ice-cold dH₂O; the cells were then moved into a clean 1.5 mL microcentrifuge tube and centrifuged again at 6000 rpm for 10 seconds at 4°C. The supernatant was removed and at this point the pellet was either processed further or frozen in liquid nitrogen and stored at -80°C until needed (up to 6 months), when it was thawed on ice before the procedure was resumed. The cell pellet was resuspended in 600 µL of TES solution (10 mM Tris-HCl pH 7.5, 10 mM EDTA, 0.5% SDS), to which 600 µL of acid phenol pH 4.3 (Thermo Fisher Scientific) was added. The suspension was vortexed vigorously for 10 seconds, until it became cloudy white. The cells were then incubated for 1 hour at 65°C, vortexing every 5-10 minutes to allow phenol to remain in solution and, therefore, to obtain efficient extraction of both large and small RNA types. After the 65°C incubation step, the samples were cooled on ice for 5 minutes and then spun at 13000 rpm for 5 minutes at 4°C. The top

aqueous phase was transferred to a 1.5 mL microcentrifuge tube containing 600 μ L of acid phenol pH 4.3 and the suspension was vortexed vigorously prior to centrifugation at 13000 rpm for 5 minutes at 4°C. The top phase was moved into a 1.5 mL microcentrifuge tube to which 600 μ L of chloroform: isoamyl alcohol (24:1) was added and, after thorough mixing, the centrifugation step was repeated. Finally, the top aqueous phase was transferred into an RNase free 1.5 mL microcentrifuge tube containing 55 μ L (1/10 of the sample volume) of 3 M sodium acetate pH 5.2. RNA was precipitated by addition of 1 mL of ice-cold 100% ethanol and by 30-minute incubation at -20°C. After precipitation, RNA was pelleted at 13000 rpm for 15 minutes at 4°C and then washed in 1 mL of 70% ice-cold ethanol. After sample centrifugation at 13000 rpm for 5 minutes at 4°C, the pellet was air-dried and resuspended in 50 μ L of DEPC treated dH₂O. Samples were stored at -20°C for up to 1 year.

2.7. Protein purification from *S. pombe* cell cultures

A rapid sodium hydroxide method was used to isolate proteins from *S. pombe* (Matsuo *et al.*, 2006). A 5 mL *S. pombe* cell culture was grown until its OD₆₀₀ reached 0.5 units (10^7 cells/mL). All the steps were carried out at room temperature. The culture was centrifuged at 3000 rpm for 5 minutes. The pellet was resuspended in 1 mL of dH₂O, transferred into a 1.5 mL microcentrifuge tube and centrifuged again. The pellet was then resuspended in 300 μ L of dH₂O, to which 300 μ L of 0.2 M sodium hydroxide was then added. The sample was vortexed briefly and then incubated 10 minutes to allow cell lysis and protein release. After the incubation step, the sample was immediately centrifuged at 13000 rpm for 2 minutes. Finally, the pellet was resuspended in

SDS sample loading buffer (60 mM Tris-HCl pH 6.8, 4% SDS, 4% β -mercaptoethanol, 0.04% bromophenol blue, 1 mM PMSF and 20% glycerol) and boiled for 5 minutes. Samples were stored at -20°C for up to 1 year and typically run on 12% SDS-PAGE gels and analysed by Western blotting (detailed in section 2.15).

2.8. PCR based gene targeting – two-step PCR approach

PCR based gene targeting approach was adopted (Krawchuk and Wahls, 1999) (Figure 2.2) to construct strains in which a small epitope tag is added to the C-terminus of a desired protein, allowing its monitoring in the cells.

This approach makes use of a plasmid, pFA6a derivative, which contains a sequence that encodes a 3xHA tag, a transcription terminator, along with a selection marker (antibiotic resistance gene) and two primer-binding sequences (PBSs) from each side, to which hybrid primers bind. Plasmid map is shown in the Appendix 8.4 while the primers used are listed in the Appendix 8.5.

Two PCR reactions were set up to produce a tagging cassette with at least 250 bp of flanking homology to the target region (Krawchuk and Wahls, 1999) (Figure 2.2). One of the primers in the first PCR reaction is short and homologous only to the target locus (w: approximately 300 bp upstream of the stop codon, z: approximately 300 bp downstream of the stop codon), while the other is a hybrid primer with homology to both the plasmid module and the target locus (pFA6-1, x in the first and pFA6-2, y in the second PCR reaction, respectively). The primers were designed so that the 3' end of the coding sequence is in frame with the epitope on the plasmid. Two separate PCR reactions were carried out, set up using Q5 DNA polymerase and *S. pombe*

genomic DNA as a template (Table 2.1 and 2.2), to create DNA fragments of approximately 300 bp of homology from each side of the desired integration site, referred to as left and right PCR products. The PCR products were used in a second PCR reaction with the plasmid module as a template and two end primers (*w* and *z*), to produce a gene tagging cassette that is flanked from each side with more than 250 bp of homology to the target locus (Krawchuk and Wahls, 1999).

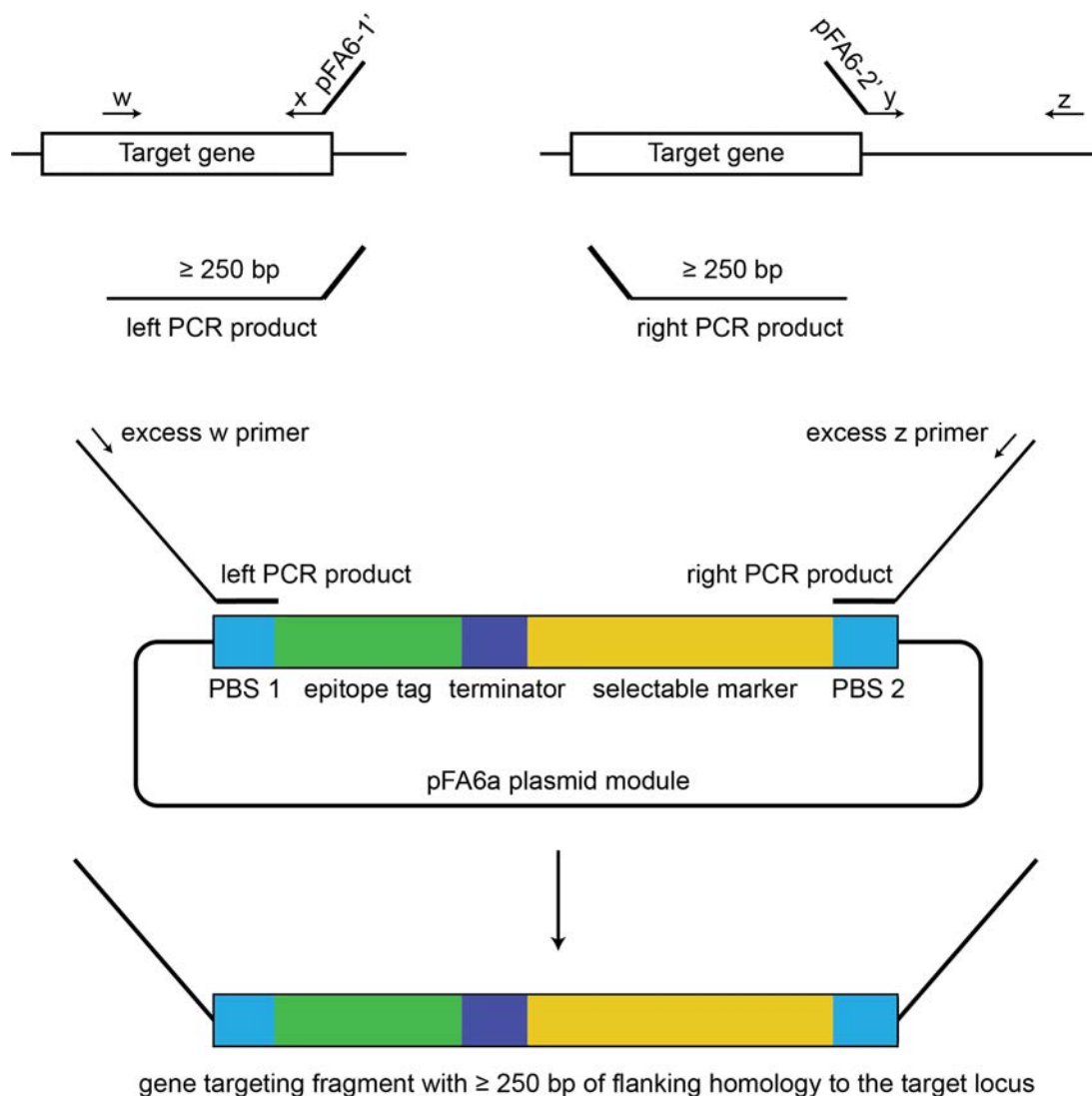


Figure 2.2 – **PCR based gene targeting – two step PCR approach.** Detailed explanation of the method is in the text. In short, two DNA fragments approximately 250 bp in length, surrounding the insertion site, whilst introducing homology to the plasmid, are amplified in the first set of two PCR reactions. These PCR products are used in a second PCR reaction with plasmid as a template and end primers, whereby the PCR products align with the plasmid due to overlapping ends the tagging cassette containing the tag, terminator sequence and selection marker, with approximately 250 bp of homology to the target gene is amplified.

Reagent	Quantity (for 25 µl)
ddH ₂ O	to 25 µL
5x Q5 buffer	5 µL
100 mM dNTP	0.2 µL
10 µM forward primer	1.25 µL
10 µM reverse primer	1.25 µL
Q5 polymerase	0.25 µL
Template	Variable: Plasmid – 50 ng Genomic DNA – 5 ng

Table 2.1 – **Q5 DNA polymerase PCR reaction set up.**

Cycles	Temp	Time
1	98°C	30 s
35	98°C	10 s
	Dependent upon primer T _m	30 s
	72°C	20-30 s / kb
1	72°C	2 min

Table 2.2 – **Thermal cycler programme set up.**

2.9. CRISPR/Cas9 gene editing in *S. pombe*

Type II CRISPR/Cas9 gene editing has only recently been implemented in *S. pombe* (Jacobs *et al.*, 2014). The system utilises a pMZ374 plasmid (obtained from Addgene plasmid database, plasmid map shown in the Appendix 8.4) expressing the Cas9 enzyme and the gRNA in which a 20-nucleotide gene-specific targeting sequence can be introduced; as well as a homologous recombination template, which incorporates a desired mutation or an epitope tag to a gene of interest. Upon their simultaneous transformation into an *S. pombe* strain, gRNA and Cas9 form a complex which binds to the desired site based on complementary base pairing. Cas9 then introduces a double strand break, repaired by the supplied homologous recombination (HR) template, creating desired changes. The approach was used to generate *ade6* and *ade2* NMD reporter genes by introducing PTCs along the gene sequence, as well as to tag *RpS2801* gene with 3xHA.

Detailed methodology is described in Chapter 6 and Appendix 8.6. In short, gene-specific gRNA sequence was designed by selecting 20 nucleotides upstream of the protospacer adjacent motif (PAM – 5'-NGG-3' sequence), which ideally should be within 3-5 nucleotides from the desired mutagenesis site. PCR approach and Gibson Assembly Master Mix (New England Biolabs) were used to clone the gene specific gRNA sequence into its placeholder on the pMZ374 plasmid. The placeholder contains *rrk1* gRNA promoter as well as additional sequences, which are necessary for gRNA expression and for its binding to the Cas9 enzyme. Homologous recombination (HR) templates were created using overlapping PCR.

The gRNA- and Cas9-expressing plasmid and the HR template were transformed simultaneously into a desired strain (a uracil auxotroph, with a deletion in the *ura4* gene, *ura4-D18*) and correct transformants were selected on EMMG uracil drop-out plates, as the plasmid contains a functional *ura4* gene and therefore confers uracil prototrophy (Grimm *et al.*, 1988) (section 2.12). Ade6 mutant strains produce pink colonies when concentration of adenine in the medium is low (Fisher, 1969). Therefore, CRISPR/Cas9 mediated *ade6* mutagenesis was examined by replica plating the cells onto EMMG uracil drop-out plates with low adenine concentration (10 mg/L adenine) and confirmed by sequencing of resulting pink colonies. Sequencing was used to screen *ade2* mutagenesis, as *ade2* mutants are phenotypically undistinguishable from normal cells. CRISPR/Cas9 mediated protein tagging was examined by diagnostic colony PCR reactions (section 2.13) and confirmed via SDS-PAGE/Western blotting (section 2.15).

2.10. pREP41 expression vector

An *S. pombe* shuttle pREP41 expression vector was used for regulated episomal expression of a gene of interest. The plasmid contains a multiple cloning site in which a desired gene can be inserted. For example, the gene of interest can encode the green fluorescent protein (GFP), in which case the vector is as represented in Appendix 8.4. The plasmid contains a medium strength *no message in thiamine* (*nmt41*) promoter that allows regulated expression of the gene (Matsuyama *et al.*, 2004; Maundrell, 1990). The gene is transcribed at a moderate level in the absence of thiamine in the medium.

Therefore, the gene is expressed solely when the plasmid containing strain is grown in EMM since YES contains enough thiamine to block its expression.

The plasmid confers leucine prototrophy as it contains a functional LEU2 gene, originating from *S. cerevisiae*, encoding an enzyme involved in leucine biosynthesis pathway. Therefore, this plasmid is used to transform *S. pombe* strains that contain a mutation in the *leu1* gene (*leu1-32* variant) since LEU2 successfully complements *leu1-32* mutation (Beach and Nurse, 1981). As parental strains cannot synthesise leucine, successfully transformed strains are selected on EMM leucine drop-out plates, on which the parental strains cannot grow.

The vector was used for episomal expression of NMD reporter genes in desired *S. pombe* strains. A functional GFP gene served as a control while a GFP gene containing a premature translation termination codon (PTC) at a specific position was used as an NMD reporter gene. Nonsense mutation (TAA triplet) was introduced using overlapping PCR, either at codon position 6, 27 or 140. The gene was then cloned into the plasmid multiple cloning site. These plasmids were obtained from Brogna lab plasmid stocks and designed by a Brogna lab member (Jikai Wen PhD Thesis). Yeast strains were transformed with the plasmids using the described LiAc method (section 2.12). The strains were used to study NMD by means of Northern blotting (section 2.16) of either whole cell culture steady state RNA or RNA from polysome fractionation samples (section 2.14).

2.11. Agarose gel electrophoresis and DNA gel extraction

DNA and RNA samples were resolved in agarose gels to allow separation according to molecular weight. Gels were made by dissolving the appropriate quantity of agarose for the % of gel required in TAE buffer (40 mM Tris base, 40 mM acetic acid and 2 mM EDTA). Typically, 1% agarose gels were used. Ethidium bromide was added just prior to casting the gel to a concentration of 0.5 µg/mL. Samples were prepared by dissolving 1 µL of a DNA or an RNA sample in 9 µL of dH₂O and 2 µL of 6x purple gel loading dye without SDS (New England Biolabs). Upon sample loading, gels were run at 90 V for approximately 30 minutes in TAE buffer.

DNA was extracted from the gel using either QIAquick gel extraction kit (QIAGEN) or Monarch DNA Gel Extraction Kit (New England Biolabs), according to manufacturer's instructions. In short, DNA was extracted from a gel by cutting the desired band out under a UV light and melting the gel slice in three times the volume of Dissolving buffer at 55°C. The sample was then loaded into a column placed into a collection tube. After one-minute centrifugation at 13000 rpm at RT the sample was washed twice in Wash buffer containing ethanol. The column was transferred into a 1.5 mL microcentrifuge tube and eluted with 20-50 µL of sterile dH₂O. Alternatively, if only specific PCR amplification occurred and no non-specific bands were observed on the gel, they were purified using Monarch DNA Purification Kit (New England Biolabs) instead. Similarly, the DNA sample was dissolved in Binding buffer, loaded onto a purification column and then washed and eluted, as previously described. The

concentration was estimated using a NanoDrop spectrophotometer (ND-1000 NanoDrop) and agarose gel electrophoresis.

2.12. Transformation of *S. pombe* strains

The strain to be transformed was grown in 10 mL of YES medium until its OD₆₀₀ was between 0.4 and 0.6 units (mid-log phase). An aliquot equivalent to an OD₆₀₀ of 0.1 was then centrifuged and the pellet was diluted in 10 mL of MB medium, after which the cells were grown overnight until the OD₆₀₀ again reached 0.5 units. Growth in MB medium together with the lithium acetate method was used to make the cells competent for transformation. The cells were collected by spinning at 3000 rpm for 5 minutes at room temperature, washed in 10 mL of sterile water and collected again by centrifugation at 3000 rpm. The pellet was then resuspended in 1 mL of sterile dH₂O, transferred into a sterile 1.5 mL microcentrifuge tube and spun at 6000 rpm for 10 seconds. The subsequent pellet was resuspended in 200 µL of lithium acetate (LiAc) buffer (100 mM Lithium acetate pH 4.9, 10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0) and the mixture was centrifuged again at 6000 rpm for 10 seconds. The cells were resuspended in 50 µL of LiAc buffer to which 240 µL of 50% PEG₃₃₅₀ solution was added. It is imperative that the PEG₃₃₅₀ solution is prepared immediately prior to transformation. Subsequently, 60 µL of LiAc buffer was added to the mixture, followed by 2 µL of 10mg/mL of ssDNA (previously boiled for 5 minutes and cooled on ice) and 1 µg of plasmid or 5 µg of linear DNA (such as a gene targeting cassette), or both (in case of CRISPR-Cas9 mutagenesis). The sample was thoroughly vortexed and then incubated at 30°C for 30 min at 160 rpm in a shaking incubator. The incubation was followed by a

15-minute heat shock at 42°C after which the cells were pelleted at 6000 rpm for 1 min. The pellet was resuspended in 500 µL of TE buffer (10 mM Tris-Cl, pH 7.5, 1 mM EDTA) and, in case of antibiotic resistance selection, 100 µL of the sample was plated onto a YES plate and grown overnight at 30°C. The following day cells were replica plated onto a selective plate containing the antibiotic. In case the plasmid or the DNA fragment used for transformation conferred leucine or uracil prototrophy, the cells were plated directly onto a selective EMM leucine or uracil drop-out plate, respectively. After plating, the strains were grown at 32 °C until colonies appeared, which would typically occur after 2-3 days. Correctly transformed strains were screened by colony PCR using an appropriate set of primers.

2.13. Colony PCR

Colony PCR reactions were prepared using GoTaq DNA polymerase (Promega) as shown in Table 2.3. Single colonies were picked using a pipette tip and resuspended in each individual reaction tube. Thermal cycler programme was carried out as demonstrated in Table 2.4.

Reagent	Quantity (for 25 µl)
ddH ₂ O	16.175 µL
5x GoTaq buffer	5 µL
25 mM MgCl	1.5 µL
100 mM dNTP	0.2 µL
10 µM forward primer	1 µL
10 µM reverse primer	1 µL
GoTaq polymerase	0.125 µL
Template	Colony

Table 2.3 – **Colony PCR reaction set up.**

Cycles	Temp	Time
1	95°C	10 min
35	95°C	30 s
	Dependent upon primer T _m	30 s
	72°C	1 min / kb
1	72°C	5 min

Table 2.4 – **Thermal cycler programme set up.**

2.14. Polysome profiling

2.14.1. Preparation of sucrose gradients

All solutions were prepared in diethylpyrocarbonate (DEPC)-treated, RNase free dH₂O. DEPC-treated dH₂O is prepared by adding DEPC to dH₂O to a concentration of 0.1% after which the mixture is incubated overnight and autoclaved at 120°C for 30 minutes the following day. All the steps were carried out at 4°C or on ice to prevent RNA degradation. Sucrose gradients (10-50%) were prepared by mixing a 10% and 50% sucrose solution (prepared in 10 mM Tris acetate pH 7.4, 70 mM ammonium acetate, 4 mM magnesium acetate, 100 µg/mL cycloheximide), using a standard Gradient Mixer. Gradients were directly dispensed into SW41 rotor tubes. Alternatively, the gradients were prepared using a Biocomp Gradient Master Base Unit (Wolf Laboratories), according to manufacturer's instructions. In short, the two end solutions were first degassed by filtration and then layered directly into the SW41 centrifuge tube. The 10% solution was firstly dispensed using a thin cannula attached to a syringe until the tube was half-full. A clean cannula was attached to a fresh syringe, which was then filled with the heavy 50% solution and the cannula was inserted to the very bottom of the tube. The 50% solution was dispensed slowly, to minimise mixing, until the half-full mark. The tube was then capped and placed into a magnetic tube holder of the machine. The gradients were then made by tilted tube rotation, using the 1-step long 10-50% sucrose pre-set programme of the machine. Prior to usage, the gradients were stored at 4°C and just before loading the samples for ultracentrifugation, the gradients were further chilled on ice.

2.14.2. Cell lysis

An *S. pombe* 50 mL cell culture was grown until its OD₆₀₀ reached 0.5 units (mid-log phase). The cells were then treated with 100 µg/mL of cycloheximide for 5 minutes at 30°C and the cultures were immediately poured into 50 mL centrifuge tubes containing ice, to instantly stop the cell metabolism. The cells were harvested by centrifugation at 3000 rpm and washed once in wash buffer (20 mM HEPES pH 7.4, 2 mM magnesium acetate, 2 mM magnesium chloride, 100 mM potassium acetate, 100 µg/mL cycloheximide). The sample was pelleted again and resuspended in 600 µL of lysis buffer (wash buffer with added 1 mM dithiothreitol (DTT), 1 mM phenylmethanesulfonyl fluoride (PMSF), 1x cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail (Roche) and either 10 mM Ribonucleoside Vanadyl Complex (RVC) (New England Biolabs) or 20 units of RiboLock (Thermo Fisher Scientific) as an RNase inhibitor). The suspension was transferred into a pre-cooled 2 mL screw cap tube containing approximately 1 mL of either acid-washed glass or Zirconia silica beads (0.5 mm diameter, Biospec products, distributed by Cole-Palmer). Cells were lysed in a tissue homogenizer (Precellys 24, Bertin instruments) using three 15-second beads shaking pulses at 5500 rpm, with 2-minute cooling intervals on ice. The bottom side of the 2 mL screw-cap tube was then pierced three times using a 25G needle. The screw-cap tube was fitted into a 5 mL syringe placed in a 15 mL tube and centrifuged at 3000 rpm for 5 minutes at 4°C to recover the cell lysate. The lysate was subsequently transferred into an RNase free 1.5 mL microcentrifuge tube and cleared by centrifugation at 13000 rpm for 5 minutes at 4°C. The supernatant was moved into a clean RNase free 1.5 mL tube and

the process of clearing at 13000 rpm for 10 minutes at 4°C was repeated two more times. If RVC was used as an RNase inhibitor, 10 µL of the cleared lysate underwent phenol extraction to remove it as it affects absorbance reading. The absorbance at 260 nm (abs_{260}) was subsequently measured using a NanoDrop spectrophotometer (ND-1000 NanoDrop), as abs_{260} of a sample is proportional to its RNA concentration, only if the ratio of OD_{260} to OD_{280} is above 1.7 units.

In case of RNase treatment to the lysate, RNase inhibitor was omitted from the lysis buffer and 10 µL of 1 mg/mL RNase was added to the cleared lysate and incubated on ice for 30 minutes prior to loading onto the sucrose gradient. In case of EDTA treatment, 30 mM EDTA was added to the cleared lysate, incubated on ice for 30 minutes and loaded onto the sucrose gradient, which also contained 30 mM EDTA. Finally, in case of puromycin treatment, cycloheximide was added solely to the cell culture, 100 µg/mL of puromycin was added to the cleared lysate together with potassium chloride to a concentration of 375 mM and incubated at room temperature for 20 minutes. The sample was loaded onto a sucrose gradient of the following composition: 20 mM Tris acetate, 70 mM ammonium acetate, 375 mM potassium chloride. Puromycin treatment can also be performed in cells by its addition to a cell culture to a concentration of 250 µg/mL, incubation at 30°C for 20 minutes and subsequent addition to the wash and lysis buffer, whilst cycloheximide is omitted from the whole procedure. Again, 375 mM potassium chloride was added to the lysate and the sample was run on a sucrose gradient lacking magnesium (20 mM Tris acetate, 70 mM ammonium acetate, 375 mM potassium chloride).

2.14.3. Sedimentation of translation complexes

The tubes containing sucrose gradients were first carefully balanced, upon which a volume corresponding to 20 abs_{260} units of each cell lysate was loaded onto each gradient. The tubes were placed into the buckets of a Beckman SW41 rotor and centrifuged at 38000 rpm for 150 minutes at 4°C.

After centrifugation, the sample was retrieved from the bottom of the tube using a steel capillary attached to tubing (previously cleaned of any RNase traces with 0.2 M NaOH and rinsed with DEPC dH_2O), by action of a peristaltic pump (P-1, Pharmacia). The speed of the pump was set to 1.2 mL/min. The absorbance at 254 nm was continually measured using a flow-through UV spectrophotometer (ISCO UA-6 detector) and recorded either on a chart recorder (Pharmacia Biotech LKB 70 REC 102, speed setting: 1 cm/min) or using an N56FU Digital Multimeter or PicoLog 1000 series data logger with appropriate software to record the data. Fractions (0.8-0.9 mL per tube, 12 fractions) were collected using a fraction collector (FRAC100, Pharmacia). The samples were stored at -20°C for up to one year.

2.14.4. RNA and protein extraction from polysomal fractions

Phenol/chloroform method was used to extract RNA from polysomal fractions. A 150 μL aliquot of each fraction was transferred into an RNase free 1.5 mL tube to which 450 μL of RNA extraction buffer (20 mM Tris pH 7.5, 10 mM EDTA, 0.5% SDS, all RNase free) was added. The suspension was mixed either with 600 μL of phenol pH 4.2 or with 600 μL of phenol pH 7.9 (Thermo Fisher Scientific) (to obtain higher RNA yield), vortexed thoroughly until it appeared cloudy white and incubated for 30 minutes at 65°C, with occasional vortexing.

Sample was then chilled on ice for 5 minutes and centrifuged at 13000 rpm for 10 minutes at 4°C. Aqueous phase was transferred to a clean 1.5 mL microcentrifuge tube into which 600 µL of phenol: chloroform: isoamyl alcohol (25:24:1) was added. After thorough vortexing, the centrifugation step was repeated. Aqueous phase was transferred into a clean 1.5 mL microcentrifuge tube and mixed thoroughly with 600 µL of chloroform: isoamyl alcohol (24:1), followed by centrifugation at 13000 rpm for 10 minutes at 4°C. Top phase of the last centrifugation step was transferred into a clean RNase free 1.5 mL microcentrifuge tube to which 55 µL (1/10 sample volume) of 3M sodium acetate pH 5.2 and 1 mL of 100% ethanol were added. To visualise the RNA pellet better, 20 µg of glycogen (Sigma) was added to each sample. The mix was incubated for 30 minutes at -20°C for RNA to precipitate and then centrifuged at 13000 rpm for 15 minutes at 4°C. The pellet was washed with RNase free 70% ethanol, which removes salt excess. RNA samples were air dried and dissolved in 20 µL of RNase free, DEPC treated dH₂O.

Methanol/chloroform method was used to precipitate the proteins from polysomal fractions. To 150 µL of a fraction 600 µL of methanol was added. After thorough mixing by inverting the tube three times, 150 µL of chloroform was added to the suspension, which was then vortexed. Finally, the solution was mixed with 450 µL of dH₂O by vortexing, until it appeared cloudy white. Sample was then immediately centrifuged at 13000 rpm for 5 minutes at RT. A white disc of proteins would form between the bottom organic and the upper aqueous phase. The top aqueous layer was discarded, whilst not disturbing the interphase, and 650 µL of methanol was added to the sample. The tube was

inverted three times to wash the proteins that detach from the interphase and then centrifuged at 13000 rpm for 5 minutes at RT. After centrifugation, all liquid was removed from the tube using a Pasteur pipette and the pellet was air-dried. Precipitated proteins were dissolved in 35 μ L of SDS loading buffer and boiled for 5 minutes. Protein samples were run on a 12% SDS polyacrylamide gel and analysed by Western blotting.

2.15. SDS-PAGE and Western blotting

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels were prepared as described in Molecular Cloning, 4th edition (Sambrook and Green, 2012). Typically, 12% gels were used. Tris-tricine gel electrophoresis and non-urea peptide separations (Bonafacino *et al.*, 2011), specialised for analysing small molecular weight proteins, were used for RpS2801-HA detection, as this protein is solely 7.66 kDa of size. The gels were run in SDS running buffer (25 mM Tris-HCl, 190 mM Glycine, 0.1% SDS, pH 8.3) at 100 V for 90 minutes at room temperature.

The proteins were transferred onto a nitrocellulose membrane (Schleicher & Schuell) using a standard wet electroblot apparatus (Mini Trans-Blot Electrophoretic Transfer Cell, Bio-Rad). The transfer was carried out in transfer buffer (25 mM Tris-HCl, 190 mM Glycine, 20% Methanol) at 150 V for 120 minutes at 4°C.

Blocking was carried out for 1 hour at RT in 10 mL of TBS (25 mM Tris-HCl, 137 mM NaCl) 0.1% Tween 20 (TBST) solution containing 5 % of milk. Antibodies used in this study are mouse anti-puromycin clone 5B12, rat anti-HA clone 3F10, polyclonal rabbit anti-HA, mouse M2 anti-FLAG and mouse B-2 anti-GFP

antibody. The membrane was incubated overnight at 4°C with typically 1:5000 dilution of primary antibody prepared in TBST. The following day, three 10-minute membrane washes were first carried out in TBST. The membrane was then incubated for 1 hour at RT with a 1:10000 TBST dilution of the secondary horseradish peroxidase (HRP) conjugated antibody (Sigma). The secondary antibody incubation was followed by three 10-minute membrane washes in TBST. ECL Western Blotting Detection Reagent (GE Healthcare) was used to develop the signal, according to manufacturer's instructions. The membrane was incubated with the substrate solution for 3 minutes after which it was carefully removed and the membrane wrapped in cling film. The signal was then visualized using a ChemiDoc imaging system and GeneSnap software.

For secondary antibodies conjugated with a fluorescent dye, the procedure was carried out as previously explained with the exception of blocking being dissolved in TBS, and TBST used in the washing steps being supplemented with sodium chloride to the final concentration of 500 mM. The secondary antibody incubation was followed by five 5-minute membrane washes in TBST while the last wash prior to detection was carried out in TBS, which eliminates any residual Tween 20 that could interfere with the detection. Additionally, secondary antibody incubation was performed in the dark and the signal was developed immediately after the last wash, using the appropriate developing software (Odyssey Imaging Software, LI-COR).

2.16. Northern blotting

2.16.1. Gel electrophoresis and blot set-up

All the solutions were either made in DEPC-treated dH₂O or were prepared in dH₂O and then subsequently treated with DEPC. A 1% agarose formaldehyde gel was prepared as shown in Table 2.5.

	Medium sized gel tank (total vol of gel =120 mL)	Large sized gel tank (total vol of gel =300 mL)
Agarose	1.2 g	3 g
10xMOPS	12 mL	60 mL
Formaldehyde	21.45 mL	53.6 mL
DEPC H ₂ O	86.25 mL	185.6 mL

Table 2.5 – **Agarose formaldehyde gel composition for Northern blot analysis.**

Samples to be analysed by Northern blotting were prepared in RNase free 1.5 mL microcentrifuge tubes. RNA concentration in the samples was estimated by measuring absorbance at 260 nm using a NanoDrop spectrophotometer. A volume corresponding to an RNA quantity of 5 µg was added to a solution of 23% formaldehyde, 50% formamide and 1x MOPS buffer (0.2 M MOPS, 20 mM sodium acetate, 10 mM EDTA pH 8.0), adjusted with DEPC treated dH₂O to the final volume of 30 µL. The sample was then incubated for 15 minutes at 65°C after which it was immediately incubated 5 minutes on ice. The gel was pre-run at 80 V in MOPS buffer while the sample was on ice. Prior to loading the sample onto the gel, 1 µL of gel loading dye (2.5 mg/mL bromophenol blue, 2.5 mg/mL xylene cyanol, 39% glycerol, 10 mM EDTA pH 8.0) was added to it and the gel was typically run at 80 V for approximately 2 hours.

After the run was completed, the gel underwent two 20-minute washes in DEPC treated dH₂O followed by a 30-minute wash in 20x SSC buffer (3 M sodium chloride, 300 mM sodium citrate). The blot apparatus was cleaned using 0.2 M NaOH and hot water and rinsed with DEPC treated dH₂O. The blotting apparatus was set up while the washes were carried out. The tray was filled with 20x SSC buffer, on top of which a glass plate was placed. Two pieces of 3 mm filter paper exactly the size of the apparatus were positioned on top of the plate with their ends dipping into the tray with the buffer, serving as a bridge. The bridge was covered with 20x SSC and a glass test tube was used to flatten the surface and remove any air bubbles. The gel was taken out of the 20x SSC solution after the last washing step was completed, and it was cut above the wells into which the samples were loaded. The gel was carefully transferred onto the bridge and any air bubbles between the bridge and the gel were gently removed using a glass test tube. A neutral nylon Amersham Hybond-N membrane (GE Healthcare) was slowly placed on top of the gel followed by two 3 mm filter papers soaked in 20x SSC buffer. The blot was then sealed on all sides using cling film and five dry filter papers were positioned on top, followed by a stack of paper towels. Finally, another glass plate was put on top and the blotting was carried out overnight.

2.16.2. RNA UV crosslinking and hybridisation with the probe

After blotting completion, the membrane was placed on a cling film with the RNA side positioned upwards. The top left corner of the membrane was cut to label membrane orientation, while the date and sample details were written on the top right corner for identification. The membrane was then placed onto a

filter paper, which was previously made wet by soaking in DEPC treated dH₂O. RNA was cross-linked to the membrane by UV light (254 nm) exposure at 120 J. It is important for the membrane to be kept wet, but not exceedingly, during UV crosslinking. The membrane was then washed in DEPC dH₂O for 5 minutes and stained in fresh 0.3% methylene blue solution, prepared in DEPC dH₂O, for 5 minutes, to assess RNA loading and quality. After 5-minute incubation with methylene blue, the membrane was de-stained in DEPC dH₂O until the bands were clearly visualised. Methylene blue staining was imaged using a gel doc system and the Gene Snap software.

Approximately 1 mL of 10 mg/mL ssDNA was boiled for 5 minutes and then immediately placed on ice at least 5 minutes prior to usage. Membrane was positioned, with RNA facing inwards, into a glass hybridisation tube to which 30 mL of Hybsol solution (1.5x SSPE solution (225 mM NaCl, 15 mM NaH₂PO₄·xH₂O, 15 mM Na₂EDTA), 7% SDS, 10% PEG₈₀₀₀) was added. With the addition of 250 µg/mL of heparin and 100 µg/mL of ssDNA, pre-hybridisation was immediately carried out for 4 hours at 68°C.

2.16.3. Random priming – probe synthesis and purification

The probe was synthesised by means of a random priming reaction (Table 2.6).

Probe	50 ng
5x labelling buffer (Promega) containing random primers	5 µL
Sterile dH ₂ O added to a total volume of 25 µL	
The above reaction was boiled for 3 minutes (denaturation) and cooled to RT (primer annealing step)	
The reaction was cooled on ice for 5 minutes and the following was added to it:	
dNTP (A ⁻) mix, prepared as follows: 0.5 µl (100 mM dCTP) + 0.5 µl (100 mM dGTP) + 0.5 µl (100 mM dTTP) + 98.5 µl sterile ddH ₂ O	1 µL
Klenow fragment (NEB), 5000 U/mL	0.5 µL
³² P dATP	2.5 µL
The mix was incubated either overnight at room temperature or 2 hours at 37°C.	

Table 2.6 – **Random priming reaction set-up.**

Probe was then purified using a MiniQuick Spin DNA Column (Roche), according to manufacturer's instructions. In short, the column matrix was first resuspended evenly by vortexing the column at low speed for 3-5 seconds. The top cap was first removed from the column after which the bottom was snapped off. The column was spun at 1000x g (3200 rpm in an Eppendorf bench top 5415 R centrifuge) for 1 minute at RT, to pack the column and to remove any residual buffer. After centrifugation, the column was used immediately, as any delay would cause the column matrix to dry out, reducing its performance. The column was placed into a clean 2 mL screw cap tube after which the sample was carefully applied to the centre of the column bed. The sample must not be applied on the side, as it would bypass the separating matrix. The tube was

immediately centrifuged at 1000x g for 4 minutes at RT. The recovered purified probe was subsequently boiled for 5 minutes, during which time the pre-hybridisation solution was removed and replaced with fresh 20 mL of Hybsol solution containing 250 µg/mL of heparin and 100 µg/mL of ssDNA. With the addition of the probe to the bottle, the hybridisation was carried out overnight at 68°C.

2.16.4. Membrane washes and exposure

The membrane was washed 4 times in 100 mL of 2x SSC and 0.1% SDS solution at 68°C. The first two washes were carried out for 2 and 5 minutes, while the last two were 30 minutes each. The membrane was then washed once again in 100 mL of 0.2x SSC and 0.1% SDS at 68°C for 30 minutes. Finally, it was wrapped in cling film and the radioactivity was measured using a Geiger counter. If the measured radioactivity was below 20 bpm, the exposure onto an intensifying screen was done overnight in an X-ray cassette. However, if detected radioactivity was above 20 bpm, the exposure was performed for 4 hours. The signal was developed using Molecular Imager and Quantity One software (Bio-Rad).

2.17. Protein affinity purification from polysomal fractions

Protein affinity purification was attempted from polysomal fractions of UPF1-FLAG and eIF4AIII-FLAG tagged strains using two protocols.

According to the first protocol, 50 μ L of magnetic Dynabeads was aliquoted into a 1.5 mL microcentrifuge tube. The beads were washed twice on a magnetic rack and then resuspended in 200 μ L of PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 , 1.8 mM KH_2PO_4 , pH 7.4). Subsequently, 5 μ L of anti-FLAG antibody was added to the beads and the beads were then coated with the antibody by a 10-min incubation at RT, with gentle rotation (Intelli-Mixer Rotator Mixer). The beads were placed on a magnetic rack, the solution was removed and the beads were resuspended in 50 μ L of PBS and stored on ice until required. Soluble fractions from two polysomes were then pooled together and diluted 5x in lysis buffer (20 mM HEPES, pH 7.4, 110 mM KOAc, 0.5% Triton X-100, 0.1% Tween 20, 10 mM MnCl_2 , 1x cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail (Roche), 50 units/mL RiboLock (Thermo Fisher Scientific)) to which Dyna beads, pre-coated with anti-FLAG antibody, were added. Sample was incubated for one hour at 4°C, with gentle rotation. The unbound fraction was separated from the beads on a magnetic rack and the beads were washed 6 times in 1 mL wash buffer (lysis buffer, supplemented with 4 mM MgCl_2 and 50 units/mL RiboLock (Thermo Fisher Scientific)), 10 minutes each time, at 4°C, with rotation. The proteins were eluted from the beads by boiling the sample in 50 μ L of SDS loading buffer.

Alternatively, soluble fractions from two polysomes were pooled together, to which 5 μ L of anti-FLAG antibody was then added. The suspension was

incubated overnight at 4°C, with gentle agitation. The following day, upon addition of 20 µL of Dyna beads, the sample was incubated for two hours at 4°C, again with gentle agitation. The unbound fraction was separated from the beads using a magnetic rack and the beads were washed 4 times in DEPC-treated, RNase free wash buffer (20 mM Tris pH 7.4, 100 mM NaCl, 1% Triton X-100, 0.1% SDS, 1x cOmplete™ Mini EDTA-free Protease Inhibitor Cocktail (Roche), 50 units/mL RiboLock (Thermo Fisher Scientific)). The proteins were eluted from the beads by boiling the sample in 50 µL of SDS loading buffer.

The quality of affinity purification was determined by silver staining of the elution fraction and Western blotting of input, unbound and elution fraction aliquots, taken throughout the procedure.

2.18. Silver staining and mass spectrometry analysis of stained proteins

Proteins to be identified by mass spectrometry (affinity purification elution or polysome profiling samples) were run on a 12% SDS-PAGE gel and stained using a Pierce Silver Stain kit (Thermo Scientific). The resolving gel was set overnight to ensure complete polymerisation and prevent any interference of acrylamide monomers with mass spectrometry analysis. All the steps were carried out in a single clean plastic tray, with gentle agitation, at room temperature, in 25 mL of each solution used. Gloves were used throughout, to prevent contamination, primarily with keratin. The gel was first washed in ultrapure dH₂O twice, 5 minute each, followed by a 15-minute fixation in Fixing solution (30% ethanol, 10% acetic acid). The solution was replaced, and the gel was fixed again for 15 minutes. Gel was then washed twice in 10% ethanol and

twice in ultrapure dH₂O, 5 minutes each. Just before use, sensitiser was prepared by addition of 50 µL of Sensitizer to 25 mL of ultrapure dH₂O. The gel was sensitised for 1 minute and then washed twice in ultrapure dH₂O, 1 minute each wash. The stain was prepared by adding 250 µL of Silver Stain Enhancer into 25 mL of Silver Stain. The gel was immediately incubated in the stain for 15 minutes, quickly washed twice in ultrapure dH₂O, 20 seconds each, after which the developer solution (prepared just before usage by dissolving 250 µL of Silver Stain Enhancer in 25 mL of Silver Stain Developer) was added. The gel was incubated in the developer until the bands were clearly distinguishable. The gel was washed briefly with Stop Solution (5% acetic acid) after which the solution was replaced with fresh 5% acetic acid and stored at 4°C. Alternatively, gel was washed twice in ultrapure dH₂O, 10 minutes each, and the desired bands were excised from the gel using a clean scalpel and placed into 1.5 mL microcentrifuge tubes, creating gel plugs. Trypsin digestion and mass spectrometry analysis was carried out by the Proteomics facility at the School of Biosciences of the University of Birmingham.

2.19. GFP fluorescence imaging

A 1 mL UPF1-GFP tagged cell culture was grown to an OD₆₀₀ of 0.5. The cells were fixed by addition of formaldehyde to a final concentration of 3.7% and a 30-minute incubation in a 30°C shaking incubator at 160 rpm. The cells were then pelleted at 3000 rpm, for 5 minutes at RT. The pellet was washed three times in 1 mL of PBS, pelleting the cells between washes, at 3000 rpm, 5 minutes each, at RT. Cell pellet was resuspended in 10-50 µL of PBS, depending on its size. A microscope slide was wiped using 70% ethanol, onto which 5 µL of cell suspension was added. A smear of cells was created, by dragging a coverslip across the cells. While the cells were drying, 5 µL of 80% glycerol containing 0.1 µg/ml DAPI stain was pipetted onto a clean coverslip. The slide was inverted and carefully lowered onto the coverslip, making sure that it adheres well to the slide. The slides were imaged using a Nikon Eclipse Ti epifluorescence microscope; with DAPI staining imaging the nuclei in blue while the green signal demonstrated UPF1 distribution in the cells.

2.20. *Escherichia coli* methods

E. coli was used for cloning and plasmid preparation. The plasmids used in the project as well as related plasmid maps are presented in the Appendices 8.3 and 8.4. The plasmids were retrieved from Brogna lab stocks, acquired from addgene plasmid database or generated via Gibson Assembly Cloning.

2.20.1. *E. coli* strains

E. coli strain used for plasmid preparation was XL1-Blue of the genotype recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIq ZΔM15 Tn10 (Tetr)]. NEB® 5-alpha Competent (High efficiency) *E. coli* cells, a derivative of DH5α strain, were used for cloning using Gibson Assembly Cloning kit.

2.20.2. *E. coli* media and culturing

Protocols to prepare all the media used for culturing are detailed in the Appendix 8.2. Typically, *E. coli* cells were grown in LB liquid and solid media. A liquid culture was grown overnight at 37°C, by resuspending a single colony in 5 mL of LB media. The subsequent culture was then scaled up to the desired volume and grown further to a required OD₆₀₀. Plated cells were grown overnight on LB agar plates in a 37°C incubator until colonies appeared. In case ampicillin resistance selection had to be performed, ampicillin would be added to LB medium to a concentration of 100 µg/mL. Also, upon transformation, cells were recovered in either SOC or NZY medium, as explained in section 2.20.4.

2.20.3. Gibson Assembly Cloning

Gibson assembly reaction enables assembly of multiple overlapping DNA fragments into a single DNA molecule. It contains three key enzymes – an exonuclease, which digests DNA 5' ends and creates 3' single-stranded overhangs that allow annealing of fragments that share complementarity in that region; DNA polymerase that extends the 3' ends of annealed fragments, and DNA ligase, which seals the nicks. In this way either multiple linear DNA fragments can be assembled, or a desired DNA sequence can be cloned into a plasmid vector.

Cloning using Gibson assembly was carried out to introduce a target-specific gRNA sequence into a pMZ374 vector, as outlined in section 2.9 and detailed in Chapter 6. PCR insert containing gene-specific gRNA was generated using overlapping PCR. Plasmid backbone was created by restriction enzyme digestion using *SpeI* and *StuI* enzymes (New England Biolabs), as explained in section 2.20.6. As these DNA fragments possess overlapping ends approximately 100 bp of length, Gibson assembly was set up on ice, by combining 3 μ L of 2x Gibson Assembly Master Mix, 0.02-0.5 pmols of each DNA fragment and deionized dH₂O up to 6 μ L. The sample was incubated 15 minutes at 50°C in a thermocycler and then either stored at -20°C or used immediately to transform NEB 5-alpha competent *E. coli* cells.

2.20.4. *E. coli* transformation

After the chemically competent cells were thawed on ice, 1 μ L of chilled Gibson assembled DNA product was added to the cells and mixed gently by pipetting up and down. The mixture was incubated on ice for 30 minutes, which was followed by heat shock at 42°C for 30 seconds. The sample was then immediately placed on ice. After a 2-minute incubation, 950 μ L of room temperature SOC media was added to the cells, followed by an hour incubation at 37°C, in a shaking incubator, at 200 rpm. Finally, 100 μ L of the cell suspension was plated on selective LB plates, typically containing ampicillin to a concentration of 100 μ g/mL, and grown overnight until colonies appeared.

In case of purchased plasmids (Addgene) or ones retrieved from Brogna lab glycerol stocks, their amplification was carried out by transforming 50 ng of the plasmid into 50 μ L of XL-1 Blue competent *E. coli* cells. The mixture was incubated on ice for 30 minutes, the cells were heat-shocked at 42°C for 90 seconds and then incubated on ice for 2 minutes. 450 μ L of NZY media was added to the cells, followed by an hour incubation at 37°C. One tenth of the cells were plated onto selective LB ampicillin plates.

2.20.5. Large scale plasmid preparation

E. coli transformants were grown overnight at 37°C, by resuspending a single colony into 5 mL of LB media. This culture was either used to prepare small amount of plasmid DNA using mini plasmid preparation or diluted and grown further in 50-100 mL for midi or maxi preparations. Typically, Invitrogen or QIAGEN kits were used, according to the instructions. In short, plasmid preparations involved equilibration of the filtration cartridge, cell harvesting and

resuspension, RNase digestion and lysis. After precipitation, the lysate was transferred into the column and clarified by draining it through the column by gravity flow. After several washing steps, the DNA was eluted from the column, precipitated using 100% isopropanol and washed with 70% ethanol. Air-dried DNA pellet was resuspended in 200 µL of TE buffer and the DNA concentration was estimated using a Nano Drop spectrophotometer, while the quality was checked by restriction enzyme digestion and agarose gel electrophoresis of the resulting fragments.

2.20.6. Restriction enzyme digestion

Restriction enzymes used either to check quality of prepared plasmids or to generate fragments for Gibson Assembly cloning were obtained from New England Biolabs. The reactions were set up as shown in Table 2.7. The buffer was chosen based on company's recommendation, and it depended on the enzyme used. In case of double digestion reactions, 1 µL of each enzyme was added into the reaction while the buffer was selected based on the Double Digest Finder tool on the NEB website. Upon set-up, restriction enzyme reaction was incubated at 37°C, for an hour.

Reagent	Quantity (for 50 µl)
ddH ₂ O	to 50 µL
10x NEBuffer	5 µL
Plasmid DNA	1 µg
Restriction Enzyme	1 µL (10 units)

Table 2.7. – **Restriction enzyme digest set up.**

RESULTS

Chapter 3: Re-evaluating methods of studying translation

3.1 Introduction

Translating ribosomes exist within cells either as monosomes, where a single ribosome is loaded on an mRNA, or as polysomes, where multiple ribosomes translate a single mRNA molecule. Monosomes are thought to be initiating ribosomes, in which the ribosome is still near the start codon, and hence has only short nascent peptides attached to it. On the other hand, polysomes will have multiple ribosomes, which are engaged in translation elongation and, consequently, can be at any position on the coding region, therefore producing nascent peptides of varied length. Recent reports challenged this view, suggesting that the majority of monosomes are in fact elongating ribosomes of specific transcripts that are fully translated by a single ribosome at a time (Heyer and Moore, 2016). Transcripts that are fully translated by monosomes are more susceptible to NMD (Heyer and Moore, 2016) and I was interested to address this link between monosomal translation and NMD further. The starting hypothesis was that the translation of many transcripts would be affected in NMD mutants such as the UPF1 deletion strain. To be able to test this, I had to set up 1) the polysome fractionation method and 2) the puromycylation technique, to monitor translation directly, in *S. pombe*.

Polysome fractionation uses sucrose gradients to separate translation complexes by means of high-speed ultracentrifugation, and can be used to analyse global translation as well as to investigate the function of protein synthesis-related factors (Figure 3.1) (McCartyStafford and Brown, 1968;

Chassé *et al.*, 2017). Puromycylation, on the other hand, utilises the antibiotic puromycin, which in structure partly mimics the 3'-end of amino-acid loaded tRNA (Pestka, 1971). As such, puromycin enters the vacant A sites of translating ribosomes and becomes incorporated into the nascent peptide in a reaction catalysed by the ribosome (Azzam and Algranati, 1973). Puromycin incorporation triggers peptide release followed by release of ribosomal subunits and translation arrest. Puromycylated peptides, which can be detected using a puromycin antibody, therefore, allow identification of those proteins being synthesised in the cells at a given time (Pestka, 1971; DavidBennink and Yewdell, 2013; AvinerGeiger and Elroy-Stein, 2013). By pairing polysome fractionation with puromycylation, I was able to assess the cellular translation status.

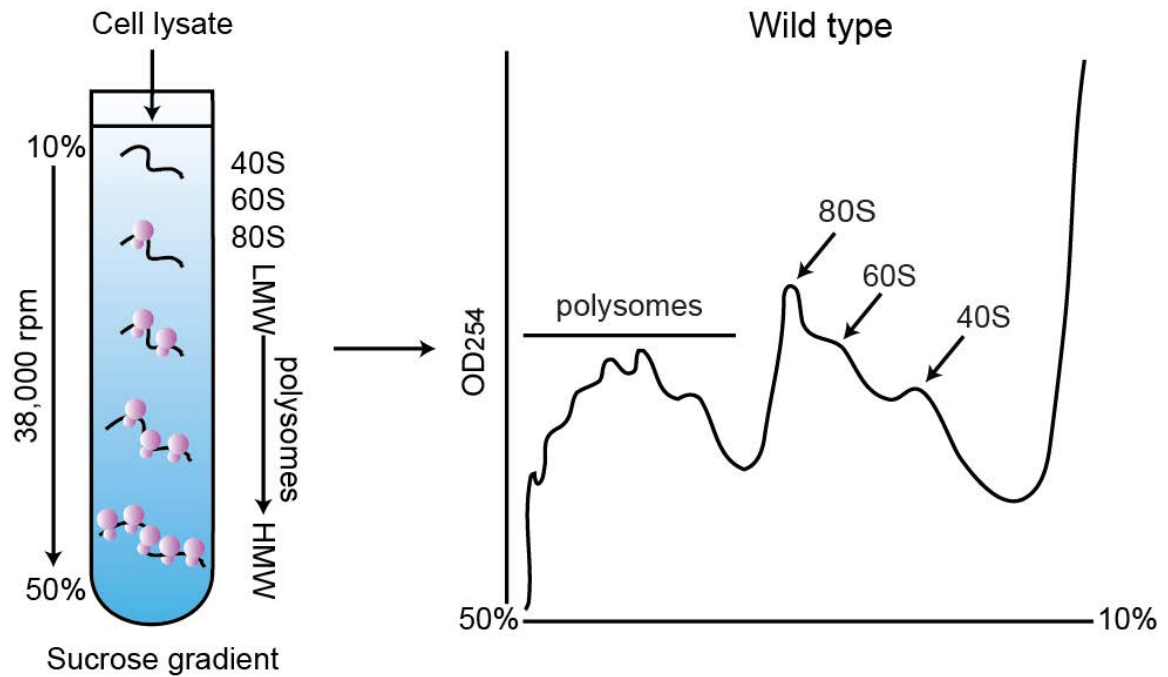


Figure 3.1 – **Polysome fractionation.** Schematic of polysome fractionation where free mRNAs and mRNAs loaded with different numbers of ribosomes migrate in a tube filled with the sucrose gradient, as indicated. The gradients are then fractionated during which RNA is monitored, the readout of which is shown on the right. 40S – small ribosomal subunit; 60S – large ribosomal subunit; 80S – monosome; LMW – low molecular weight; HMW – high molecular weight.

Upon using the method to study cellular translation profile, I detected molecules of high molecular weight synthesised by the monosome, suggesting that single ribosomes can translate full proteins, as reported (Heyer and Moore, 2016). On the other hand, since puromycylation is thought to be catalysed only by the ribosome, the expectation is that puromycylated peptides would be detected solely in ribosome-containing samples. Strikingly, approximately 70% of puromycylation was detected in samples that were void of ribosomes. The reaction was driven in a translation-independent manner, the nature of which remains unclear. I hypothesised that since puromycin interacts with the ester linkage between the growing peptide chain and tRNA in the ribosome catalytic centre, it may also react, independently of the ribosome, with peptidyl-tRNAs that might be released prematurely during translation. The observations presented in this chapter suggest that this might be the case, and that the translation of many proteins can be terminated by an alternative mechanism, which does not involve hydrolysis of the last tRNA, contrary to general understanding.

3.2 Results and discussion

3.2.1 Emetine does not stabilise polysomes in *S. pombe*, in contrast to cycloheximide

Prior to polysome fractionation, cells are typically treated with translation elongation inhibitors that trap ribosomes on translating mRNAs, to allow analysis of the cellular translation state at a given time (Pestka, 1971). Two potent inhibitors, cycloheximide and emetine, are readily used in mammalian cells to prevent ribosome dissociation from mRNAs (Pestka, 1971). Cycloheximide acts by binding to the E site within the large ribosomal subunit, impairing eEF2-mediated tRNA translocation, thus blocking translational elongation (Grollman, 1966; Schneider-Poetsch *et al.*, 2010). Emetine inhibits translation by binding to the E-site within the small ribosomal subunit and, similarly to cycloheximide, inhibiting mRNA/tRNA translocation (Grollman, 1966; Wong *et al.*, 2014). I used both translation inhibitors during optimisation to assess under what conditions optimal polysome fractionation is achieved.

Both translation inhibitors should result in polysome stabilisation. However, when looking at the effects of emetine, I found that, even in its lower concentrations, 80S and 60S peaks were notably increased, comparatively to the no drug treatment (Figure 3.2A), suggesting polysome breakdown. This effect was more prominent as I increased the concentration of the drug (Figure 3.2C; D). Prolonged emetine treatment (Tina McLeod PhD Thesis, University of Birmingham, 2016) leads to a more prominent polysome breakdown suggesting that emetine, contrary to its observed effect in mammalian cells (Pestka *et al.*, 1972), does not stabilise polysomes in *S. pombe*.

On the other hand, I found that incubation with 100 µg/mL of cycloheximide for 5 minutes leads to polysome stabilisation and gives optimal polysome profiles (Figure 3.2B). Therefore, subsequent polysome fractionation was carried out using cycloheximide at this concentration.

Additionally, other conditions, such as salt and magnesium concentration in the lysis buffer, the number and duration of beads-shaking pulses (to ensure high cell break efficiency), lysate clearing conditions, duration and speed of gradient centrifugation, as well as fractionation parameters were also adjusted during standardisation, to obtain optimal polysome profiles. The key to successful polysome fractionation, apart from RNase-free reagents and inhibition of endogenous RNases, was carrying out all the steps quickly and effectively, at 4°C.

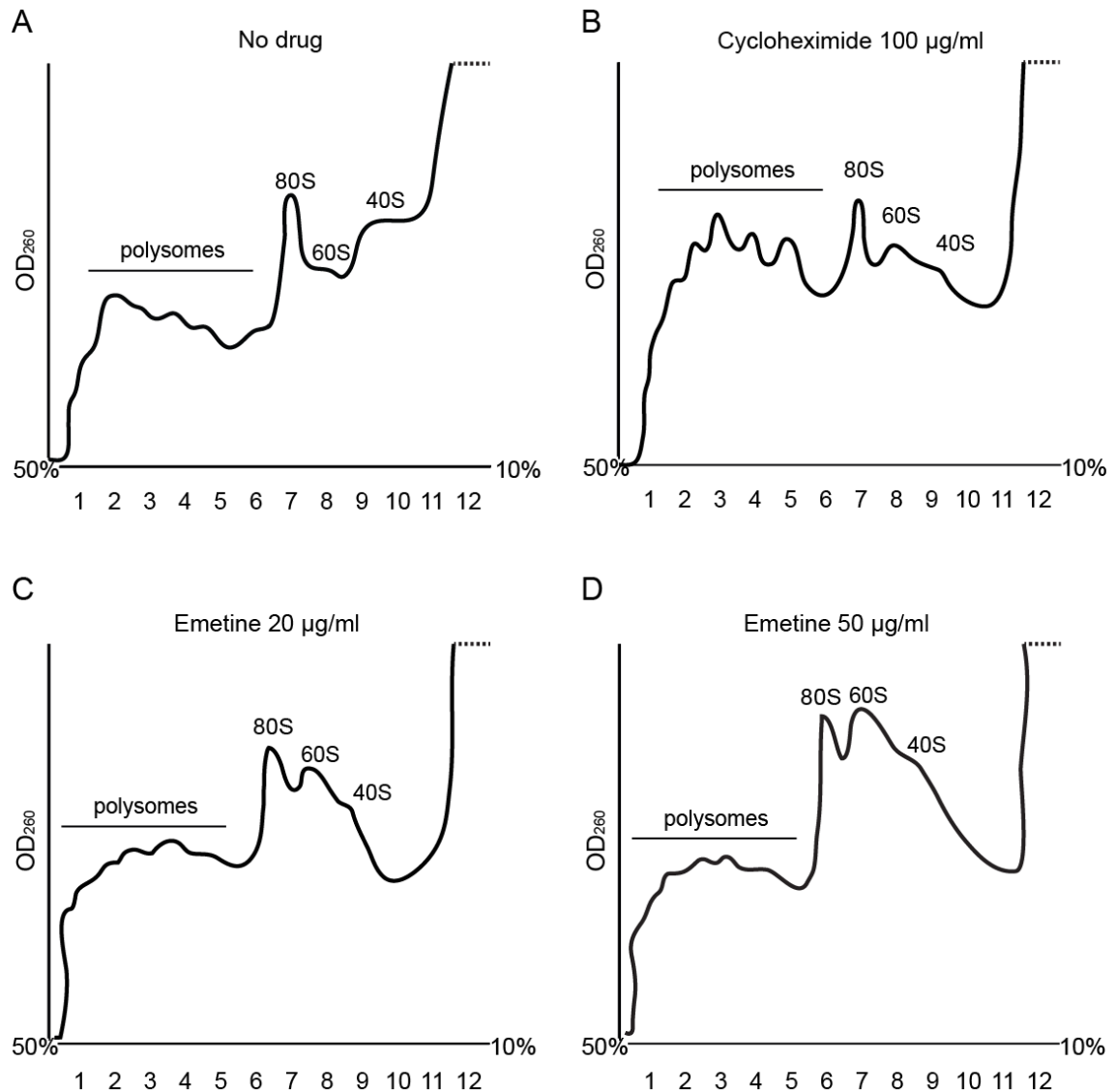


Figure 3.2 – **Cycloheximide stabilises polysomes in *S. pombe***. Polysome trace of cell extracts (A) untreated with translation inhibitors, (B) treated with cycloheximide (100 µg/mL) for 5 minutes, (C) treated with emetine (20 µg/mL) for 15 minutes or (D) treated with emetine (50 µg/mL) for 15 minutes, prior to cell lysis. Traces represent the absorbance profile (OD₂₅₄) recorded during the fractionation. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated.

3.2.2 Puromycin is incorporated into polypeptides in *S. pombe*

To use puromycylation to monitor translation in *S. pombe*, I first performed a 15-minute puromycin treatment at a concentration of 250 µg/mL, followed by SDS-PAGE/Western blotting using an anti-puromycin antibody (Tina McLeod PhD thesis, 2016) (DavidBennink and Yewdell, 2013). Quantification is expressed as the mean percentage \pm standard deviation (SD) of puromycin signal in cycloheximide- and anisomycin- pre-treated samples relative to the puromycin signal in the control, derived from three biological replicates. A ladder of puromycylated peptides is observed in the puromycin-treated sample (Figure 3.3, lane 3). Detected polypeptides demonstrate specific puromycylation as no signal is observed in the non-treated control (Figure 3.3, lane 2). The signal is dependent on translation as pre-treatment with 250 µg/mL anisomycin that competes with puromycin in binding to the ribosome A site, inhibiting peptide bond formation and, hence, ribosome-dependent puromycylation (Grollman, 1967), reduces the puromycin signal to $11 \pm 6\%$ (Figure 3.3, lane 5). Cycloheximide does not have an inhibitory effect on puromycylation as the puromycin signal was reduced only to $73 \pm 11\%$ (Figure 3.3, lane 4), contrary to previous observations ((BaligaCohen and Munro, 1970), Tina McLeod PhD thesis, 2016) but in line with a recent study that established the ribopuromycylation (RPM) technique (DavidBennink and Yewdell, 2013). The group concluded that pre-treatment with cycloheximide or emetine immobilises the ribosomes on mRNAs but does not inhibit puromycylation. Puromycylated peptides remain associated with the ribosomes. RPM, linked with

immunofluorescence, was used to visualise ribosomes loaded with nascent, puromycylated peptides, to identify cellular loci in which translation is occurring.

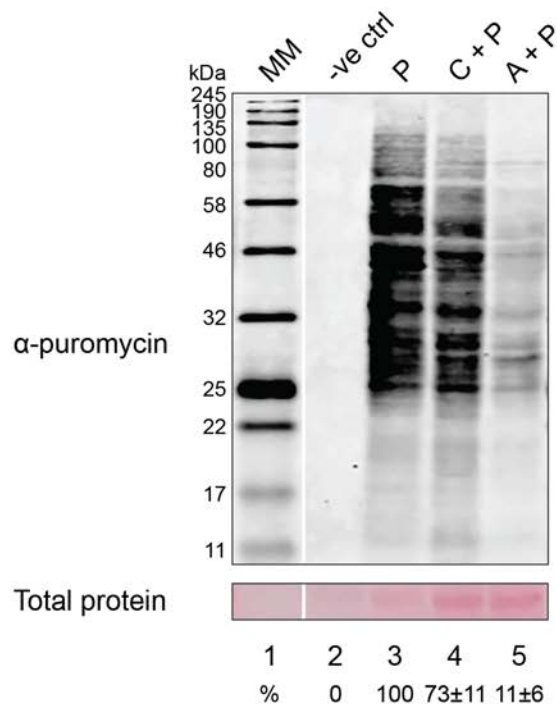


Figure 3.3 – **Puromycin is incorporated into polypeptides in *S. pombe*.** Top panel shows SDS-PAGE/Western blotting of protein extracts from wild type cells (-ve ctrl) untreated with translation inhibitors, (P) treated with puromycin (250 µg/mL) for 15 minutes, (C + P) treated with cycloheximide (100 µg/mL) for 15 minutes prior to 20-min puromycin treatment (250 µg/mL), (A + P) treated with anisomycin (250 µg/mL) for 15 minutes prior to 20-min puromycin treatment (250 µg/mL). Puromycylated peptides were detected using an anti-puromycin antibody. A representative area of Ponceau S staining demonstrating protein loading is shown in the bottom panel, which was used for normalisation during quantification of the puromycin signal. The intensity of puromycin signal for each lane was determined using Image J software and was expressed as a percentage of the total signal. The signal in cycloheximide- and anisomycin-treated samples was expressed as a percentage relative to the sample solely treated with puromycin, which represents 100%. The results are expressed as a mean \pm SD, derived from three biological replicates.

3.2.3 Puromycylation can be driven independently of ribosomes

As detailed in the Introduction, to assess whether there is any difference in the pattern of proteins synthesised in polysomes of higher versus lower molecular weight, as well as in single ribosomes, I performed puromycylation in samples obtained by polysomal cell fractionation.

I used a wild type strain expressing RpL25 tagged with HA in the polysome fractionation experiment (Figure 3.4A). Upon puromycin treatment (50 µg/mL), proteins extracted from 100 µL of each fraction were subjected to SDS-PAGE/Western blotting analysis using an anti-puromycin antibody, with each fraction being loaded onto the gel in equal proportion (Figure 3.4B). RpL25-HA was also detected using an anti-HA antibody to track ribosome distribution and to assess polysome quality and loading. The quantification detailed below is expressed as a mean of three biological replicates \pm SD.

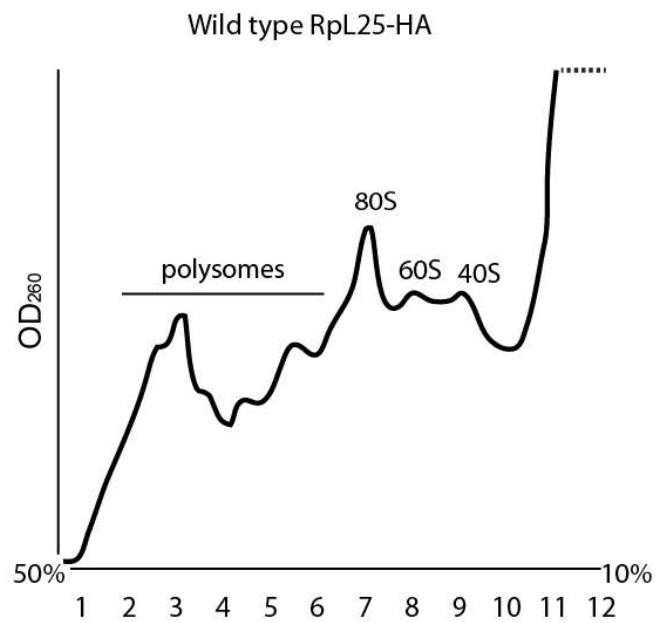
Polypeptides of a wide range of molecular weights were detected in the monosomal fraction (Figure 3.4B, lane 7). What was particularly intriguing is that proteins of high molecular weight were observed, suggesting that proteins of full molecular weight can be synthesised by a single ribosome, and monosomes might not represent solely initiating ribosomes. Another interpretation is that a considerable proportion of ribosomes present in the monosomal fraction represent terminating ribosomes, which may be stalled at the end of the mRNA sequence.

Puromycylation is considered to be dependent on translation, as puromycin incorporation into the nascent peptide is catalysed by the translating ribosome (Blobel and Sabatini, 1971; Azzam and Algranati, 1973). Therefore,

puromycylated peptides should be observed solely in ribosome-containing fractions, polysomes and monosomes. Unexpectedly and strikingly, $71 \pm 7\%$ of puromycin signal was detected in light fractions, which are void of ribosomes (Figure 3.4B, lanes 11-12). A very low puromycin signal was observed in polysomal fractions with $6 \pm 2\%$ (Figure 3.4B, lanes 1-6) and monosomal fractions (Figure 3.4B, lane 7) with $19 \pm 5\%$ of total puromycin signal detected. The signal was calculated from three biological repeats, and represents the mean percentage \pm SD. For each category, the data is expressed as a percentage of the total puromycylation observed, normalised against the control (RpL25-HA signal in polysomes and Ponceau S staining in the light fractions).

To ensure that the signal I observed was in fact translation-dependent puromycylation, prior to treating polysomal fractions with puromycin, I performed 250 $\mu\text{g/mL}$ anisomycin treatment. Indeed, I observed a complete reduction of the signal in fractions corresponding to polysomes and monosomes (Figure 3.4C, lanes 1-7), however, the signal in light fractions remained unchanged (Figure 3.4C, lanes 11-12). Therefore, puromycylation observed in fractions containing ribosomes is in fact driven by translation, while the nature of puromycylation in light fractions differs and is not catalysed by a translating ribosome.

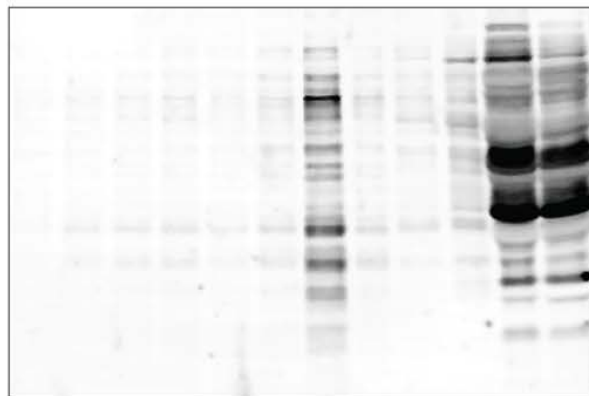
A



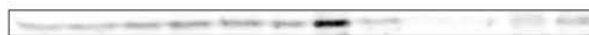
B

Puromycin (50 μ g/ml)
treatment to the fractions

α -puromycin



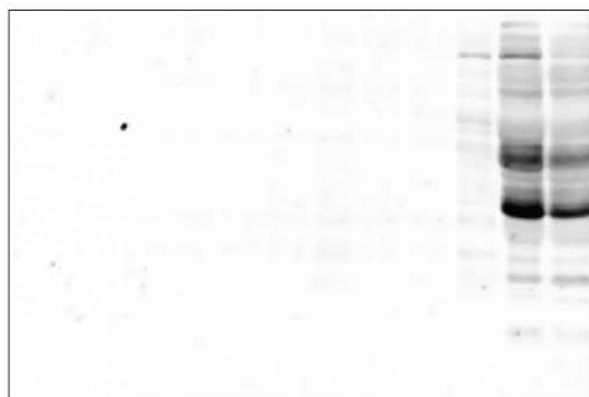
RpL25-HA



C

Anisomycin (250 μ g/ml)
followed by
puromycin (50 μ g/ml)
treatment
to the fractions

α -puromycin



RpL25-HA



Figure 3.4 – Puromycylation can occur in ribosome-free fractions. (A) Representative trace of polysome fractionation of yeast cells expressing RpL25-HA. (B) Top panel shows Western blotting of protein extracts from individual polysome fractionations incubated with puromycin (50 µg/mL). Bottom panel shows the RpL25-HA band, detected using an anti-HA antibody (this is restricted to polysomal and 60S fractions, as expected). (C) Top panel shows Western blotting of protein extracts from individual polysome fractions incubated with anisomycin (250 µg/mL) followed by puromycin (50 µg/mL) treatment. Again, bottom panel shows RpL25-HA distribution, detected using an anti-HA antibody (restricted to polysomal and 60S fractions). The intensity of puromycin signal for each lane was determined using Image J software and was expressed as a percentage of the total signal. The percentages from lanes 1 to 6 were added together to give the combined signal for the polysomal fractions, lanes 7-10 were similarly combined to give the total signal for the ribosomal subunits and lanes 11-12 were combined to give the total signal for the light fractions. The signal in anisomycin-treated sample was expressed as a percentage relative to the sample solely treated with puromycin, which represents 100%. The results are presented in the text as a mean of three independent biological replicates \pm SD.

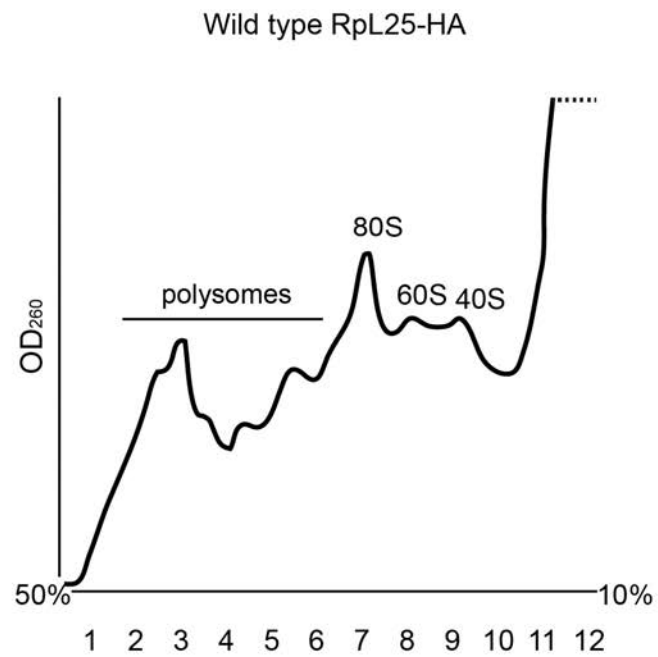
As an unexpectedly low signal of puromycylated proteins was observed in ribosome-containing fractions. I considered that due to the protein quantity in light fractions being very high, the puromycylation in polysomal fractions could be masked, due to the saturation of the signal. Therefore, I loaded proteins corresponding to 200 μ L of ribosome-containing fractions while only one fifth of the light fractions, in an attempt to improve detection.

Indeed, increasing the amount of proteins resulted in a detectable signal (Figure 3.5B, lanes 1-7). As previously observed, puromycylation detected in ribosome-containing fraction is dependent on translation, as pre-treatment with anisomycin leads to a reduction in the signal to $10 \pm 4\%$ (Figure 3.5C, lanes 1-7). The reduction was quantified as a mean percentage of the signal observed in the anisomycin-puromycin treated sample, relative to the solely puromycin-treated sample \pm SD, derived from three biological replicates. Puromycylation detected in the light, ribosome-free fractions, does not depend on translation since it is not impaired by pre-treatment with anisomycin (Figure 3.5B; C, lanes 11-12).

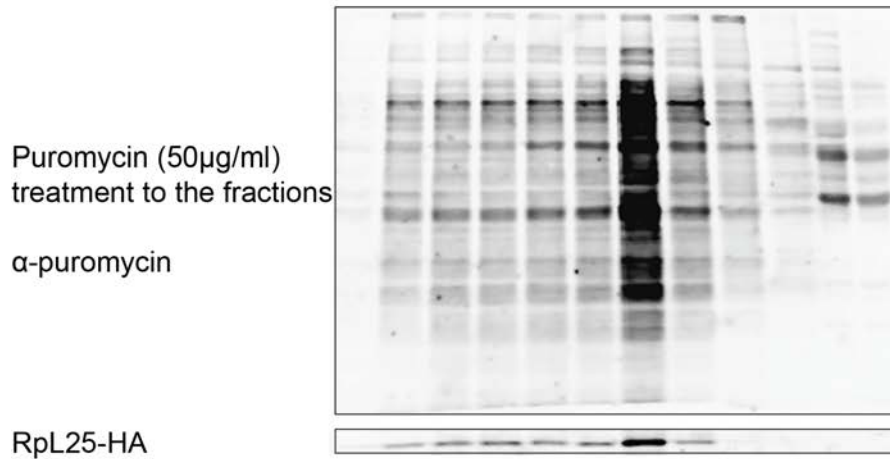
The amount of puromycin and anisomycin-puromycin treated samples used for SDS-PAGE/Western blotting analysis was comparable. However, to make sure that the reason for the reduction in the signal is not due to variability in loading, blotting or detection, corresponding samples of the two treatments were loaded on a single gel (Figure 3.6). When compared on the same gel, indeed, as previously shown, the puromycin signal from the anisomycin-puromycin treated sample was strongly reduced relative to the samples solely treated with puromycin (Figure 3.6A, lanes 1-6 compared to lanes 8-13). A similar effect was

observed when comparing the monosome samples (Figure 3.6B, lane 1 compared to lane 8). However, the signal in light fractions remained unchanged (Figure 3.6B, lanes 5-6 compared to 12-13), confirming a different nature of puromycylation, driven independently of the ribosome.

A



B



C

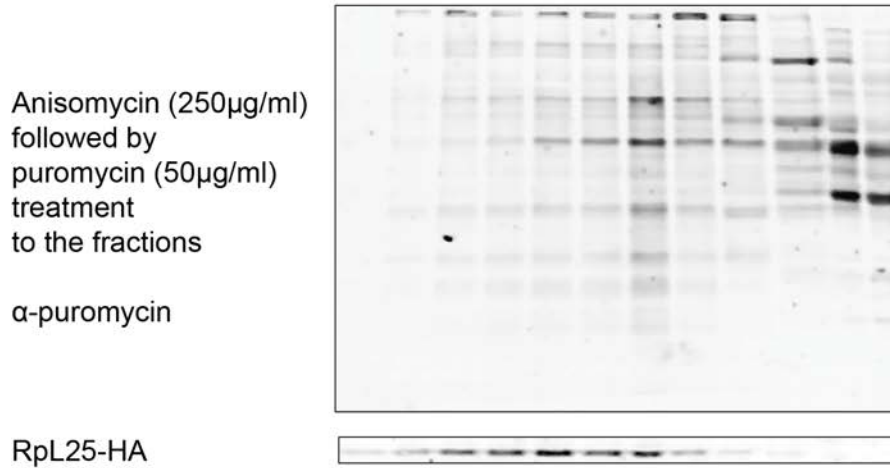


Figure 3.5 – Puromycylation can be driven independently of ribosomes. (A) Representative trace of polysome fractionation of yeast cells expressing RpL25-HA. (B) Top panel shows Western blotting of protein extracts from individual polysome fractionations incubated with puromycin (50 µg/mL) where only 20% of the light fractions (fractions 11 and 12) was loaded onto the gel. (C) Top panel shows Western blotting of protein extracts from polysome fractionations incubated with anisomycin (250 µg/mL) followed by puromycin (50 µg/mL). Again, only 20 % of the light fractions (fractions 11 and 12) was loaded onto the gel. Bottom panel shows RpL25-HA distribution in both cases, detected using an anti-HA antibody. The intensity of puromycin signal for each lane was determined using Image J software and was expressed as a percentage of the total signal. The percentages from lanes 1 to 6 were added together to give the combined signal for the polysomal fractions, lanes 7-10 were similarly combined to give the total signal for the ribosomal subunits and lanes 11-12 were combined to give the total signal for the light fractions. The signal in anisomycin-treated sample was expressed as a percentage relative to the sample solely treated with puromycin, which represents 100%. The results are presented in the text as a mean of three independent biological replicates \pm SD.

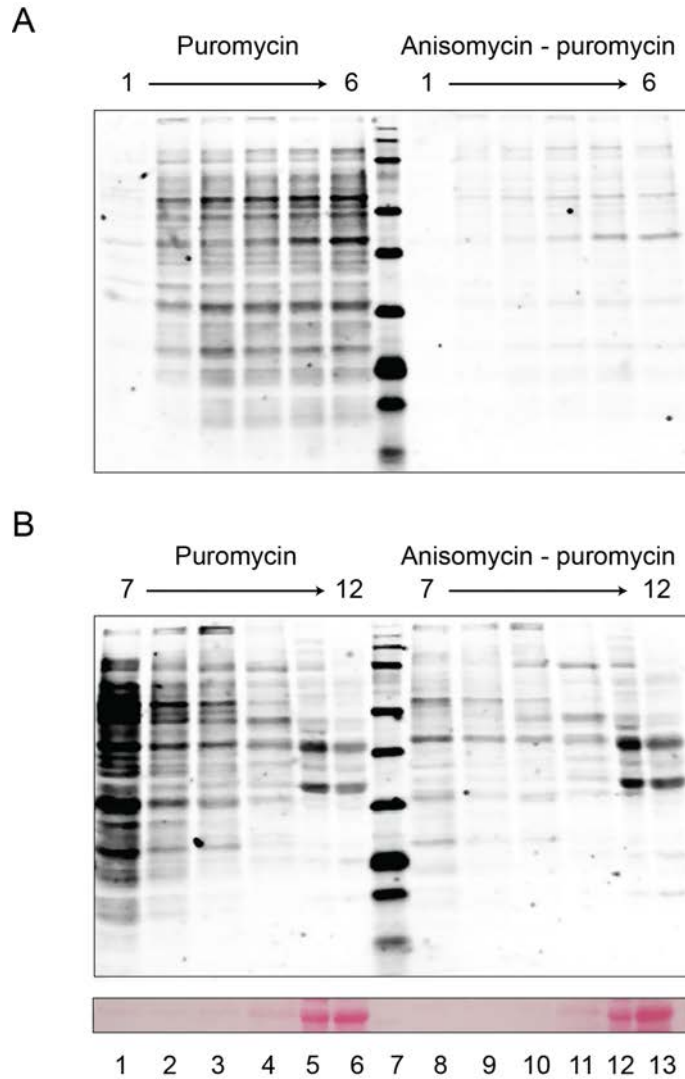


Figure 3.6 – **Puromylation can be driven independently of ribosomes.** (A) Western blotting of protein extracts from individual polysome fractionations incubated with puromycin (50 µg/mL) and anisomycin (250 µg/mL) followed by puromycin treatment (50 µg/mL), as indicated. The fractions compared are ribosome-containing fractions (1-6, 1 being polysomes of highest and 6 polysome of lowest molecular weight). (B) Western blotting of protein extracts from individual polysome fractionations incubated with puromycin (50 µg/mL) and anisomycin (250 µg/mL) followed by puromycin treatment (50 µg/mL). The fractions in question are monosome, ribosomal subunits and light fractions (from 7 to 12). Bottom panel shows a representative area of Ponceau S staining, demonstrating protein loading.

3.2.5 Puromycylation potentially occurs in light fractions on peptidyl-tRNAs

As indicated by the data in sections 3.2.3 and 3.2.4, puromycylation may not solely be driven by translation. Not all proteins within the light fractions are puromycylated, as puromycylated peptides detected in the light fractions by Western blotting differ from the total proteins observed by Ponceau S staining. Some feature of the polypeptide seems to allow puromycin incorporation. I hypothesised that potentially what I am detecting in light fractions are newly synthesised peptides that arise due to errors in translation where stalled ribosomes may terminate early and the polypeptides are released from the ribosomes with a tRNA still attached to them (Figure 3.7). As such, they could potentially become a substrate for puromycin incorporation as puromycin typically binds peptidyl-tRNA, a reaction catalysed by the ribosome peptidyl transferase centre.

Free peptidyl-tRNAs have been reported primarily in bacteria, but also indicated in yeast and higher eukaryotes, as well as in *in vitro* translation studies (Menninger, 1976; Cao and Geballe, 1998; Bresler *et al.*, 1966). Peptidyl-tRNA hydrolase, an enzyme that catalyses rapid degradation of such peptidyl-tRNAs, was identified in bacteria and its homologues have also been described in eukaryotes (Menninger, 1976; GrossCrow and White, 1992; Das and Varshney, 2006).

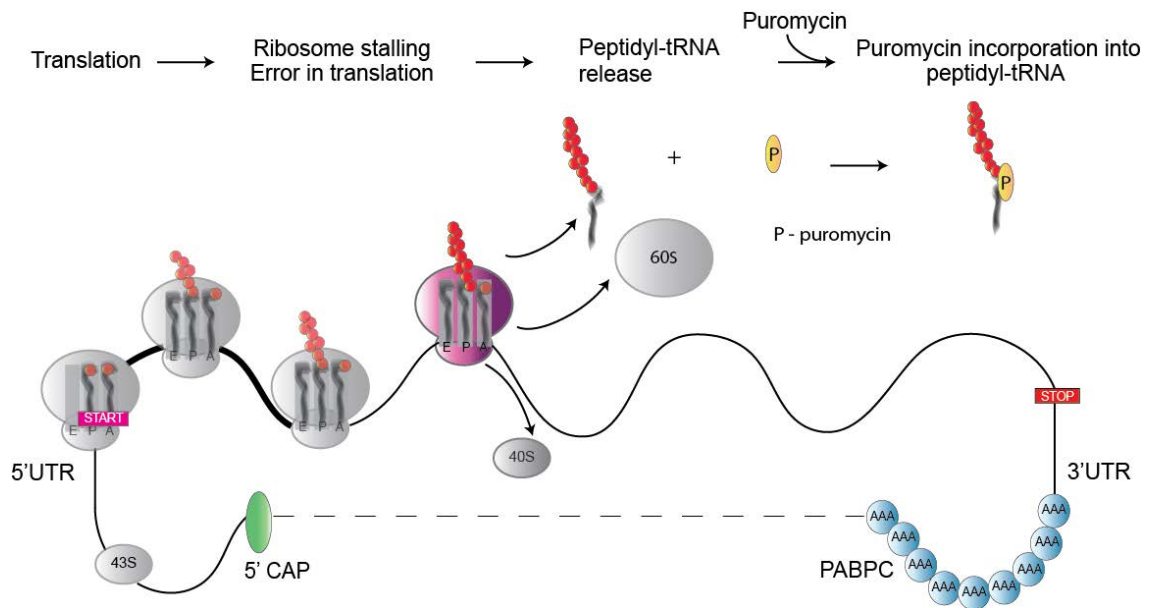


Figure 3.7 – **Peptidyl-tRNAs that might be generated during translation could serve as templates for puromylation.** During translation, ribosome stalling leads to the release of the nascent polypeptide with a tRNA still attached to it. Such products can become substrates for puromycin incorporation. 5' CAP – 5' capping complex; UTR – untranslated region; START – start codon; 40S – small ribosomal subunit; 60S – large ribosomal subunit; P – puromycin; STOP – stop codon; PABPC – poly(A)-binding protein, cytosolic.

In vitro, peptidyl-tRNAs are shown to be very stable in neutral and acid pH environments, while basic pH leads to a considerable reduction of their half-life (Bresler *et al.*, 1966; Bresler *et al.*, 1968). The stability increases with the length of the peptide chain (Bresler *et al.*, 1968). Primarily, the peptidyl-tRNAs tested were short peptides, and it is hypothesised that in cells, only such peptides could be generated (Cruz-Vera *et al.*, 2003). However, I found that the puromycylated products in the light fractions are long peptides, and potentially full-length proteins. If the reason why I observe puromycylation in light fractions is indeed due to a tRNA remaining attached to newly synthesised peptides generated by the ribosome, this could mean that even full-length proteins can be generated in such a way.

To test this hypothesis, due to the different stability of peptidyl-tRNAs in different pH conditions, I decided to alter the pH of the light fractions prior to puromycin treatment, either toward acid pH by HCl addition or toward basic pH via NaOH addition. Due to the previously explained nature of peptidyl-tRNA behaviour, I expected an increase in puromycylation in an acid environment, and a reduction in a basic environment (Bresler *et al.*, 1968; Cao and Geballe, 1998; Johansson *et al.*, 2011). I performed the same experiment in the polysome-containing fractions, as ribosome-dependent puromycylation is reportedly abolished at extreme pH values (Johansson *et al.*, 2011).

The results shown in Figure 3.8 demonstrate that puromycylation efficiency depends on the pH of the buffer. In the case of both ribosome-dependent and independent puromycylation, peak puromycylation efficiency occurs at pH 8.5. The puromycylation reaction occurs only between pH 6.5 and 9.5 (Figure 3.8B;

C, lanes 3-6, 9). Beyond this pH range, puromycylation is impaired in polysomal fractions while approximately 20% of puromycylation is observed in light fractions, relative to puromycylation detected at physiological conditions of pH 7.5 (Figure 3.8B; C, lanes 3-8, 10-13). The quantification (Figure 3.8D) is derived solely from a single biological replicate, based on the gels shown (Figure 3.8B; C).

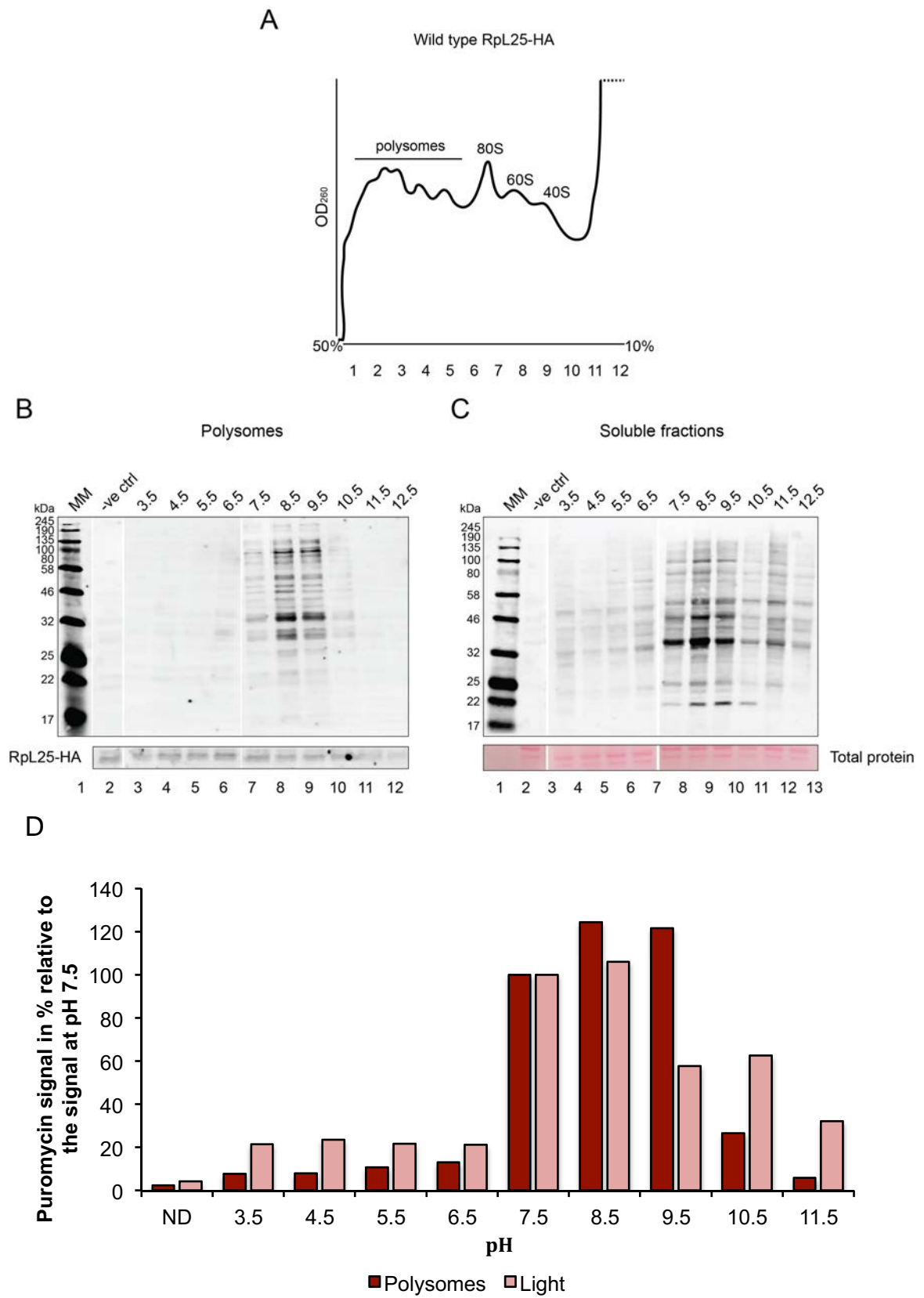


Figure 3.8 – Puromycylation kinetics depends on the pH of the buffer and the reaction occurring in light fractions is distinctly different to the reaction catalysed by the ribosome. (A) Polysome fractionation trace of yeast cells expressing RpL25-HA. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. The first 6 fractions were pooled to generate the polysomal sample, while fractions 11 and 12 were combined to generate the light fractions sample. The pool of each was then split to the required number of samples, 200 μ L for each of the polysomal and 100 μ L for each of the light, and treatment was carried out as described below. (B) Top panel shows Western blotting of protein extracts from polysome fractions upon indicated pH change and subsequent puromycin (50 μ g/mL) treatment. Bottom panel shows the RpL25-HA band, detected using an anti-HA antibody, serving as a loading control. (C) Top panel shows Western blotting of protein extracts from light, ribosome-free fractions upon indicated pH change and subsequent puromycin (50 μ g/mL) treatment. Bottom panel shows Ponceau S staining, serving as a loading control. The numbers above the lanes indicate the pH upon which puromycin treatment was performed. MM indicates molecular marker. Each gel contains a negative, non-puromycin-treated control, indicated by '-ve control', which shows no signal in all samples tested. (D) The intensity of puromycin signal for each lane was determined using Image J software and was expressed as a percentage relative to the signal observed at pH 7.5, which represents 100%. The data is derived from a single biological replicate.

Further, I altered the pH, in polysomal and light fractions in parallel, for 30 minutes, reverted back to physiological pH of 7.5 and then performed puromycylation (Figure 3.9B; C). This allows stabilisation of peptidyl-tRNA in acid conditions and degradation of peptidyl-tRNA in basic conditions, but the puromycylation conditions are the same and are carried out at an optimal pH in all samples. I observed that puromycylation in polysomal fractions becomes impaired at pH values above 9.5 and below 5.5 (Figure 3.9B, lanes 5-7, 10-11). However, the light fractions display a different pattern where puromycylation is efficient at all pH values tested (Figure 3.9C, lanes 2-7, 9-12). These results further support the hypothesis that the two reactions are distinctly different as the ribosome catalytic centre seems to become impaired at extreme pH values (Johansson *et al.*, 2011). Peptidyl-tRNAs are stabilised in acid and degraded in alkaline conditions. Therefore, as predicted, puromycylation is reduced after treatment at basic pH, but increased after acidic pH treatment (Figure 3.9C, lanes 3-7, 9-12). The effects were minor, with the puromycylation difference being only 10% (80% puromycylation detected after alkaline and approximately 90% after acid pH treatment in light fractions, relative to puromycylation observed at pH 7.5), the reason potentially being that long peptidyl-tRNAs are very stable.

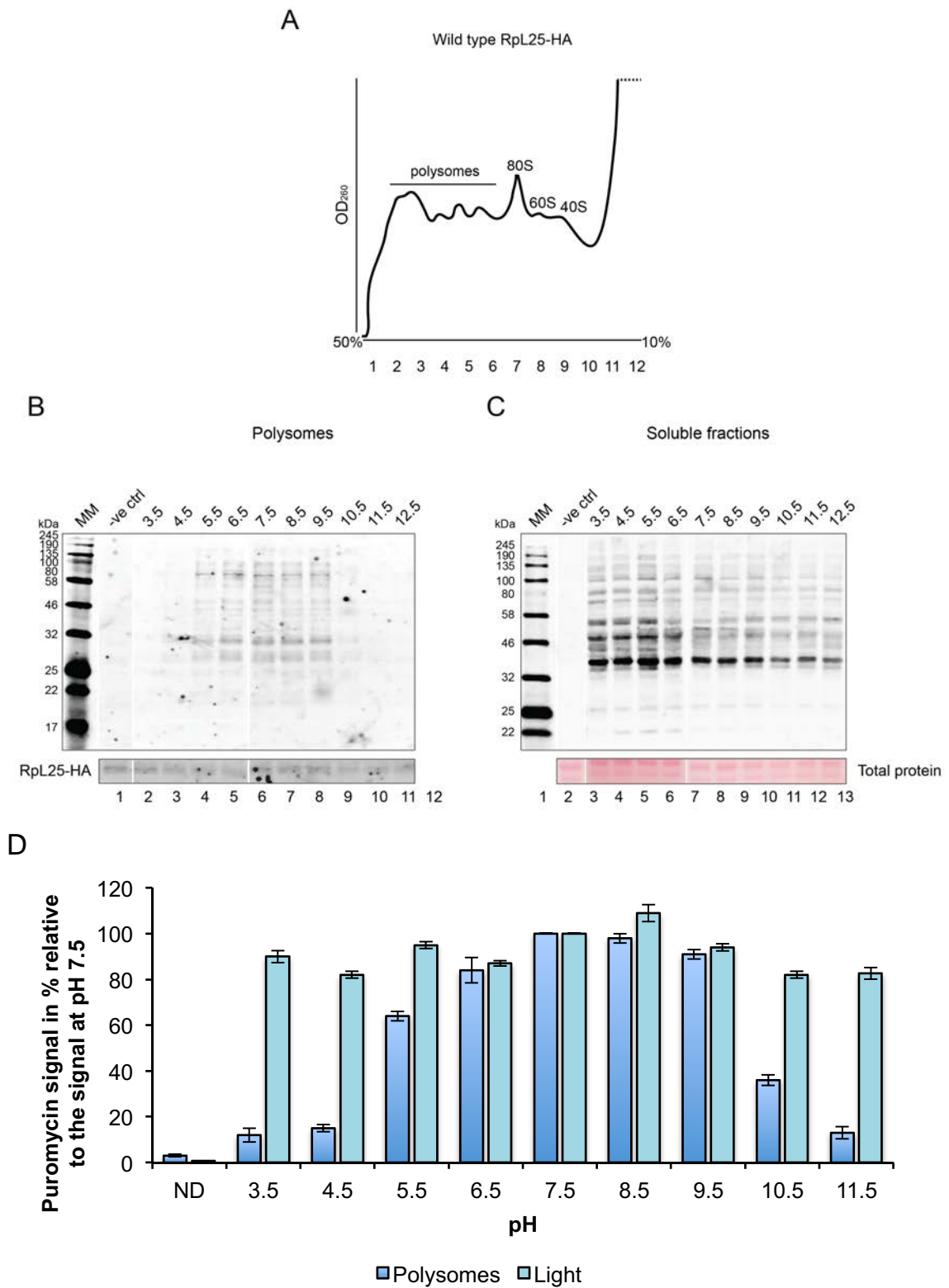
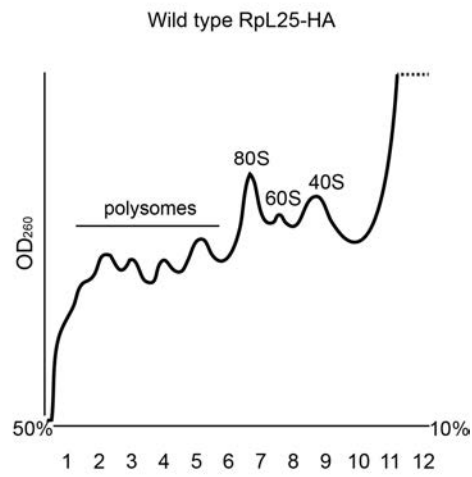


Figure 3.9 – Puromycylation upon pH change and reversion to pH 7.5 shows a different pattern in polysomal versus light fractions. **(A)** Polysome fractionation trace of yeast cells expressing RpL25-HA. The gradient was split into 12 fractions, as indicated. The first 6 fractions were pooled to generate the polysomal sample, while fractions 11 and 12 were combined to generate the light fractions sample. The pool of each was then split into the required number of samples, 200 μ L for each of the polysomal and 100 μ L for each of the light, and treatment was carried out as described below. **(B)** Top panel shows Western blotting of protein extracts from polysome fractions upon a 30-min treatment at an indicated pH (numbers above the lanes) after which pH was reverted to 7.5 and puromycin (50 μ g/mL) treatment was carried out. Bottom panel shows the RpL25-HA band, detected using an anti-HA antibody, serving as a loading control. **(C)** Top panel shows Western blotting of protein extracts from light, ribosome-free fractions upon a 30-min treatment at an indicated pH (numbers above the lanes) after which pH was reverted to 7.5 and subsequent puromycin (50 μ g/mL) treatment was performed. Bottom panel shows a representative area of Ponceau S staining, serving as a loading control. MM indicates molecular marker. Each gel contains a negative, non-puromycin-treated control, indicated by '-ve control', which shows no signal in all samples tested. **(D)** The intensity of puromycin signal for each lane was determined using Image J software and was expressed as a percentage relative to the signal observed at pH 7.5, which represents 100%. The data is derived as a mean percentage \pm SD, derived from three biological replicates.

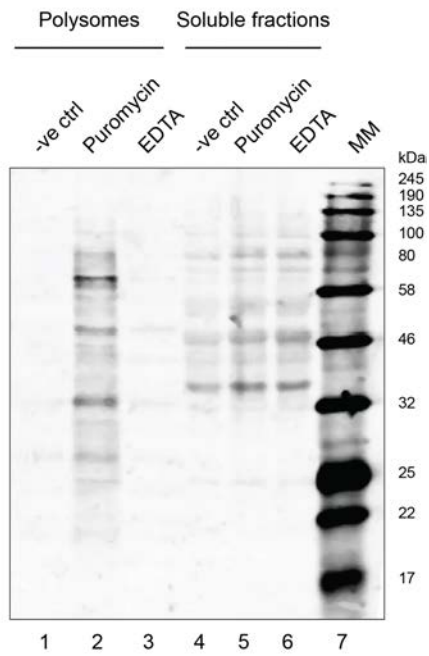
3.2.6 Ribosome-independent puromycylation is not impaired upon EDTA treatment

To further investigate the nature of ribosome-independent puromycylation, I incubated the samples with EDTA prior to puromycylation. Mg^{2+} ions are crucial for maintaining the assembly of ribosomes. EDTA is a Mg^{2+} chelating agent, which reduces the availability of Mg^{2+} ions and therefore compromises ribosome integrity (Miall and Walker, 1969; Nolan and Arnstein, 1969). In the absence of Mg^{2+} , ribosomal subunits dissociate, and the catalytic centre is impaired as a result (Miall and Walker, 1969). I performed EDTA-puromycin treatment in both polysome-containing and ribosome-free samples (Figure 3.10). As expected, EDTA treatment led to a complete reduction of the puromycin signal in the reaction driven by the ribosome (Figure 3.10B, lane 3; C). However, in the ribosome-free sample, the signal was not impaired and was comparable to the untreated sample (Figure 3.10B, lanes 5 and 6; C). Quantification of the results is presented on the graph (Figure 3.10C), derived from three biological replicates where the puromycin signal in different conditions is expressed relative to the sample solely treated with puromycin. The observations suggest that the reaction is either not enzymatic or that an enzyme in question that catalyses this reaction is not inhibited by EDTA.

A



B



C

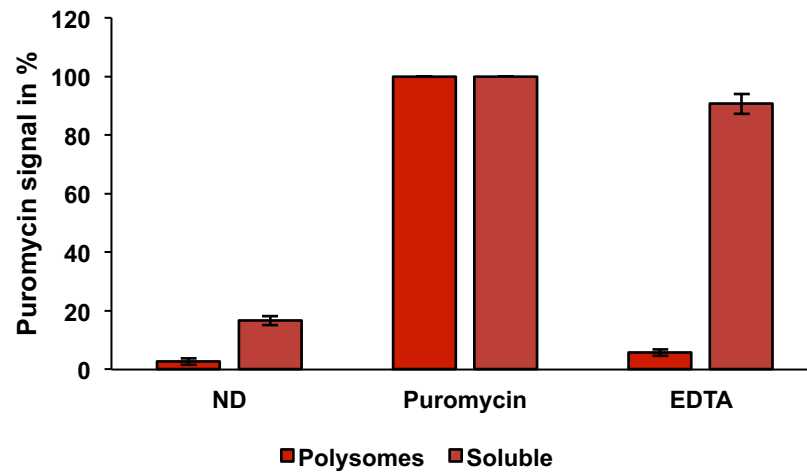


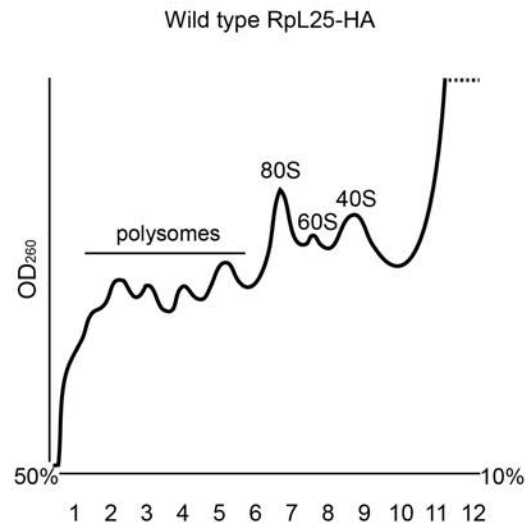
Figure 3.10 – Ribosome-independent puromycylation is not impaired by EDTA treatment. **(A)** Trace of polysome fractionation of yeast cells expressing RpL25-HA. The gradient was split into 12 fractions. The first 6 fractions were combined to generate the polysomal sample, while fractions 11 and 12 were combined to generate the light fractions sample. The pool of each was then split into the required number of samples, 200 μ L for each of the polysomal and 100 μ L for each of the light, and treatment was carried out as described below. **(B)** Western blotting of protein extracts from polysomal and light fractions samples which were either untreated (-ve ctrl), treated with puromycin (50 μ g/mL) (Puromycin) or treated with 30 mM EDTA which was followed by puromycin treatment (50 μ g/mL) (EDTA). MM indicates molecular marker. **(C)** The intensity of puromycin signal for each lane was determined using Image J software and was expressed as a percentage relative to the samples solely treated with puromycin that represent 100% (Puromycin, lanes 2 and 5). The results are derived as a mean percentage \pm SD from three biological replicates.

3.2.7 Ribosome-independent puromycylation is impaired upon hydroxylamine treatment

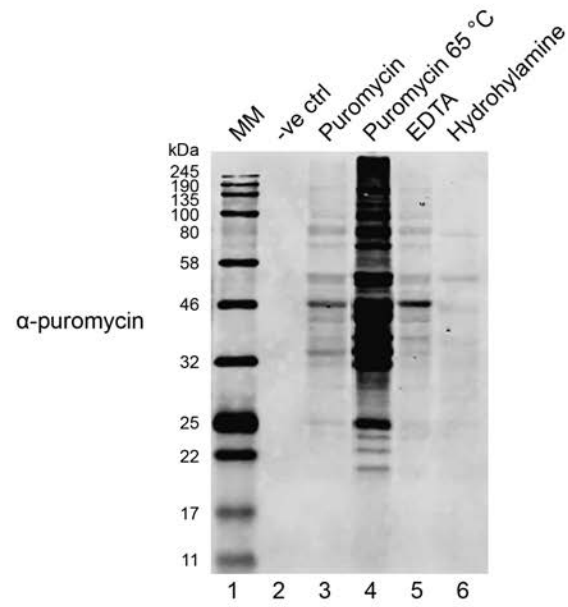
I further tested whether the nature of ribosome-dependent puromycylation is enzymatic by performing the reaction upon 30-min 65°C incubation (Figure 3.11B, lane 4, C). During this treatment most cellular enzymes are inactivated and the reaction would be impaired if catalysed by an enzyme. On the contrary, I found that the reaction is stimulated at higher temperature, even enhanced, showing $261 \pm 21\%$ of the signal, relative to puromycylation performed at RT (Figure 3.11B, lane 3; C).

To directly assess the hypothesis that ribosome-independent puromycylation occurs due to a presence of peptidyl-tRNAs to which puromycin can bind, I performed hydroxylamine treatment prior to puromycylation (Figure 3.11B, lane 6; C). Hydroxylamine triggers rapid peptidyl-tRNA degradation of short, free peptidyl-tRNAs, reducing their half-life to only 16 minutes, while having little to no effect on ribosome bound peptidyl-tRNAs (Bresler *et al.*, 1968; JimenezMonro and Vazquez, 1970). Indeed, upon 30-min 1M hydroxylamine treatment, subsequent puromycylation becomes strongly reduced, with $23 \pm 4\%$ of the puromycin signal remaining, relative to control conditions (Figure 3.11B, lane 6, C). With peptidyl-tRNA degradation, puromycylation is significantly impaired, which suggests that ribosome-independent puromycylation might indeed occur as a result of puromycin covalently binding to peptidyl-tRNAs, which could arise due to translation errors.

A



B



C

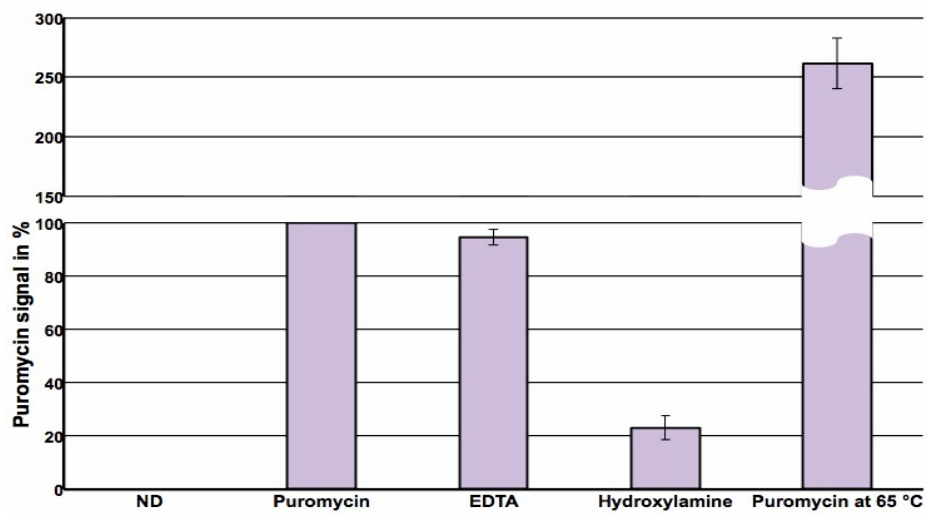
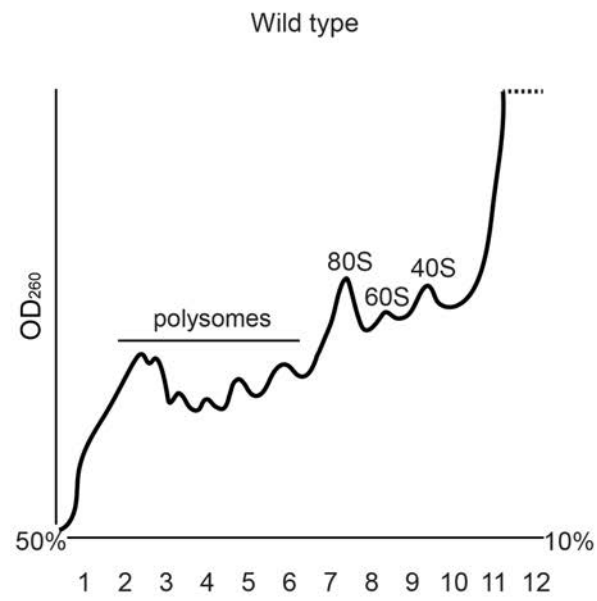


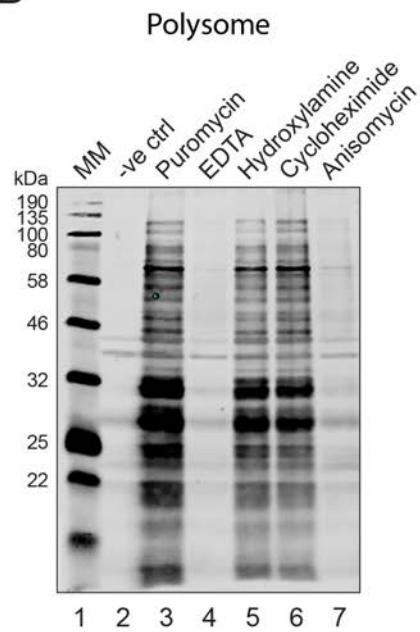
Figure 3.11 – **Ribosome-independent puromycylation is impaired by hydroxylamine treatment.** (A) Polysome fractionation trace of yeast cells expressing RpL25-HA. The gradient was split into 12 fractions. Fractions 11 and 12 were combined to generate the light fractions sample. This pool was then split into the required number of samples, 100 μ L for each, and treatment was carried out as described below. (B) Western blotting of protein extracts from light fractions samples which were either (-ve ctrl) untreated, (Puromycin) treated with puromycin (50 μ g/mL) for 20 minutes at RT, (Puromycin 65°C) treated with puromycin (50 μ g/mL) immediately upon 30-min treatment at 65°C, (EDTA) treated with 30 mM EDTA for 30 minutes at RT, which was followed by puromycin treatment (50 μ g/mL), (Hydroxylamine) treated with 1M hydroxylamine for 30 minutes at RT, which was followed by puromycin treatment (50 μ g/mL). MM indicates molecular marker. (C) The intensity of puromycin signal for each lane was determined using Image J software and was expressed as a percentage relative to the samples solely treated with puromycin that represent 100%, for key selected samples. The results represent a mean percentage \pm SD, derived from three biological replicates.

Results presented in this chapter are summarised in Figure 3.12, where, upon polysome fractionation, I performed key treatments prior to conducting puromycylation, both in ribosome-containing and in light fractions, and ran the samples on the same gel for an easy comparison. As previously shown, ribosome-dependent puromycylation is sensitive to EDTA and anisomycin, while it is unaffected by cycloheximide pre-treatment. Hydroxylamine treatment had no effect on the reaction, confirming that hydroxylamine does not degrade ribosome-bound peptidyl-tRNAs and also does not have any deleterious effects on the proteins in the sample. On the other hand, ribosome-independent puromycylation is only sensitive to hydroxylamine treatment, while treatments that inhibit translation or interrupt ribosome structure have no effect. The data suggests that the nature of the two reactions is distinctly different, ribosome-independent puromycylation probably occurring due to a non-enzymatic chemical reaction between peptidyl-tRNA and puromycin.

A



B



C

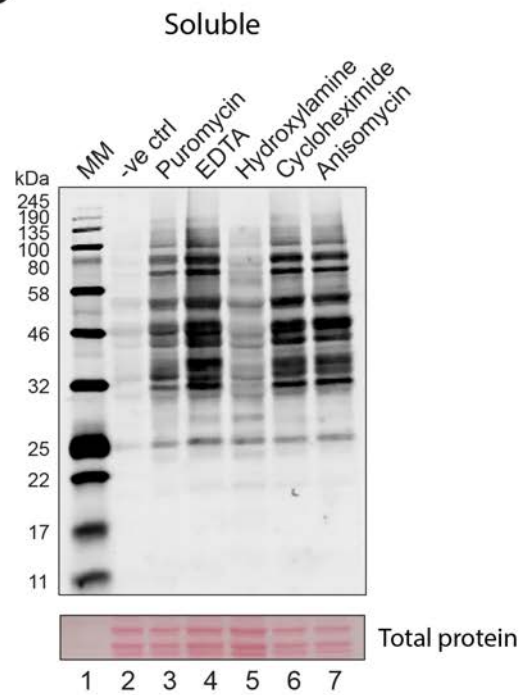


Figure 3.12 – The nature of puromycylation occurring in ribosome-containing and light polysomal fractions is distinctly different. (A) Polysome fractionation trace of wild type yeast cells. The gradient was split into 12 fractions. Fractions 1-6 were combined to generate the polysome sample, while fractions 11 and 12 were combined to generate the light fractions sample. The pools were then split into the required number of samples, 200 μ L for each of the polysomal and 100 μ L for each of the light and treatments were carried out as described below. (B) and (C) Western blotting of protein extracts from (B) polysome or (C) light fraction samples, which were either (-ve ctrl) untreated, (Puromycin) treated with puromycin (50 μ g/mL) for 20 minutes at RT, (EDTA) treated with 30 mM EDTA for 30 minutes at RT, which was followed by puromycin treatment (50 μ g/mL), (Hydroxylamine) treated with 1M hydroxylamine for 30 minutes at RT, which was followed by puromycin treatment (50 μ g/mL), (Cycloheximide) treated with cycloheximide (100 μ g/mL) for 15 minutes prior to puromycin treatment (50 μ g/mL) and (Anisomycin) treated with anisomycin (250 μ g/mL) for 15 minutes prior to puromycin treatment (50 μ g/mL). MM indicates molecular marker. Bottom panel in (C) shows a representative area of Ponceau staining, demonstrating protein loading.

3.2.8 Proteins puromycylated in a ribosome-independent manner represent highly expressed proteins

To investigate which proteins become puromycylated in a ribosome-independent manner, I attempted to optimise conditions for affinity purification of puromycylated peptides, followed by mass spectrometry. As affinity purification was not previously attempted from polysomal fractions, but only from total cell culture, and the α -puromycin antibody was limited, I first used UPF1-FLAG and eIF4AIII-FLAG tagged strains, to optimise the procedure. During optimisation, the main parameters adjusted were salt and detergent concentration, incubation with the beads, and washing steps, yet none of the attempts led to satisfactory affinity purification, as I could not detect the protein of interest in the elution (data not shown). Due to time constraints, further optimisation was not carried out. I believe that increasing the protein concentration used for the experiment would improve the experiment significantly.

Even though the affinity purification was not successful, I decided to try and identify proteins represented by the strongest puromycylated band detected by Western blotting in light fractions (Figure 3.13). The fraction was run on an SDS-PAGE gel, the proteins were visualised using Coomassie Brilliant Blue staining and the band of interest was carefully excised from the gel using a clean scalpel (Figure 3.13, the band indicated by a red rectangle). The proteins were trypsinised and the products subjected to mass spectrometry analysis. The puromycylated peptides were searched by adding the molecular weight of puromycin to the parameters. A non-treated sample served as a negative control. Proteins identified with the highest confidence are listed in the table

(Figure 3.13). Puromycylated peptides were not identified which is not entirely unexpected, as puromycylation was detected by Western blotting that considerably amplifies the signal while puromycylation probably occurs only in a small proportion of the total proteins. Also, none of the identified peptides covered the C-terminus (or close) of the proteins where we mostly expected tRNA to be attached and puromycylation to take place, which could be one of the main reasons why no puromycylation would be detected by mass spectrometry but could still occur.

Identified proteins are highly expressed proteins of the cells. This is expected, as, if my hypothesis is correct, translation of proteins expressed at an elevated level in the cell would be more prone to mistakes and peptidyl-tRNAs would occur more frequently.

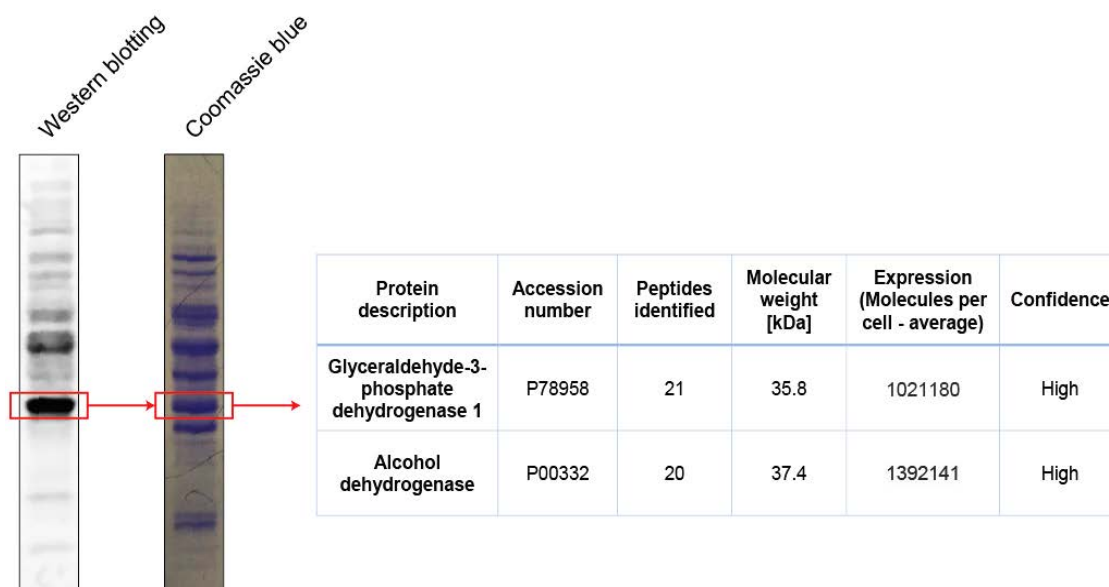


Figure 3.12 – **Polypeptides detected to be puromycylated at a very high level in light fractions represent highly expressed proteins of the cell.** Western blotting of protein extract from light ribosome-free fraction generated by polysome profiling and treated with puromycin (50 $\mu\text{g/mL}$) is shown on the left. Coomassie Brilliant Blue staining of the same sample is shown on the right. The band corresponding to the very strong puromycin signal, indicated by the red rectangle was excised from the Coomassie Blue-stained gel and analysed by mass spectrometry. The experiment was carried out in triplicate. Two proteins of the corresponding molecular weight were identified with high confidence and are described in the table.

3.3 Conclusions

Although the view is changing (Heyer and Moore, 2016), single ribosomes associated with mRNAs are considered to represent primarily initiating ribosomes translating only short peptides, while efficient translation occurs only when mRNAs associate with multiple ribosomes. In contrast, by utilising puromycylation performed upon polysome fractionation, I detected proteins of a wide range of molecular weights, particularly proteins of high molecular weight, in the monosomal sample. The results indicate that even complete polypeptides can be translated by a single ribosome or that a substantial proportion of translating polypeptides remain associated with a single, terminating ribosome.

As puromycylation is a reaction dependent on the ribosome, the prediction was that it would only be detected in ribosome-containing fractions, polysomes and monosomes. Interestingly, approximately 70% of puromycylation performed upon polysome fractionation occurs in the light, ribosome-free fractions. Low signal of puromycylation observed in the monosomal fractions as well as in fractions corresponding to polysomes is translation-dependent as it is impaired by anisomycin pre-treatment. On the other hand, puromycylation detected in light fractions was not affected, suggesting that its nature is different to ribosome-driven puromycylation.

The protein profile of puromycylated polypeptides does not match the total cell protein profile, indicating that puromycin cannot bind covalently to any protein in the cell. I hypothesised that puromycylation in light fraction occurs due to puromycin covalently binding to cellular peptidyl-tRNAs, which could arise due to errors in translation (Cao and Geballe, 1998). Due to the different stability of

peptidyl-tRNAs in acid versus basic conditions (Bresler *et al.*, 1968), I performed pH treatments prior to puromycylation to assess whether this would affect puromycylation. Puromycin reaction carried out at pH 7.5 after a 30-min treatment at an acid or basic pH, demonstrated that ribosome-dependent puromycylation occurs only within the pH range of 5.5-9.5, which is probably due to the ribosome catalytic centre being destroyed at extreme pH conditions. Interestingly, ribosome-independent puromycylation is unaffected by pH treatment and it is not impaired even in extreme pH conditions. A very minor effect was observed, with more puromycylation detected after acid pH and less after basic pH treatment, indicating that polypeptides undergoing puromycylation are more stable at acid and less stable at alkaline pH. These results go in line with my predictions that these polypeptides are in fact peptidyl-tRNAs. The minor effect could be explained by long peptidyl-tRNAs having a very high stability. Additionally, as expected, EDTA treatment was shown to efficiently inhibit ribosome-driven puromycylation (Nolan and Arnstein, 1969), however, it had no effect on ribosome-independent puromycylation. Ribosome-independent puromycylation was also enhanced upon 65°C treatment. These observations suggest that the process is either not enzymatic or that the enzyme activity is unaffected by EDTA or high-temperature treatment. Finally, treatment with hydroxylamine, a chemical that stimulates peptidyl-tRNA degradation (Bresler *et al.*, 1968; JimenezMonro and Vazquez, 1970), strongly reduced ribosome-independent puromycylation, indicating that the puromycylation observed in ribosome-free fractions could be the result of

puromycin binding covalently to peptidyl-tRNAs, occurring during translation, potentially during ribosome stalling.

Highly puromycylated proteins in light fractions were identified as highly expressed proteins. High expression would lead to a higher occurrence of errors and consequently, peptidyl-tRNAs, which would result in high puromycylation, should my hypothesis be correct. Unfortunately, I could not successfully optimise conditions for protein affinity purification of polysomal fractions, to assess the exact nature and location where puromycin incorporates. This will be important to pursue in future, to further test this hypothesis.

The results described in this chapter suggest that the principle of puromycylation, a universally used method to study translation, should be reviewed. Also, proteins synthesised by a single monosome could represent very long polypeptides, changing the view on how translation occurs in cells. Additionally, conclusions from preliminary experiments suggest that a considerable proportion of newly synthesised proteins are released from ribosomes with the last tRNA still attached to them at the C-terminal, and that this linkage can react with puromycin independently of the ribosome.

Chapter 4: Beyond NMD: Role of UPF1 in global translation and stress response

4.1 Introduction

UPF1 is the most evolutionarily conserved protein involved in NMD (Kerényi *et al.*, 2008), an ATP-dependent RNA helicase that associates with mRNAs *in vitro* as well as *in vivo* (Fiorini *et al.*, 2015; Cheng *et al.*, 2007; Zünd *et al.*, 2013). UPF1 binds mRNA indiscriminately, it can contact any sequence but it is not evenly distributed on the RNA because the scanning ribosome would detach it from the coding region and UPF1 would concentrate in the 3'UTR (Zünd *et al.*, 2013). UPF1 interacts with eIF3, which stimulates ribosome recycling upon translation termination. Therefore, UPF1 may be involved in triggering the process of ribosome release upon translation termination via its interaction with eIF3 and via its remodelling activity. UPF3, on the other hand, is a component of the EJC complex (Le Hir *et al.*, 2001), recently shown to play a role in translation termination and stimulating ribosome recycling (Neu-Yilik *et al.*, 2017). This is interesting in the possible context that NMD may occur due to more general roles that NMD factors have in gene expression, specifically in translation termination-related processes. Therefore, I investigated further the mode of UPF1 and UPF3 global mRNA association in *S. pombe*, and assessed their role in translation, which could consequently influence NMD.

I used the polysome fractionation approach paired with Western blotting to look at how both UPF1 and UPF3 associate with mRNAs relative to translation. To investigate whether these factors affect translation globally, and to examine the link between monosomal translation and NMD, I performed polysome

fractionation analysis followed by puromycylation in the resulting fractions in wild type versus UPF1 and UPF3 deletion mutants. The goal was to assess whether translation of any proteins becomes deregulated in the absence of these factors. As proteins translated fully by a single ribosome are more affected by NMD (Heyer and Moore, 2016), I anticipated that the most prominent changes would be observed in monosomal translation. UPF1 was also linked to stress granules and P bodies (Sheth and Parker, 2006; Cherkasov *et al.*, 2015), therefore, to further investigate possible cytoplasmic roles of UPF1 beyond NMD, I biochemically assessed UPF1 association with ribosome-free heavy cytoplasmic granules using polysome fractionation, as well as cytologically, using fluorescent microscopy.

I found that UPF1 associates with translating mRNAs partially, with mode of association being via direct RNA binding and protein interactions, potentially via small ribosomal subunit. UPF1 also associates with free mRNAs as well as with other cytoplasmic compartments. UPF3, on the other hand, is primarily linked with translating mRNAs in the cell. Investigating the role of UPF1 and UPF3 in translation indicates that these factors may regulate translation of a subset of proteins, with proteins synthesised by the monosome being affected the most, particularly in the UPF3 deletion strain, in line with my predictions. Finally, exposing cells to high salt or heat shock stress leads to a reversible re-distribution of UPF1 from the pool of translating mRNAs into cytoplasmic granules, as demonstrated by polysome fractionation and by visible granules detected using GFP-tagged UPF1. Mass spectrometry analysis of these cellular compartments suggests that these are stress granules.

4.2 Results

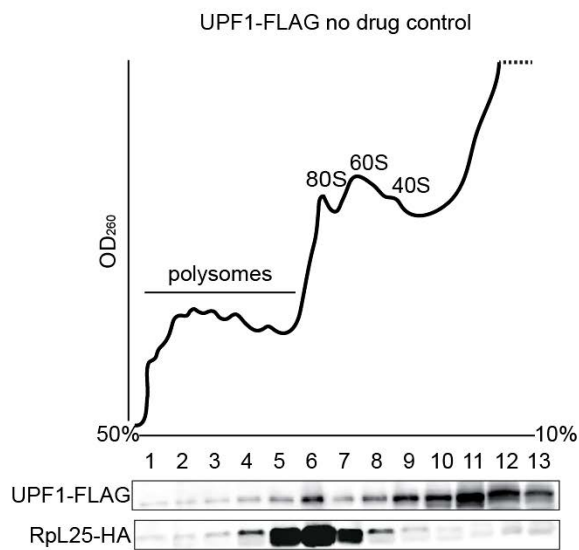
4.2.1 UPF1 associates with translating mRNAs

As reviewed in the Introduction, UPF1 is an ATP-dependent RNA helicase thought to associate with mRNAs globally and to play a key role in NMD by selectively disassembling target mRNPs (FranksSingh and Lykke-Andersen, 2010). I investigated the mode of UPF1 mRNA association globally in *S. pombe* by utilising the polysome fractionation methodology.

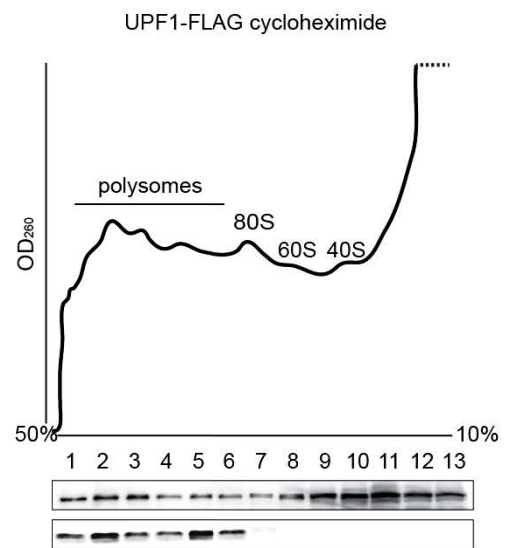
I used a strain in which endogenous UPF1 was FLAG-tagged to perform polysome fractionation coupled with Western blotting to investigate UPF1 mRNA association relative to translation. I also assessed the distribution of Rpl25-HA in the same strain to assess the quality of polysome fractionation. Quantification of UPF1 and Rpl25 distribution expressed as a percentage of the signal observed in ribosome-containing fractions, fractions representing ribosomal subunits as well as light fractions out of the total signal, for all conditions detailed below, is shown in Figure 4.1, panel G.

When the polysome fractionation was carried out in cells that were either untreated with a translation inhibitor, or the polysomes were stabilised using cycloheximide, UPF1-FLAG associated with fractions corresponding to translating ribosomes (from $31 \pm 2\%$ in the no drug sample to $46 \pm 3\%$ when polysomes were stabilised using cycloheximide) (Figure 4.1A; B, lanes 1-7; G), suggesting that only a proportion of UPF1 binds translating mRNAs. A considerable majority of UPF1 is not present in fractions corresponding to translating mRNAs but was instead located in fractions corresponding to individual ribosomal subunits (mostly the 40S). UPF1, however, primarily

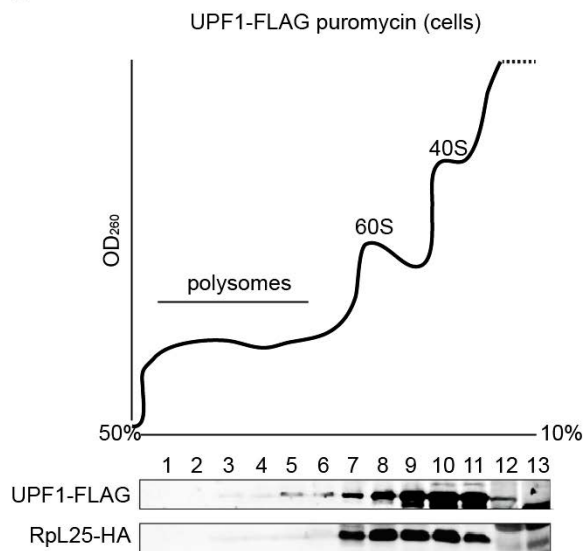
A



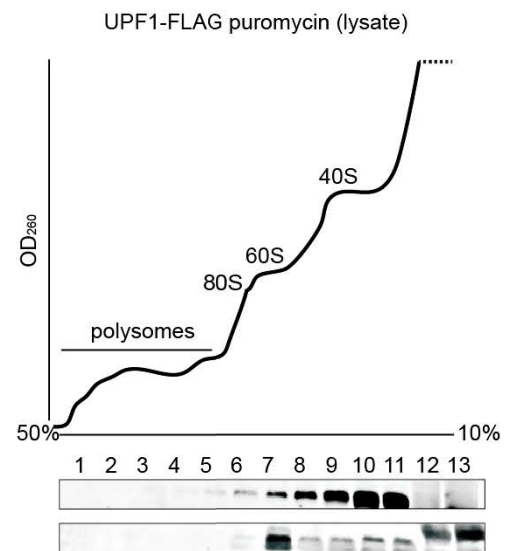
B



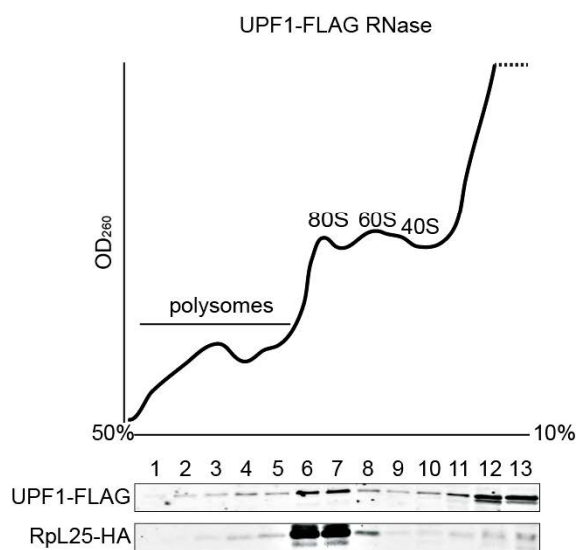
C



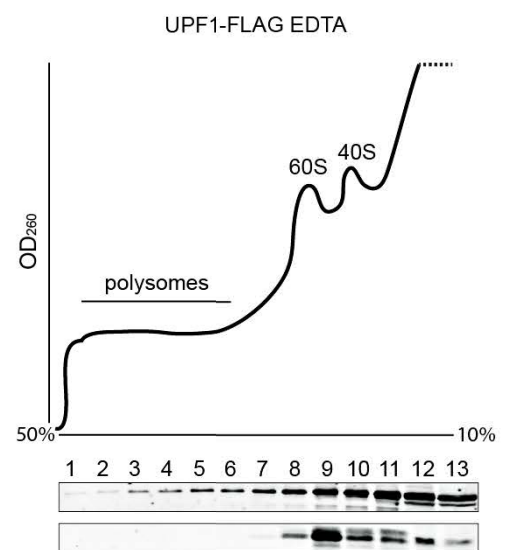
D



E



F



G

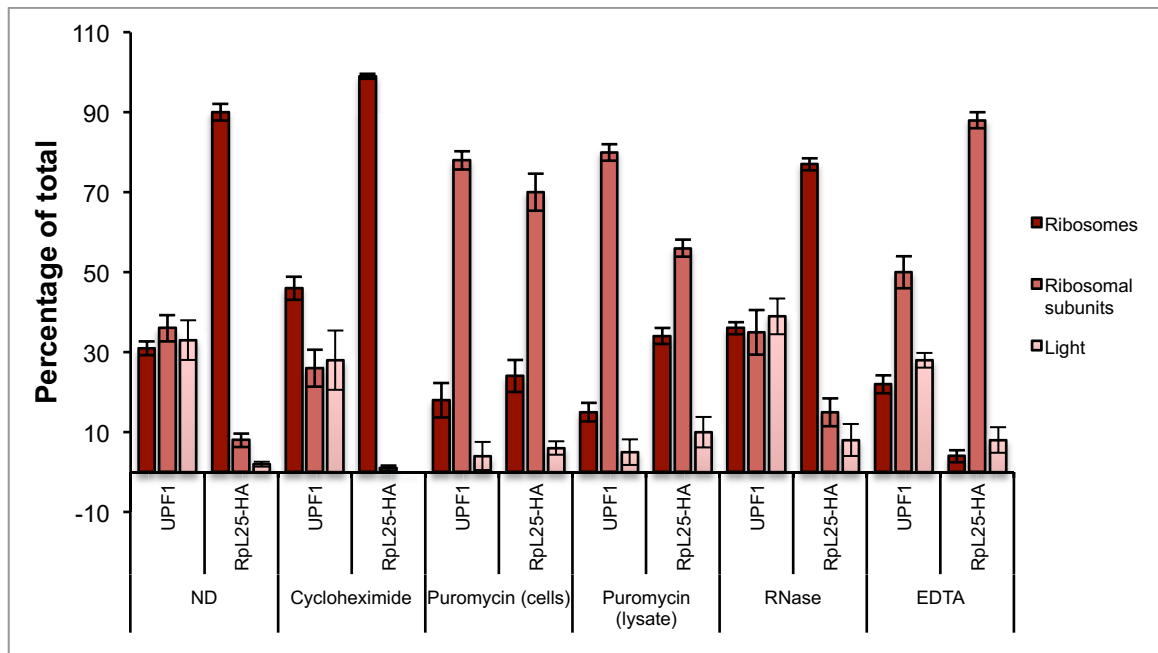


Figure 4.1 – UPF1 only partially associates with translating mRNAs in *S. pombe*.

Representative polysome traces of UPF1-FLAG and RpL25-HA tagged strains in which (A) the sample was untreated with translation inhibitors, (B) the cells were treated with cycloheximide (100 µg/mL) for 5 minutes, (C) the cells were treated with puromycin (250 µg/mL) for 20 minutes, (D) the cell lysate was treated with puromycin (100 µg/mL) on ice for 20 minutes, (E) the lysate was treated with RNase (20 µg/mL) on ice for 30 minutes and (F) the lysate was treated with EDTA (30 mM) on ice for 30 minutes. Traces represent the absorbance profile (OD₂₅₄) recorded during the fractionation. The gradient was split into 13 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. Top panel in each sample shows Western blotting of protein extracts from individual polysome fractionations detecting UPF1-FLAG using an anti-FLAG antibody, while bottom panel shows the RpL25-HA band, detected using an anti-HA antibody. (G) The intensity of UPF1 signal for each lane was determined using Image J software and was expressed as a percentage of the total signal. The percentages from lanes 1 to 6 were added together to give the combined signal for the ribosome-containing fractions, lanes 7-10 were similarly combined to give the total signal for the ribosomal subunits and lanes 11-12 were combined to give the total signal for the light fractions. The means of three independent biological repeats ± SD are demonstrated in the graph.

remains at the top of the gradient with the low molecular weight proteins and small RNAs such as the 5.8S rRNA and tRNAs (Figure 4.1A; B, lanes 8-13). UPF1 distribution did not follow the distribution pattern of the ribosomal protein, which strictly follows that of the ribosome (Figure 4.1A; B, compared in G).

The presence of UPF1 in heavy polysomal fractions in itself only indicates that some of UPF1 associates with heavy cellular compartments. To ensure that these are indeed polysomes, I performed puromycin treatment either to the cells prior to lysis (Figure 4.1C) or to the cell lysate (Figure 4.1D), which, as described in detail in Chapter 3, leads to the dissociation of ribosomal subunits and translation arrest. If UPF1 indeed associates with polysomes, its distribution would shift towards lighter fractions upon puromycin treatment. Ribosomal subunits dissociated from mRNAs in the presence of puromycin can join back together in the low salt conditions of the lysis buffer, resulting in a high 80S peak, consisting of empty ribosomes (Azzam and Algranati, 1973). To prevent this and to keep the subunits separated, I supplemented the sample with 500mM KCl after puromycylation (Blobel and Sabatini, 1971). The polysome trace clearly shows that the polysomes have been efficiently destroyed, resulting in a high peak of ribosomal subunits (Figure 4.1C; D). Dissociation of polysomes was further substantiated by the ribosomal protein distribution. RpL25-HA, a protein of the large ribosomal subunit, shows a clear shift towards light ribosomal fractions, being predominantly present in the fractions that correspond to the large ribosomal subunit (from $56 \pm 2\%$ when the lysate was treated with puromycin to $70 \pm 5\%$ when puromycylation was carried out in the cells and throughout), as expected (Figure 4.1C; D, lanes 7-11; G).

When looking at the UPF1 distribution, on the other hand, I observed a shift towards the light fractions, with the protein, interestingly, following the pattern of primarily the small ribosomal subunit ($80 \pm 2\%$ of UPF1 detected in the fractions corresponding to ribosomal subunits in both cases) (Figure 4.1C; D, lanes 7-11; G). UPF1, therefore, potentially forms interactions with the ribosomal subunits (mainly with the small ribosomal subunit) that are preserved upon puromycylation. Alternatively, this could indicate ribosome-independent complexes with which UPF1 associates.

To investigate whether the association of UPF1 is driven via RNA binding or protein interactions, I performed a mild RNase treatment to the cell lysate prior to performing fractionation. If UPF1 predominantly binds the RNA, due to RNA clipping, I would observe a considerable shift of UPF1 towards lighter fractions. The polysome trace demonstrated that the RNase treatment led to considerable dissociation of the polysomes while the 80S and individual ribosomal subunits peaks were higher (Figure 4.1E). The trace was mirrored by the ribosomal protein distribution, with its considerable shift towards lighter fractions when compared to the distribution when polysomes are stabilised with cycloheximide (Figure 4.1E, lanes 1-6, compared to corresponding lanes in B; G – from $99 \pm 0.6\%$ (cycloheximide-treated sample) to $77 \pm 2\%$ (RNase-treated sample) signal present in the polysomes and from $1 \pm 0.6\%$ to $15 \pm 4\%$ in the ribosomal subunits). The highest signal was present in the fractions corresponding to the monosome (Figure 4.1E, lanes 6-7; G). UPF1 distribution was mirroring the polysome breakdown, and, when looking at individual fractions, a very minor

shift of UPF1 was also noted, from the heaviest polysomal fraction (Figure 4.1E, lane 1), suggesting that UPF1, in part, associates directly to mRNAs.

Finally, I performed an EDTA treatment to the cell lysate, since EDTA, as detailed in Chapter 3, chelates Mg^{2+} ions and leads to ribosome and, consequently, polysome dissociation (Nolan and Arnstein, 1969). Indeed, EDTA treatment did lead to polysome destruction, demonstrated by the polysome trace where the polysome peaks have disappeared, and the peaks of ribosomal subunits were more prominent (Figure 4.1F), as well as by the complete shift of Rpl25-HA towards the lighter fractions ($90 \pm 2\%$ of the signal present in fractions corresponding to ribosomal subunits) (Figure 4.1F, lanes 8-10; G). Although I observed that UPF1 also shifts towards the light fractions, it was less prominent with a considerable amount still present in heavy cellular compartments ($22 \pm 3\%$ remaining in heavy fractions while $50 \pm 4\%$ associating with ribosomal subunit-containing fractions) (Figure 4.1, panel F, fractions 3-8; G).

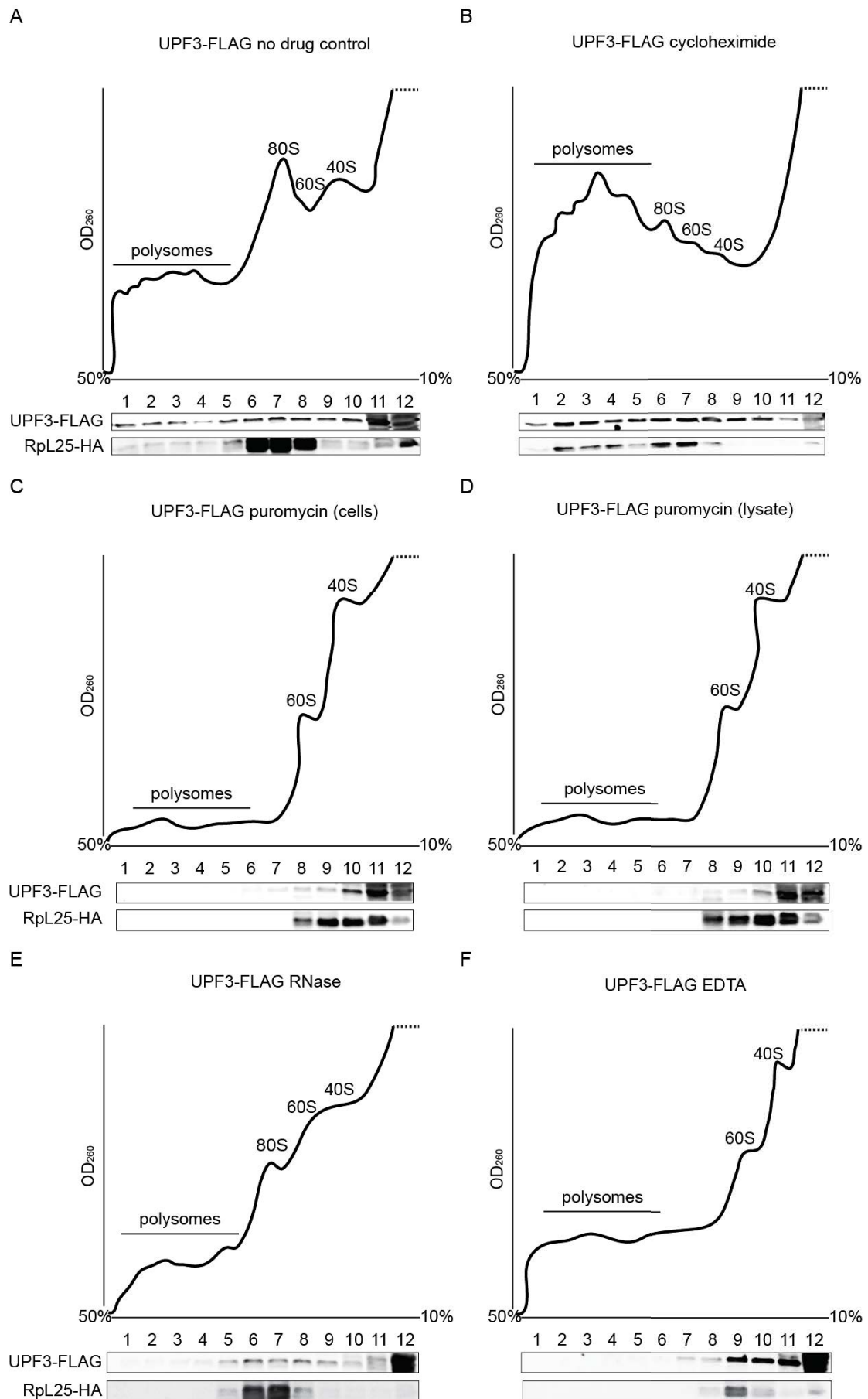
Altogether, the results suggest that UPF1 only partially associates with translating mRNA, which is in part driven by its binding to mRNA, and in part by associating with ribosomal subunits. However, a considerable proportion of UPF1 is present in the cell independently of translating ribosomes, associating with free mRNAs, co-fractionating with ribosomal subunits as well as with heavy ribosome-free cellular compartments.

4.2.2 UPF3 associates with translating mRNAs

UPF3 is a component of the EJC complex, playing a key role in NMD (Le Hir *et al.*, 2001). Recent reports indicate that its role also lies in the later stages of translation termination and ribosome recycling (Neu-Yilik *et al.*, 2017). I investigated the global association of UPF3 relative to translation by using the polysome fractionation method, as described in Chapter 4.2.1. Quantification of UPF3 and RpL25 distribution expressed as a percentage of the signal observed in ribosome-containing fractions, fractions representing ribosomal subunits as well as light fractions out of the total signal, for all conditions detailed below, is shown in Figure 4.2, panel G.

Most of UPF3 associates with the heavy polysomal fractions both in the untreated control and in the cycloheximide-treated sample ($69 \pm 2\%$ of the signal in untreated; $67 \pm 2\%$ in cycloheximide-treated sample) (Figure 4.2A; B, lanes 1-7; G). It is also present in fractions corresponding to ribosomal subunits ($25 \pm 3\%$ in untreated, $29 \pm 2\%$ in cycloheximide) (Figure 4.2A; B, lanes 8-10; G) while low signal was observed in fractions containing free mRNAs (Figure 4.2A, lanes 11-12; G), suggesting that the protein is mostly present in translating mRNA complexes.

Again, to confirm that the UPF3 detected in heavy polysomal fractions is indeed bound to polysomes, I performed puromycin treatment either to the cells or to the cell lysate prior to fractionation, to dissociate polysomes (Figure 4.2C; D). The polysomes were broken down in both cases, demonstrated by high 60S and 40S peaks of the polysome trace and the complete RpL25-HA shift to the fractions corresponding to the large ribosomal subunits (84-90% of the signal)



G

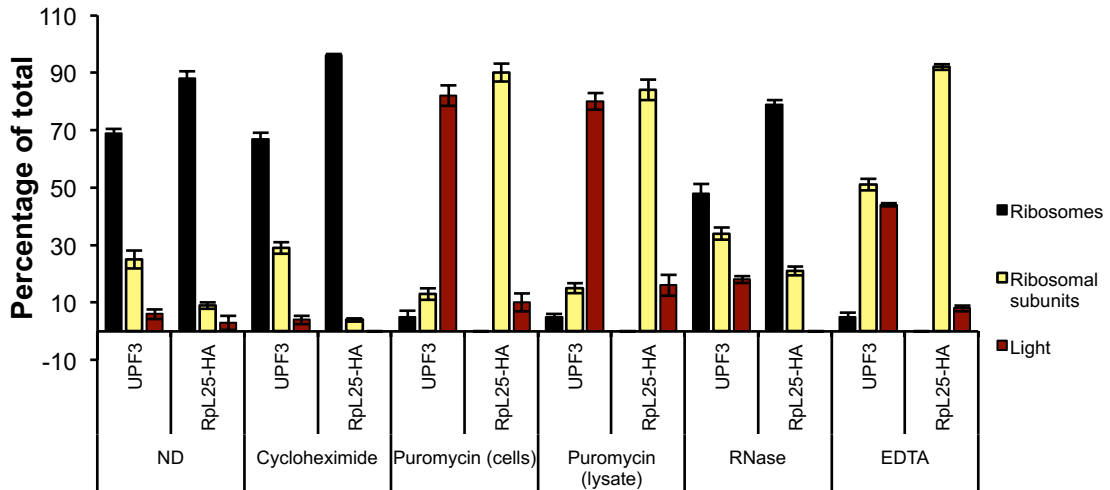


Figure 4.2 – **UPF3 associates with translating mRNAs in *S. pombe***. Representative polysome traces of UPF3-FLAG and Rpl25-HA tagged strains in which (A) the sample was untreated with translation inhibitors, (B) the cells were treated with cycloheximide (100 $\mu\text{g/mL}$) for 5 minutes, (C) the cells were treated with puromycin (250 $\mu\text{g/mL}$) for 20 minutes, (D) the lysate was treated with puromycin (100 $\mu\text{g/mL}$) on ice for 20 minutes, (E) the lysate was treated with RNase (20 $\mu\text{g/mL}$) on ice for 30 minutes and (F) the lysate was treated with EDTA (30 mM) on ice for 30 minutes. Traces represent the absorbance profile (OD_{254}) recorded during the fractionation. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. Top panel in each sample shows Western blotting of protein extracts from individual polysome fractionations detecting UPF3-FLAG using an anti-FLAG antibody, while bottom panel shows the Rpl25-HA band, detected using an anti-HA antibody. (G) The intensity of UPF3 signal for each lane was determined using Image J software and was expressed as a percentage of the total signal. The percentages from lanes 1 to 6 were added together to give the combined signal for the ribosome-containing fractions, lanes 7-10 were similarly combined to give the total signal for the ribosomal subunits and lanes 11-12 were combined to give the total signal for the light fractions. The means of three independent biological repeats \pm SD are demonstrated in the graph.

(Figure 4.2C; D, lanes 8-11; G), as expected. UPF3 re-distributed primarily into light fractions with polysome breakdown (80% of the total signal) (Figure 4.2C; D, lanes 10-12; G) with very low signal detectable within ribosomal subunits, primarily bound to the 40S (15%) (Figure 4.2C; D, lanes 7-10; G), suggesting that its presence in heavy cytoplasmic compartments is due to its association either directly with translating mRNAs or with mRNP-associated factors other than ribosomes.

Upon mild RNase treatment, again, I observed minor polysome breakdown (Figure 4.2E), resulting in a significant shift of UPF3 from heavy polysomal fractions (Figure 4.2E, lanes 1-4) – from $69 \pm 2\%$ in cycloheximide-treated sample to $48 \pm 3\%$ in RNase-treated sample of the total signal present in polysomes and monosomes (Figure 4.2G). EDTA treatment resulted in a shift of UPF3 towards light fractions ($51 \pm 2\%$ of the signal present in the ribosomal subunits-containing and $44 \pm 1\%$ present in light fractions) (Figure 4.2F, lanes 9-12: G), further demonstrating that this protein associates predominantly with translating mRNAs in the cytoplasm.

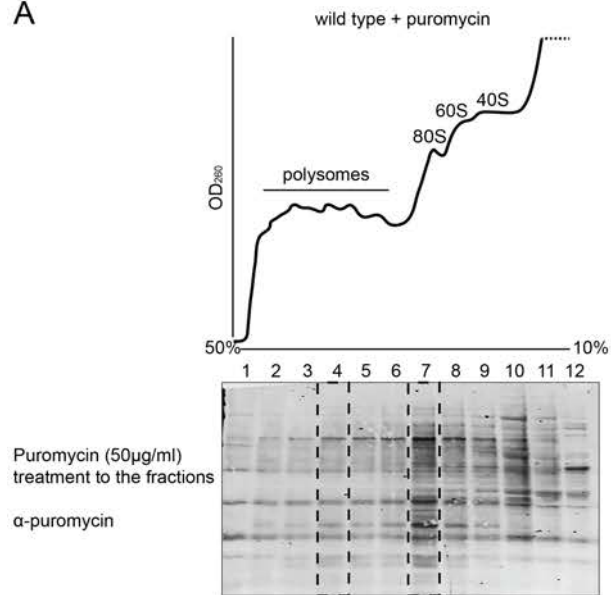
4.2.3 UPF1 and UPF3 have a role in translation efficiency

As NMD particularly affects mRNAs that are fully translated by a single ribosome, I aimed to investigate the link between monosomal translation and NMD further (Heyer and Moore, 2016). I hypothesised that in the absence of the NMD factors, UPF1 and UPF3, translation of some transcripts may be altered, particularly those translated by a single ribosome.

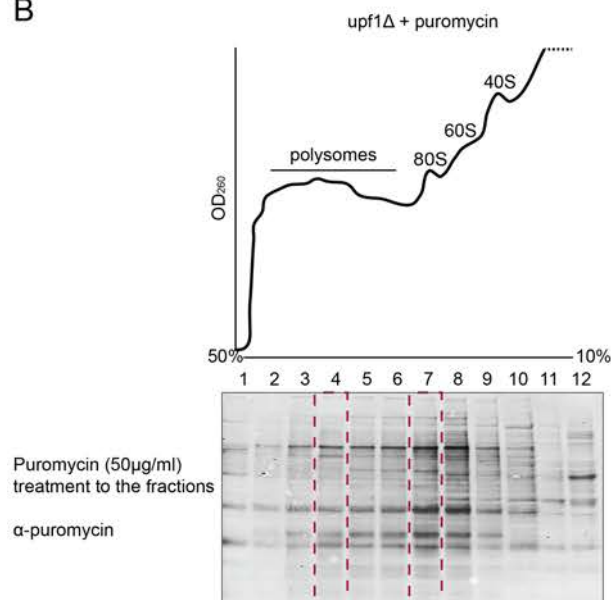
To address these questions, I investigated proteins synthesised by single versus multiple ribosomes between wild type, *upf1*Δ and *upf3*Δ strains, by performing polysome fractionation followed by puromycylation (as detailed in Chapter 3) (Figure 4.3A; B; C). When comparing the proteins synthesised by polysomes and monosomes between the three strains (Figure 4.3A; B; C, lanes that are compared are labelled with rectangles with dashed borders), I identified the bands that differ, by visual analysis (Figure 4.3A; B; C) and by comparing the representative lane intensity profiles generated using ImageJ software (Figure 4.3D; E; F; G). In particular, when looking at the monosomal translation, a number of bands were different, particularly in the *upf3*Δ strain, suggesting that monosomal translation is mostly affected/regulated by the NMD factors (Figure 4.3F; G).

The results suggest that the two proteins may have a role in regulating translation of a subset of cellular proteins, particularly those that are synthesised by a single ribosome. It would be interesting to identify which proteins these are.

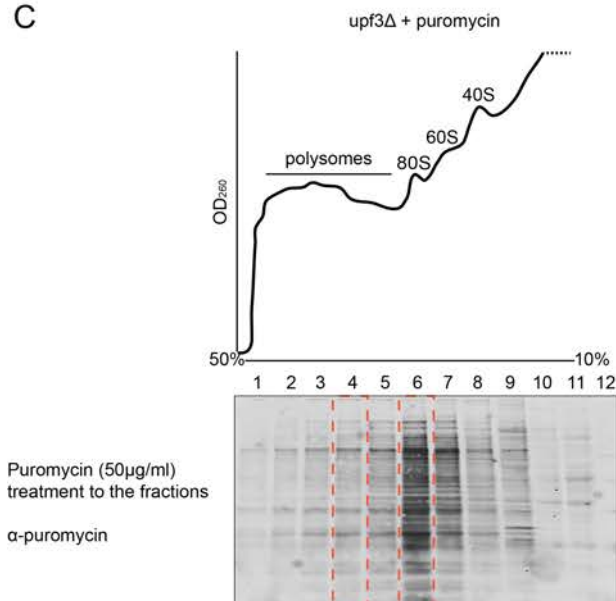
A

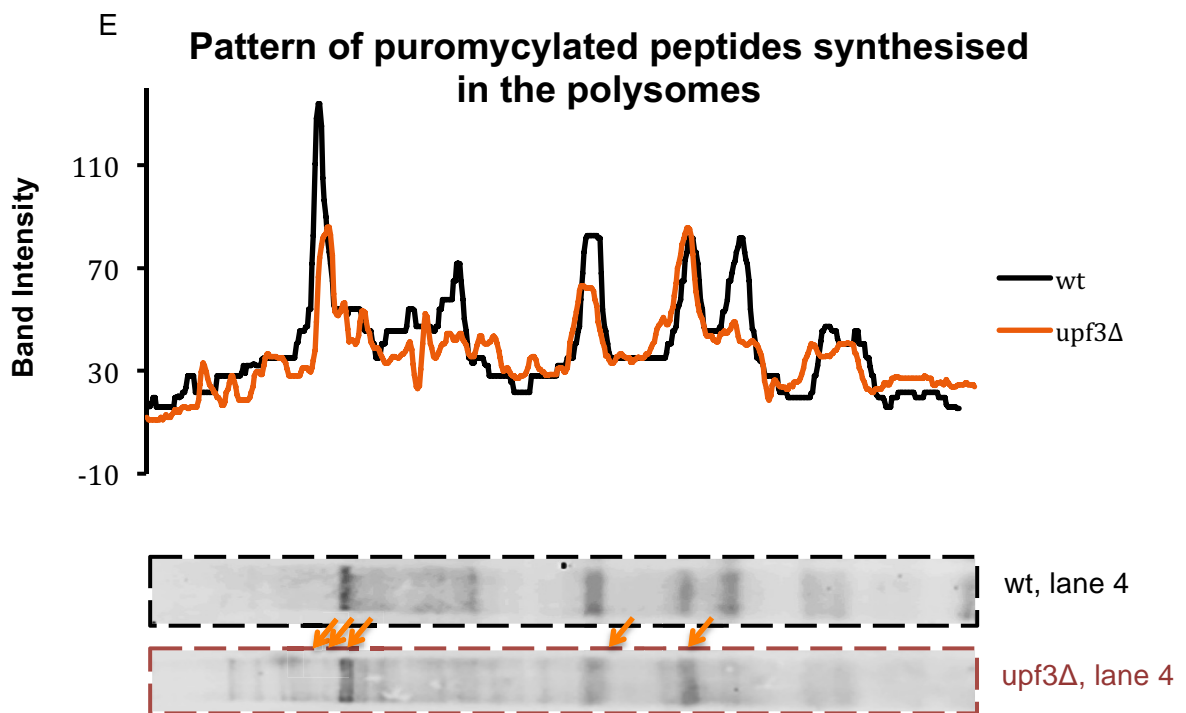
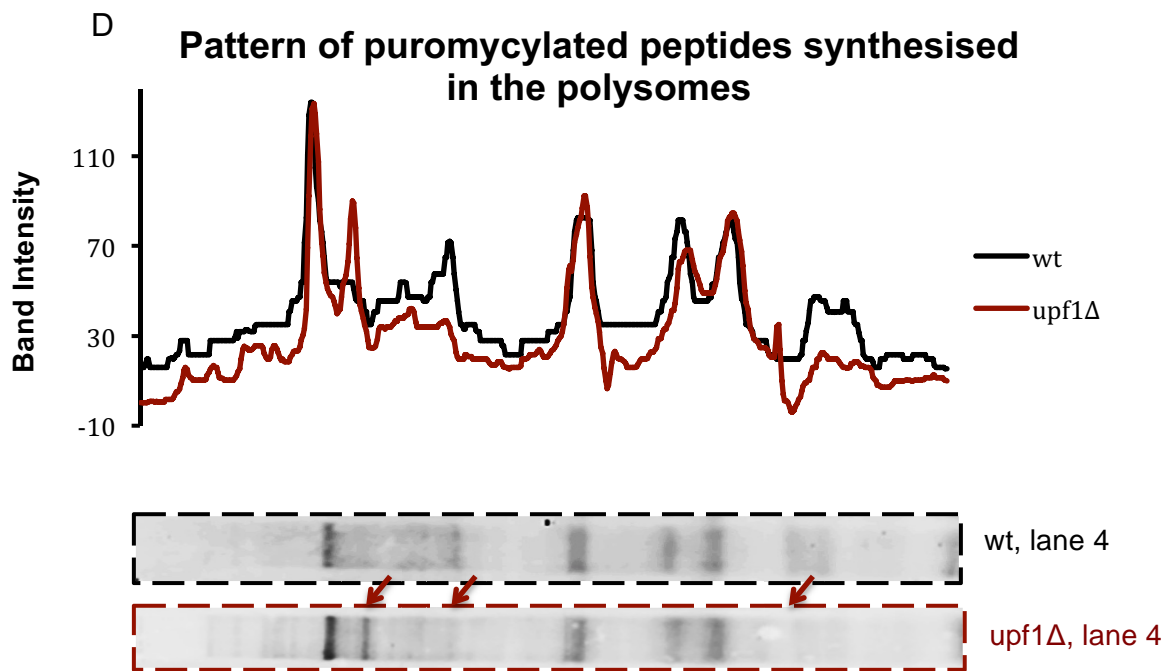


B



C





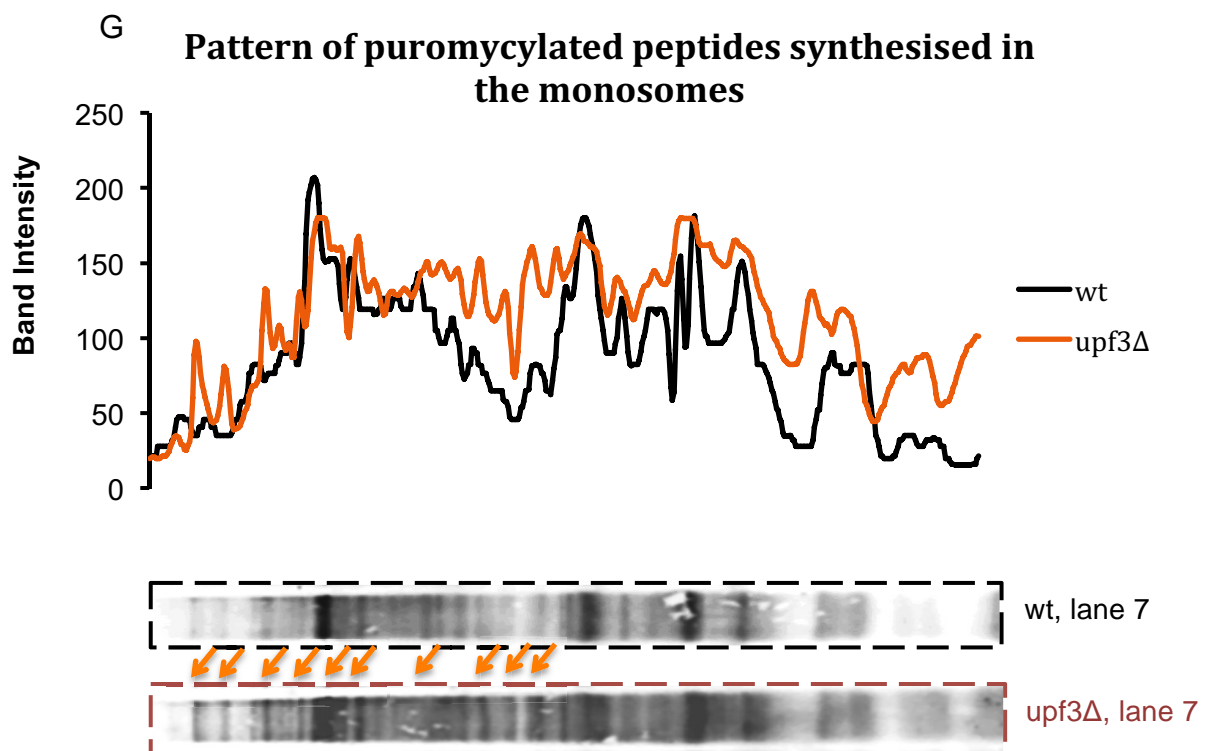
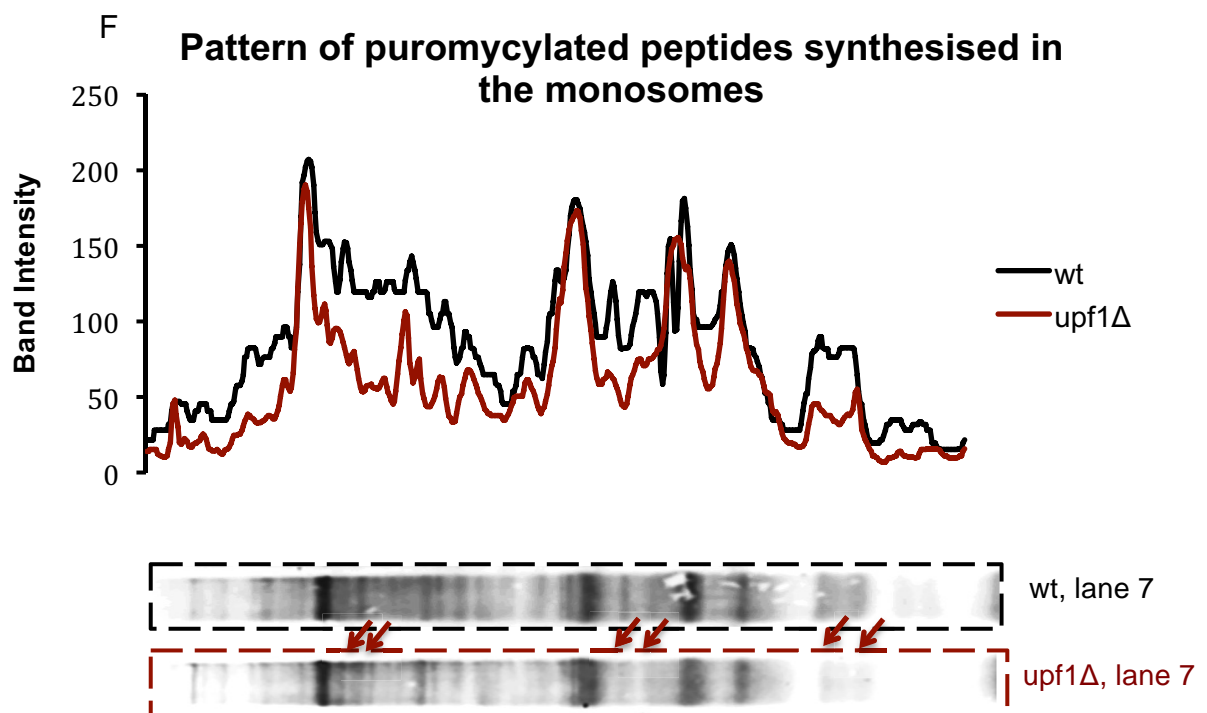


Figure 4.3 – NMD factors might have a minor effect in translation of individual proteins. Representative polysome fractionation traces of cell extracts from **(A)** wild type, **(B)** *upf1Δ* and **(C)** *upf3Δ* strains. Traces represent the absorbance profiles (OD₂₅₄) recorded during the fractionation. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. Below each of the traces, its corresponding Western blot is shown, derived from protein extracts of individual polysomal fractions, which were treated with puromycin (50 µg/mL) immediately upon collection, and the signal was detected using an α-puromycin antibody. The lanes of which the intensity profiles were compared between the strains (analysed in D-G) are indicated with rectangles with dashed borders – black (wild type), red (*upf1Δ*) and orange (*upf3Δ*). **(D)** Intensity lane profiles of puromycylated peptides synthesised by polysomes, compared between wild type (black) and *upf1Δ* (dark red) strain. The key bands/areas containing proteins that are differentially translated between the two strains are indicated with red arrows. **(E)** Intensity lane profiles of puromycylated peptides synthesised by polysomes, compared between wild type (black) and *upf3Δ* (orange) strain. The key bands/areas containing proteins that are differentially translated between the two strains are indicated with orange arrows. **(F)** Intensity lane profiles of puromycylated peptides synthesised by monosomes, compared between wild type (black) and *upf1Δ* (dark red) strain. The key bands/areas containing proteins that are differentially translated between the two strains are indicated with red arrows. **(G)** Intensity lane profiles of puromycylated peptides synthesised by monosomes, compared between wild type (black) and *upf3Δ* (orange) strain. The key bands/areas containing proteins that are differentially translated between the two strains are indicated with orange arrows.

4.2.4 UPF1 is directed into heavy ribosome-free cytoplasmic compartments upon treatment with high salt

As observed in Figure 4.1, UPF1 binds to translating mRNAs, however, it also localises in heavy cellular compartments that are void of ribosomes (Figure 4.1C; D; F). I investigated under which conditions UPF1 becomes strongly re-distributed from the pool of translating mRNAs into the detected heavy cytoplasmic granules. Furthermore, I attempted to identify their composition.

I hypothesised that the granules in question could potentially be stress granules or P bodies, to which UPF1 has previously been brought into connection (Sheth and Parker, 2006; Hubstenberger *et al.*, 2017; Cherkasov *et al.*, 2015). Therefore, I first tested the effect of stress on UPF1 distribution relative to translation by treating the cells with high salt (500 mM KCl) during the experiment. To eliminate the possibility that any UPF1 present in heavy polysomal fractions is due to UPF1 associating with polysomes, I treated the cell lysate with puromycin to induce polysome breakdown (Figure 4.4B). Under these conditions, RpS2801, a protein of the small ribosomal subunit, completely shifts into fractions corresponding to ribosomal subunits ($97 \pm 1\%$) (Figure 4.4B, lanes 9-11; C), demonstrating that puromycin treatment efficiently destroys polysomes. On the other hand, UPF1 does not shift to the light fractions, but remains stably associated with heavy fractions, which are now void of ribosomes ($58 \pm 2\%$ of the total signal detected in heavy fractions) (Figure 4.4B, fractions 1-6; C). These observations suggest that UPF1 indeed re-localises from translating mRNAs into heavy ribosome-free cytoplasmic compartments upon high salt treatment.

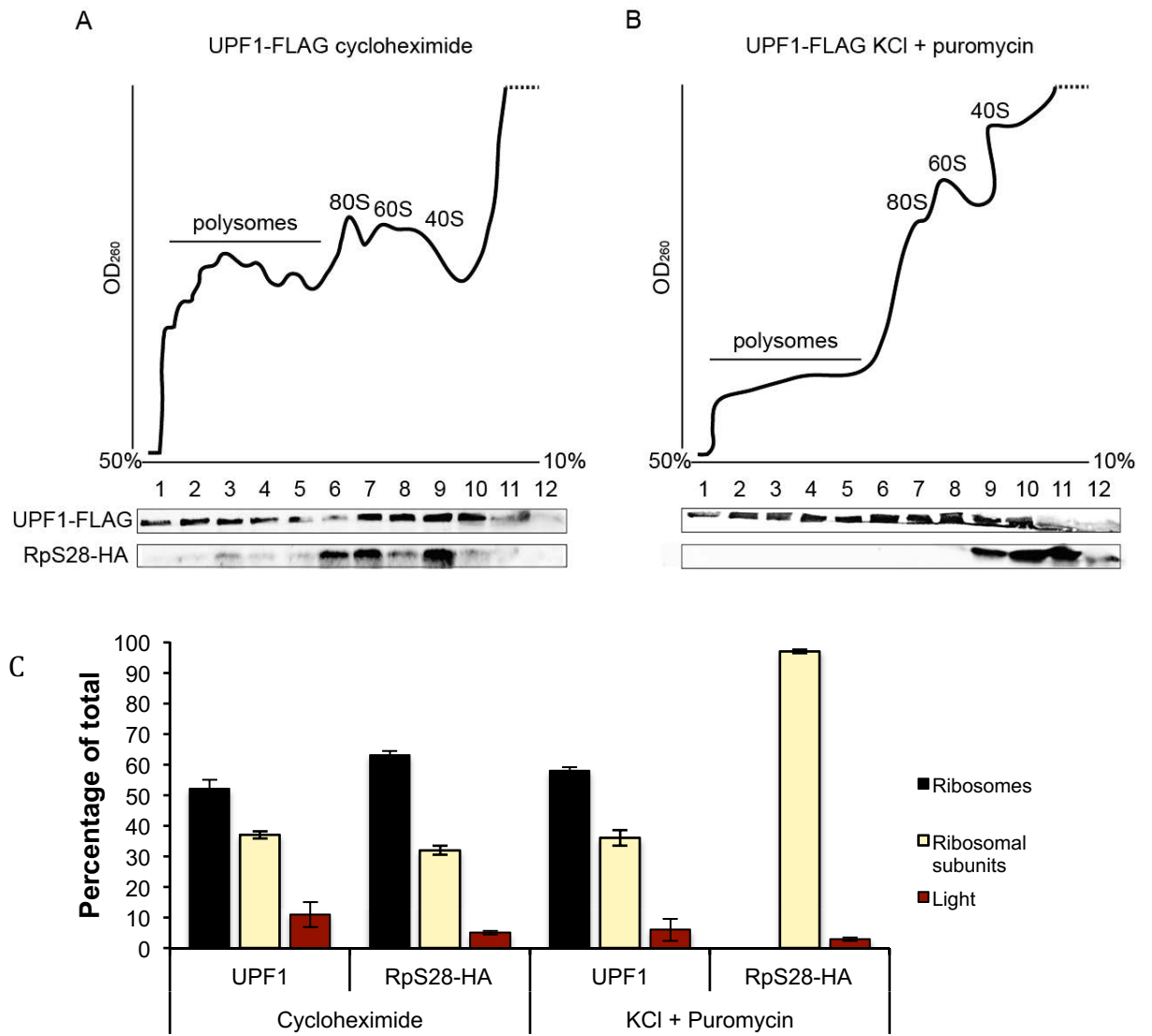


Figure 4.4 – UPF1 localises in heavy, ribosome-free cytoplasmic compartments upon high-salt stress. Polysome traces of UPF1-FLAG and RpS2801-HA tagged strains in which (A) the cells were treated with cycloheximide (100 µg/mL) for 5 minutes, (B) the cells were treated with 500 mM potassium chloride and throughout polysome fractionation and the lysate was treated with puromycin (100 µg/mL) for 20 minutes. Traces represent the absorbance profile (OD₂₅₄) recorded during the fractionation. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. Top panel in each sample shows Western blotting of protein extracts from individual polysome fractionations detecting UPF1-FLAG using an anti-FLAG antibody, while bottom panel shows the RpS28-HA band, detected using an anti-HA antibody. (C) The intensity of UPF1 signal for each lane was determined using Image J software and was expressed as a percentage of the total signal. The percentages from lanes 1 to 6 were added together to give the combined signal for the polysomal fractions, lanes 7-10 were similarly combined to give the total signal for the ribosomal subunits and lanes 11-12 were combined to give the total signal for the light fractions. The means of three independent biological repeats ± SD are demonstrated in the graph.

4.2.5 Over-expressed UPF1-GFP shows the same distribution pattern as that of endogenous UPF1-FLAG

As I demonstrated biochemically that UPF1-FLAG localises in heavy cellular fractions upon stress (Figure 4.4B), I intended to further substantiate this observation by visualising the distribution of UPF1 microscopically under these stress conditions. A strain in which UPF1-GFP is episomally expressed had already been designed in the Brogna lab (Sandip De PhD thesis, 2011). Therefore, I visualised UPF1 by monitoring GFP fluorescence and comparing UPF1 distribution in normal versus stress conditions. Prior to performing microscopy, since UPF1 is overexpressed in this strain, I checked whether its distribution upon polysome fractionation mirrors that of endogenous UPF1, to assess whether I could extrapolate any results I obtain to the endogenous gene.

In the control sample, untreated with translation inhibitors, as well as in the cycloheximide treated sample, UPF1-GFP is present throughout polysomal fractions, reflecting the distribution of the endogenous gene (Figure 4.5A; B; G compared to Figure 4.1A; B; G). In parallel, I carried out high salt (500 mM KCl throughout) (Figure 4.5C) as well as heat shock (30 minutes at 39°C) (Figure 4.5D) treatments upon which I detected a strong accumulation of UPF1-GFP in heavy cytoplasmic fractions (Figure 4.5C; D, lanes 1-6; E), particularly upon high salt treatment (Figure 4.5D, lanes 1-6; E), despite the polysome breakdown observed in polysome traces in both cases. The observations suggest that UPF1-GFP is directed to heavy cytoplasmic compartments, which are void of ribosomes, upon stress conditions.

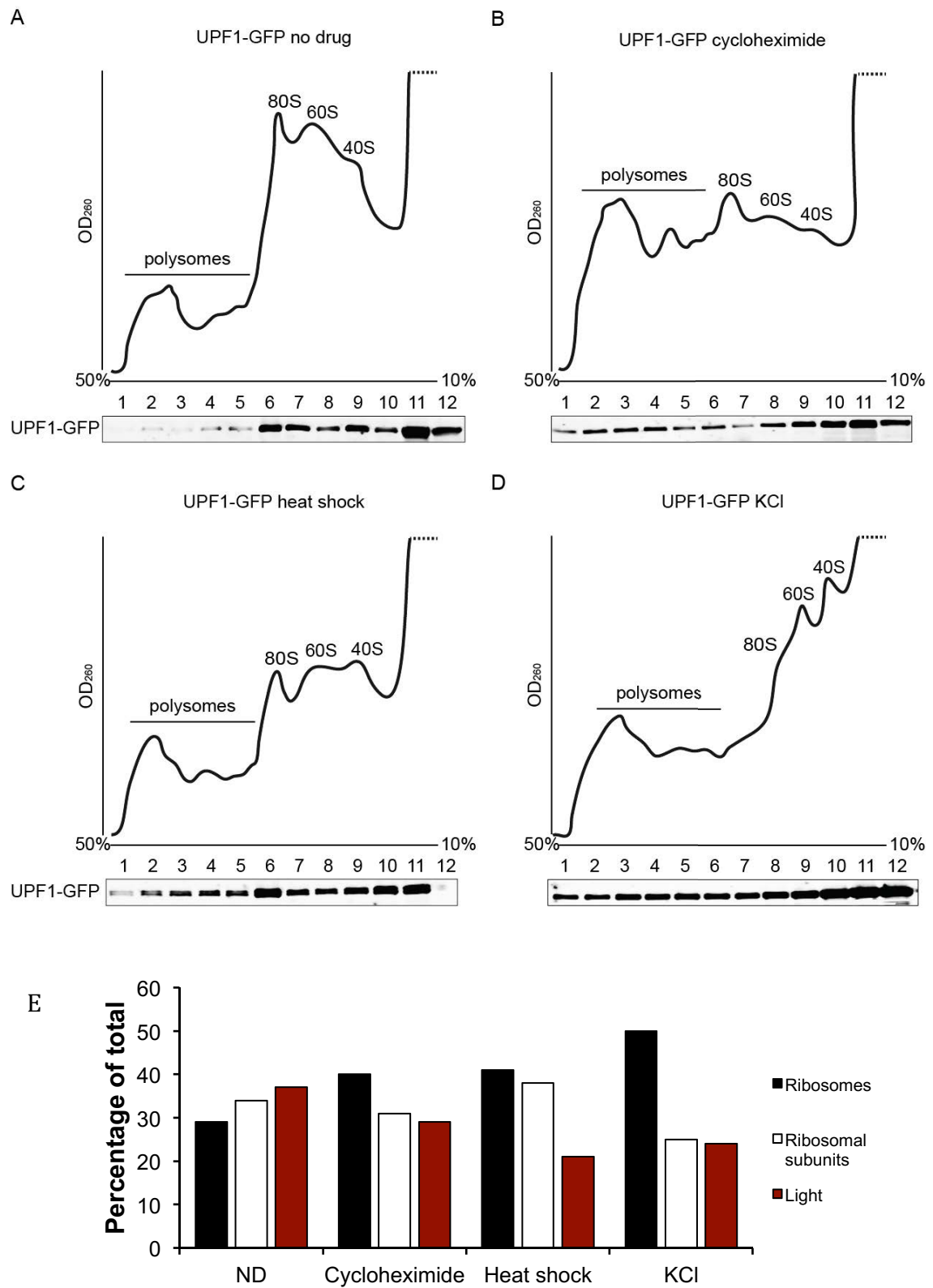


Figure 4.5 – Episomally expressed UPF1-GFP polysome distribution pattern mimics the endogenous UPF1 distribution. Polysome traces of an episomally-expressed UPF1-GFP tagged strain in which (A) the sample was untreated with translation inhibitors, (B) the cells were treated with cycloheximide (100 µg/mL) for 5 minutes, (C) the cells were heat-shocked at 39°C for 30 minutes, (D) the cells were treated with 500 mM KCl for 20 minutes and throughout polysome fractionation. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. The panel below each of the traces shows Western blotting of protein extracts from individual polysome fractionations, detecting UPF1-GFP using an anti-GFP antibody. (E) The intensity of UPF1 signal for each lane was determined using Image J software and was expressed as a percentage of the total signal. The percentages from lanes 1 to 6 were added together to give the combined signal for the polysomal fractions, lanes 7-10 were similarly combined to give the total signal for the ribosomal subunits and lanes 11-12 were combined to give the total signal for the light fractions. The results are summarised in the graph.

4.2.6 UPF1-GFP localises in cytoplasmic granules upon heat-shock treatment

Since UPF1-GFP has the same distribution pattern upon polysome fractionation as endogenous UPF1-FLAG, I used the strain to assess UPF1 distribution in the cell, visualising GFP fluorescence using fluorescent microscopy, in physiological and stress conditions.

UPF1-GFP is homogenously distributed in the cytoplasm of untreated cells (Figure 4.6A, 1), and cycloheximide-treated cells (Figure 4.6A, 2). Upon high salt (500 mM KCl) treatment for 15 minutes, I did not observe any obvious accumulation of UPF1 in cytoplasmic foci (Figure 4.6A, 3). This could be due to the treatment being mild, resulting in changes that are not visible via fluorescent microscopy. Heat shock treatment (30 minutes at 39°C) in turn led to strong accumulation of UPF1-GFP in distinct cytoplasmic foci (Figure 4.6A, 4). This suggests that UPF1 detected in heavier polysomal fractions in Figure 4.5C is associated with cytoplasmic granules under heat shock conditions, which could indicate that UPF1 found in heavier polysomal fractions under high-salt stress conditions (Figures 4.4B and 4.5D) is also associated with such granules.

A member of the Brogna lab has investigated the same phenomenon in *D. melanogaster* S2 cells and observed the same pattern of UPF1 distribution (Anand Singh, unpublished). From being homogenously distributed in the cytoplasm under normal conditions (Figure 4.6B, top images), UPF1 re-localises into distinct granules under heat shock (Figure 4.6B, bottom images), indicating that the UPF1 distribution in response to stress is conserved from yeast to fruit fly.

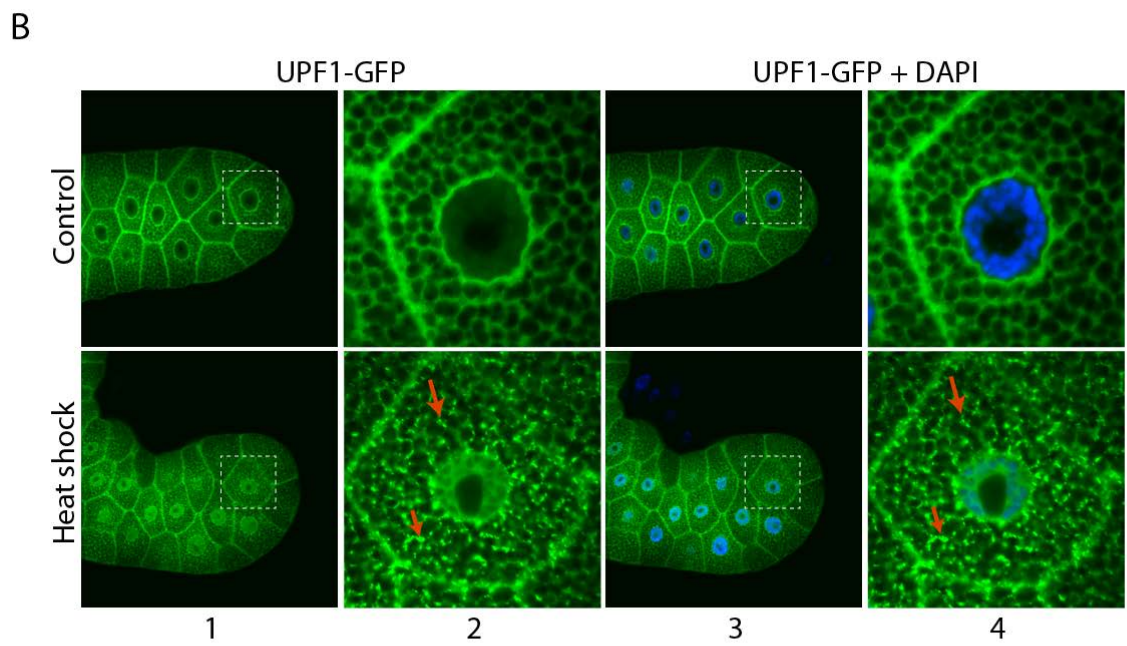
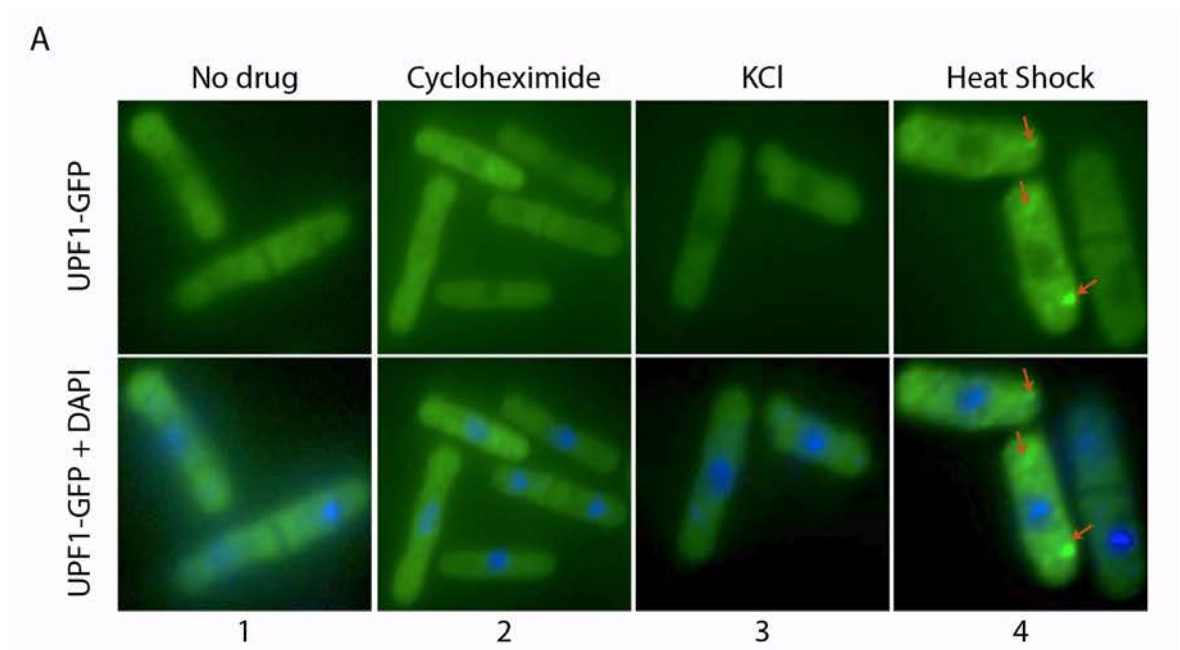


Figure 4.6 – **UPF1-GFP localises in cytoplasmic granules upon heat-shock.** (A) UPF1 GFP fluorescence (green) of *S. pombe* control cells (1), cycloheximide-treated cells (2), cells treated with 500 mM KCl for 15 minutes (3) and cells heat-shocked at 39°C for 30 minutes (4), as indicated. Counter-staining of chromosomes with DAPI (blue), merged with the GFP signal, is shown in the bottom row. Observed UPF1 granules are indicated with red arrows. (B) UPF1 GFP fluorescence of *D. melanogaster* S2 cells under control conditions (top) and under heat shock treatment at 42°C for 30 minutes. Counter-staining of chromosomes with DAPI (blue), merged with the GFP signal, is shown on the right, as indicated. Observed UPF1 granules are indicated with red arrows (Anand Singh, unpublished data).

4.2.7 UPF1-GFP localises in cytoplasmic granules under higher concentrations of KCl

As I did not visualise any accumulations of UPF1 in distinct foci upon 500 mM KCl treatment, I hypothesised that the treatment is too mild, and the granules observed biochemically in polysome fractionation are probably too small to be detected using fluorescent microscopy. Therefore, I performed a time-course treatment using higher salt concentrations (Figure 4.7). Indeed, upon 1 M KCl (Figure 4.7, 2nd row images) and particularly upon 2 M KCl treatment (Figure 4.7, 3rd row images), a clear accumulation of UPF1 in cytoplasmic granules was visualised. Prolonged 30-min treatment led to increased granule formation (Figure 4.7, 3). The observation suggests that the effects are more prominent, and the granules to which UPF1 associates increase in size, the more severe the stress conditions are.

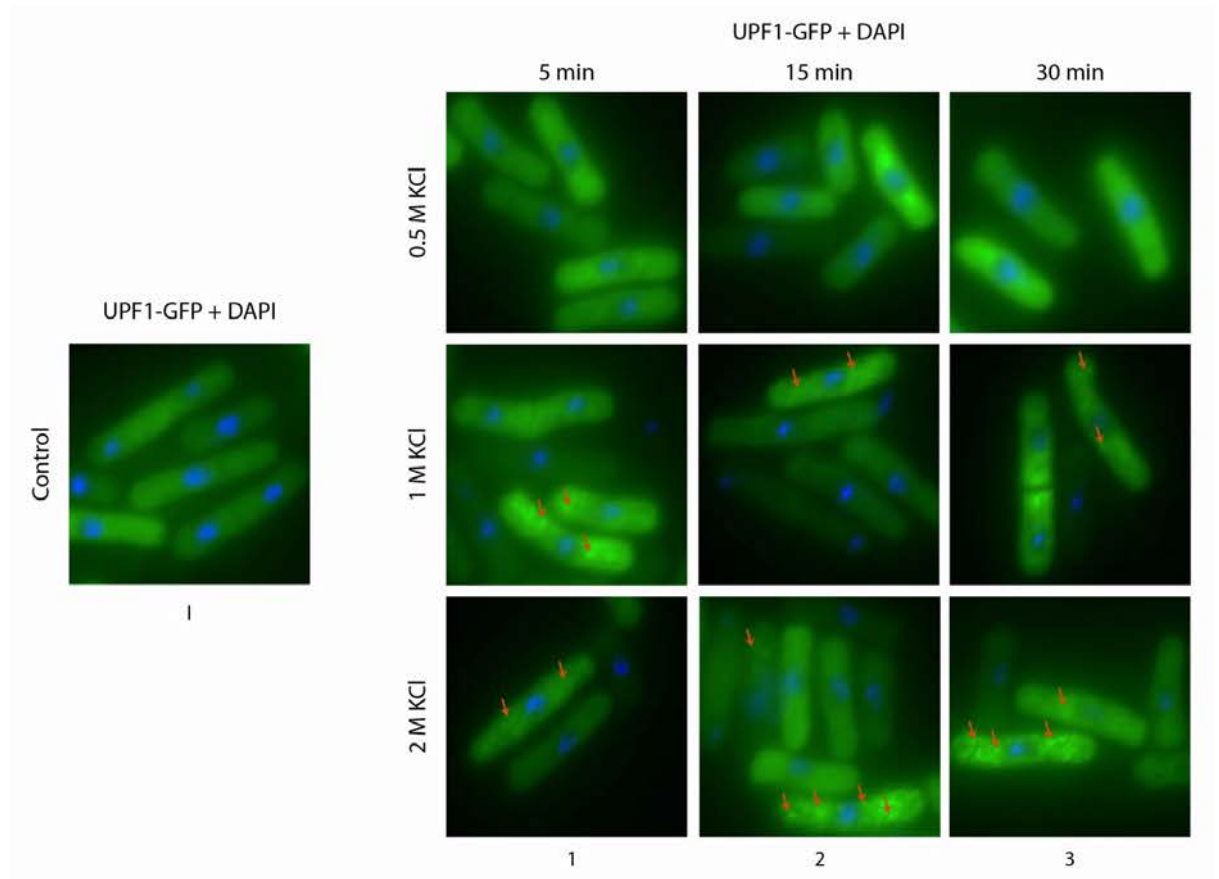


Figure 4.7 – **UPF1-GFP localises in cytoplasmic granules upon higher concentrations of KCl.** (I) UPF1-GFP fluorescence (green) of *S. pombe* control cells; UPF1-GFP fluorescence of cells treated with 0.5 M KCl (top row), 1 M KCl (middle row) and 2 M KCl (bottom row) carried out for 5 min (1), 15 min (2) and 30 min (3), prior to fixation and GFP visualisation, as indicated. Chromosome DAPI counter-staining (blue) was carried out in all samples. Observed UPF1 granules are indicated with red arrows.

4.2.8 UPF1-GFP localisation in cytoplasmic granules is reversible

As detailed in previous sections, UPF1 re-distributes from translating mRNAs into cytoplasmic granules upon stress. Further, I assessed whether these accumulations are reversible once the stress conditions are overcome.

Therefore, I first induced stress conditions by incubating the cells at 39°C for 30 minutes and then performed recovery, by returning them back to 30°C for another 30 minutes, taking aliquots prior to, after treatment and after recovery, for GFP fluorescence monitoring (Figure 4.8A). As expected from previous results, under heat shock conditions, UPF1 localises in cytoplasmic granules (Figure 4.8A, 2nd row images). Upon recovery, the granules are no longer visible in the cells and UPF1 was again homogeneously distributed throughout the cytoplasm (Figure 4.8A, 3rd row images), suggesting that UPF1 reversibly accumulates in distinct cellular foci, being released from accumulations upon recovery. The release could be due to UPF1 solely being released from the granules, although the granules themselves most probably disassemble once the stress conditions are overcome.

A similar pattern of distribution was observed in *Drosophila*, where UPF1, from localising in discrete cytoplasmic foci under heat shock (Figure 4.8B, 2nd row images), again becomes homogeneously distributed in the cytoplasm after recovery (Figure 4.8B, 3rd row images) (Anand Singh, unpublished data).

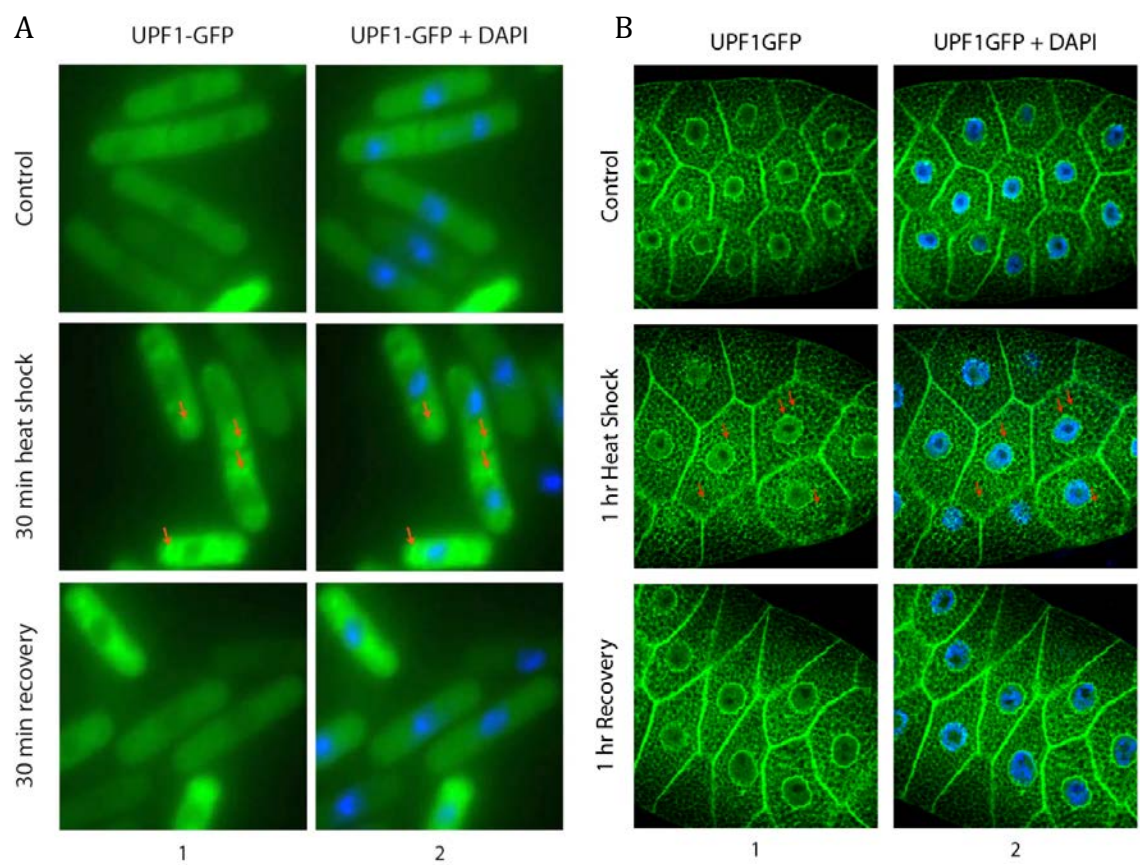


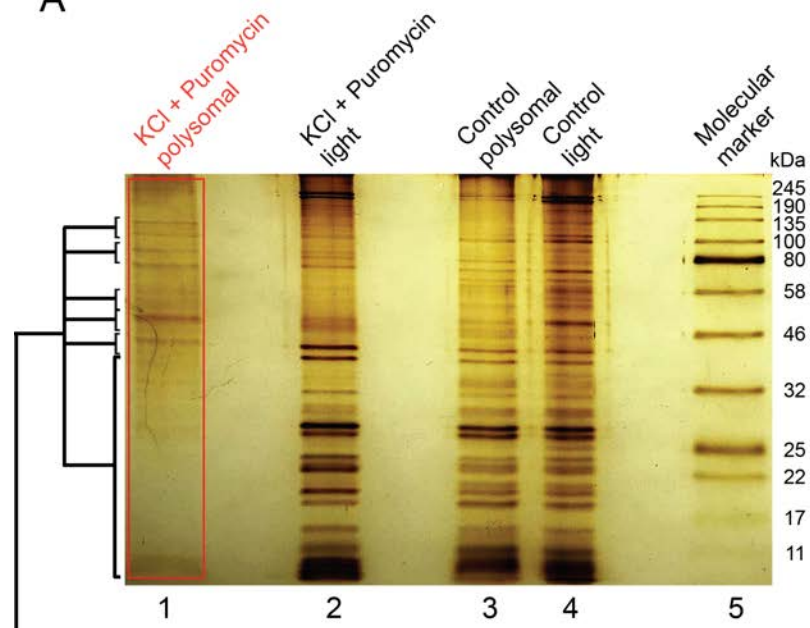
Figure 4.8 – **UPF1-GFP cytoplasmic granule localisation is reversible.** (A) UPF1 GFP fluorescence (green) of *S. pombe* control cells (top row), cells heat-shocked at 39°C for 30 min (middle row) and cell recovered from heat shock at 30°C for 30 min (bottom row), as indicated. Chromosome DAPI counter-staining merged with the GFP signal is shown on the right (2). (B) UPF1 GFP fluorescence (green) of *D. melanogaster* control cells (top row), cells heat-shocked at 42°C for 1 hour (middle row) and cell recovered from heat shock at 25°C for 1 hour (bottom row), as indicated. Chromosome DAPI counter-staining merged with the GFP signal is shown on the right (2) (Anand Singh, unpublished data). Red arrows indicate observed UPF1 granules in both experiments.

4.2.9 The composition of heavy cytoplasmic ribosome-free compartments suggests that UPF1 associates with stress granules

To investigate the composition of stress-induced UPF1-associating granules, I performed mass spectrometry analysis on the heavy polysomal fractions of the sample treated with 500 mM KCl and with puromycin upon lysis (Figure 4.4B). Puromycin dissociates the polysomal complexes so that all that remains in heavier fractions are ribosome-free heavy cytoplasmic compartments formed under high salt stress.

I ran the sample on an SDS-PAGE gel together with a control, cycloheximide-treated heavy polysome fractionation sample, as a comparison. Alongside these, I loaded light fractions from both samples, again, as a comparison. I excised the protein bands that appear in the stress-induced sample and not in the control, to identify proteins that associate only with the granules, and are not present under normal conditions (Figure 4.9A). After trypsin digestion and mass spectrometry analysis of the resulting peptides, the proteins were classified into gene ontology groups using DAVID bioinformatics tool, with the highly enriched categories shown in Figure 4.9B. Enriched proteins identified in the granules strongly indicate that the cellular compartments in question are stress granules. A highly enriched category represents ATP-dependent RNA helicases, one of which is UPF1. It would be interesting to investigate UPF1 association with these granules further, particularly whether UPF1 has a role in the stress granule formation.

A



B

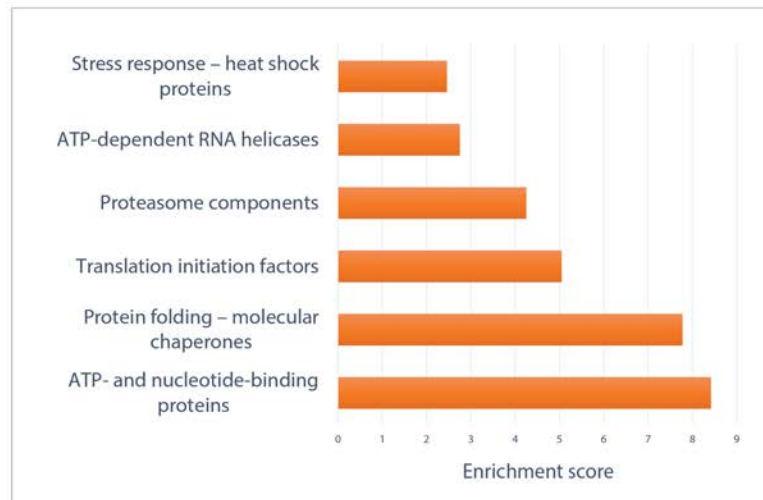


Figure 4.9 – **Proteins identified in heavy ribosome-free compartments to which UPF1 associates are components of stress granules.** (A) Silver staining of heavy (1) and light (2) polysome fraction of sample treated with high salt (500 mM KCl) and puromycin (100 µg/mL), as well as of heavy (3) and light (4) polysome fraction of control sample treated with cycloheximide (100 µg/mL). Molecular marker is shown on the right (5). Bands indicated in red in lane 1 were excised from the gel using a clean scalpel and analysed using mass spectrometry. (B) Gene ontology analysis was performed on identified proteins, grouping them into categories, based on the highest confidence (DAVID Bioinformatics online tool). The groups most highly enriched in the sample are shown in the bar chart.

4.3 Discussion

This chapter details my investigation of the association of NMD factors with mRNAs in relation to translation, as well as attempts to assess any NMD-independent roles of NMD factors, in global translation and stress response.

UPF1 was found to associate with translating mRNAs *in vivo*, throughout mRNA sequence, becoming displaced from coding sequence by translating ribosomes (Zünd *et al.*, 2013). Indeed, in control conditions, I found that UPF1 localised both with fractions corresponding to translating mRNAs as well as with ribosomal subunits and light fractions, suggesting that a considerable fraction of cytoplasmic UPF1 is in cellular compartments that are independent of translating ribosomes. Polysome UPF1 association is dependent on translation, as polysome breakdown caused by puromycin treatment leads to UPF1 moving towards light fractions. UPF1 remains associated with fractions corresponding to the 40S, suggesting that UPF1 potentially forms interactions with the small ribosomal subunit. A very low UPF1 signal was present in heavy polysome fractions upon puromycin treatment, and particularly upon EDTA treatment. This experiment indicates that, even though UPF1 shifts towards light fractions upon polysome breakdown, a fraction of UPF1 remains associated with heavy cytoplasmic complexes, which potentially represent cytoplasmic granules. Mild RNase treatment of polysomes leads to a minor shift of UPF1 towards light fractions, suggesting that association of UPF1 with the polysomes is in part driven via RNA binding.

I found that UPF3 is present throughout polysomal fractions, suggesting, similarly to UPF1, that UPF3 binds translating mRNAs in the cell, as well as

other cellular compartments. A clear shift towards light fractions was observed both upon puromycin and EDTA treatments, indicating that UPF3 present in heavy polysome fractions is indeed predominantly associated with translating mRNAs. Mild RNase treatment leads again to a shift of UPF3 towards light fractions, suggesting that UPF3 polysome association is driven either by binding the mRNAs directly or binding proteins associated with mRNAs other than ribosomes. These could potentially be proteins of the EJC as UPF3 was previously shown to interact with this complex (Le Hir *et al.*, 2001).

Since UPF1 and UPF3 bind mRNAs globally, I investigated whether these factors have a general role in translation. I assessed if there is any difference between proteins being synthesised by monosomes versus polysomes in the mutants compared to the wild type, since proteins fully translated by the monosome are more targeted by NMD (Heyer and Moore, 2016). Indeed, I observed a different pattern of puromycylation between the two strains and the wild type, with translation of a subset of proteins being altered. In particular, proteins synthesised by the monosome were affected, especially in the *upf3Δ* strain. Further, it would be interesting to identify which proteins these are. Additionally, any effects I observed could be a non-specific, indirect effect of the factors altering other cellular processes. Therefore, using an auxin-inducible degron system, for example, to quickly reduce the levels of these proteins in the cells and then performing the polysome fractionation paired with puromycylation could be a more direct way to look at the roles of NMD factors in translation.

I found UPF1 to associate with heavy cytoplasmic fractions void of ribosomes and hypothesised that UPF1 might associate with either P bodies or stress

granules, as it has previously been linked to these cellular bodies (Sheth and Parker, 2006; Cherkasov *et al.*, 2015). I first investigated whether stress conditions would shift the balance of UPF1 association with translating mRNAs towards cytoplasmic granules, as stress granules are more prominent under stress conditions. Indeed, under high salt treatment, and upon subsequent puromycin treatment to destroy polysomes and shift the ribosome components to light fractions, I observed that UPF1 remained in heavy polysomal fractions and consequently, in heavy ribosome-free cellular granules.

UPF1-GFP, episomally expressed in the cells, mimics the polysome distribution of the endogenous UPF1. Therefore, I used UPF1-GFP to visualise UPF1 distribution in the cells by monitoring GFP fluorescence in normal versus stress conditions, to be able to substantiate the biochemical results with microscopy. Both under heat shock treatment, as well as upon increased concentrations of KCl, I observed accumulation of UPF1 in distinct cytoplasmic foci. The association is reversible, as upon recovery from stress, UPF1 becomes again homogeneously distributed throughout the cytoplasm. The data obtained by another Brogna lab member (Anand Singh, unpublished) shows the same phenomenon occurring in *D. melanogaster*, suggesting that UPF1 association to granules is conserved from fission yeast to fruit flies. Mass spectrometry analysis of granule components suggests that these cellular compartments are in fact stress granules, as key proteins identified were typical stress granule components such as translation initiation factors, poly(A) binding proteins, as well as ATP-dependent RNA helicases.

These exciting observations should be investigated further, either by investigating how the absence or over-expression of UPF1 might affect granule formation. Also, it would be interesting to identify, via high-throughput sequencing, which RNAs localise in these compartments, as well as which sequences UPF1 binds on these mRNAs, which could be investigated by utilising the iCLIP approach, a method tailored to specifically identify protein-RNA interactions.

Chapter 5: What is the mechanism of NMD?

5.1. Introduction

Although many aspects of NMD functioning in the cells have been investigated over the years, the crucial mechanism underlying the recognition and degradation of PTC-containing mRNAs remains largely unclear. Several models have been proposed to help understand this phenomenon, most prominent of which are the EJC and the faux 3'-UTR models (Le Hir *et al.*, 2001; Kashima *et al.*, 2006; Amrani *et al.*, 2004). These models have in common the prediction that a specific feature of the PTC-containing mRNA, either the presence of an EJC complex downstream of the PTC, or the long distance of a PTC from the poly-A tail, labels an mRNA for degradation. This results in the recruitment of NMD factors and, consequently, the decay machinery (Kashima *et al.*, 2006; Amrani *et al.*, 2004).

In contrast to the previous models, we propose that premature and normal translation termination do not differ mechanistically, and NMD is a passive consequence of ribosome not scanning and therefore not protecting the mRNA (Broгна and Wen, 2009; BroгнаMcLeod and Petric, 2016). As UPF1 has been shown to interact with eIF3, a protein which has a role in recycling ribosomal subunits from mRNAs (Morris *et al.*, 2007), we proposed that, through this association and through its helicase activity, UPF1 might release post-termination ribosomal subunits and potentially other mRNP components, from mRNAs (Morris *et al.*, 2007; Welch and Jacobson, 1999; FranksSingh and Lykke-Andersen, 2010) (as described in the Introduction section 1.9 and in Figure 1.13, wild type). As a result, NMD may not act as an active process that

recognises and degrades mRNAs harbouring a PTC. NMD could instead occur as a passive consequence of the dissociation of ribosomal subunits and other proteins, which normally bind and protect mRNAs from the activity of nucleases, making mRNAs susceptible for degradation.

To assay whether UPF1 can act as a stimulator of ribosome dissociation from prematurely terminating transcripts, I used strains that express NMD reporter genes, harbouring a PTC at a specific location, in a wild type as well as UPF1 deletion background. I used the green fluorescent protein (GFP) gene as a reporter, which either contained a PTC at codon position 6 (PTC6), 27 (PTC27) or 140 (PTC140), while wild type GFP gene served as a control (GFP). The prediction of the ribosome release model is that if UPF1 indeed stimulates release of ribosomal subunits from mRNAs, its absence will result in more ribosomes or ribosomal subunits remaining bound to mRNAs (as detailed in the Introduction section 1.9 and Figure 1.13, UPF1 null mutant). Therefore, a PTC-containing reporter mRNA would be associated with heavier polysomal fractions in *upf1Δ* compared to wild type, if the hypothesis is correct (Figure 1.13, UPF1 null mutant) (Hu *et al.*, 2010). Reporter mRNA distribution was also assessed in *upf2Δ* and *upf3Δ* strains to assess whether any observed effect is specific to UPF1 or due to an absence of NMD in general.

In short, results presented in this chapter suggest that UPF1 and UPF3 might have a role in stimulating release of ribosomes from mRNAs, which goes in line with the findings presented in Chapter 4 where translation of a subset of proteins is altered in the absence of these NMD factors.

5.2. Results

5.2.1. NMD is suppressed in *upf1* Δ and *upf2* Δ strains

I performed initial experiments using wt, *upf1* Δ and *upf2* Δ strains expressing either GFP or GFP with a PTC at codon position 6, previously constructed by a Brogna lab member, to obtain preliminary results (Jianming Wang PhD thesis, 2015). I verified the strains by Northern blotting of total cell culture RNA. The membrane was first stained with methylene blue to visualise RNA and assess its loading and then hybridised with a GFP-specific probe (Material and Methods) (Figure 5.1). The data is representative of three biological repeats \pm standard deviation (SD) (Figure 5.1).

The expression of GFP harbouring a PTC at position 6 in the wild type strain was reduced to $20 \pm 5\%$ compared to the wild type reporter (Figure 5.1C, lane 2 versus lane 1). The mRNA levels were restored to $42 \pm 4\%$ and $94 \pm 7\%$ relative to wild type GFP reporter gene expression in *upf1* Δ and *upf2* Δ strains, respectively (Figure 5.1C, lanes 4 and 6). NMD is expected to be equally suppressed in both mutants (Wen and Brogna, 2010). Even though NMD was not fully suppressed in *upf1* Δ , it can still be concluded that absence of either UPF1 or UPF2 leads to significant stabilisation of the NMD substrate, as previously reported (Wen and Brogna, 2010).

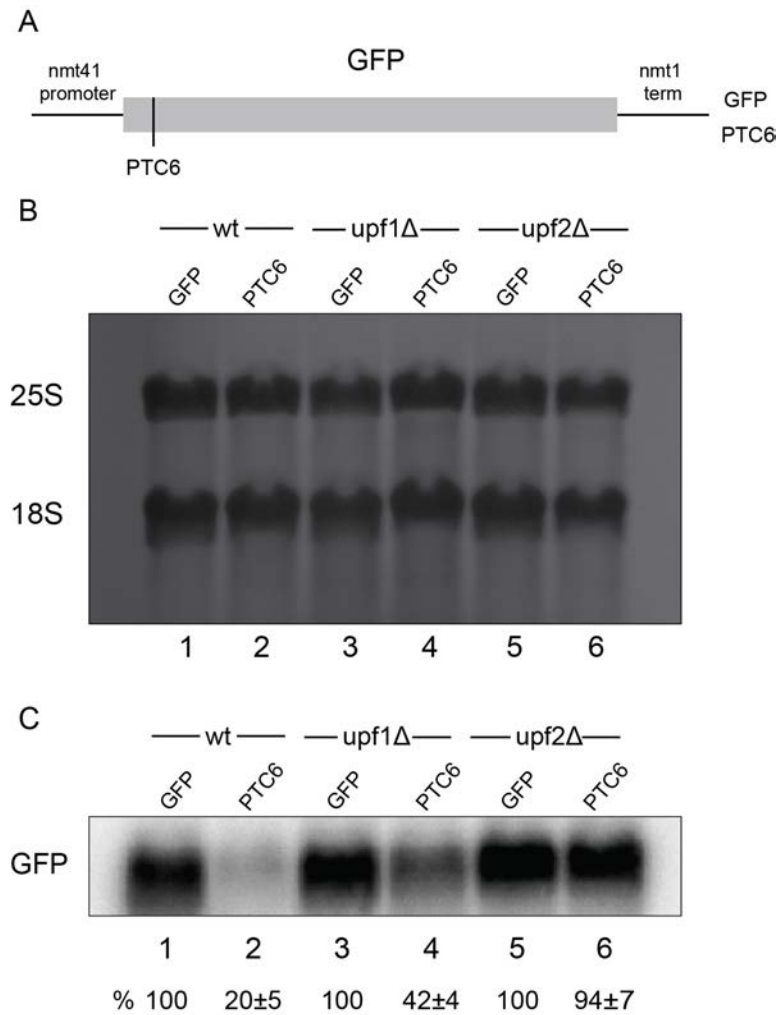


Figure 5.1 – NMD is suppressed in the absence of UPF1 and UPF2. (A) Schematic of the GFP reporter genes expressed in the analysed strains (Jianming Wang PhD Thesis, 2015). The strains either possess a complete functional GFP sequence or a GFP gene containing a PTC at codon position 6. (B) Methylene blue staining of total-cell RNA showing equal sample loading and satisfactory RNA quality. (C) Northern blotting analysis of total RNA from wild type GFP (1), wild type PTC6 (2), upf1Δ GFP (3), upf1Δ PTC6 (4), upf2Δ GFP (5) and upf2Δ PTC6 (6), hybridised with a GFP-specific probe. The data is represented as a mean percentage of GFP mRNA signal derived from three biological replicates \pm SD, where the percentage of GFP PTC6 mRNA signal in each strain is calculated relative to that of functional GFP mRNA of the corresponding strain.

5.2.2. NMD factors seem to affect ribosome release from mRNAs

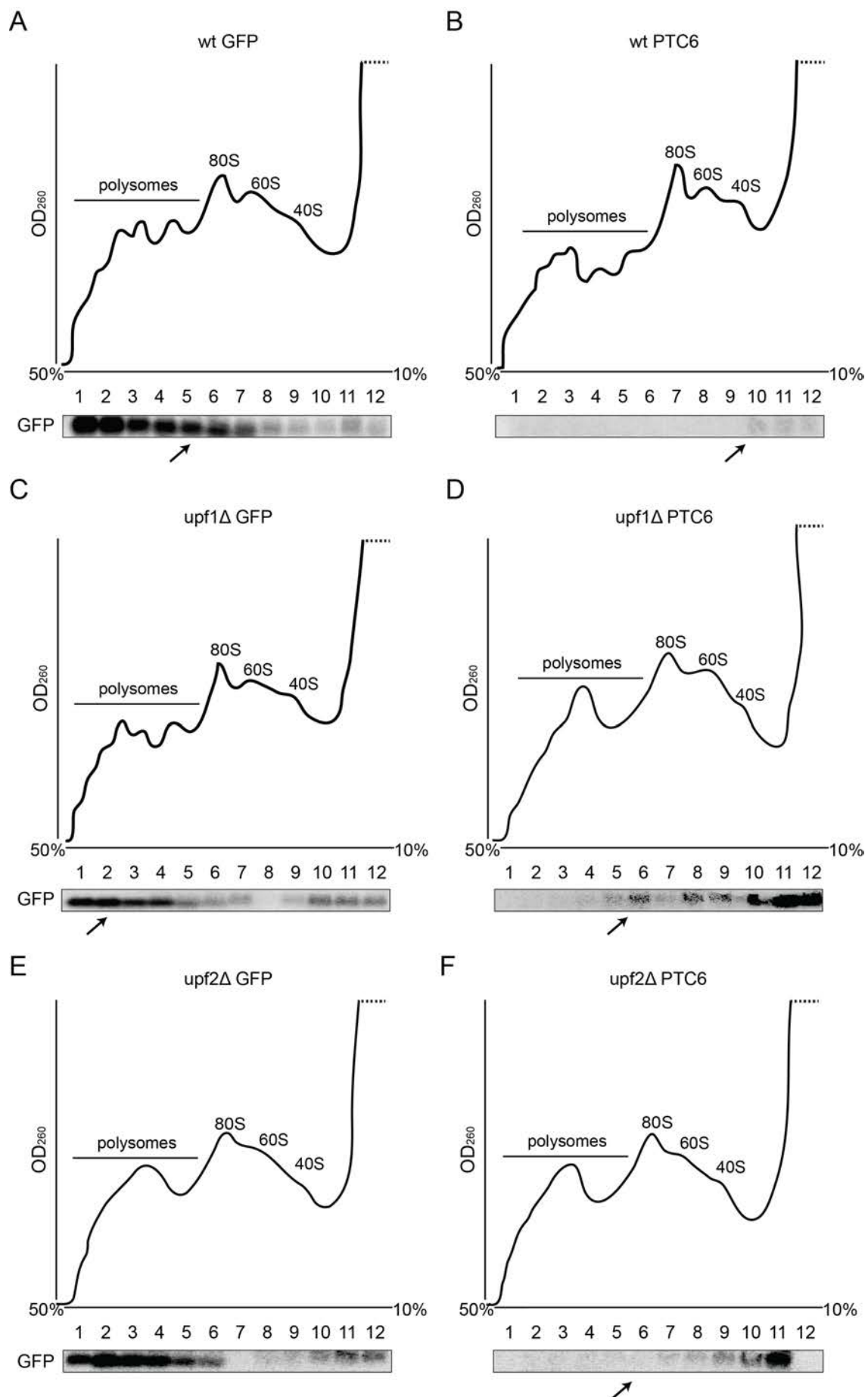
To assess whether the absence of NMD factors affects the release of ribosomes from mRNAs, I carried out polysome fractionation of the strains detailed in section 5.2.1 and analysed the RNA by Northern blotting using a GFP-specific probe, to assess the distribution of the GFP reporter mRNA relative to the number of ribosomes associated with it (Figure 5.2).

The functional GFP mRNA (containing no mutations) is associated with translating ribosomes, being present throughout polysomal fractions (Figure 5.2A, C, E, lanes 1-7) and at a very low level in the remaining fractions (Figure 5.2A, C, E, lanes 8-12), in all strains analysed. However, it is interesting that there is a shift towards polysomes of a higher molecular weight in *upf1Δ*, indicated by arrows under the blots (Figure 5.2C, lanes 1-4, relative to Figure 5.2A, lanes 1-7). This is consistent with the prediction of the ribosome release model that UPF1 may stimulate ribosome dissociation from mRNA and therefore more ribosomes are globally bound to mRNAs in its absence, perhaps regardless of whether they are NMD substrates or not, indicating a role of UPF1 in general translation.

As the PTC occurs very early in the NMD reporter gene – at codon position 6 – only one or two translating ribosomes would be expected to be bound to it before they terminate at the PTC. I detected the NMD reporter gene only in light fractions, but not in monosomes or disomes in the wild type strain (Figure 5.2B, lanes 10-12), which could be due to low levels of this mRNA. In *upf2Δ* (Figure 5.2F, lanes 7-9), and particularly in *upf1Δ* (Figure 5.2D, lanes 4-9), there are more ribosomes associated with PTC6 containing mRNA compared to the wild

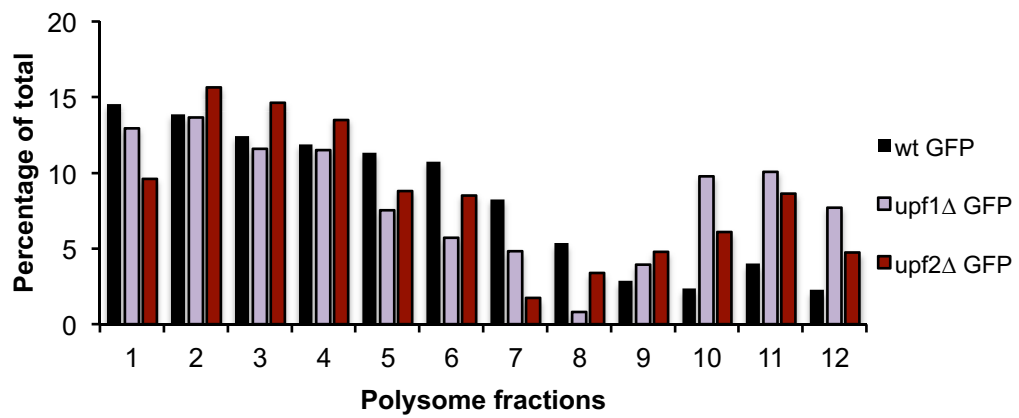
type (Figure 5.2B). The results are in line with the ribosome release model, as in the absence of NMD factors, more ribosomes are associated with the PTC-containing mRNA. The results obtained were reproducible in two biological replicates.

These preliminary results indicated that UPF1 may affect the distribution of ribosomes (or other mRNP components) on mRNAs and that in its absence, more proteins associate with mRNAs. However, NMD was not fully suppressed in *upf1Δ* strain (Figure 5.1C, lane 4), which might influence reporter mRNA distribution in polysomal fractions. For this reason, and to test whether any phenotypic effects would be observed when analysing the remaining GFP reporters (PTC27 and PTC140), I decided to transform wild type and *upf1Δ* strains (to focus on any UPF1 related effects first) with plasmids expressing all NMD reporters, and then repeated these experiments.



G

GFP mRNA distribution



H

PTC6 mRNA distribution

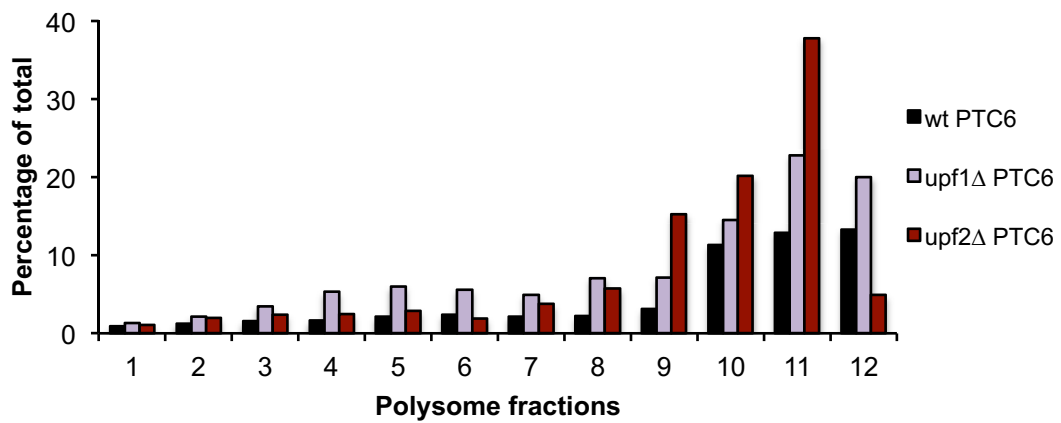


Figure 5.2 – **NMD appears to affect ribosome release from mRNA.** Polysome fractionation traces followed by Northern blotting from (A) wild type GFP, (B) wild type PTC6, (C) *upf1Δ* GFP, (D) *upf1Δ* PTC6, (E) *upf2Δ* GFP and (F) *upf2Δ* PTC6 polysomal fractions, analysed via hybridisation with a GFP-specific probe. Traces represent the absorbance profile (OD_{254}) recorded during the fractionation. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. mRNA shows higher ribosome loading in the mutants compared to the wild type. The arrows under the Northern blots indicate the observed shift of GFP mRNA towards heavier polysomal fractions, when wild type and *upf1Δ* functional GFP mRNA, as well as PTC6 mRNA, is compared between all three strains. The intensity of (G) wild type GFP mRNA and (H) PTC6-containing mRNA signal for each lane of the wild type, *upf1Δ* and *upf2Δ* strain was determined using Image J software and was expressed as a percentage of the total signal.

5.2.3. NMD is position-dependent and it is fully suppressed in newly transformed *upf1* Δ strains

As discussed in section 5.2.2, I transformed wild type and *upf1* Δ strains with plasmids expressing GFP NMD reporters (Materials and methods), to repeat polysome fractionation experiments in order to test the ribosome release model. Firstly, I analysed total-cell RNA from these strains to assess whether NMD suppression is complete in *upf1* Δ strain. The data is representative of three biological repeats, with \pm SD (Figure 5.3C).

The results are consistent with previous observations that NMD is position-dependent in *S. pombe* (Wen and Broгна, 2010). As shown in Figure 5.3, PTC6, which occurs very early in the sequence, induced strong NMD in the wild type strain, as the level of PTC6 containing mRNA was reduced to $10 \pm 3\%$ compared to that of wild type (Figure 5.3C, lane 2). PTC27 containing mRNA showed reduction to $19 \pm 5\%$ (Figure 5.3C, lane 3), while GFP containing a PTC at position 140 was unaffected by NMD with an mRNA level of $96 \pm 7\%$ relative to wild type (Figure 5.3C, lane 4). NMD is therefore more prominent the earlier a PTC occurs within a gene, which is in line with previous findings. The mRNA levels were restored in *upf1* Δ strains (Figure 5.3C, lanes 6-8) with NMD substrates stabilised to $75 \pm 10\%$ of PTC6 (Figure 5.3C, lane 6) and $84 \pm 9\%$ of PTC27 (Figure 5.3C, lane 7), while PTC140 showed $87 \pm 11\%$ signal relative to the wild type GFP level in *upf1* Δ strain (Figure 5.3C, lane 8). NMD is efficiently suppressed in *upf1* Δ strains, which can therefore be used for further experiments.

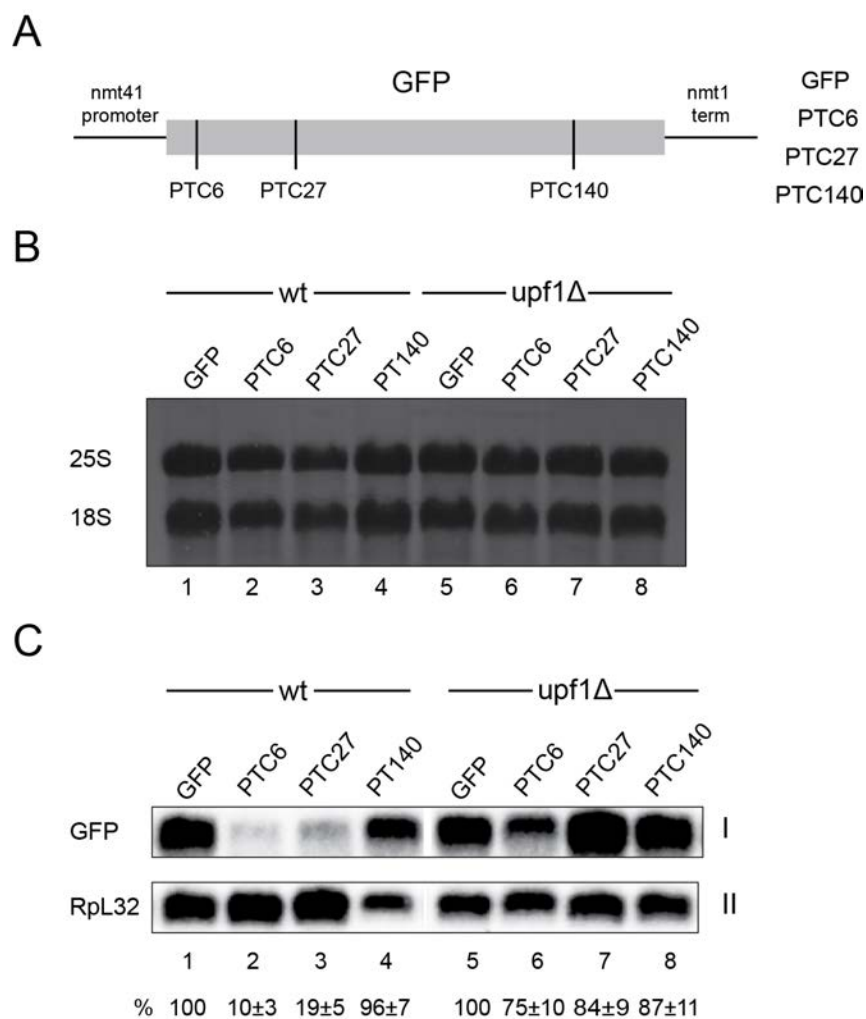


Figure 5.3 – NMD is position-dependent and suppressed in the absence of UPF1.

(A) Schematic of the reporter gene expressed in the analysed strains. The strains either possess a complete functional GFP sequence or a GFP gene containing a PTC at codon position 6, 27 or 140. (B) Methylene blue staining of total-cell RNA demonstrating satisfactory RNA quality and loading. (C) Northern blotting analysis of total-cell RNA from wild type GFP (1), wild type PTC6 (2), wild type PTC27 (3), wild type PTC140 (4); upf1Δ GFP (5), upf1Δ PTC6 (6), upf1Δ PTC27 (7), upf1Δ PTC140 (8) cells. Hybridisation was carried out with a GFP-specific (I) and RpL32-specific (II – loading control) probe. The percentages below the lanes show GFP mRNA levels relative to that of the functional GFP reporter in each strain. The data is derived from the average of three biological replicates ± SD.

5.2.4. NMD reporter mRNA was detectable only at an unexpectedly low level in polysome fractionation samples of upf1 Δ strain

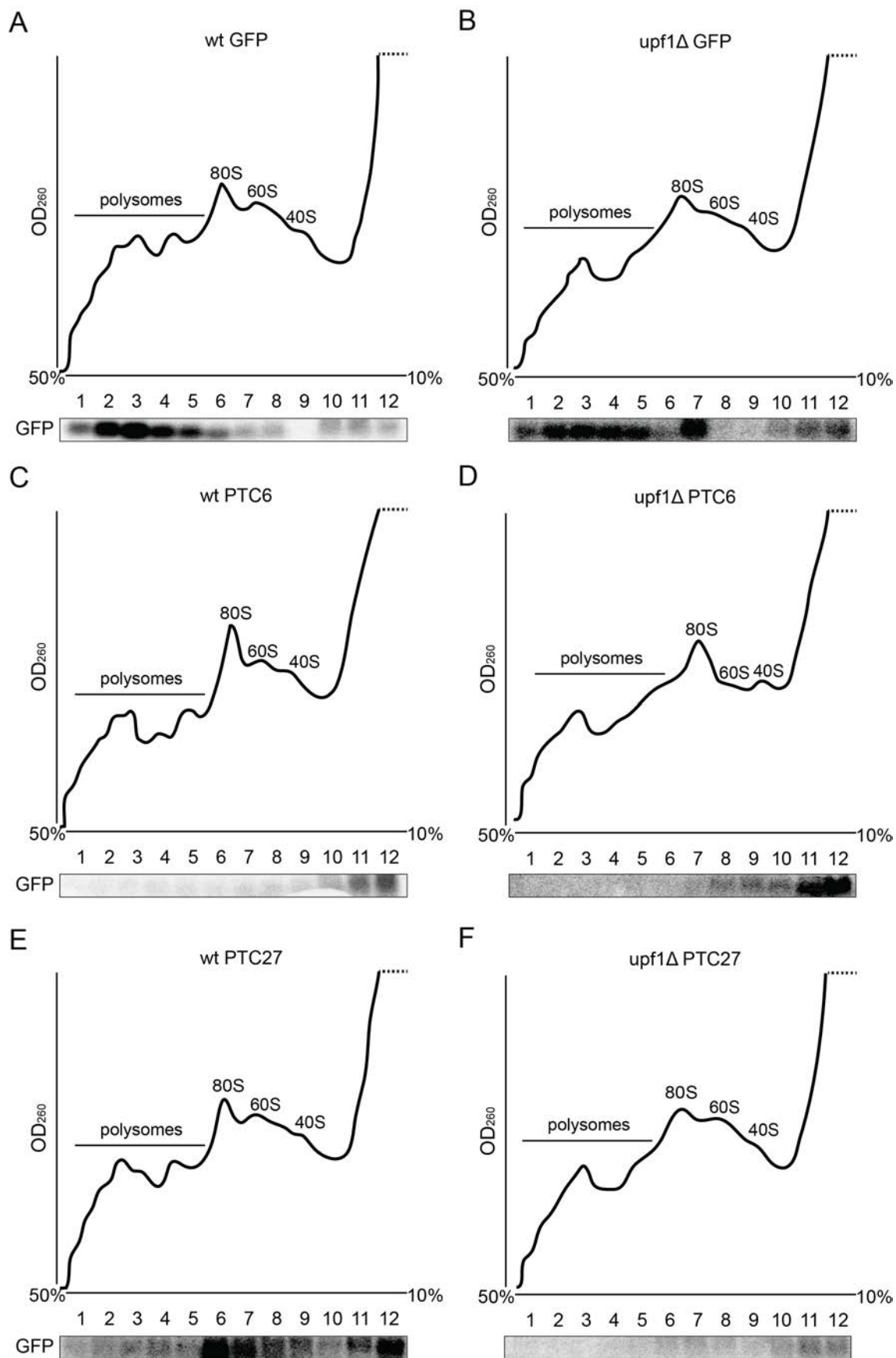
Since the NMD phenotype of the strains I obtained was verified (as detailed in section 5.2.3), I carried out polysome profile fractionation followed by Northern blotting using a GFP-specific probe to assess GFP mRNA distribution (Figure 5.4), and compare it between strains, firstly of the NMD reporters affected by NMD (PTC6 and 27) as well as functional GFP.

As previously shown, functional GFP mRNA, expressed at moderate levels, is associated with translating ribosomes, being present throughout polysomal fractions in both strains (Figure 5.4A, B, lanes 1-7). GFP mRNA containing a PTC at codon position 6 is detected in the wild type strain mostly in the light fractions (Figure 5.4C, lanes 10-12) and only at a very low level in fractions corresponding to the 80S and disome (Figure 5.4C, lanes 6-9). This result is consistent with only one or two ribosomes at most being able to engage in translation with the mRNA before they terminate upon encountering a PTC at codon position 6. Similarly, since a PTC at codon position 27 is early in the sequence, it is expected that only 3 or 4 ribosomes would be able to bind to this GFP mRNA before termination. This is indeed observed in the wild type strain by detecting GFP mRNA mostly in light fractions (Figure 5.4E, lanes 10-12) as well as in low molecular weight polysomal fractions (Figure 5.4E, lanes 4-9). Northern blotting results therefore represent expected physiological conditions well in wild type.

GFP mRNA was distributed primarily throughout polysomal fractions in the upf1 Δ strain (Figure 5.4B, lanes 1-7), and was well expressed, comparable to

the wild type. However, it was surprising that, although the total cell culture RNA analysis showed that NMD is efficiently suppressed and reporter mRNA levels are comparable to the functional GFP mRNA levels in the absence of UPF1 (Figure 5.3C, lanes 6, 7), a very low amount of PTC6 mRNA, and even less of PTC27 mRNA, was detected in polysome profiling fractions (Figure 5.4D, F). The amount of RNA was expected to be comparable to the functional GFP level (Figure 5.4A, B), however, much less was observed, which raised the question why this proportion of RNA is undetectable.

The results were reproducible in at least three biological replicates. To confirm that NMD was suppressed in the cells used for polysome fractionation, I repeated the experiment by using aliquots of the cultures to carry out the total-cell RNA analysis and performed the polysome fractionation on the remainder. In this way I made sure that both results came from a single culture. However, the outcome remained unchanged, NMD reporter mRNA was stabilised in the *upf1Δ* strain based on total-cell culture RNA, yet mRNA levels in the polysomal samples were again very low (data not shown).



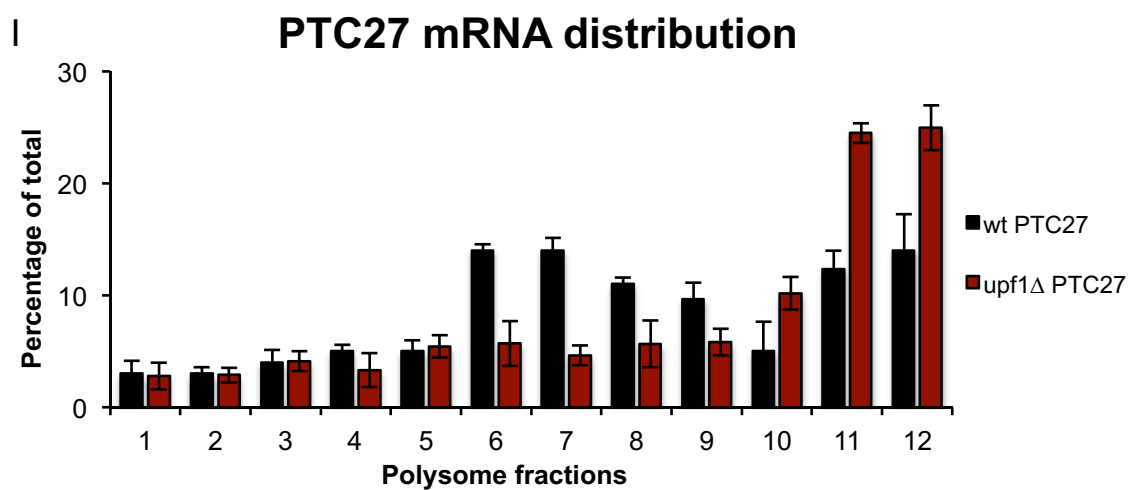
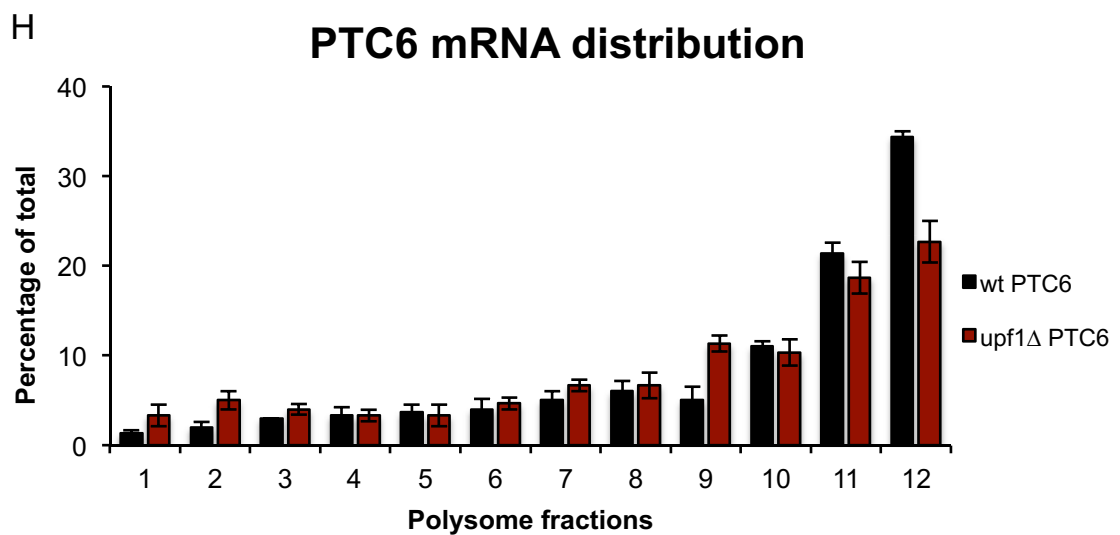
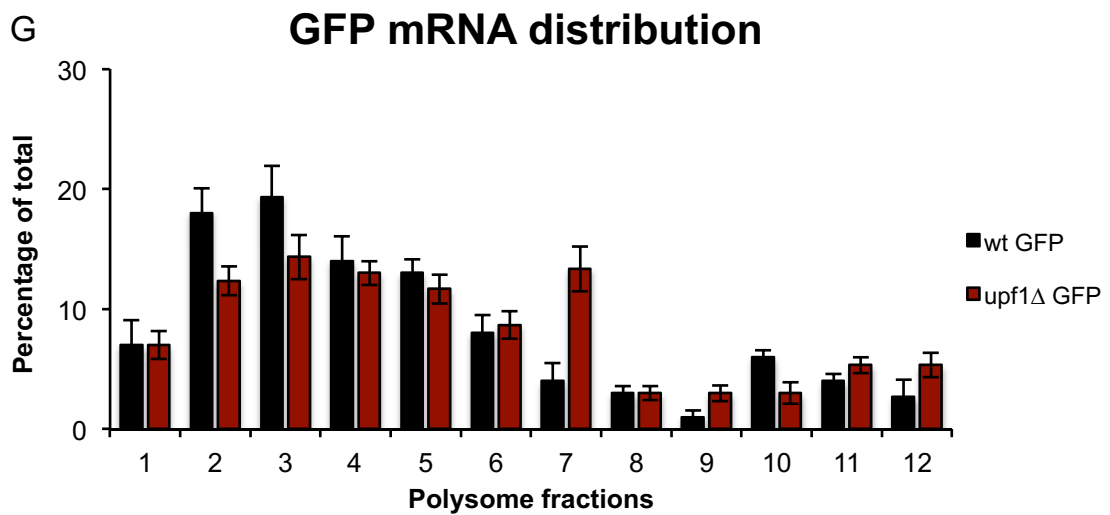


Figure 5.4 – NMD reporter mRNA was detected at an unexpectedly low level in upf1Δ polysome fractionation samples – Polysome fractionation traces followed by Northern blotting analysis of RNA from (A) wild type GFP, (B) upf1Δ GFP, (C) wild type PTC6, (D) upf1Δ PTC6, (E) wild type PTC27 and (F) upf1Δ PTC27 polysomal fractions, analysed via hybridisation with a GFP-specific probe. Traces represent the absorbance profile (OD₂₅₄) recorded during the fractionation. The gradient was split into 12 fractions: polysomal, 80S, 60S, 40S and light fractions, as indicated. The intensity of (G) wild type GFP mRNA, (H) PTC6-containing mRNA and (I) PTC27-containing mRNA signal for each lane of the wild type and upf1Δ strain was determined using Image J software and was expressed as a percentage of the total signal. The means of three independent biological repeats ± SD are demonstrated in the graph.

5.2.5. NMD reporter mRNA does not pellet during polysomal clearing stages

As a significant proportion of NMD reporter mRNA was missing upon polysome fractionation of *upf1Δ* strain, I investigated if the RNA might be lost during this procedure, whilst processing the samples. After cell lysis, the lysate was cleared from cell debris and nuclei by centrifugation. Therefore, I examined whether RNA pelleted during these clearing stages. RNA from the pellets derived from the 1st and 2nd clearing step of all examined strains was analysed by Northern blotting (Figure 5.5).

The results show that a proportion of ribosomal RNA pellets during the cell lysate clearing stages. This was detected in both the wild type and to a lesser extent in *upf1Δ*, visualised by methylene blue staining (Figure 5.5A, lanes 1-3 and 7-9, compared to lanes 4-6 and 10-12). The difference is even more prominent during the 2nd clearing stage (Figure 5.5A, lanes 10-12), which indicates that overall RNA in *upf1Δ* may be more prone to degradation.

GFP reporter mRNAs are detected in clearing pellets of wild type strains, particularly the functional GFP (Figure 5.5B, I, lanes 1-3, 7-9). which is expected, considering it is more abundant compared to the NMD reporters. On the contrary, very little GFP mRNA is detected in the clearing pellets of the *upf1Δ* strain (Figure 5.5B, I, lanes 4-6, 10-12), which, therefore, cannot account for mRNA missing from polysomal fractions. Interestingly, overall RNA quality is low in *upf1Δ* strain, as demonstrated by Methylene blue staining (Figure 5.5A) and hybridisation with Rpl32 where very little to no Rpl32 is detected (Figure 5.5B, II, lanes 4-6, 10-12). RNA seems to be more susceptible to degradation in *upf1Δ* in these conditions and that is why it is lost during the procedure.

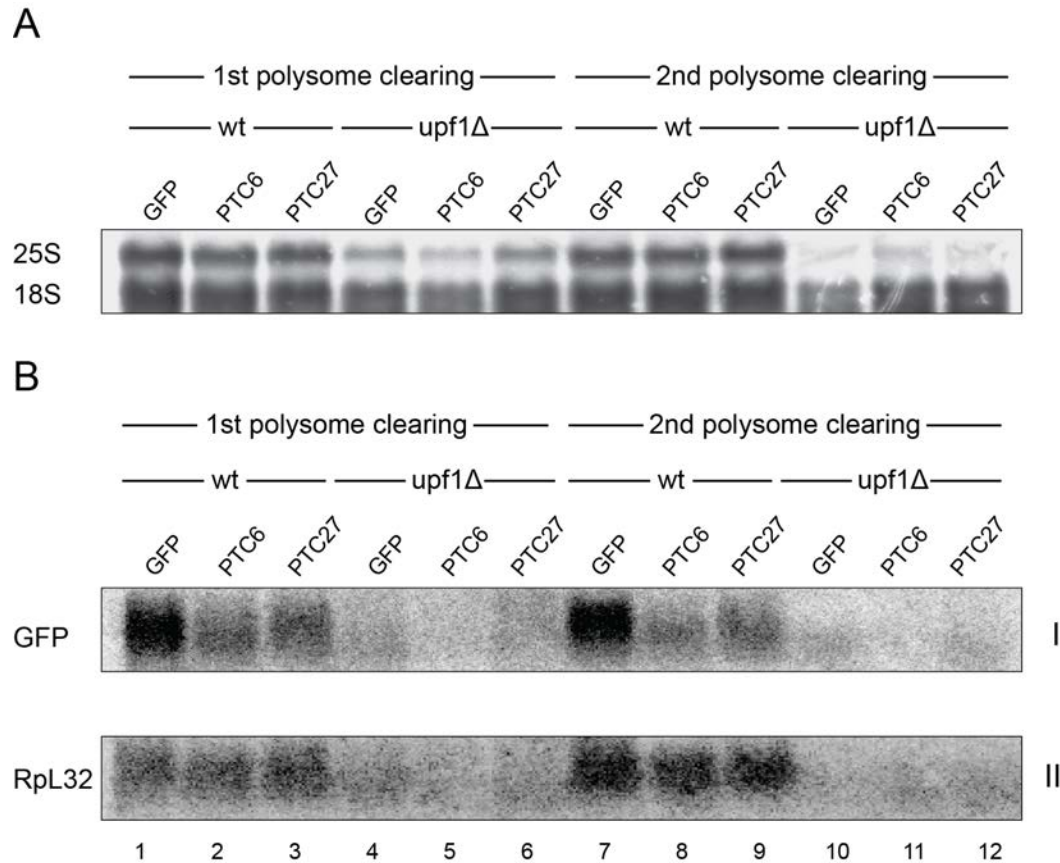


Figure 5.5 – **NMD reporter mRNA does not pellet during clearing stages.** (A) Methylene blue staining of RNA isolated from 1st and 2nd polysome clearing stages, from wild type and upf1Δ strains. (B) Northern blotting analysis of RNA hybridised with (I) GFP-specific and (II – loading control) RpL32-specific probe of: 1st (1-6) and 2nd (7-12) polysome clearing from wild type and upf1Δ strains, as indicated.

5.2.6. Modifications to the lysis buffer lead to an improved RNA yield and demonstrate that UPF1 and UPF3 may have a minor effect on ribosome release

Polysome fractionation conditions result in satisfactory RNA quality in the wild type strain, in which the quantity and observed association of NMD reporter genes' mRNA with polysomal fractions was as expected. However, I attempted to improve the procedure to address the absence of detectable GFP mRNA in *upf1Δ* strain.

I wanted to investigate whether the mRNP composition of NMD substrates becomes modified in *upf1Δ* strain in such a way that RNA extraction efficiency is low or makes RNA particularly prone to degradation. These may be the reasons why the RNA could not be recovered from polysomal fractions, which could be dependent on the lysis buffer composition used in the polysome profiling procedure. Therefore, I introduced modifications to the lysis buffer, based on literature (Hu *et al.*, 2010). Major changes included addition of 500 µg/mL heparin to further stabilise RNA, as well as addition of 1% Triton X-100 after cell lysis to release any RNA that is potentially in complexes or granules. These changes were made to assess whether recovery of the reporter mRNA would be enhanced (Figures 5.6-5.9). The modifications were previously shown to successfully recover RNA from *upf1Δ Saccharomyces cerevisiae* strain (Hu *et al.*, 2010).

UPF3 was recently demonstrated to play a role in stimulating ribosome recycling upon translation termination (Neu-Yilik *et al.*, 2017). Hence, this could be one of the key factors that contribute to UPF1-mediated ribosome release, as predicted by the model. I tested the role of UPF3 by generating *upf3Δ* strains

expressing the described reporter genes. As before, total cell RNA was analysed by Northern blotting and hybridisation with a GFP-specific and Rpl32-specific probe (as a loading control), which showed efficient NMD suppression in both *upf1Δ* and *upf3Δ* strains (data not shown). The same culture was then used for polysome fractionation. The data is presented for each of the reporter RNAs in all three strains, to allow a more readily comparison (Figures 5.6-5.9).

As previously observed, wild type GFP mRNA is present throughout polysomal fractions in all three strains analysed (Figure 5.6A, B, C, lanes 1-7). In case of GFP PTC6, mRNA is present predominantly in the fractions corresponding to the ribosomal subunits as well as the monosome in the wild type strain (Figure 5.7A, lanes 7-10). In the mutant strains, the GFP PTC6 mRNA has a similar distribution, although the presence of relatively more mRNA in fraction 6 in the mutants demonstrates a minor shift towards the heavier polysomal fractions, in comparison to the wild type (Figure 5.7A, B, C, lane 6). Interestingly, relatively more mRNA is detected in the light fractions and those fractions that correspond to the ribosomal subunits in both *upf1Δ* and *upf3Δ* strains (Figure 5.7B, C, lanes 9-12).

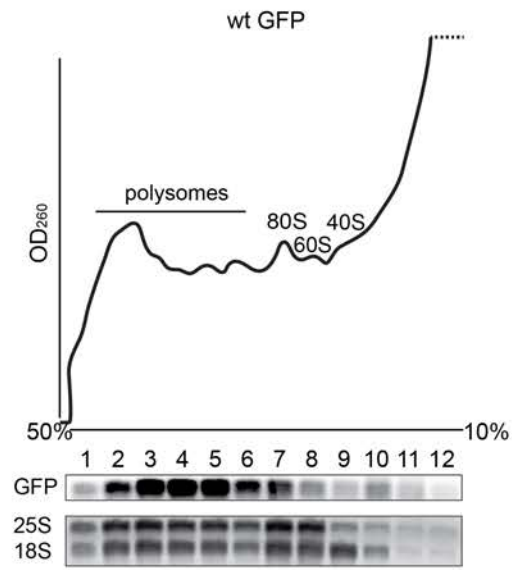
Comparable results were observed in case of GFP mRNA containing a PTC at position 27. In the wild type strain, mRNA is present predominantly within the fractions corresponding to ribosomal subunits, monosomes and disomes (Figure 5.8A, lanes 7-10). However, in the mutants, there is proportionally more mRNA in the heavier polysomal fraction (Figure 5.8B, C, lane 6) and, particularly in *upf3Δ*, there is a shift towards the heaviest polysomal fractions (Figure 5.8C, fractions 1-5, when compared to the wild type strain – Figure

5.8A). Again, a notably higher proportion of mRNA is found in association with the ribosomal subunits and residing in soluble fractions.

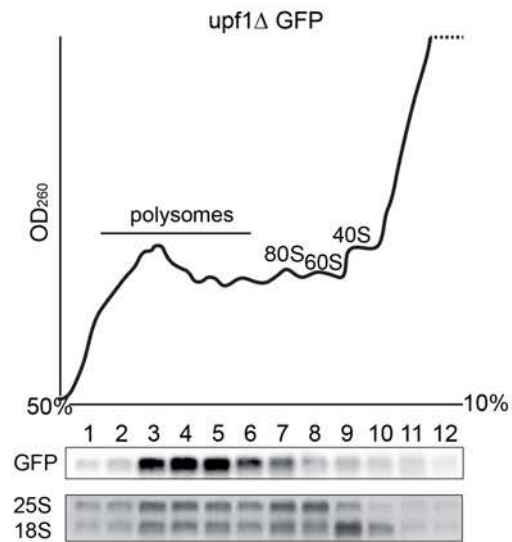
The distribution of GFP gene containing a PTC at position 140 (Figure 5.9) follows a very similar pattern to the functional GFP gene (Figure 5.9) in all three strains. This is consistent with the fact that NMD is not observed in this reporter mRNA (Figure 5.3C, lane 4).

The analysis using a modified lysis buffer was performed only once, due to time constraints and time-consuming nature of the experiments. The results are presented as they demonstrate an improvement in the recovery of previously undetectable GFP reporter mRNA, however, more biological replicates are necessary to assess reproducibility of the experiments.

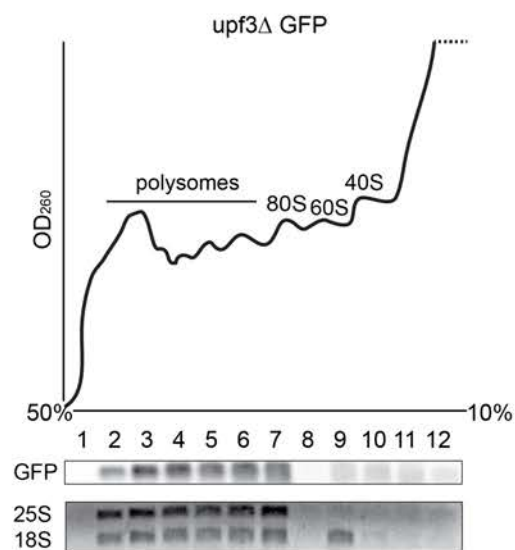
A



B



C



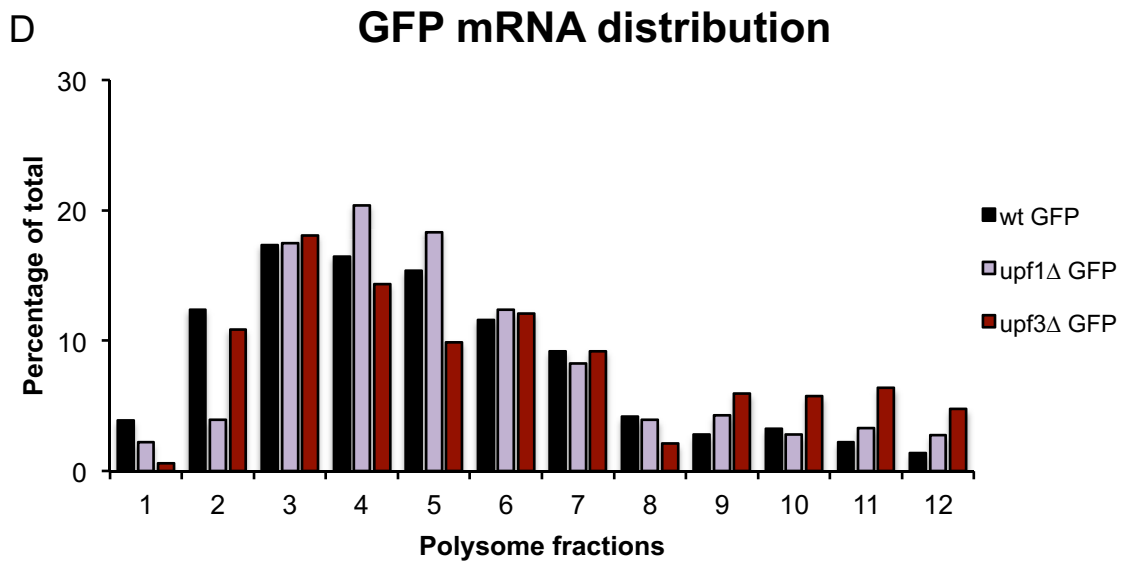
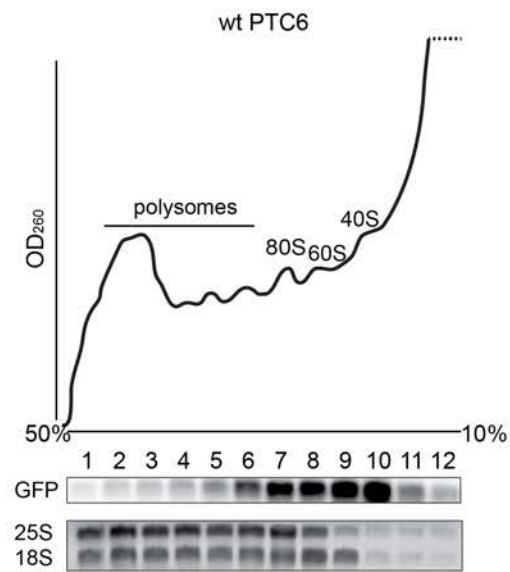
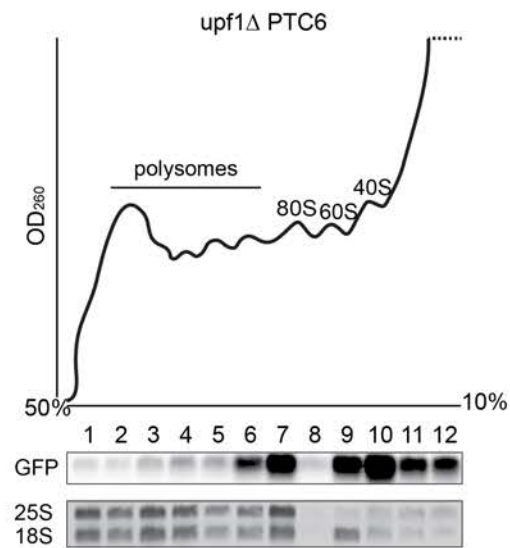


Figure 5.6 – **GFP mRNA is present primarily in polysomal fractions, moderately expressed.** Polysome fractionation traces followed by Northern blotting analysis of RNA from (A) wild type GFP, (B) upf1Δ GFP, (C) upf3Δ GFP polysomal fractions, hybridised with a GFP-specific probe. Methylene blue staining is shown bellow each Northern blotting panel, demonstrating polysome profiling and RNA quality as well as loading. (D) The intensity of wild type GFP mRNA for each lane of the wild type, upf1Δ and upf3Δ strain was determined using Image J software and was expressed as a percentage of the total signal.

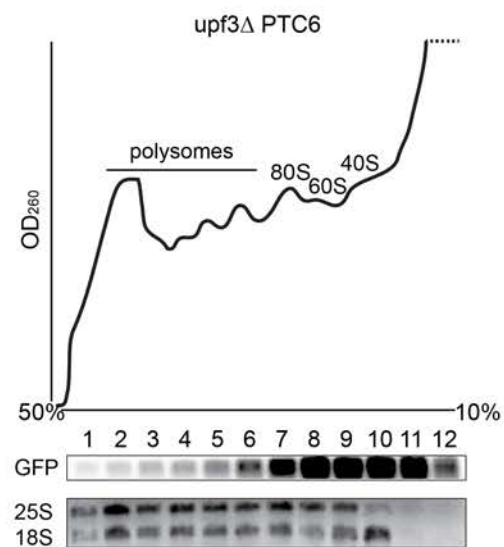
A



B



C



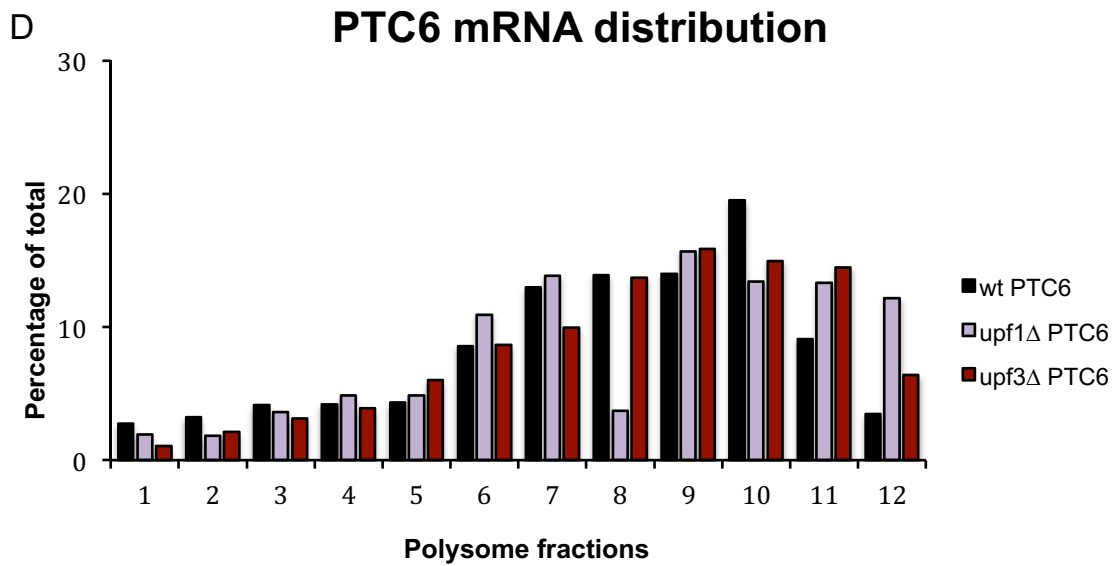
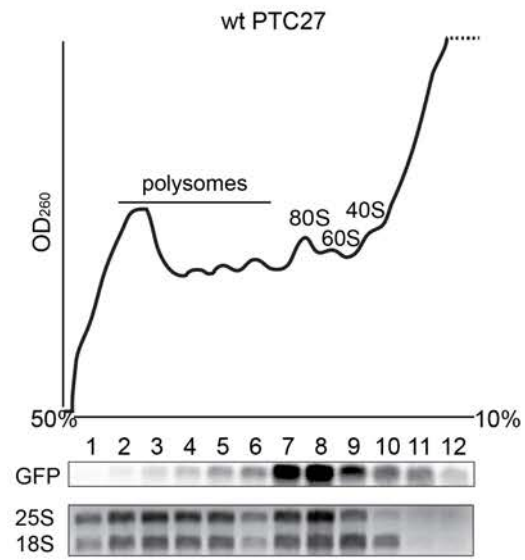
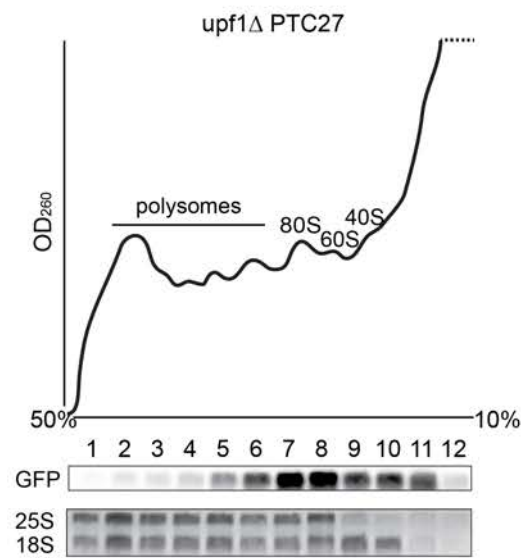


Figure 5.7 – **PTC6 GFP mRNA shows a minor shift towards heavier polysomal fractions in upf1Δ and upf3Δ strain.** Polysome fractionation traces followed by Northern blotting analysis of RNA from (A) wild type PTC6, (B) upf1Δ PTC6, (C) upf3Δ PTC6 polysomal fractions, hybridised with a GFP-specific probe. Methylene blue staining is shown bellow each Northern blotting panel, demonstrating polysome profiling and RNA quality as well as loading. mRNA is missing from fraction 8 (panel B, fraction 8) of the upf1Δ strain as the RNA pellet was lost during extraction. (D) The intensity of the PTC6-containing mRNA for each lane of the wild type, upf1Δ and upf3Δ strain was determined using Image J software and was expressed as a percentage of the total signal.

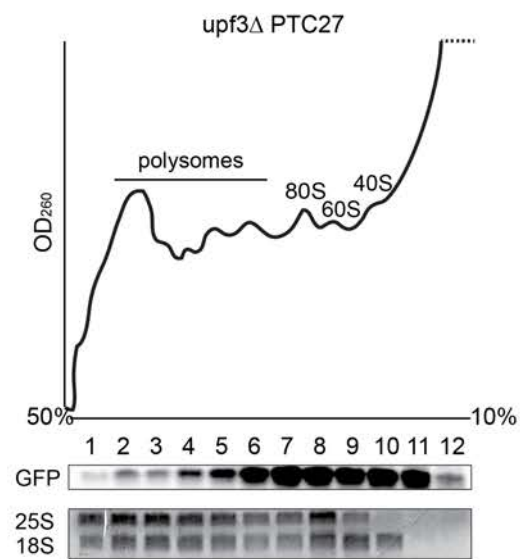
A



B



C



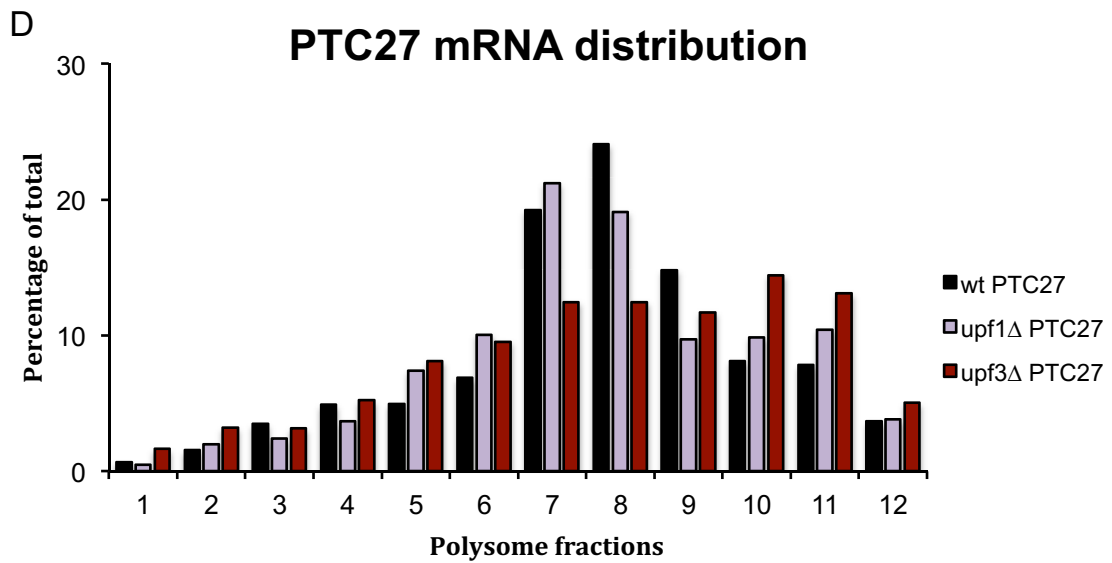
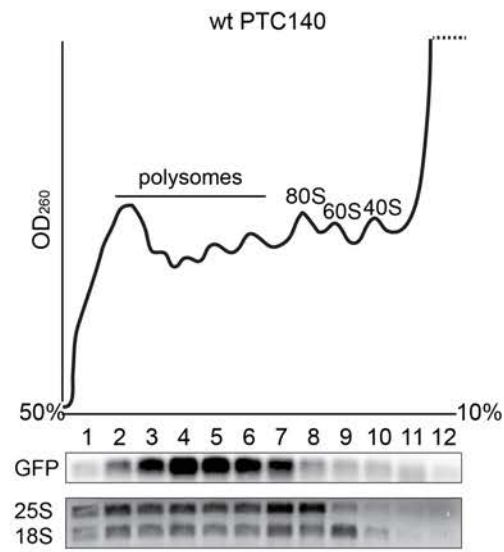
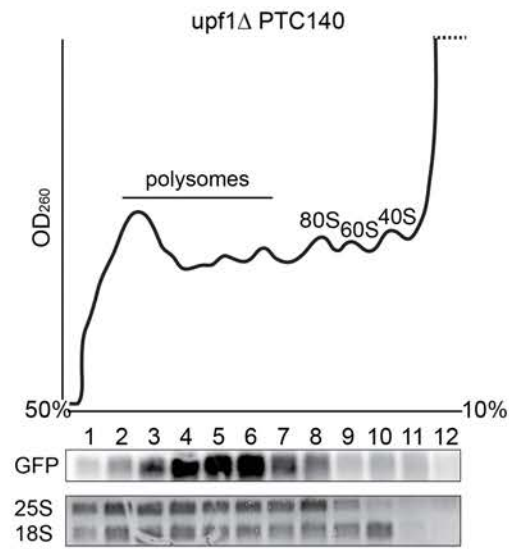


Figure 5.8 – **PTC27 GFP mRNA shows a minor shift towards heavier polysomal fractions in upf1Δ and particularly in upf3Δ strain.** Polysome fractionation traces followed by Northern blotting analysis of RNA from (A) wild type PTC27, (B) upf1Δ PTC27, (C) upf3Δ PTC27 polysomal fractions, hybridised with a GFP-specific probe. Methylene blue staining is shown bellow each Northern blotting panel, demonstrating polysome profiling and RNA quality as well as loading. (D) The intensity of the PTC27-containing mRNA for each lane of the wild type, upf1Δ and upf3Δ strain was determined using Image J software and was expressed as a percentage of the total signal.

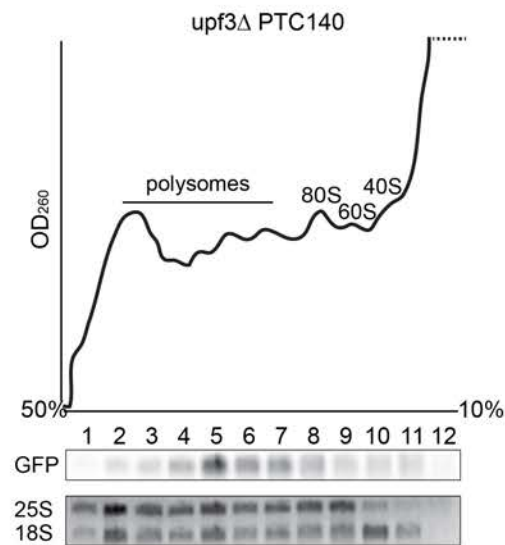
A



B



C



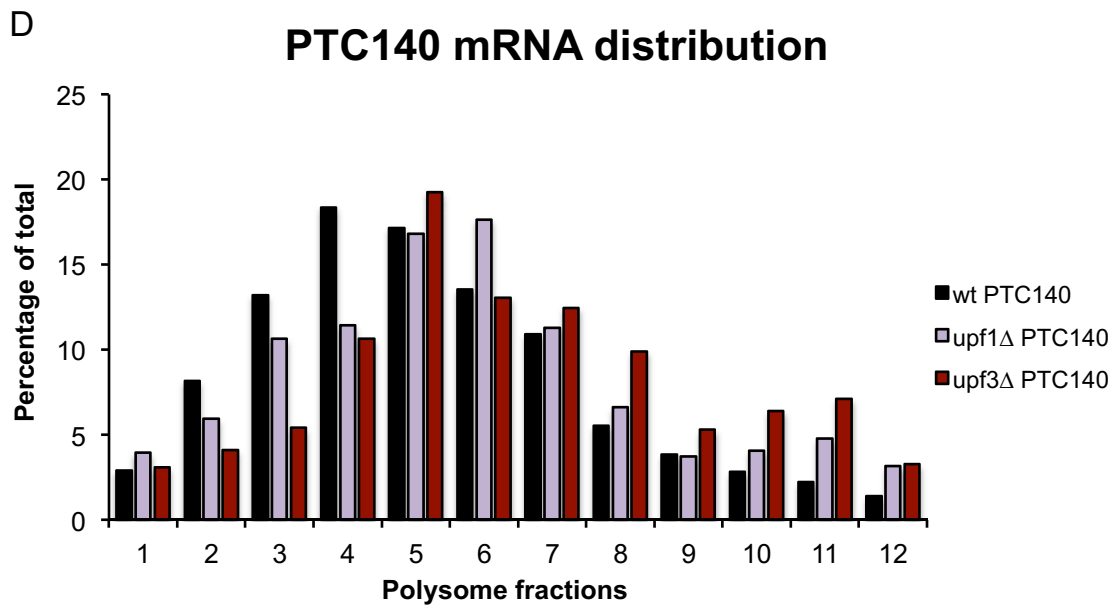


Figure 5.9 – **PTC140 GFP mRNA distribution follows a similar pattern as the functional GFP mRNA.** Polysome fractionation traces followed by Northern blotting analysis of RNA from (A) wild type PTC140, (B) upf1Δ PTC140, (C) upf3Δ PTC140 polysomal fractions, hybridised with a GFP-specific probe. Methylene blue staining is shown below each Northern blotting panel, demonstrating polysome profiling and RNA quality as well as loading. (D) The intensity of the PTC140-containing mRNA for each lane of the wild type, upf1Δ and upf3Δ strain was determined using Image J software and was expressed as a percentage of the total signal.

5.3. Discussion

This chapter details attempts to assess the ribosome release model of position-dependent NMD. The model proposes that in the case of premature translation termination, the release of ribosomes and other mRNA-associated components from mRNAs leads to the destabilisation of the mRNP and triggers subsequent degradation of the mRNA (Broгна and Wen, 2009; BroгнаMcLeod and Petric, 2016). I assessed this hypothesis by utilising NMD mutant strains that express NMD reporter genes and performed polysome fractionation to analyse their mRNA distribution throughout polysomal fractions. If, as the model predicts, UPF1 and potentially other NMD factors, affect the release of ribosomes from mRNA, in their absence, we would detect relatively more mRNA associated with heavier polysomal fractions.

Initial experiments have demonstrated that PTC6 GFP mRNA shifts towards heavier polysomal fractions in *upf1Δ*, and to a lesser extent in *upf2Δ* strain, when compared to its distribution in the wild type strain. The wild type GFP mRNA was also reproducibly associated with heavier polysomal fractions in *upf1Δ* mutant relative to the wild type strain. These results were in line with the prediction of the ribosome release model.

Upon obtaining new strains in which NMD suppression in the NMD mutants was more readily observed, the mRNA levels and the distribution observed in the polysomal samples were as expected in the wild type strain. PTC6 and PTC27 mRNA could potentially be associated with one or two ribosomes before translation terminates and the mRNA is degraded. This is exactly what we observed, with the mRNA being at low levels, present predominantly within the

light fractions, and fractions corresponding to the ribosomal subunits, monosome and disome. However, a much lower proportion of RNA was detected in the mutant strains in which the quantity was expected to be comparable to the wild type GFP. I hypothesised that this mRNA may be undetectable as it could potentially form complexes/granules. Alternative means of extracting the RNA, although improving RNA yield, did not improve PTC-containing mRNA detection. I also assessed whether these mRNAs became trapped within the nuclei or heavy cellular particles or if they are particularly loaded with the ribosomes and hence form a pellet at the very bottom of the ultracentrifugation tube, however, none of these was the case.

The RNA quality of the *upf1Δ* strain was lower than that of the wild type. I considered that the mRNA could potentially be particularly prone to degradation in this strain under the conditions used. Therefore, I modified the polysome fractionation protocol based on literature (Hu *et al.*, 2010). I introduced heparin to the lysis buffer, which binds to and stabilises RNA, and further inhibits RNase activity, and added Triton X-100 upon lysis to release any RNA trapped in granules or fat bodies. These modifications significantly improved RNA yield from polysome samples. I observed a minor shift towards heavier fractions which is in line with the prediction of the ribosome release model. However, I found that a considerable proportion of mRNA associates with the soluble fractions and ribosomal subunits, which suggests that either translation initiation is impaired or that the mRNA remains associated with ribosomal subunits upon termination. It also needs to be considered that using heparin can affect the polysome fractionation quality in yeast as heparin competes with ribosomes to

bind mRNA. This could potentially lead to polysome dissociation, which could in turn affect mRNA distribution.

In conclusion, the results indicate that the NMD factors could affect ribosome loading on NMD-targeted mRNAs. Additional experiments are needed to further substantiate and clarify the data presented in this chapter. One potential approach could be polysome fractionation paired with ribosome profiling where ribosome footprinting could identify whether the positions of ribosomes on PTC-containing mRNA and their abundance change in NMD mutants compared to the wild type strain. Also, as GFP is an exogenous gene, it could behave differently to endogenous genes. Therefore, I created endogenous NMD-reporter genes using CRISPR/Cas9 approach (explained in Chapter 6) to repeat the experiments detailed in this Chapter. I obtained the strains; however, I could not perform the analysis due to time constraints. Further efforts should be made to investigate this model in the future.

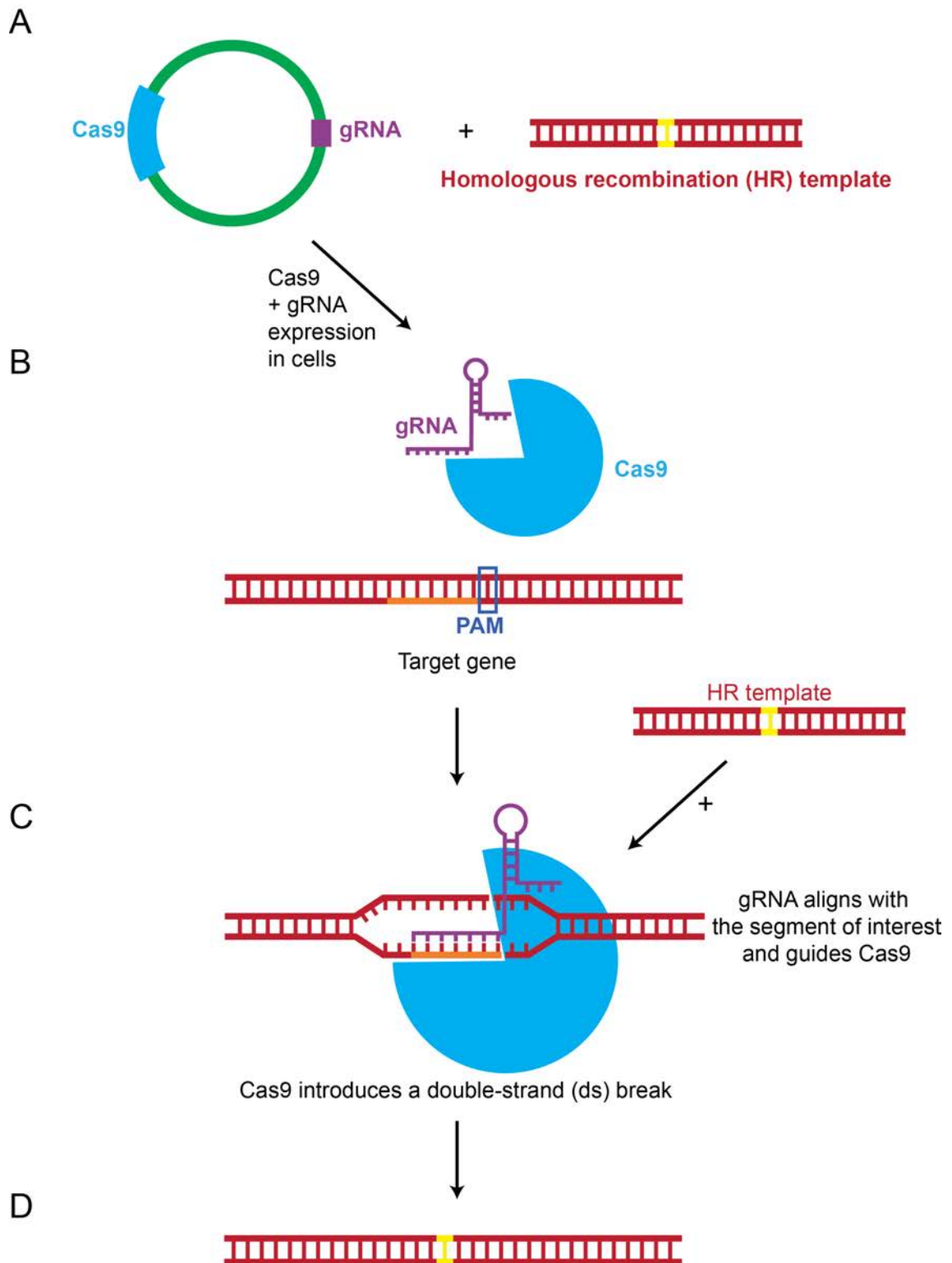
Chapter 6: CRISPR/Cas9 approach as a novel and efficient method to manipulate *S. pombe* genes

6.1 Introduction

Fission yeast *S. pombe* has been extensively used as a model organism, due to the ease of its use in laboratory conditions, fast growth rate and a fairly easy manipulation of its genome. *S. pombe* genome contains a higher percentage of intron-containing genes relative to the budding yeast *S. cerevisiae*, therefore, in studies looking at eukaryotic gene expression, particularly gene splicing, it is often the model of preference compared to *S. cerevisiae* (Kupfer *et al.*, 2004). An array of methods to modify genes, to introduce mutations, gene deletions or protein tagging, has been developed over the years, ranging from PCR-based gene targeting to allele replacement (pop-in pop-out method) (Bähler *et al.*, 1998; Krawchuk and Wahls, 1999; Gao *et al.*, 2014). CRISPR/Cas9 gene editing is an up and coming method to introduce changes in the gene of interest, with minimal alterations to the genome, and has only recently been implemented in *S. pombe* (Jacobs *et al.*, 2014). The aim of my work was to optimise CRISPR/Cas9 approach in our laboratory conditions, based on published data, to generate reagents for studying NMD and translation, and to compare its efficiency to more traditional methodologies.

As outlined in Materials and Methods, the CRISPR/Cas9 approach utilises gRNA that forms a complex with the Cas9 enzyme, leading it to the target gene by aligning with the genome just upstream of the protospacer adjacent motif (PAM), thereby allowing Cas9 to perform a double strand break (Figure 6.1) (Cong *et al.*, 2013; Mali *et al.*, 2013). The break can be repaired via non-

homologous end joining that introduces random mutations or, if a template homologous to the target site but containing the desired mutations is supplied, the double strand break is repaired via homology-directed repair (HDR) whilst introducing the desired alteration (Figure 6.1) (Cong *et al.*, 2013; Jacobs *et al.*, 2014).



The homology-directed repair (HDR) introduces the desired mutation/tag

Figure 6.1 – **CRISPR/Cas9 gene editing approach overview.** (A) Schematic of the two components required for CRISPR/Cas9 editing in *S. pombe* – a plasmid expressing Cas9 (blue) and gRNA (purple) required to direct Cas9 to the region of interest, and a homologous recombination (HR) template (red) containing the desired change (yellow) (point mutation, insertion, deletion or protein tag). (B) Schematic of Cas9 and gRNA expressed in the cells. Target gene is shown in red with protospacer adjacent motif (PAM) indicated in blue while the 20-nucleotide sequence upstream of the PAM indicated in orange is the sequence that will be recognised by the gRNA. (C) gRNA and Cas9 form a complex that is guided to the target gene by gRNA aligning with the target sequence. Cas9 then introduces a double strand break, which is repaired by the HR template supplied to the cells (shown on the right), incorporating the desired alteration. (D) Schematic of the gene after homology-directed repair and successful editing. Introduced changes prevent the gRNA from aligning to the locus again, and Cas9 from binding to this region, which prevents further alterations to the target gene.

In this chapter, I present how to most effectively create gRNA and homologous templates to use in CRISPR/Cas9 gene editing, to introduce specific mutations and to perform protein tagging in *S. pombe*, based on published data (Jacobs *et al.*, 2014). I used the CRISPR/Cas9 approach to generate endogenous NMD reporters by introducing PTCs at different positions along the sequences of *ade6* and *ade2* genes, which can further be used in studying NMD in *S. pombe* cells (Figure 6.2A). Ade6 was chosen as a candidate, as the mutants have a red colony phenotype, which would significantly help with screening and estimating the efficiency of mutagenesis. Ade2 was selected as it contains introns, and could therefore be used in studying splicing-dependent NMD. I also tested the efficiency of CRISPR/Cas9 editing in tagging proteins by creating strains expressing a small ribosomal protein RpS2801 tagged with 3xHA either at its N- or C-terminus (Figure 6.2B). The strains can be utilised in translation studies in the Brogna lab. Detailed gRNA and HR template design for *ade2* and *ade6* editing and RpS2801-HA tagging is outlined in the Appendix 8.7.

High efficiency, flexibility, minimal edits to the genome, and relatively fast strain generation could make the CRISPR/Cas9 approach a method of preference in the future.

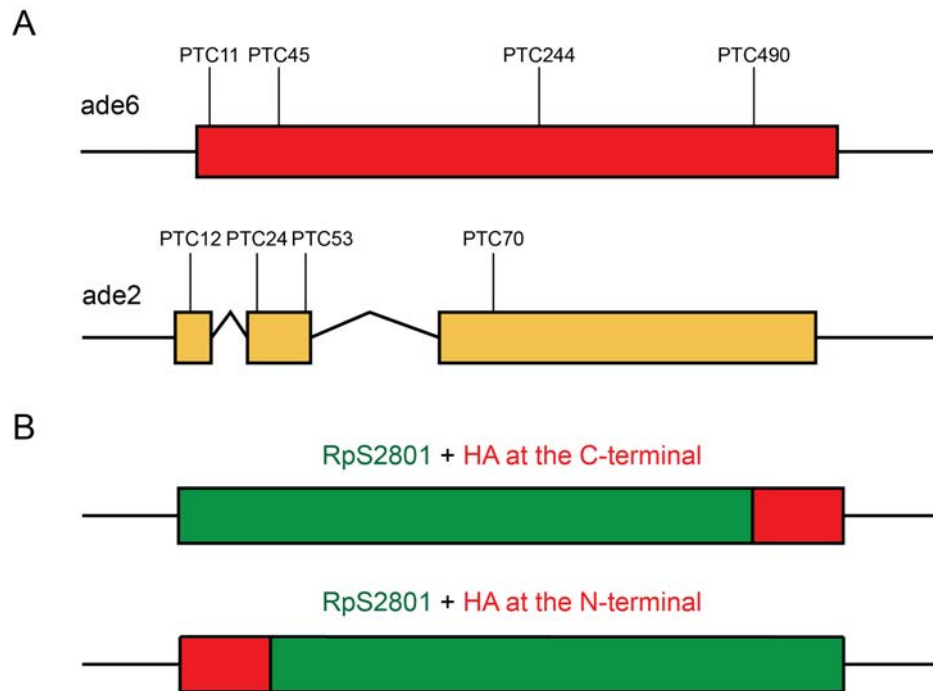


Figure 6.2 – **Schematic of the desired target gene edits to be generated by CRISPR/Cas9 approach** (A) Schematic of *ade6* (red) and *ade2* (yellow) genes with desired mutations – PTCs – indicated along the sequence. (B) Schematic of the *RpS2801* gene tagged with HA, either at the N- or the C-terminus. Only the tag is introduced and no selectable marker or any additional sequences are present.

6.2 Results

6.2.1 CRISPR/Cas9 gRNA design

As detailed in the Introduction, gRNA represents a key component of the CRISPR/Cas9 gene editing system as it forms a complex with the Cas9 nuclease and leads it to the desired gene locus, enabling double strand breakage and successful editing. One of the key steps of CRISPR/Cas9 approach, therefore, is gRNA design and its efficient expression in the cells.

When designing the gRNA, the key sequence to consider is the protospacer adjacent motif (PAM) that enables Cas9 to efficiently bind and cleave the desired target site. The PAM motif is a 3-nucleotide sequence 5'-NGG-3' in the gene of interest, N being any nucleotide. Gene-specific target gRNA is designed as a 20 nucleotide sequence just upstream the PAM (Figure 6.3) (Ran *et al.*, 2013). When gRNA is expressed in the cells, it forms a ribonucleoprotein complex with Cas9 that binds DNA and anneals with the target locus. Sufficient target homology, and the presence of the PAM motif just downstream of the alignment, trigger Cas9-induced DNA cleavage (Ran *et al.*, 2013).

It is important for the desired change to be close (within 5-10 nucleotides) of the PAM, or ideally adjacent to the PAM, so that once successful editing has occurred, the gRNA is no longer capable of aligning perfectly with the target and Cas9 is unable to cleave and alter the site any further (Cong *et al.*, 2013). The necessity of the PAM motif to be close to the editing site can be a limiting factor when designing specific mutations. However, as it is a fairly frequent sequence in the genome it allows a degree of flexibility (Ran *et al.*, 2013).

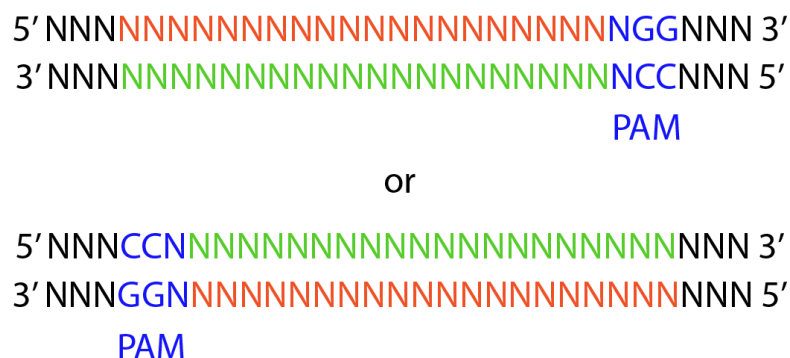


Figure 6.3 – **Gene-specific gRNA sequence design.** Firstly, a 5'-NGG-3' PAM sequence is selected (blue) close to the desired mutagenesis site, ideally at or 5-10 nucleotides within the desired alteration. If there are none such PAMs, the closest one is selected and a silent mutation is introduced at or close to the chosen PAM within the HR template, to abrogate further Cas9 cleavage after mutagenesis has occurred. Once the PAM is selected, a 20-nucleotide sequence just upstream of the PAM is chosen as the target gRNA, as indicated (orange represents its forward strand, while green its reverse strand).

Alternatively, introducing an additional silent mutation close to or at the PAM, within the HR template, is enough to render the target immune to further Cas9 cleavage, once the editing has been accomplished (Jacobs *et al.*, 2014).

Once the desired gene-specific gRNA sequence is chosen, it has to be expressed in the cells. Apart from the 20 nucleotides aligning to the target gene, additional sequences are required for the gRNA to fold properly and to form a complex with Cas9. Jacobs *et al.* designed a plasmid, pMZ374, which contains the *rrk1* promoter under which gRNA will be expressed, the *rrk1* leader sequence, a placeholder for the gene specific gRNA to be cloned in and the remaining sequences important for Cas9 binding, allowing efficient gRNA expression in *S. pombe* (Jacobs *et al.*, 2014).

I utilised two approaches to clone the gene-specific target sequence into the gRNA placeholder of the pMZ374 plasmid. Detailed gRNA design for *ade2* and *ade6* mutagenesis and RpS2801-HA tagging is outlined in the Appendix 8.7.

The first approach involved cloning using PCR, by amplifying the whole plasmid with two hybrid primers, where each contains a 20 nucleotide gene-specific gRNA sequence (as shown in Figure 6.3) and 20 nucleotides aligning within the placeholder, so that the gene specific gRNA introduced is in frame with the leader and remaining gRNA sequences (Figure 6.4A). The primers amplify the plasmid in opposite directions, generating a linear DNA molecule, which is, upon DpnI treatment to remove template DNA, transformed into *E. coli*, where it becomes circularised (Figure 6.4A). Plasmids resulting from plasmid preparation in *E. coli* were first checked by restriction enzyme digestion to ensure that no major rearrangements occurred and the correct insertion was

then verified by sequencing. Only gRNAs designed to generate *ade2* gene containing a PTC at position 12 and *ade6* gene containing a PTC at position 45 were cloned into the pMZ374 vector in this manner (Figure 6.4B). This approach was not successful for any other gRNA, probably due to the length of the plasmid being approximately 11.5 kb and the PCR reactions of such long DNA not being efficient.

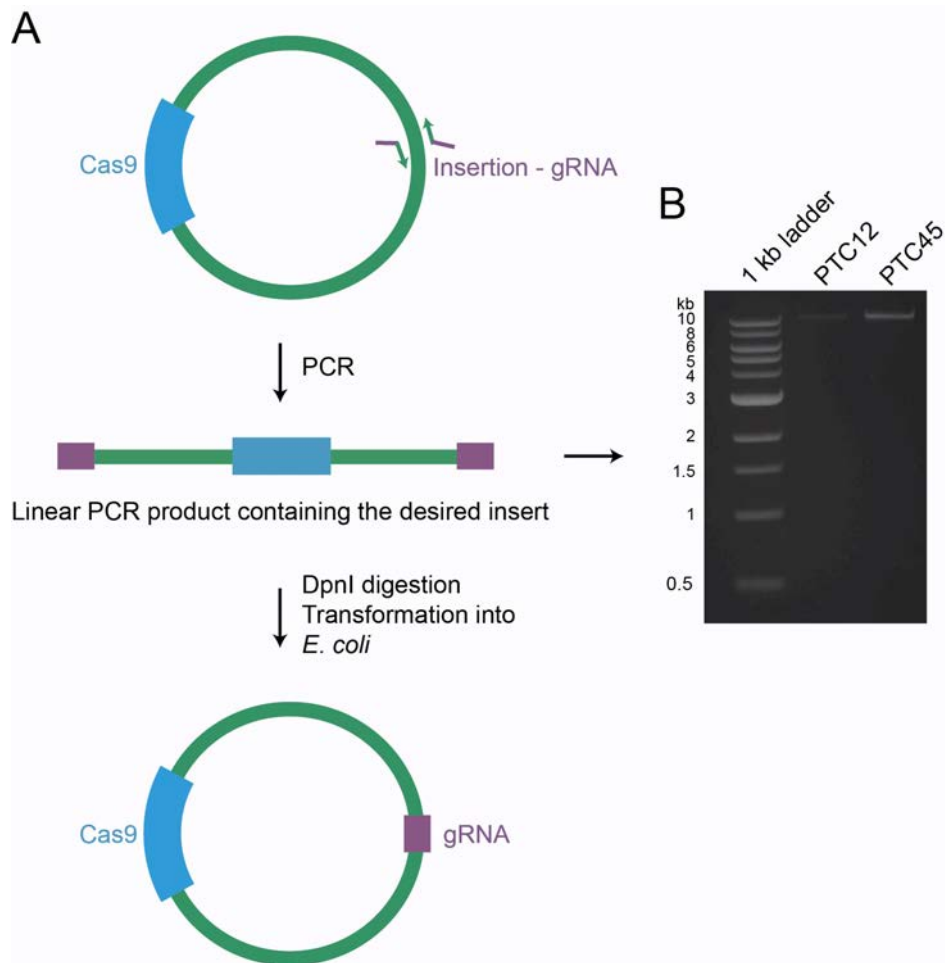


Figure 6.4 – **Cloning the gene-specific gRNA sequence into its placeholder on pMZ374 plasmid using PCR approach.** (A) Schematic of the procedure. Two hybrid primers aligning with the plasmid placeholder and containing a 20-bp gene specific gRNA sequence at the 5' end are used in a PCR reaction to amplify the whole plasmid, introducing the required sequence. The reaction is treated with the DpnI enzyme to eliminate the template. The PCR product is re-circularised upon its transformation into *E. coli* cells. The plasmid is amplified by plasmid preparation and its correct integration is verified by sequencing. (B) Two gRNAs, one to generate PTC12-containing *ade2* and second to create PTC45-containing *ade6*, were generated as described, and were visualised by agarose gel electrophoresis and ethidium bromide staining as a band of approximately 11.5 kb, shown on the right.

The second approach involved cloning using Gibson assembly of two DNA fragments, one generated by restriction enzyme digestion and second by PCR. The pMZ374 plasmid was digested using *SpeI* and *StuI* restriction enzymes creating an approximately 10 kb long desired fragment (Figure 6.5A). The hybrid primers described above were used in two PCR reactions that amplify fragments of the plasmid surrounding the integration site so that the ends of the resulting PCR products overlap with the digested plasmid (Figure 6.5B, PCR1). The two PCR products were joined together in another PCR reaction using end primers (Figure 6.5B, PCR2). In this way, I generated a DNA fragment that contains gene specific gRNA at the desired location, and a 100-bp nucleotide overlap with the plasmid backbone at each end. Due to the overlapping ends, the plasmid backbone generated by restriction enzyme digestion and the PCR product are joined together using Gibson Assembly Master Mix (NEB) (Figure 6.5C), which is then transformed into *E. coli*. Upon plasmid preparation, the plasmids were first checked using restriction enzyme digestion to test whether any rearrangements occurred. Correct integration of gene specific gRNA into its placeholder, in frame with the gRNA leader sequence on one side and the remaining gRNA sequences on the other, was tested via sequencing. Using this approach, gRNA-expressing plasmids for all the remaining *ade2* and *ade6* gene mutations (Figure 6.6A; B) as well as two gRNAs for N- and C-terminal RpS2801 protein tagging were generated. All the plasmids screened by sequencing (three plasmids for each of the gRNAs) had correct integration, making this approach exceptionally efficient.

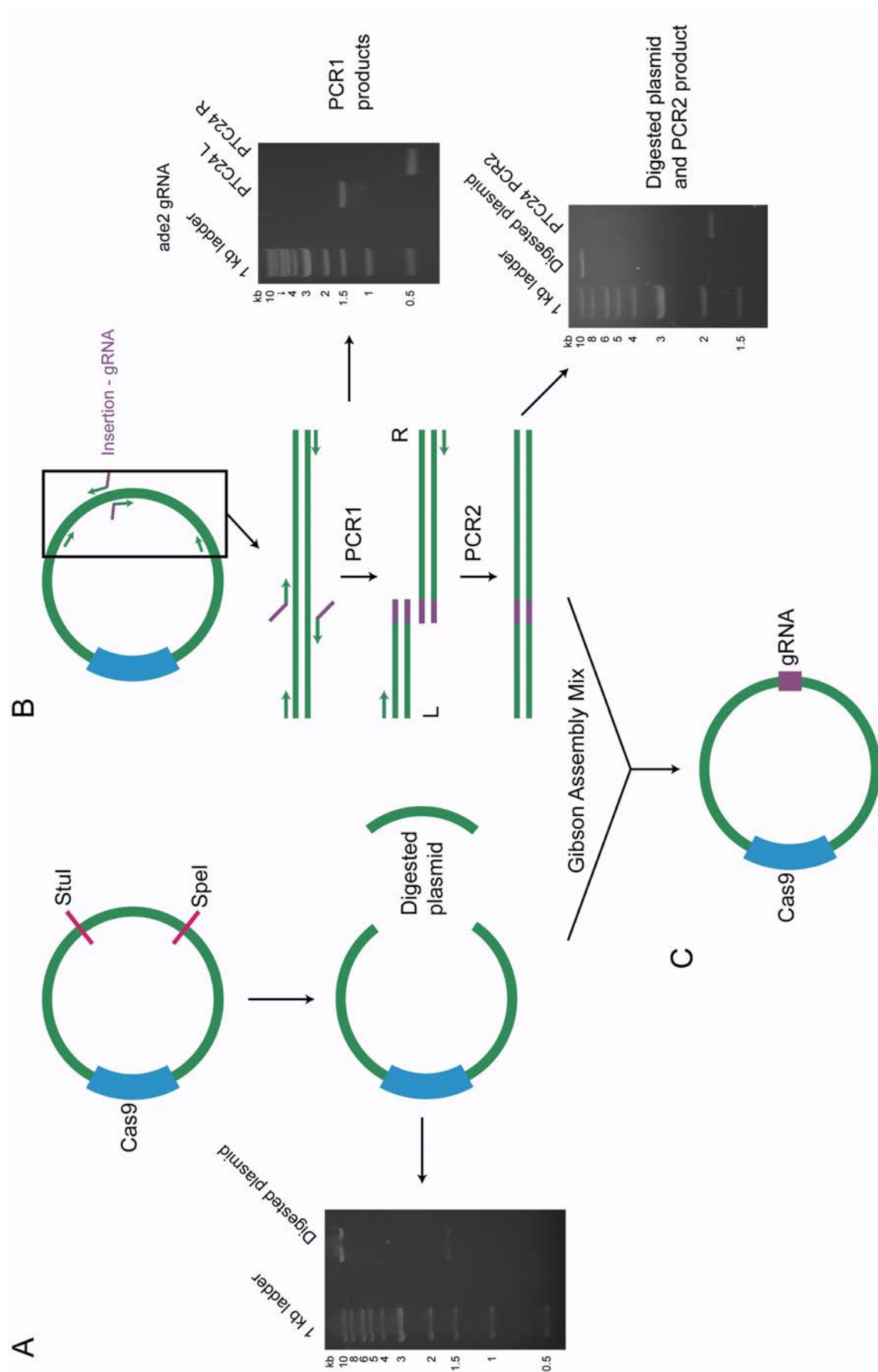


Figure 6.5 – **Cloning the gene-specific gRNA sequence into its placeholder on the pMZ374 plasmid using Gibson Assembly Master Mix.** (A) Schematic of restriction enzyme digestion of the pMZ374 plasmid using *Stu*I and *Spe*I restriction enzymes to cut the region of the desired insertion site out of the plasmid backbone. The two resulting fragments are visualised by agarose gel electrophoresis and ethidium bromide staining on the left, with the fragment of approximately 10 kb in length being the one desired, purified from the gel to be further used in the Gibson assembly reaction. (B) Schematic of PCR reactions used to introduce gene specific gRNA into its pMZ374 plasmid placeholder and generate the second fragment to be used in the Gibson assembly mix. Two hybrid primers containing (from 5' to 3' end) a 20-nucleotide gene specific gRNA sequence and 20 nucleotides that align with the integration site on the plasmid are used in two PCR reactions with flanking primers, amplifying the two regions surrounding the placeholder (PCR1), generating left and right PCR products (L and R). Agarose gel electrophoresis of such fragments used for PTC24-containing *ade2* gRNA construction is shown on the right, as an example. PCR1 reaction products share 20 bp of homology and are used in another PCR reaction (PCR2), which joins them together and creates a second DNA fragment used in the Gibson assembly reaction. Agarose gel electrophoresis visualising the product of PCR2 reaction for PTC24-containing *ade2* gRNA construction, together with the digested plasmid is shown on the right, as indicated. (C) The digested plasmid and product of the PCR2 reaction, that share approximately a 100 bp overlap at their ends were finally joined together using Gibson Assembly Master Mix, creating a plasmid expressing both Cas9 and gRNA targeting the gene of interest.

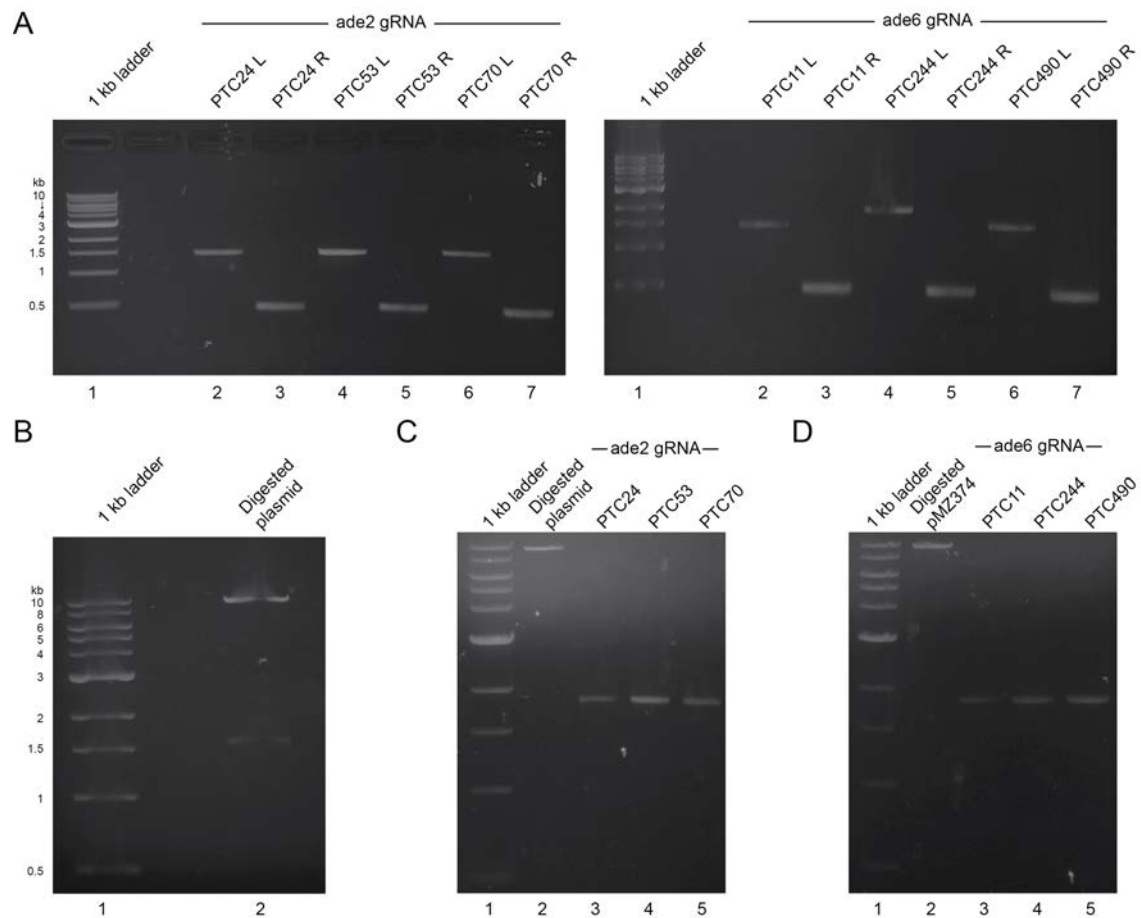


Figure 6.6 – **Cloning gene-specific gRNA sequence for *ade2* and *ade6* gene editing into pMZ374 plasmid using Gibson Assembly Master Mix.** Agarose gel electrophoresis and ethidium bromide staining of (A) left and right PCR1 products, as detailed above (Figure 6.4B), for *ade2* (left) and *ade6* (right) gene editing, as indicated, (B) two pMZ374 plasmid fragments generated by *Stu*I and *Spe*I restriction enzyme digestion (2), (C) purified pMZ374 restriction enzyme digestion fragment (2) and PCR2 products (Figure 6.4B), used in a Gibson assembly reaction to generate a plasmid expressing gRNA and Cas9 for creating PTC24 (3), PTC53 (4) and PTC70-containing (5) *ade2* strains, (D) purified pMZ374 restriction enzyme digestion fragment (2) and PCR2 products, used in a Gibson assembly reaction to generate gRNA and Cas9 expressing plasmid for creating PTC11 (3), PTC244 (4) and PTC490-containing (5) *ade6* strains.

6.2.2 Creating HR templates for CRISPR/Cas9 induced gene mutagenesis

To create specific mutations, it is necessary to supply the cells with a DNA molecule that is homologous to the target gene, but contains the desired alterations, termed as the homologous recombination (HR) template. If this were not supplied, the double strand break introduced by Cas9 would be repaired by non-homologous end joining (NHEJ), resulting in the introduction of random mutations. NHEJ is only favourable if the goal is to inactivate the gene rather than to introduce specific changes.

To create HR templates for *ade2* and *ade6* gene mutagenesis, I used the overlapping PCR approach. The first round of two PCR reactions was carried out, using genomic DNA as a template, amplifying the surrounding region of approximately 300 bp, whilst introducing the desired alteration (Figure 6.7). The first PCR reaction used a forward primer aligning to a sequence approximately 300 bp upstream of the desired mutation and a 50-nucleotide long reverse primer perfectly complementary to the target gene from its 3' end (to ensure its hybridisation with the target sequence), while containing the desired mutation close to its 5' end. The same principle was used to generate the second PCR product – a 50-nucleotide forward primer containing the mutation at its 5' end and otherwise perfectly aligning with the target gene, paired with the primer hybridising with a sequence approximately 300 bp downstream of the mutation site. In this way, two PCR products (left and right, L and R) are generated, containing 30 bp-overlapping ends, which are used in the final PCR reaction that joins them together into a complete HR template including the desired mutation (Figure 6.7). Sequencing the HR template confirms whether the

desired change is indeed introduced. Products of the first and second round of PCR reactions used to create *ade2* and *ade6* mutagenesis HR templates, visualised by agarose gel electrophoresis, are shown in Figure 6.8.

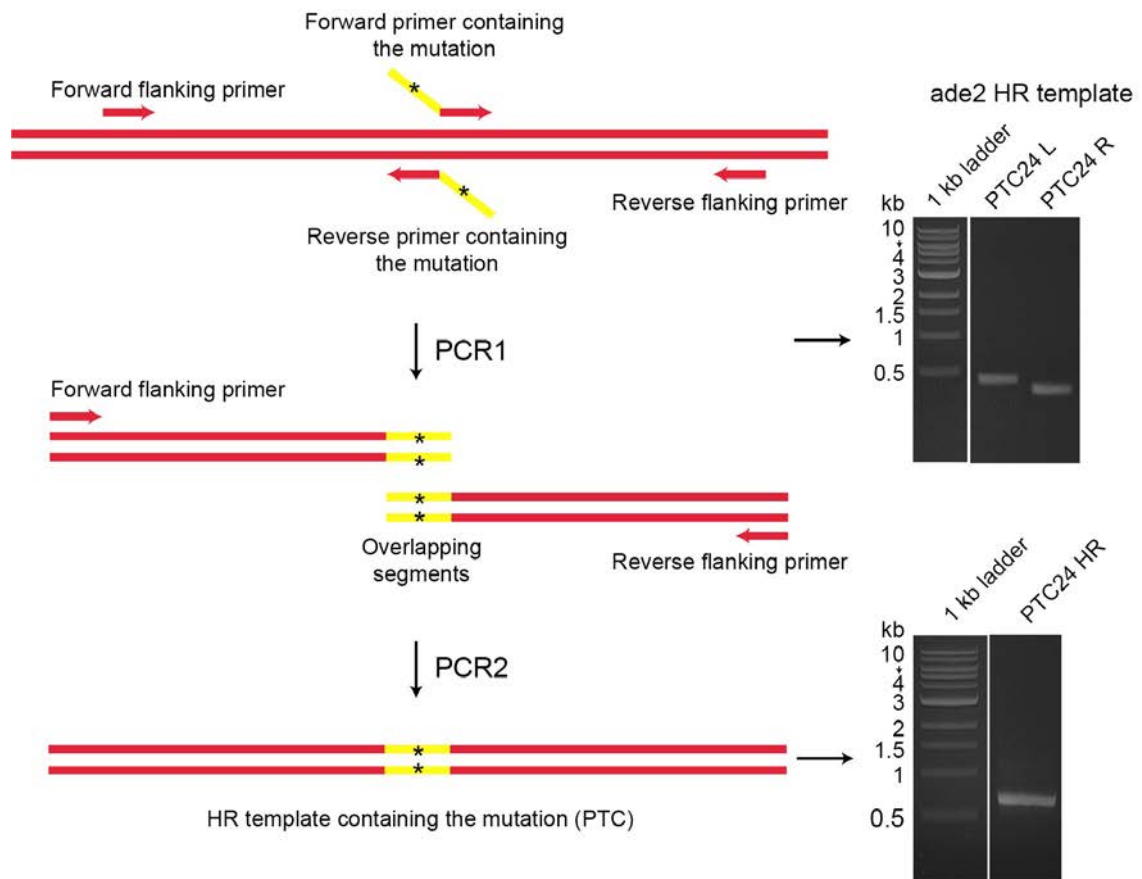


Figure 6.7 – **Creating HR templates for *ade6* and *ade2* gene editing, using a PCR-based approach** Schematic of PCR reactions used to create an HR template containing a desired point mutation (indicated with an asterisk). A 50-nucleotide long reverse primer designed to contain its 3' end pairing with the gene of interest while also containing a desired mutation towards its 5' end was used in the left PCR reaction with another flanking primer aligning approximately 300 bp upstream of the site of interest. The same principle was used on the right side, as shown. Two PCR reactions generate fragments that have a 30 bp overlap, which allows them to be joined together via another PCR reaction, using flanking primers, finally creating an approximately 600 bp-long HR template, containing the desired mutation. Agarose gel electrophoresis and ethidium bromide staining, displaying products of the first and second PCR reactions to generate PTC24-containing *ade2* HR template, is shown on the right, as an example.

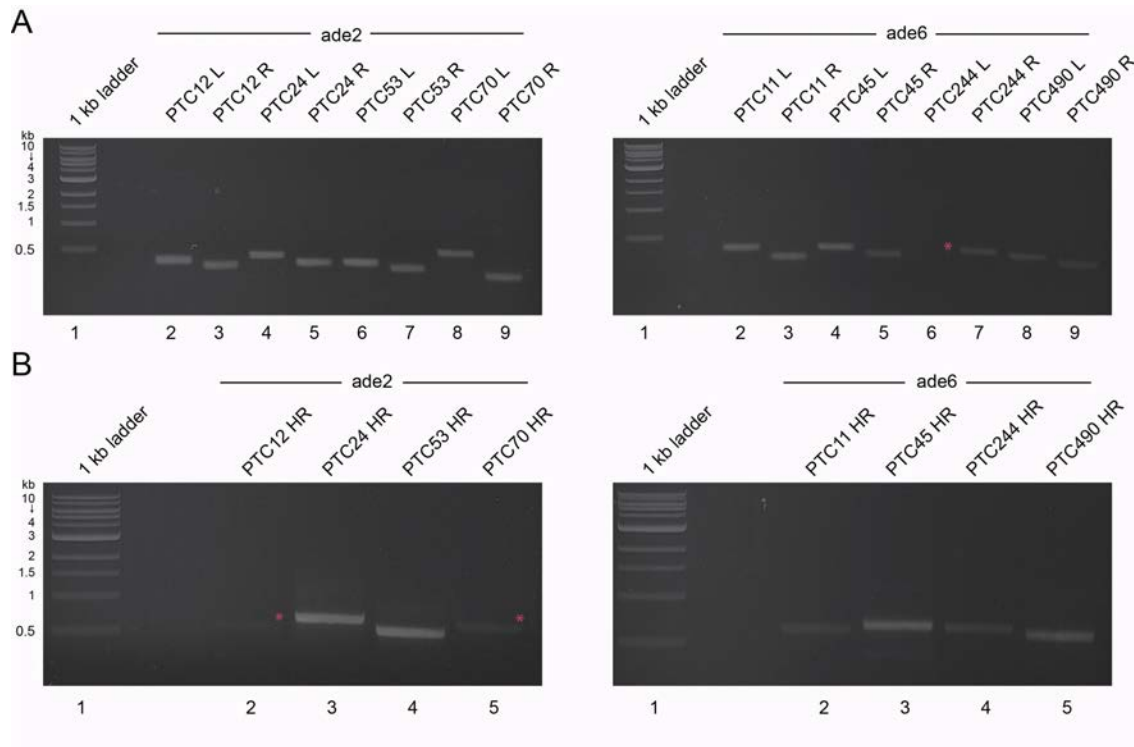


Figure 6.8 – **Generating HR templates for *ade2* and *ade6* gene editing** Agarose gel electrophoresis and ethidium bromide staining of (A) left and right PCR products obtained in the first PCR reaction, as detailed above (Figure 6.7, PCR1), for *ade2* (left) and *ade6* (right) gene editing, (B) PCR products obtained by joining the left and right fragments in the second PCR reaction (Figure 6.7, PCR2), representing HR templates for creating *ade2* (left) and *ade6* (right) gene mutants – PTC12 (2), PTC24 (3), PTC53 (4), PTC70 (5)-containing *ade2* strains (left) as well as PTC11 (2), PTC45 (3), PTC244 (4), PTC490 (5)-containing *ade6* strains (right). Bands of a very low signal are indicated with a red asterisk.

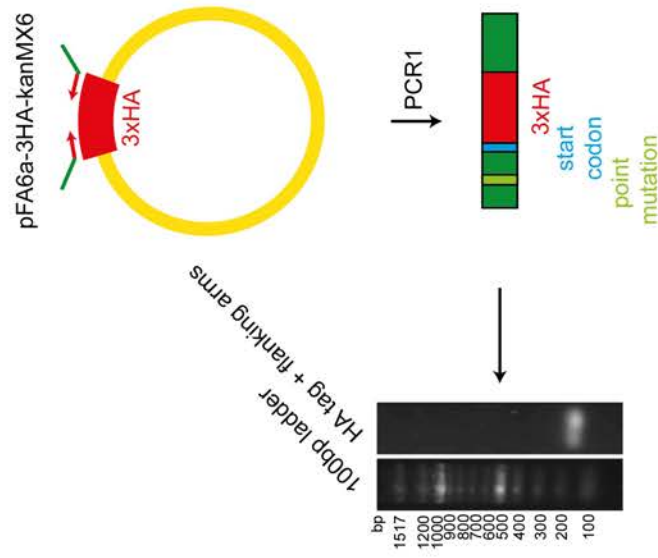
6.2.3 Creating HR templates for CRISPR/Cas9 induced protein tagging

To be able to tag target genes, it is again necessary to supply the cells with a DNA molecule that is homologous to the gene of interest and contains the tag in frame with the coding sequence. To generate HR templates for protein tagging, I again implemented an overlapping PCR approach.

To create HR templates for N-terminal protein tagging, I firstly amplified the 3xHA tag (3xHA) from a pFA6a-3HA-kanMX6 plasmid, by using two hybrid primers (Figure 6.9A, PCR1). The forward primer contained, from its 5' to 3' end, 20 nucleotides aligning with the sequence just upstream and containing the start codon of the target gene and 20 nucleotides aligning to the start of the HA tag, while the reverse primer contained 20 nucleotides aligning to the sequence just downstream of the start codon and 20 nucleotides homologous to the very end of the HA tag. Additionally, the forward primer contained a silent mutation within the range of the chosen PAM motif, to make sure that upon introducing the tag, Cas9 can no longer cut and alter the gene of interest. In such a way, a PCR product, containing the HA-tag flanked with gene-specific sequences, is generated, where start codon is perfectly in frame with the HA tag and the coding sequence (Figure 6.9A). On the other hand, using genomic DNA as a template, two PCR reactions were carried out amplifying approximately 300 bp of the desired RpS2801 gene surrounding the insertion site (Figure 6.9B, PCR2). In such a way, a 300-bp DNA fragment upstream and including the start codon and a 300-bp DNA fragment downstream of the start codon were generated. The three PCR products containing the indicated overlapping ends were used in the final PCR reaction (Figure 6.9B, PCR3),

generating RpS2801 N-terminal tagging HR template that contains the HA tag just downstream of the start codon, in frame with the gene coding sequence, together with the silent point mutation in the gene's 5' UTR, to render the target gene immune to further alternations by Cas9 once the tag is introduced. The products of all described PCR reactions, visualised by agarose gel electrophoresis and ethidium bromide staining are shown in Figure 6.9. Upon successfully obtaining the HR template, it was further verified by sequencing.

A



B

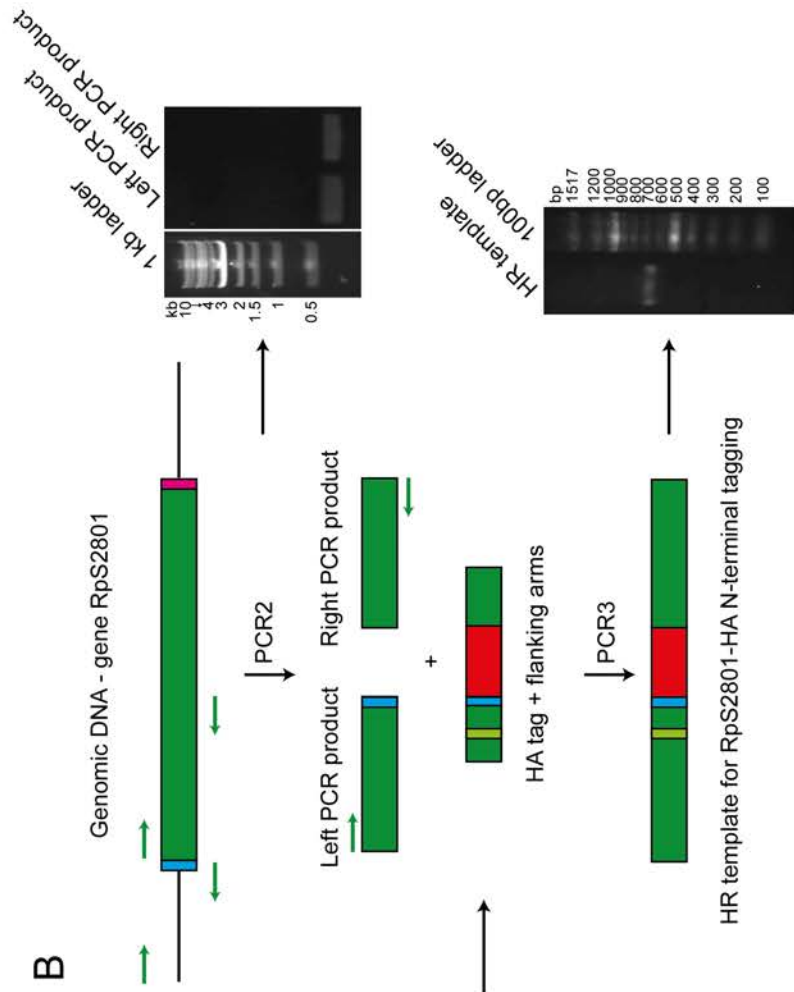


Figure 6.9 – Creating HR templates for N-terminal protein tagging, using a PCR-based approach Schematic of PCR reactions used to create HR template containing a desired tag (3xHA in this case). **(A)** Schematic of the pFA6a-3HA-kanMX6 plasmid used to amplify HA in a PCR reaction (PCR1), using hybrid primers containing a 20-nucleotide sequence aligning with HA (red) and a 20-nucleotide sequence of the gene of interest at its 5' end (green). The gene specific sequence of the forward primer contains 20 nucleotides upstream and including the start codon (blue) as well as a silent mutation (light green) to ensure Cas9 will not cut the region again after the desired change is introduced. The gene specific sequence of the reverse primer contains 20 nucleotides downstream of the gene start codon. Agarose gel electrophoresis and ethidium bromide staining of HA amplified with RpS2801 gene specific flanking arms is shown on the left. **(B)** Genomic DNA is used as a template in two PCR reactions (PCR2), amplifying approximately 300 bp from each side of the integration site. The left PCR product is amplified up to and including the start codon, while the right PCR product generates an approximately 300 bp-long fragment, just downstream of the start codon. Two such PCR2 reaction products amplifying RpS2801 gene fragments are visualised using agarose gel electrophoresis and ethidium bromide staining shown on the right. PCR products from PCR1 and PCR2 reactions contain overlapping ends and are used in the final PCR reaction (PCR3), creating an HR-template used for N-terminal RpS2801-HA tagging, as visualised in agarose gel electrophoresis – bottom right.

To create HR templates for C-terminal protein tagging, I again amplified the 3xHA tag using two hybrid primers, one of which contained, from the 5' to 3' end, 20 nucleotides aligning with the sequence just upstream of the gene stop codon and 20 nucleotides aligning to the start of the HA tag, while the reverse primer contained 20 nucleotides aligning to the sequence including, and just downstream of, the stop codon, as well as 20 nucleotides homologous to the very end of the HA tag. The reverse primer contained a point mutation within the range of the chosen PAM motif; to render the target gene immune to further alternations once the tag is introduced (Figure 6.10A). In such a way, a PCR product containing the HA tag flanked with gene-specific sequences is generated, where the coding sequence is perfectly in frame with the HA tag and the stop codon (Figure 6.10A). Using genomic DNA as the template, two PCR reactions were carried out amplifying approximately 300 bp of the desired RpS2801 gene surrounding the insertion site. A 300-bp DNA fragment upstream of the stop codon as well as a 300-bp DNA fragment downstream and including the stop codon was generated (Figure 6.10B). The three PCR products contain overlapping ends and were used in the final PCR reaction, generating an RpS2801 C-terminal tagging HR template that contains HA tag in frame with the gene coding sequence, terminated by a stop codon, together with the silent point mutation in the gene 3' UTR (Figure 6.10B). Using this method, I created the HR template for RpS2801 C-terminal tagging. The products of all PCR reactions, visualised by agarose gel electrophoresis and ethidium bromide staining are shown in Figure 6.10. Upon successfully obtaining the HR template, it was further verified by sequencing.

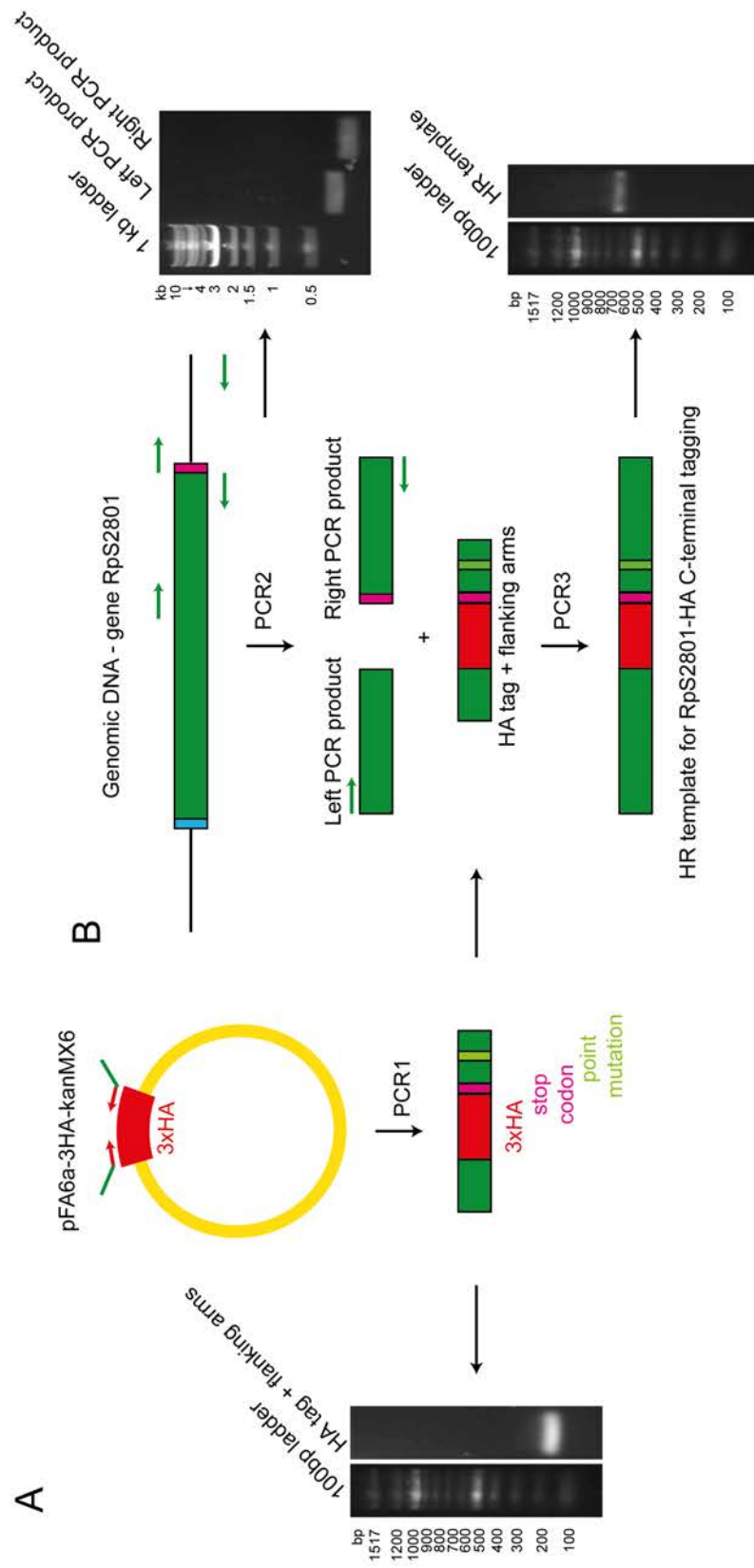


Figure 6.10 – Creating HR templates for C-terminal protein tagging, using a PCR-based approach Schematic of PCR reactions used to create the HR template containing the desired tag (3xHA in this case). **(A)** Schematic of the pFA6a-3HA-kanMX6 plasmid used to amplify HA in a PCR reaction (PCR1), using hybrid primers containing a 20-nucleotide sequence aligning with HA (red) and a 20-nucleotide sequence of the gene of interest at the 5' end (green). The gene specific sequence of the forward primer contains 20 nucleotides upstream of the stop codon while the gene specific sequence of the reverse primer contains 20 nucleotides downstream of and including the gene stop codon, as well as a silent mutation (light green) to ensure that the region will not be cut again by Cas9 after the desired change is introduced. Agarose gel electrophoresis of HA amplified with specific flanking arms of the RpS2801 gene is shown on the left. **(B)** Genomic DNA is used as a template in two PCR reactions (PCR2), amplifying approximately 300 bp from each side of the integration site. The left PCR product is amplified upstream of the stop codon, while the right PCR product generates an approximately 300 bp-long fragment downstream of, and including, the stop codon. Two such PCR2 reaction products amplifying the RpS2801 gene are visualised using agarose gel electrophoresis and ethidium bromide staining shown on the right. PCR products from PCR1 and PCR2 reactions contain overlapping ends and are used in the final PCR reaction (PCR3), creating the HR-template used for C-terminal RpS2801-HA tagging, as visualised using agarose gel electrophoresis – bottom right.

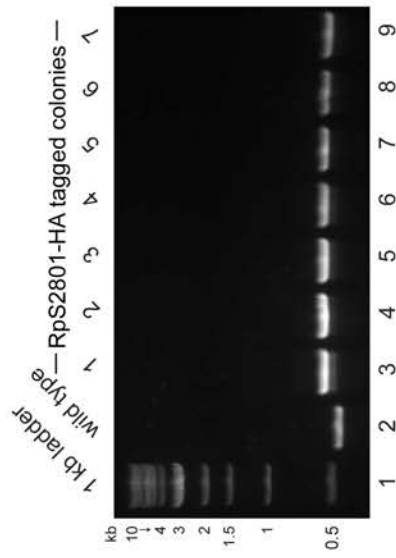
6.2.4 CRISPR/Cas9 gene editing approach efficiency

Upon transforming the Cas9 and gRNA-expressing plasmid, as well as the appropriate HR template, the colonies were screened in several ways. In case of *ade6* editing, the red colony phenotype indicated that the editing was successful (Figure 6.11A; B; C; D). However, to determine that the desired change was introduced, the strains were verified by colony PCR paired with sequencing (Figure 6.11E). As *ade2* mutations do not cause any visible phenotypic effects, the correct mutagenesis was analysed by colony PCR and sequencing (Figure 6.11F).

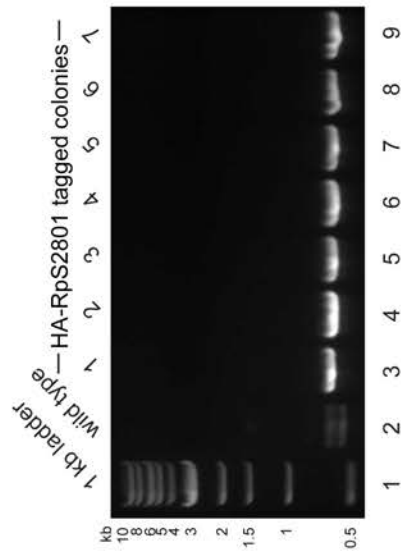
Figure 6.11 – **Verifying ade2 and ade6 mutant strains.** (A) PTC11-, (B) PTC45-, (C) PTC244- and (D) PTC490-containing ade6 mutant strains, obtained using the CRISPR/Cas9 gene editing approach, visualised as red colonies, while the white colonies represent the cells in which the editing process did not occur. The strains were further verified by colony PCR paired with sequencing. (E) An example sequencing trace is shown for ade6 gene mutagenesis (PTC11-containing gene) with the introduced PTC highlighted in blue. (F) Ade2 mutant strains lack visible phenotypic characteristics, so the correct mutagenesis was verified solely by colony PCR and sequencing. The graph illustrates an example sequencing trace for ade2 gene mutagenesis (PTC24-containing gene) with the introduced PTC highlighted in blue.

To screen for 3xHA protein tagging, the region of the insertion was typically amplified using colony PCR, and the size of the product was compared to that of the original, parental product. If the tag was introduced, the size of the DNA fragment should be larger than the original fragment by the size of the tag. All seven colonies tested, both in case of N-terminal and C-terminal protein tagging, contained the insertion, meaning that the tagging efficiency of the CRISPR/Cas9 approach was 100% (Figure 6.12A; B). Protein extracts from five of these colonies were analysed by SDS-PAGE/Western blotting, to determine whether the protein is indeed successfully expressed with the tag (Figure 6.12C). RpS2801 was successfully HA-tagged in 60% of the colonies tested (Figure 6.12C, lanes 5, 7, 8, 10, 11, 13). As ribosomal protein genes are typically present in the cells in two copies in *S. pombe* (Coddington and Fluri, 1977; Gross and Käufer, 1998), the second, untagged copy is probably preferentially expressed in the remaining 40% of cells in which RpS2801-HA was not detected (Figure 6.12C, lanes 4, 6, 9, 12).

A



B



C

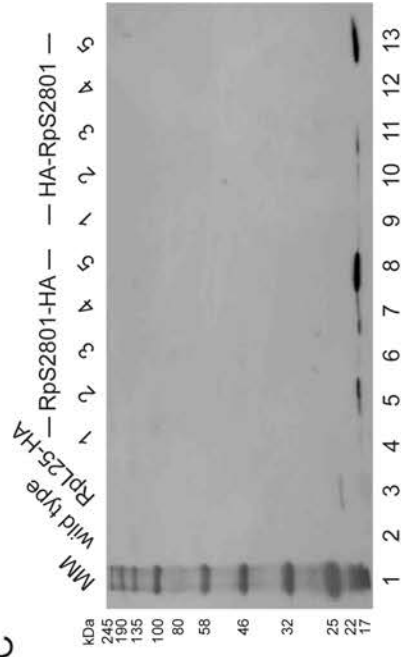


Figure 6.12 – **Verifying RpS2801 strains tagged with HA.** (A) Agarose gel electrophoresis of colony PCR amplifying the region of 3xHA insertion in RpS2801-HA C-terminally tagged strains. Seven colonies were tested (3-9) and the size of the product was compared to that of the wild type (2), parental strain. The shift in size towards higher molecular weight of approximately 100 bp, of the size of the 3xHA tag, suggests that the tag was integrated in all 7 colonies analysed. (B) Agarose gel electrophoresis of colony PCR used to amplify the region of 3xHA insertion in RpS2801-HA N-terminally tagged strains. Seven colonies were tested (3-9) and the size of the product was compared to that of the wild type, parental strain (2). The shift in size towards higher molecular weight of approximately 100 bp, of the size of the HA tag, suggests that the tag was integrated in all 7 colonies analysed. Two faint bands in the wild type strain (2) are due to a cross-contamination with a colony of the tagged strain, as the tip containing a residue of the transformed strain was accidentally dipped into this PCR reaction. (C) Western blotting of protein extracts of wild type as a negative control (2), RpL25-HA tagged strain as a positive control (3), RpS2801-HA C-terminally tagged strains derived from colonies 1-5 (lanes 4-8) and RpS2801-HA N-terminally tagged strains derived from colonies 1-5 (lanes 9-13). Bands corresponding to C-terminally HA tagged RpS2801 were detected in cells derived from colonies 2, 4 and 5 while bands corresponding to N-terminally HA tagged RpS2801 were detected in cells derived from colonies 2, 3 and 5.

The efficiency of CRISPR/Cas9 gene editing approach to generate the desired strains – mutants as well as tagged proteins is summarised in Table 6.1. The mutagenesis efficiency was calculated as a number of successful mutants out of the total number of colonies screened for each of the strains.

Mutation	Efficiency
ade6 PTC at position 11	87.27% - the fraction of red colonies out of the total number; 75% of tested colonies had the correct mutation upon sequencing
ade6 PTC at position 45	57.14% - the fraction of red colonies out of the total number; 25% of colonies tested had the correct mutation upon sequencing
ade6 PTC at position 244	87.02% - the fraction of red colonies out of the total number; 80% of colonies tested had the correct mutation upon sequencing
ade6 PTC at position 490	89.95% - the fraction of red colonies out of the total number; 75% of colonies tested had the correct mutation upon sequencing
ade2 PTC at position 12	Mutagenesis was unsuccessful – 0%
ade2 PTC at position 24	75% of colonies tested had the correct mutation upon sequencing
ade2 PTC at position 53	33.33% of colonies tested had the correct mutation upon sequencing
ade2 PTC at position 70	66.67% of colonies tested had the correct mutation upon sequencing
RpS2801, C-terminally tagged with HA	100% of colonies tested contained an insert in the gene while 60% expressed the protein tagged with HA
RpS2801, N-terminally tagged with HA	100% of colonies tested contained an insert in the gene while 60% expressed the protein tagged with HA

Table 6.1 – **CRISPR/Cas9 editing efficiency.** The table details the efficiency of the approach based on the percentage of correct mutants out of the total number of colonies screened.

6.3 Discussion

This chapter details implementation of CRISPR/Cas9 approach in our laboratory conditions to create reagents, such as endogenous NMD reporter genes and small ribosomal protein-tagged strains, to study NMD and translation in *S. pombe*.

Two key components required for successful CRISPR/Cas9 editing in *S. pombe* are the pMZ374 plasmid expressing Cas9 and gRNA targeting the locus of interest, as well as an HR template that introduces the desired modification. The plasmid already contains the *rrk1* gRNA expression system into which gene-specific gRNA needs to be cloned (Jacobs *et al.*, 2014). By far the best method to clone this sequence in our laboratory conditions was by using the Gibson Assembly Master mix, detailed in section 6.2.1. This is a multistep process, requiring several PCR reactions, restriction enzyme digestion and also a Gibson assembly reaction; however, this multistep process was very efficient, achieving 100% correct integration. This method is indeed less laborious than cloning the fragment by amplifying the complete plasmid, as, due to the plasmid's length (11.5 kb), the PCR reaction can often be inefficient and can require extensive troubleshooting. HR templates were generated using the overlapping PCR approach, as explained in sections 6.2.2 and 6.2.3. Whilst optimising the PCR reactions, increasing the length of the overlap between the fragments that are joined together provides the cleanest PCR results.

Evaluating CRISPR/Cas9 efficiency when looking at the *ade6* mutagenesis phenotypically as a fraction of red colonies out of the total number, showed that it ranged from as little as 57.41% to a remarkable 89.95%, with an average of

80.35%, meaning that this approach is extremely successful in generating mutant strains. When the red colonies were screened for correct mutations induced by homology directed repair, via sequencing, between 25% and 80% of colonies (average 63.75%) contained the desired mutations. The remaining displayed other small insertions or deletions suggesting that Cas9 did introduce the cut, but that this was repaired by the NHEJ mechanism. Modifying the quantity of the HR template supplied to the cells could therefore significantly increase editing efficiency. Regarding *ade2* gene mutagenesis, the colonies do not have a visible phenotype, and were therefore solely screened by sequencing. The efficiency of correct integration ranged from 33.33% to 75% (average 58.33%), with the remaining colonies typically preserving the parental gene or containing other mutations. Several attempts were made to create a PTC12-containing *ade2* gene, however, they were unsuccessful; all colonies screened showed that the parental gene was preserved, suggesting that the designed gRNA most probably triggered Cas9-induced cleavage inefficiently. Therefore, the key to successful editing depends upon gRNA selection as well as balancing the amount of HR template added to the cells. The efficiency of protein tagging was remarkably 100%, with all colonies tested having the desired 3xHA insert. Only 60% of cells expressed the protein tagged with HA, while in the remaining cells the second untagged copy of the gene is probably preferentially expressed.

CRISPR/Cas9 shows comparable efficiencies to more traditional methodologies such as PCR-based gene targeting. The advantages of CRISPR/Cas9 system are the flexibility, the ability to introduce specific changes of interest however

small, and to inactivate only parts of proteins rather than having to carry out complete deletions. Pop in pop out allele replacement can also be used to introduce specific changes into the genome; however, in our experimental conditions, its efficiency in *ade6* mutagenesis was much lower than that of the CRISPR/Cas9 approach – with the average efficiency being only 20.33%, peaking at 55.32%.

The main advantage of using CRISPR/Cas9 system compared to traditional approaches was demonstrated to be in protein tagging, as large selection cassettes are not introduced in the genome, no linker sequences are required and the changes to the genome are minimal. Additionally, when performing N-terminal tagging, the gene stays under control of its endogenous promoter, maintaining its endogenous expression, rather than the promoter region being changed, which would alter the expression levels of the protein of interest and potentially modify its function.

Design of gRNA used for mutagenesis is crucial and, to prevent the gRNA being inefficient at inducing Cas9-mediated DNA cleavage, several gRNAs should typically be designed and tested, which in turn might lengthen the procedure and increase its costs. Additionally, Cas9 overexpression in *S. pombe* can be toxic to the cells and is selected against (Jacobs *et al.*, 2014). I did observe that the CRISPR/Cas9 mutant strains exhibited slower growth compared to the wild type. Cas9-expressing plasmid contains the *ura4* gene as a selectable marker, which allows the plasmid to be eliminated from the cells using 5-FOA counter-selection. Therefore, a fast 5-FOA counter-selection, which would eliminate the Cas9-expressing plasmid, should be carried out

shortly upon selecting the correct mutant strains. This would prevent the toxic effect of prolonged Cas9 expression. Further important aspects to consider are off-target effects that have been reported in the literature, but can be minimised by careful gRNA design. Off-target gRNA binding sites can be identified by ChIP-seq of gRNA complexes with catalytically inactive Cas9 while GUIDE-seq can profile all double strand breaks that could be introduced genome-wide by Cas9 (Zhang *et al.*, 2015).

In summary, the results presented in this chapter demonstrate that CRISPR/Cas9 is an efficient method of gene editing, due to introducing minimal alterations to the genome, its high efficiency, flexibility, as well as fast strain creation.

Chapter 7: Conclusions

7.1. Ribosome release model

With my PhD research, I aimed to improve general understanding of the translation dependent process of nonsense-mediated mRNA decay (NMD). This is widely considered to be an active mRNA surveillance process requiring the concerted action of several proteins to recognise and to target “aberrant” transcripts for rapid degradation (Broгна and Wen, 2009). Here, I investigated an alternative hypothesis that NMD, rather than being the active process so far envisaged, occurs as a passive consequence of the release of ribosomes and other mRNA-associated proteins from an mRNA upon translation termination. A long stretch of the coding sequence normally protected by translating ribosomes would then become exposed and susceptible to nucleolytic attacks (BroгнаMcLeod and Petric, 2016). This hypothesis was formulated as the ribosome release model of NMD (BroгнаMcLeod and Petric, 2016). We postulated that mRNAs are generally more stable in cells depleted of NMD factors, because these proteins play a broad role in mRNP remodelling, including, at least in the case of UPF1, a role in the release of post-termination ribosomal subunits. We based this model on the extensive evidence indicating that UPF1 is an RNA helicase and an mRNP remodelling enzyme (FranksSingh and Lykke-Andersen, 2010; Fiorini *et al.*, 2015), and on the report that UPF1 can associate with eIF3 (Isken *et al.*, 2008); an important ribosome-recycling factor (PisarevHellen and Pestova, 2007).

I tested the release model directly, by examining whether PTC-containing transcripts would remain associated with post-termination ribosomes, and as

such, would sediment in heavier fractions upon polysome fractionation in cells depleted of NMD factors than in normal cells. The results of this experiment were, however, not conclusive. Although, in line with the model, PTC containing transcripts were found in heavier polysome fractions in NMD mutants, the shift was modest. Furthermore, the results could not clearly be interpreted because, for reasons that remain unexplained, most of the NMD reporter mRNA was lost during polysome fractionation. The results remain thus largely inconclusive requiring further experiments and probably adopting an alternative approach. Nevertheless, the release model may be generally correct, as it was reported that NMD reporters indeed associate with heavy polysomes in UPF1 deletion strain in *S. cerevisiae* (Hu *et al.*, 2010). Additionally, it was recently reported that UPF3 in fact stimulates ribosome release after peptide disassociation (Mühlemann and Karousis, 2017; Gao and Wilkinson, 2017; Neu-Yilik *et al.*, 2017). These observations appear to confirm the key predictions of the release model.

7.2. NMD factors partially associate with polysomes and affect translation

Our model of NMD makes the prediction that UPF1 and other NMD factors might affect the general process of translation, as detailed above. I tested this by comparing the total translation yield of wild type versus UPF1 and UPF3 deletion mutants. For this purpose, I used polysome fractionation paired with puromycylation, which allowed me to compare proteins being synthesised by a certain number of ribosomes at a given time, between the normal cells and the mutants.

The effect of UPF1 and UPF3 on translation differs, with more proteins having de-regulated translation in the absence of UPF3 than in the absence of UPF1. The data indicates that, even though UPF1 is involved, UPF3 has a predominant role in the general process of translation. As proteins translated fully by a single ribosome are more affected by NMD, I anticipated finding the most prominent effects in monosomal translation (Heyer and Moore, 2016). Indeed, the majority of proteins whose translation was altered in UPF3 deletion strain relative to normal conditions are those that are fully translated by a single ribosome. Therefore, UPF1 and particularly UPF3 seem to regulate translation of a subset of transcripts. It would be important to identify these proteins as their characteristics could help explain the reason behind their translation being altered in the absence of UPF3 and UPF1.

Additionally, I assessed the global distribution of NMD factors with mRNAs, relative to translation. I found that UPF1 and UPF3 do bind translating mRNAs; however, I discovered that each has a different mode of association. UPF1 only partially binds translating mRNAs, with a large proportion of UPF1 also co-fractionating with the small ribosomal subunit and associating with other cellular compartments other than translation circuits. UPF1 potentially interacts with the 40S, as upon polysome breakdown induced by puromycin, it shifts towards the light fractions, however, majority still remains bound to the small ribosomal subunit. UPF3, on the other hand, associates primarily with translating mRNAs, with its mode of association being either to mRNA directly, or via mRNP components other than the ribosome. Potentially, its association is via the EJC, as previously observed (Le Hir *et al.*, 2001).

7.3. UPF1 associates with stress granules

By analysing the distribution of UPF1 in the cells relative to translation, I also observed that UPF1 associates with other cellular compartments rather than translating mRNAs, in contrast to previous studies that have linked UPF1 almost exclusively to the translating mRNA pool. I found UPF1 association with heavy cellular entities, which do not represent translating polyribosomes, particularly interesting. I tested whether these may represent stress granules by looking at their formation under stress conditions both biochemically, using polysome fractionation, as well as microscopically, by observing the distribution of UPF1-GFP in cells. I indeed detected an accumulation of UPF1 in cytoplasmic granules upon heat shock and upon increasing high salt stress conditions. UPF1 almost entirely re-localises to these cytoplasmic compartments upon stress, however, its association is reversible and UPF1 is released from these cytoplasmic foci after the stress conditions are overcome. The protein composition of these particles, identified by mass spectrometry analysis, indicates that these are stress granules. However, it is unclear whether UPF1 may play a role in their formation, which would be very interesting to further investigate. As I identified many factors associating with the granules, I could observe their localisation in the cells upon stress conditions in normal versus UPF1 deletion cells. If UPF1 plays a key role in the formation of the granules, I would not observe them in *upf1Δ*. A model detailing the cytoplasmic roles of UPF1 is presented in Figure 7.1.

Figure 7.1 – **Model detailing the cytoplasmic role of UPF1.** UPF1 binds translating mRNAs, either via association with ribosomal subunits or directly with mRNAs, potentially improving ribosome scanning via its helicase activity (top panel). UPF1 also plays a role in triggering degradation of PTC-containing mRNAs (middle panel) as well as sequestering mRNAs that are no longer being translated in the cells into ribosome-free cytoplasmic granules (bottom left panel). Upon stress, global translation is suppressed and in such conditions, UPF1 almost entirely re-localises into such granules, increasing their number and size (bottom right panel). 5' CAP – 5' capping complex; 43S – pre-initiation complex 43S; UTR – untranslated region; START – start codon; STOP – stop codon; PABPC, poly(A) binding protein, cytosolic; eRF – eukaryotic release factor; PT – post-termination; eIF – eukaryotic initiation factor; 40S – small ribosomal subunit; 60S – large ribosomal subunit.

7.5. Monosomes translate full length proteins

Monosomes, single ribosomes translating individual mRNA molecules, have long been considered as initiating ribosomes that generate only short polypeptides, while multiple translating ribosomes, polysomes, were thought to be the sites of efficient translation which could translate full polypeptides. A recent report (Heyer and Moore, 2016) challenged this view, demonstrating that the majority of monosomes represent elongating ribosomes that can translate specific transcripts fully, requiring only a single ribosome at a time. The data that I obtained in this study goes in line with these findings, as polypeptides of high molecular weight have been observed in the monosomal fraction, with the signal being dependent on translation.

7.7. Ribosome-independent puromycylation uncovers that complete proteins can be released from the ribosome as peptidyl-tRNAs

Puromycin, as reviewed earlier, is an antibiotic that binds to the ribosome A site and gets incorporated into the nascent polypeptide chain, in a reaction catalysed by the ribosome (Pestka, 1971; DavidBennink and Yewdell, 2013). However, upon polysome fractionation and subsequent puromycin treatment, I observed puromycylation occurring in fractions that are void of ribosomes, a reaction not affected by translation inhibition with anisomycin, or by EDTA treatment. The reaction is therefore not driven by the ribosome.

The proteins that I found to be puromycylated independently of ribosomes do not represent all proteins present in the sample, but a reproducible subset of proteins. This suggests that a specific feature of these proteins renders them susceptible to puromycin incorporation. As the reaction of puromycylation

normally occurs between a peptidyl-tRNA and puromycin in the ribosome active site, I hypothesised that during translation a fraction of polypeptides might be released from the ribosome with a tRNA still attached to them, possibly due to ribosome stalling, and that these peptidyl-tRNAs can react with puromycin in a ribosome-independent manner. The release of peptidyl-tRNAs during translation was largely reported in bacteria, where these products are eliminated by the action of peptidyl-tRNA hydrolase (Cruz-Vera *et al.*, 2004; Kössel, 1970). It is unclear how frequently this process occurs in eukaryotic organisms, where peptidyl-tRNAs were mainly reported to arise during the NGD process, due to ribosome stalling (Harigaya and Parker, 2010). Two homologues of the bacterial peptidyl-tRNA hydrolase (Pth and Pth2) were identified in eukaryotes, with Pth2 demonstrated to possess such enzymatic activity, suggesting that its substrates indeed occur in the cells. Thereby, it is possible that peptidyl-tRNAs that arise during translation are indeed substrates for ribosome-independent puromycylation.

I tested this hypothesis by experiments detailed in Chapter 3. Destabilising peptidyl-tRNAs by alkaline treatment has led to a reduction in puromycylation, going in line with the hypothesis. Probably the most direct evidence in favour of the hypothesis comes from treatment with hydroxylamine, a chemical that destroys peptidyl-tRNAs (Bresler *et al.*, 1968), which indeed abolished ribosome-independent puromycylation. Proteins that undergo ribosome-independent puromycylation were identified as highly expressed proteins in the cell, which is expected, since the more frequently translation occurs in the cell, the higher the probability that peptidyl-tRNAs might arise.

The discovery of ribosome-independent puromycylation not only requires a re-evaluation of using this method to report solely translation sites in the cell but also, most importantly, indicates the existence of previously overlooked eukaryotic translation by-products, peptidyl-tRNAs that significantly contribute to overall gene expression.

It is necessary to confirm directly that the substrates for ribosome-independent puromycylation are indeed peptidyl-tRNAs, by performing protein affinity purification of puromycylated proteins and subsequent mass spectrometry analysis, to assess the exact nature and location where puromycin incorporates.

References

- Amrani, N., Ganesan, R., Kervestin, S., Mangus, D. A., Ghosh, S. and Jacobson, A. (2004) 'A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay', *Nature*, 432(7013), pp. 112-8.
- Armache, J. P., Jarasch, A., Anger, A. M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., Marquez, V., Mielke, T., Thomm, M., Berninghausen, O., Beatrix, B., Söding, J., Westhof, E., Wilson, D. N. and Beckmann, R. (2010) 'Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-Å resolution', *Proc Natl Acad Sci U S A*, 107(46), pp. 19748-53.
- Atkinson, G. C., Baldauf, S. L. and Hauryliuk, V. (2008) 'Evolution of nonstop, no-go and nonsense-mediated mRNA decay and their termination factor-derived components', *BMC Evol Biol*, 8, pp. 290.
- Aviner, R., Geiger, T. and Elroy-Stein, O. (2013) 'Novel proteomic approach (PUNCH-P) reveals cell cycle-specific fluctuations in mRNA translation', *Genes Dev*, 27(16), pp. 1834-44.
- Azzalin, C. M. and Lingner, J. (2006a) 'The double life of UPF1 in RNA and DNA stability pathways', *Cell Cycle*, 5(14), pp. 1496-8.
- Azzalin, C. M. and Lingner, J. (2006b) 'The human RNA surveillance factor UPF1 is required for S phase progression and genome stability', *Curr Biol*, 16(4), pp. 433-9.
- Azzam, M. E. and Algranati, I. D. (1973) 'Mechanism of puromycin action: fate of ribosomes after release of nascent protein chains from polysomes', *Proc Natl Acad Sci U S A*, 70(12), pp. 3866-9.
- Baliga, B. S., Cohen, S. A. and Munro, H. N. (1970) 'Effect of cycloheximide on the reaction of puromycin with polysome-bound peptidyl-tRNA', *FEBS Lett*, 8(5), pp. 249-252.

- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. and Horvath, P. (2007) 'CRISPR provides acquired resistance against viruses in prokaryotes', *Science*, 315(5819), pp. 1709-12.
- Barta, A., Steiner, G., Brosius, J., Noller, H. F. and Kuechler, E. (1984) 'Identification of a site on 23S ribosomal RNA located at the peptidyl transferase center', *Proc Natl Acad Sci U S A*, 81(12), pp. 3607-11.
- Barthelme, D., Dinkelaker, S., Albers, S. V., Londei, P., Ermler, U. and Tampé, R. (2011) 'Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1', *Proc Natl Acad Sci U S A*, 108(8), pp. 3228-33.
- Beach, D. and Nurse, P. (1981) 'High-frequency transformation of the fission yeast *Schizosaccharomyces pombe*', *Nature*, 290(5802), pp. 140-2.
- Becker, T., Armache, J. P., Jarasch, A., Anger, A. M., Villa, E., Sieber, H., Motaal, B. A., Mielke, T., Berninghausen, O. and Beckmann, R. (2011) 'Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome', *Nat Struct Mol Biol*, 18(6), pp. 715-20.
- Behrmann, E., Loerke, J., Budkevich, T. V., Yamamoto, K., Schmidt, A., Penczek, P. A., Vos, M. R., Bürger, J., Mielke, T., Scheerer, P. and Spahn, C. M. (2015) 'Structural snapshots of actively translating human ribosomes', *Cell*, 161(4), pp. 845-57.
- Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. and Yusupov, M. (2011) 'The structure of the eukaryotic ribosome at 3.0 Å resolution', *Science*, 334(6062), pp. 1524-9.
- Blobel, G. and Sabatini, D. (1971) 'Dissociation of mammalian polyribosomes into subunits by puromycin', *Proc Natl Acad Sci U S A*, 68(2), pp. 390-4.
- Bolotin, A., Quinquis, B., Sorokin, A. and Ehrlich, S. D. (2005) 'Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin', *Microbiology*, 151(Pt 8), pp. 2551-61.

- Bonafacino, J. S., Gallagher, S. R., Schamel, W. W. A., Fuller, S. A. and Hurrell, J. G. (2011) 'Electrophoresis and Immunoblotting', *Curr Protoc Cell Biol*, Chapter 6.
- Bresler, S., Grajevskaja, R., Kirilov, S. and Saminski, E. (1968) 'Stability of peptidyl-tRNA - the intermediate of protein synthesis', *Biochim Biophys Acta*, 155(2), pp. 465-75.
- Bresler, S., Grajevskaja, R., Kirilov, S., Saminski, E. and Shutov, F. (1966) 'Peptidyl-transfer RNA: an intermediate in protein biosynthesis', *Biochim Biophys Acta*, 123(3), pp. 534-45.
- Broгна, S., McLeod, T. and Petric, M. (2016) 'The Meaning of NMD: Translate or Perish', *Trends Genet*, 32(7), pp. 395-407.
- Broгна, S., Ramanathan, P. and Wen, J. (2008) 'UPF1 P-body localization', *Biochem Soc Trans*, 36(Pt 4), pp. 698-700.
- Broгна, S. and Wen, J. (2009) 'Nonsense-mediated mRNA decay (NMD) mechanisms', *Nat Struct Mol Biol*, 16(2), pp. 107-13.
- Buratowski, S. (2009) 'Progression through the RNA polymerase II CTD cycle', *Mol Cell*, 36(4), pp. 541-6.
- Burge, C. B., Tuschl, T. and Sharp, P. A. (1999) *Splicing of Precursors to mRNAs by the Spliceosomes*. Second edition (ed. Gesteland RF et al.) edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Burks, G. L., McFeeters, H. and McFeeters, R. L. (2016) 'Expression, purification, and buffer solubility optimization of the putative human peptidyl-tRNA hydrolase PTRHD1', *Protein Expr Purif*, 126, pp. 49-54.
- Bähler, J., Wu, J. Q., Longtine, M. S., Shah, N. G., McKenzie, A., Steever, A. B., Wach, A., Philippsen, P. and Pringle, J. R. (1998) 'Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*', *Yeast*, 14(10), pp. 943-51.

- Cali, B. M., Kuchma, S. L., Latham, J. and Anderson, P. (1999) 'smg-7 is required for mRNA surveillance in *Caenorhabditis elegans*', *Genetics*, 151(2), pp. 605-16.
- Cao, J. and Geballe, A. P. (1998) 'Ribosomal release without peptidyl tRNA hydrolysis at translation termination in a eukaryotic system', *RNA*, 4(2), pp. 181-8.
- Carter, M. S., Li, S. and Wilkinson, M. F. (1996) 'A splicing-dependent regulatory mechanism that detects translation signals', *EMBO J*, 15(21), pp. 5965-75.
- Cech, T. R. (2000) 'Structural biology. The ribosome is a ribozyme', *Science*, 289(5481), pp. 878-9.
- Celik, A., Kervestin, S. and Jacobson, A. (2015) 'NMD: At the crossroads between translation termination and ribosome recycling', *Biochimie*, 114, pp. 2-9.
- Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H. and Conti, E. (2011) 'Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2', *Mol Cell*, 41(6), pp. 693-703.
- Chamieh, H., Ballut, L., Bonneau, F. and Le Hir, H. (2008) 'NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity', *Nat Struct Mol Biol*, 15(1), pp. 85-93.
- Chassé, H., Boulben, S., Costache, V., Cormier, P. and Morales, J. (2017) 'Analysis of translation using polysome profiling', *Nucleic Acids Res*, 45(3), pp. e15.
- Chavatte, L., Seit-Nebi, A., Dubovaya, V. and Favre, A. (2002) 'The invariant uridine of stop codons contacts the conserved NIKSR loop of human eRF1 in the ribosome', *EMBO J*, 21(19), pp. 5302-11.
- Cheng, Z., Muhlrads, D., Lim, M. K., Parker, R. and Song, H. (2007) 'Structural and functional insights into the human Upf1 helicase core', *EMBO J*, 26(1), pp. 253-64.

- Cherkasov, V., Grousl, T., Theer, P., Vainshtein, Y., Glässer, C., Mongis, C., Kramer, G., Stoecklin, G., Knop, M., Mogk, A. and Bukau, B. (2015) 'Systemic control of protein synthesis through sequestration of translation and ribosome biogenesis factors during severe heat stress', *FEBS Lett*, 589(23), pp. 3654-64.
- Chiu, S. Y., Serin, G., Ohara, O. and Maquat, L. E. (2003) 'Characterization of human Smg5/7a: a protein with similarities to *Caenorhabditis elegans* SMG5 and SMG7 that functions in the dephosphorylation of Upf1', *RNA*, 9(1), pp. 77-87.
- Clerici, M., Deniaud, A., Boehm, V., Gehring, N. H., Schaffitzel, C. and Cusack, S. (2014) 'Structural and functional analysis of the three MIF4G domains of nonsense-mediated decay factor UPF2', *Nucleic Acids Res*, 42(4), pp. 2673-86.
- Coddington, A. and Fluri, R. (1977) 'Characterisation of the ribosomal proteins from *Schizosaccharomyces pombe* by two-dimensional polyacrylamide gel electrophoresis: demonstration that a cycloheximide resistant strain, cyh1, has an altered 60S ribosomal protein', *Mol Gen Genet*, 158(1), pp. 93-100.
- Colgan, D. F. and Manley, J. L. (1997) 'Mechanism and regulation of mRNA polyadenylation', *Genes Dev*, 11(21), pp. 2755-66.
- Collart, M. A. and Oliviero, S. (2001) 'Preparation of yeast RNA', *Curr Protoc Mol Biol*, Chapter 13, pp. Unit13.12.
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. and Zhang, F. (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science*, 339(6121), pp. 819-23.
- Cosson, B., Berkova, N., Couturier, A., Chabelskaya, S., Philippe, M. and Zhouravleva, G. (2002a) 'Poly(A)-binding protein and eRF3 are associated in vivo in human and *Xenopus* cells', *Biol Cell*, 94(4-5), pp. 205-16.
- Cosson, B., Couturier, A., Chabelskaya, S., Kiktev, D., Inge-Vechtomov, S., Philippe, M. and Zhouravleva, G. (2002b) 'Poly(A)-binding protein acts in translation termination via eukaryotic release factor 3 interaction and does not influence [PSI(+)] propagation', *Mol Cell Biol*, 22(10), pp. 3301-15.

Cruz-Vera, L. R., Hernandez-Ramon, E., Perez-Zamorano, B. and Guarneros, G. (2003) 'The rate of peptidyl-tRNA dissociation from the ribosome during minigene expression depends on the nature of the last decoding interaction', *J Biol Chem*, 278(28), pp. 26065-70.

Cruz-Vera, L. R., Magos-Castro, M. A., Zamora-Romo, E. and Guarneros, G. (2004) 'Ribosome stalling and peptidyl-tRNA drop-off during translational delay at AGA codons', *Nucleic Acids Res*, 32(15), pp. 4462-8.

Culbertson, M. R., Underbrink, K. M. and Fink, G. R. (1980) 'Frameshift suppression *Saccharomyces cerevisiae*. II. Genetic properties of group II suppressors', *Genetics*, 95(4), pp. 833-53.

Czaplinski, K., Ruiz-Echevarria, M. J., Paushkin, S. V., Han, X., Weng, Y., Perlick, H. A., Dietz, H. C., Ter-Avanesyan, M. D. and Peltz, S. W. (1998) 'The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs', *Genes Dev*, 12(11), pp. 1665-77.

Czaplinski, K., Weng, Y., Hagan, K. W. and Peltz, S. W. (1995) 'Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation', *RNA*, 1(6), pp. 610-23.

Das, G. and Varshney, U. (2006) 'Peptidyl-tRNA hydrolase and its critical role in protein biosynthesis', *Microbiology*, 152(Pt 8), pp. 2191-5.

Das, S. and Maitra, U. (2001) 'Functional significance and mechanism of eIF5-promoted GTP hydrolysis in eukaryotic translation initiation', *Prog Nucleic Acid Res Mol Biol*, 70, pp. 207-31.

David, A., Bennink, J. R. and Yewdell, J. W. (2013) 'Emetine optimally facilitates nascent chain puromycylation and potentiates the ribopuromycylation method (RPM) applied to inert cells', *Histochem Cell Biol*, 139(3), pp. 501-4.

De Pereda, J. M., Waas, W. F., Jan, Y., Ruoslahti, E., Schimmel, P. and Pascual, J. (2004) 'Crystal structure of a human peptidyl-tRNA hydrolase reveals a new fold and suggests basis for a bifunctional activity', *J Biol Chem*, 279(9), pp. 8111-5.

- Dever, T. E. and Green, R. (2012) 'The elongation, termination, and recycling phases of translation in eukaryotes', *Cold Spring Harb Perspect Biol*, 4(7), pp. a013706.
- Doma, M. K. and Parker, R. (2006) 'Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation', *Nature*, 440(7083), pp. 561-4.
- Eliseev, B., Yeramala, L., Leitner, A., Karuppasamy, M., Raimondeau, E., Huard, K., Alkalaeva, E., Aebersold, R. and Schaffitzel, C. (2018) 'Structure of a human cap-dependent 48S translation pre-initiation complex', *Nucleic Acids Res*, 46(5), pp. 2678-2689.
- Fiorini, F., Bagchi, D., Le Hir, H. and Croquette, V. (2015) 'Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities', *Nat Commun*, 6, pp. 7581.
- Fisher, C. R. (1969) 'Enzymology of the pigmented adenine-requiring mutants of *Saccharomyces* and *Schizosaccharomyces*', *Biochem Biophys Res Commun*, 34(3), pp. 306-10.
- Fourati, Z., Roy, B., Millan, C., Coureux, P. D., Kervestin, S., van Tilbeurgh, H., He, F., Usón, I., Jacobson, A. and Graille, M. (2014) 'A highly conserved region essential for NMD in the Upf2 N-terminal domain', *J Mol Biol*, 426(22), pp. 3689-3702.
- Franks, T. M., Singh, G. and Lykke-Andersen, J. (2010) 'Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay', *Cell*, 143(6), pp. 938-50.
- Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M. and Kisselev, L. (1996) 'Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase', *RNA*, 2(4), pp. 334-41.
- Fromant, M., Ferri-Fioni, M. L., Plateau, P. and Blanquet, S. (2003) 'Peptidyl-tRNA hydrolase from *Sulfolobus solfataricus*', *Nucleic Acids Res*, 31(12), pp. 3227-35.

Gao, J., Kan, F., Wagnon, J. L., Storey, A. J., Protacio, R. U., Davidson, M. K. and Wahls, W. P. (2014) 'Rapid, efficient and precise allele replacement in the fission yeast *Schizosaccharomyces pombe*', *Curr Genet*, 60(2), pp. 109-19.

Gao, Z. and Wilkinson, M. (2017) 'An RNA decay factor wears a new coat: UPF3B modulates translation termination', *F1000Res*, 6, pp. 2159.

Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W. and Kulozik, A. E. (2003) 'Y14 and hUpf3b form an NMD-activating complex', *Mol Cell*, 11(4), pp. 939-49.

Gerbi, S. (1996) Expansion segment: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA. Ribosomal RNA : Structure, Evolution, Processing, and Function in Protein Synthesis New York: CRC Press, p. 71–87.

González, C. I., Ruiz-Echevarría, M. J., Vasudevan, S., Henry, M. F. and Peltz, S. W. (2000) 'The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay', *Mol Cell*, 5(3), pp. 489-99.

Grimm, C., Kohli, J., Murray, J. and Maundrell, K. (1988) 'Genetic engineering of *Schizosaccharomyces pombe*: a system for gene disruption and replacement using the *ura4* gene as a selectable marker', *Mol Gen Genet*, 215(1), pp. 81-6.

Grollman, A. P. (1966) 'Structural basis for inhibition of protein synthesis by emetine and cycloheximide based on an analogy between ipecac alkaloids and glutarimide antibiotics', *Proc Natl Acad Sci U S A*, 56(6), pp. 1867-74.

Grollman, A. P. (1967) 'Inhibitors of protein biosynthesis. II. Mode of action of anisomycin', *J Biol Chem*, 242(13), pp. 3226-33.

Gromadski, K. B., Schümmer, T., Strømgaard, A., Knudsen, C. R., Kinzy, T. G. and Rodnina, M. V. (2007) 'Kinetics of the interactions between yeast elongation factors 1A and 1B α , guanine nucleotides, and aminoacyl-tRNA', *J Biol Chem*, 282(49), pp. 35629-37.

Gross, M., Crow, P. and White, J. (1992) 'The site of hydrolysis by rabbit reticulocyte peptidyl-tRNA hydrolase is the 3'-AMP terminus of susceptible tRNA substrates', *J Biol Chem*, 267(3), pp. 2080-6.

- Gross, M., Starn, T. K., Rundquist, C., Crow, P., White, J., Olin, A. and Wagner, T. (1992) 'Purification and initial characterization of peptidyl-tRNA hydrolase from rabbit reticulocytes', *J Biol Chem*, 267(3), pp. 2073-9.
- Gross, T. and Käufer, N. F. (1998) 'Cytoplasmic ribosomal protein genes of the fission yeast *Schizosaccharomyces pombe* display a unique promoter type: a suggestion for nomenclature of cytoplasmic ribosomal proteins in databases', *Nucleic Acids Res*, 26(14), pp. 3319-22.
- Guan, Q., Zheng, W., Tang, S., Liu, X., Zinkel, R. A., Tsui, K. W., Yandell, B. S. and Culbertson, M. R. (2006) 'Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast', *PLoS Genet*, 2(11), pp. e203.
- Harigaya, Y. and Parker, R. (2010) 'No-go decay: a quality control mechanism for RNA in translation', *Wiley Interdiscip Rev RNA*, 1(1), pp. 132-41.
- Harlen, K. M. and Churchman, L. S. (2017) 'The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain', *Nat Rev Mol Cell Biol*, 18(4), pp. 263-273.
- He, F., Brown, A. H. and Jacobson, A. (1997) 'Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway', *Mol Cell Biol*, 17(3), pp. 1580-94.
- He, F. and Jacobson, A. (1995) 'Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen', *Genes Dev*, 9(4), pp. 437-54.
- Heyer, E. E. and Moore, M. J. (2016) 'Redefining the Translational Status of 80S Monosomes', *Cell*, 164(4), pp. 757-69.
- Hinnebusch, A. G. and Lorsch, J. R. (2012) 'The mechanism of eukaryotic translation initiation: new insights and challenges', *Cold Spring Harb Perspect Biol*, 4(10).
- Hirokawa, G., Nijman, R. M., Raj, V. S., Kaji, H., Igarashi, K. and Kaji, A. (2005) 'The role of ribosome recycling factor in dissociation of 70S ribosomes into subunits', *RNA*, 11(8), pp. 1317-28.

- Hodgkin, J., Papp, A., Pulak, R., Ambros, V. and Anderson, P. (1989) 'A new kind of informational suppression in the nematode *Caenorhabditis elegans*', *Genetics*, 123(2), pp. 301-13.
- Hoshino, S., Imai, M., Kobayashi, T., Uchida, N. and Katada, T. (1999) 'The eukaryotic polypeptide chain releasing factor (eRF3/GSPT) carrying the translation termination signal to the 3'-Poly(A) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein', *J Biol Chem*, 274(24), pp. 16677-80.
- Hu, W., Petzold, C., Collier, J. and Baker, K. E. (2010) 'Nonsense-mediated mRNA decapping occurs on polyribosomes in *Saccharomyces cerevisiae*', *Nat Struct Mol Biol*, 17(2), pp. 244-7.
- Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J. B., Munier, A., Fradet, M., Daunesse, M., Bertrand, E., Pierron, G., Mozziconacci, J., Kress, M. and Weil, D. (2017) 'P-Body Purification Reveals the Condensation of Repressed mRNA Regulons', *Mol Cell*, 68(1), pp. 144-157.e5.
- Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. and Weissman, J. S. (2009) 'Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling', *Science*, 324(5924), pp. 218-23.
- Ioannou, M., Coutso Georgopoulos, C. and Synetos, D. (1998) 'Kinetics of inhibition of rabbit reticulocyte peptidyltransferase by anisomycin and sparsomycin', *Mol Pharmacol*, 53(6), pp. 1089-96.
- Isken, O., Kim, Y. K., Hosoda, N., Mayeur, G. L., Hershey, J. W. and Maquat, L. E. (2008) 'Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay', *Cell*, 133(2), pp. 314-27.
- Ivanov, P. V., Gehring, N. H., Kunz, J. B., Hentze, M. W. and Kulozik, A. E. (2008) 'Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways', *EMBO J*, 27(5), pp. 736-47.

- Jackson, R. J., Hellen, C. U. and Pestova, T. V. (2010) 'The mechanism of eukaryotic translation initiation and principles of its regulation', *Nat Rev Mol Cell Biol*, 11(2), pp. 113-27.
- Jacobs, J. Z., Ciccaglione, K. M., Tournier, V. and Zaratiegui, M. (2014) 'Implementation of the CRISPR-Cas9 system in fission yeast', *Nat Commun*, 5, pp. 5344.
- Jimenez, A., Monro, R. E. and Vazquez, D. (1970) 'Interaction of Ac-Phe-tRNA with E. coli ribosomal subunits. 2. Resistance of the sparsomycin-induced complex to hydroxylamine action', *FEBS Lett*, 7(2), pp. 109-111.
- Johansson, M., leong, K. W., Trobro, S., Strazewski, P., Åqvist, J., Pavlov, M. Y. and Ehrenberg, M. (2011) 'pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA', *Proc Natl Acad Sci U S A*, 108(1), pp. 79-84.
- Kadlec, J., Izaurralde, E. and Cusack, S. (2004) 'The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3', *Nat Struct Mol Biol*, 11(4), pp. 330-7.
- Kapp, L. D. and Lorsch, J. R. (2004) 'The molecular mechanics of eukaryotic translation', *Annu Rev Biochem*, 73, pp. 657-704.
- Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G. and Ohno, S. (2006) '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(3), pp. 355-67.
- Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A., Gyula, P., Lacomme, C., Barta, E., Nagy, F. and Silhavy, D. (2008) 'Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay', *EMBO J*, 27(11), pp. 1585-95.
- Kessler, M. M., Henry, M. F., Shen, E., Zhao, J., Gross, S., Silver, P. A. and Moore, C. L. (1997) 'Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast', *Genes Dev*, 11(19), pp. 2545-56.

- Kim, Y. K., Furic, L., Desgroseillers, L. and Maquat, L. E. (2005) 'Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay', *Cell*, 120(2), pp. 195-208.
- Klauser, A. A. and van Hoof, A. (2012) 'Degradation of mRNAs that lack a stop codon: a decade of nonstop progress', *Wiley Interdiscip Rev RNA*, 3(5), pp. 649-60.
- Komarnitsky, P., Cho, E. J. and Buratowski, S. (2000) 'Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription', *Genes Dev*, 14(19), pp. 2452-60.
- Korneeva, N. L., Lamphear, B. J., Hennigan, F. L. and Rhoads, R. E. (2000) 'Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1', *J Biol Chem*, 275(52), pp. 41369-76.
- Krawchuk, M. D. and Wahls, W. P. (1999) 'High-efficiency gene targeting in *Schizosaccharomyces pombe* using a modular, PCR-based approach with long tracts of flanking homology', *Yeast*, 15(13), pp. 1419-27.
- Kuersten, S. and Goodwin, E. B. (2005) 'Linking nuclear mRNP assembly and cytoplasmic destiny', *Biol Cell*, 97(6), pp. 469-78.
- Kupfer, D. M., Drabenstot, S. D., Buchanan, K. L., Lai, H., Zhu, H., Dyer, D. W., Roe, B. A. and Murphy, J. W. (2004) 'Introns and splicing elements of five diverse fungi', *Eukaryot Cell*, 3(5), pp. 1088-100.
- Kössel, H. (1970) 'Purification and properties of peptidyl-tRNA hydrolase from *Escherichia coli*', *Biochim Biophys Acta*, 204(1), pp. 191-202.
- Le Hir, H., Gatfield, D., Izaurralde, E. and Moore, M. J. (2001) 'The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay', *EMBO J*, 20(17), pp. 4987-97.
- Le Hir, H., Izaurralde, E., Maquat, L. E. and Moore, M. J. (2000) 'The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions', *EMBO J*, 19(24), pp. 6860-9.

- Llorca, O. (2013) 'Structural insights into nonsense-mediated mRNA decay (NMD) by electron microscopy', *Curr Opin Struct Biol*, 23(1), pp. 161-7.
- Longman, D., Plasterk, R. H., Johnstone, I. L. and Cáceres, J. F. (2007a) 'Mechanistic insights and identification of two novel factors in the *C. elegans* NMD pathway', *Genes Dev*, 21(9), pp. 1075-85.
- Longman, D., Plasterk, R. H. A., Johnstone, I. L. and Cáceres, J. F. (2007b) 'Mechanistic insights and identification of two novel factors in the *C. elegans* NMD pathway', *Genes and Development*, 21, pp. 1075-1085.
- Losson, R. and Lacroute, F. (1979) 'Interference of nonsense mutations with eukaryotic messenger RNA stability', *Proc Natl Acad Sci U S A*, 76(10), pp. 5134-7.
- López-Perrote, A., Castaño, R., Melero, R., Zamarro, T., Kurosawa, H., Ohnishi, T., Uchiyama, A., Aoyagi, K., Buchwald, G., Kataoka, N., Yamashita, A. and Llorca, O. (2016) 'Human nonsense-mediated mRNA decay factor UPF2 interacts directly with eRF3 and the SURF complex', *Nucleic Acids Res*, 44(4), pp. 1909-23.
- Ma, Y., Zhang, L. and Huang, X. (2014) 'Genome modification by CRISPR/Cas9', *FEBS J*, 281(23), pp. 5186-93.
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. and Koonin, E. V. (2006) 'A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action', *Biol Direct*, 1, pp. 7.
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. and Church, G. M. (2013) 'RNA-guided human genome engineering via Cas9', *Science*, 339(6121), pp. 823-6.
- Maquat, L. E. and Li, X. (2001) 'Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay', *RNA*, 7(3), pp. 445-56.

- Matsuo, Y., Asakawa, K., Toda, T. and Katayama, S. (2006) 'A rapid method for protein extraction from fission yeast', *Biosci Biotechnol Biochem*, 70(8), pp. 1992-4.
- Matsuyama, A., Shirai, A., Yashiroda, Y., Kamata, A., Horinouchi, S. and Yoshida, M. (2004) 'pDUAL, a multipurpose, multicopy vector capable of chromosomal integration in fission yeast', *Yeast*, 21(15), pp. 1289-305.
- Maundrell, K. (1990) 'nmt1 of fission yeast. A highly transcribed gene completely repressed by thiamine', *J Biol Chem*, 265(19), pp. 10857-64.
- Mayfield, J. E., Robinson, M. R., Cotham, V. C., Irani, S., Matthews, W. L., Ram, A., Gilmour, D. S., Cannon, J. R., Zhang, Y. J. and Brodbelt, J. S. (2017) 'Mapping the Phosphorylation Pattern of Drosophila melanogaster RNA Polymerase II Carboxyl-Terminal Domain Using Ultraviolet Photodissociation Mass Spectrometry', *ACS Chem Biol*, 12(1), pp. 153-162.
- McCarthy, J. E. and Kollmus, H. (1995) 'Cytoplasmic mRNA-protein interactions in eukaryotic gene expression', *Trends Biochem Sci*, 20(5), pp. 191-7.
- McCarty, K. S., Stafford, D. and Brown, O. (1968) 'Resolution and fractionation of macromolecules by isokinetic sucrose density gradient sedimentation', *Anal Biochem*, 24(2), pp. 314-29.
- Meaux, S., van Hoof, A. and Baker, K. E. (2008) 'Nonsense-mediated mRNA decay in yeast does not require PAB1 or a poly(A) tail', *Mol Cell*, 29(1), pp. 134-40.
- Mehlin, H., Daneholt, B. and Skoglund, U. (1992) 'Translocation of a specific premessenger ribonucleoprotein particle through the nuclear pore studied with electron microscope tomography', *Cell*, 69(4), pp. 605-13.
- Mehlin, H., Daneholt, B. and Skoglund, U. (1995) 'Structural interaction between the nuclear pore complex and a specific translocating RNP particle', *J Cell Biol*, 129(5), pp. 1205-16.
- Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G. and Yusupov, M. (2012) 'One core, two shells: bacterial and eukaryotic ribosomes', *Nat Struct Mol Biol*, 19(6), pp. 560-7.

- Mendell, J. T., Medghalchi, S. M., Lake, R. G., Noensie, E. N. and Dietz, H. C. (2000) 'Novel Upf2p orthologues suggest a functional link between translation initiation and nonsense surveillance complexes', *Mol Cell Biol*, 20(23), pp. 8944-57.
- Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F. and Dietz, H. C. (2004) 'Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutates genomic noise', *Nat Genet*, 36(10), pp. 1073-8.
- Menninger, J. R. (1976) 'Peptidyl transfer RNA dissociates during protein synthesis from ribosomes of *Escherichia coli*', *J Biol Chem*, 251(11), pp. 3392-8.
- Miall, S. H. and Walker, I. O. (1969) 'Structural studies on ribosomes. II. Denaturation and sedimentation of ribosomal subunits unfolded in EDTA', *Biochim Biophys Acta*, 174(2), pp. 551-60.
- Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. and Soria, E. (2005) 'Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements', *J Mol Evol*, 60(2), pp. 174-82.
- Moore, M. J. and Proudfoot, N. J. (2009) 'Pre-mRNA processing reaches back to transcription and ahead to translation', *Cell*, 136(4), pp. 688-700.
- Moore, M. J., Query, C. C. and Sharp, P. A. (1993) 'Splicing of precursors to mRNA by the spliceosome .', *In RNA World*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp. 303–357.
- Morris, C., Wittmann, J., Jäck, H. M. and Jalinot, P. (2007) 'Human INT6/eIF3e is required for nonsense-mediated mRNA decay', *EMBO Rep*, 8(6), pp. 596-602.
- Muhrad, D. and Parker, R. (1999) 'Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance', *RNA*, 5(10), pp. 1299-307.

Mühlemann, O., Eberle, A. B., Stalder, L. and Zamudio Orozco, R. (2008) 'Recognition and elimination of nonsense mRNA', *Biochim Biophys Acta*, 1779(9), pp. 538-49.

Mühlemann, O. and Jensen, T. H. (2012) 'mRNP quality control goes regulatory', *Trends Genet*, 28(2), pp. 70-7.

Mühlemann, O. and Karousis, E. D. (2017) 'New functions in translation termination uncovered for NMD factor UPF3B', *EMBO J*, 36(20), pp. 2928-2930.

Nag, N., Lin, K. Y., Edmonds, K. A., Yu, J., Nadkarni, D., Marintcheva, B. and Marintchev, A. (2016) 'eIF1A/eIF5B interaction network and its functions in translation initiation complex assembly and remodeling', *Nucleic Acids Res*, 44(15), pp. 7441-56.

Neu-Yilik, G., Raimondeau, E., Eliseev, B., Yeramala, L., Amthor, B., Deniaud, A., Huard, K., Kerschgens, K., Hentze, M. W., Schaffitzel, C. and Kulozik, A. E. (2017) 'Dual function of UPF3B in early and late translation termination', *EMBO J*, 36(20), pp. 2968-2986.

Nissen, P., Hansen, J., Ban, N., Moore, P. B. and Steitz, T. A. (2000) 'The structural basis of ribosome activity in peptide bond synthesis', *Science*, 289(5481), pp. 920-30.

Nolan, R. D. and Arnstein, H. R. (1969) 'The dissociation of rabbit reticulocyte ribosomes with EDTA and the location of messenger ribonucleic acid', *Eur J Biochem*, 9(4), pp. 445-50.

Noller, H. F., Hoffarth, V. and Zimniak, L. (1992) 'Unusual resistance of peptidyl transferase to protein extraction procedures', *Science*, 256(5062), pp. 1416-9.

Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K. R., Grimson, A., Hachiya, T., Hentze, M. W., Anderson, P. and Ohno, S. (2003) 'Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7', *Mol Cell*, 12(5), pp. 1187-200.

Park, E. and Maquat, L. E. (2013) 'Staufen-mediated mRNA decay', *Wiley Interdiscip Rev RNA*, 4(4), pp. 423-35.

- Peltz, S. W., Brown, A. H. and Jacobson, A. (1993) 'mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor', *Genes Dev*, 7(9), pp. 1737-54.
- Pestka, S. (1971) 'Inhibitors of ribosome functions', *Annu Rev Microbiol*, 25, pp. 487-562.
- Pestka, S., Rosenfeld, H., Harris, R. and Hintikka, H. (1972) 'Studies on transfer ribonucleic acid-ribosome complexes. XXI. Effect of antibiotics on peptidyl-puromycin synthesis by mammalian polyribosomes', *J Biol Chem*, 247(21), pp. 6895-900.
- Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I. and Hellen, C. U. (2001) 'Molecular mechanisms of translation initiation in eukaryotes', *Proc Natl Acad Sci U S A*, 98(13), pp. 7029-36.
- Petersen, J. and Russell, P. (2016) 'Growth and the Environment of *Schizosaccharomyces pombe*', *Cold Spring Harb Protoc*, 2016(3), pp. pdb.top079764.
- Pisarev, A. V., Hellen, C. U. and Pestova, T. V. (2007) 'Recycling of eukaryotic posttermination ribosomal complexes', *Cell*, 131(2), pp. 286-99.
- Pisarev, A. V., Skabkin, M. A., Pisareva, V. P., Skabkina, O. V., Rakotondrafara, A. M., Hentze, M. W., Hellen, C. U. and Pestova, T. V. (2010) 'The role of ABCE1 in eukaryotic posttermination ribosomal recycling', *Mol Cell*, 37(2), pp. 196-210.
- Polacek, N. and Mankin, A. S. (2005) 'The ribosomal peptidyl transferase center: structure, function, evolution, inhibition', *Crit Rev Biochem Mol Biol*, 40(5), pp. 285-311.
- Pourcel, C., Salvignol, G. and Vergnaud, G. (2005) 'CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies', *Microbiology*, 151(Pt 3), pp. 653-63.

- Prévôt, D., Darlix, J. L. and Ohlmann, T. (2003) 'Conducting the initiation of protein synthesis: the role of eIF4G', *Biol Cell*, 95(3-4), pp. 141-56.
- Pulak, R. and Anderson, P. (1993) 'mRNA surveillance by the *Caenorhabditis elegans* smg genes', *Genes Dev*, 7(10), pp. 1885-97.
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. and Zhang, F. (2013) 'Genome engineering using the CRISPR-Cas9 system', *Nat Protoc*, 8(11), pp. 2281-2308.
- Ruiz-Echevarría, M. J., González, C. I. and Peltz, S. W. (1998) 'Identifying the right stop: determining how the surveillance complex recognizes and degrades an aberrant mRNA', *EMBO J*, 17(2), pp. 575-89.
- Sambrook, J. and Green, M. (2012) *Molecular Cloning: A Laboratory Manual*. 4th edn. (3 vols). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 2012.
- Schneider-Poetsch, T., Ju, J., Eyler, D. E., Dang, Y., Bhat, S., Merrick, W. C., Green, R., Shen, B. and Liu, J. O. (2010) 'Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin', *Nat Chem Biol*, 6(3), pp. 209-217.
- Schuller, A. P. and Green, R. (2018) 'Roadblocks and resolutions in eukaryotic translation', *Nat Rev Mol Cell Biol*, 19(8), pp. 526-541.
- Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A. and Mühlemann, O. (2013) 'Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells', *Biochim Biophys Acta*, 1829(6-7), pp. 612-23.
- Shandilya, J. and Roberts, S. G. (2012) 'The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling', *Biochim Biophys Acta*, 1819(5), pp. 391-400.
- Shatkin, A. J. (1976) 'Capping of eucaryotic mRNAs', *Cell*, 9(4 PT 2), pp. 645-53.

- Sheth, U. and Parker, R. (2006) 'Targeting of aberrant mRNAs to cytoplasmic processing bodies', *Cell*, 125(6), pp. 1095-109.
- Shi, Z., Fujii, K., Kovary, K. M., Genuth, N. R., Röst, H. L., Teruel, M. N. and Barna, M. (2017) 'Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide', *Mol Cell*, 67(1), pp. 71-83.e7.
- Shum, E. Y., Jones, S. H., Shao, A., Dumdie, J., Krause, M. D., Chan, W. K., Lou, C. H., Espinoza, J. L., Song, H. W., Phan, M. H., Ramaiah, M., Huang, L., McCarrey, J. R., Peterson, K. J., De Rooij, D. G., Cook-Andersen, H. and Wilkinson, M. F. (2016) 'The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay', *Cell*, 165(2), pp. 382-95.
- Simsek, D., Tiu, G. C., Flynn, R. A., Byeon, G. W., Leppeck, K., Xu, A. F., Chang, H. Y. and Barna, M. (2017) 'The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity', *Cell*, 169(6), pp. 1051-1065.e18.
- Simón, E. and Séraphin, B. (2007) 'A specific role for the C-terminal region of the Poly(A)-binding protein in mRNA decay', *Nucleic Acids Res*, 35(18), pp. 6017-28.
- Spahn, C. M., Beckmann, R., Eswar, N., Penczek, P. A., Sali, A., Blobel, G. and Frank, J. (2001) 'Structure of the 80S ribosome from *Saccharomyces cerevisiae*--tRNA-ribosome and subunit-subunit interactions', *Cell*, 107(3), pp. 373-86.
- Stalder, L. and Mühlemann, O. (2008) 'The meaning of nonsense', *Trends Cell Biol*, 18(7), pp. 315-21.
- Suh, H., Ficarro, S. B., Kang, U. B., Chun, Y., Marto, J. A. and Buratowski, S. (2016) 'Direct Analysis of Phosphorylation Sites on the Rpb1 C-Terminal Domain of RNA Polymerase II', *Mol Cell*, 61(2), pp. 297-304.
- Sundermeier, T., Ge, Z., Richards, J., Dulebohn, D. and Karzai, A. W. (2008) 'Studying tmRNA-mediated surveillance and nonstop mRNA decay', *Methods Enzymol*, 447, pp. 329-58.

- Topisirovic, I., Svitkin, Y. V., Sonenberg, N. and Shatkin, A. J. (2011) 'Cap and cap-binding proteins in the control of gene expression', *Wiley Interdiscip Rev RNA*, 2(2), pp. 277-98.
- Venters, B. J. and Pugh, B. F. (2009) 'How eukaryotic genes are transcribed', *Crit Rev Biochem Mol Biol*, 44(2-3), pp. 117-41.
- Villa, N., Do, A., Hershey, J. W. and Fraser, C. S. (2013) 'Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome', *J Biol Chem*, 288(46), pp. 32932-40.
- Wahl, M. C., Will, C. L. and Lührmann, R. (2009) 'The spliceosome: design principles of a dynamic RNP machine', *Cell*, 136(4), pp. 701-18.
- Wang, W., Czaplinski, K., Rao, Y. and Peltz, S. W. (2001) 'The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts', *EMBO J*, 20(4), pp. 880-90.
- Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M. and Losick, R. (2014) *Molecular Biology of the Gene*. 7th edition edn.: Pearson.
- Welch, E. M. and Jacobson, A. (1999) 'An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA', *EMBO J*, 18(21), pp. 6134-45.
- Wells, S. E., Hillner, P. E., Vale, R. D. and Sachs, A. B. (1998) 'Circularization of mRNA by eukaryotic translation initiation factors', *Mol Cell*, 2(1), pp. 135-40.
- Wen, J. and Brogna, S. (2008) 'Nonsense-mediated mRNA decay', *Biochem Soc Trans*, 36(Pt 3), pp. 514-6.
- Wen, J. and Brogna, S. (2010) 'Splicing-dependent NMD does not require the EJC in *Schizosaccharomyces pombe*', *EMBO J*, 29(9), pp. 1537-51.
- Westra, E. R., Buckling, A. and Fineran, P. C. (2014) 'CRISPR-Cas systems: beyond adaptive immunity', *Nat Rev Microbiol*, 12(5), pp. 317-26.
- Wong, W., Bai, X. C., Brown, A., Fernandez, I. S., Hanssen, E., Condrón, M., Tan, Y. H., Baum, J. and Scheres, S. H. (2014) 'Cryo-EM structure of the

Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine', *Elife*, 3.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., Kurata, R., Hirano, H., Anderson, P. and Ohno, S. (2009) 'SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay', *Genes & Development*, 23, pp. 1091-1105.

Youngman, E. M., Brunelle, J. L., Kochaniak, A. B. and Green, R. (2004) 'The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release', *Cell*, 117(5), pp. 589-99.

Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. and Noller, H. F. (2001) 'Crystal structure of the ribosome at 5.5 Å resolution', *Science*, 292(5518), pp. 883-96.

Yusupova, G. and Yusupov, M. (2014) 'High-resolution structure of the eukaryotic 80S ribosome', *Annu Rev Biochem*, 83, pp. 467-86.

Zhang, J., Sun, X., Qian, Y., LaDuca, J. P. and Maquat, L. E. (1998) 'At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation', *Mol Cell Biol*, 18(9), pp. 5272-83.

Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S. and Yang, S. H. (2015) 'Off-target Effects in CRISPR/Cas9-mediated Genome Engineering', *Mol Ther Nucleic Acids*, 4, pp. e264.

Zünd, D., Gruber, A. R., Zavolan, M. and Mühlemann, O. (2013) 'Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs', *Nat Struct Mol Biol*, 20(8), pp. 936-43.

8. APPENDIX

8.1. *S. pombe* strains

Stock	Strain name	Genotype	Source
MP1	py114 (wt h+)	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E	Dr Janet F. Partridge (janet.partridge@stjude.org)(Petrie et al., 2005)
MP2	py115	h- ade6-210 arg3D his3D leu1-32 ura4DS/E	Dr Janet F. Partridge (janet.partridge@stjude.org)(Petrie et al., 2005)
MP3	h+ RpS2801-HA	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E RpS2801:HA:kanMX6	This study
MP4	h+ RpS2802-HA	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E RpS2802:HA:kanMX6	This study
MP5	upf1Δ	h+ his3D leu1-32 ura4D18 arg? upf1::KanMX6	Broгна lab (Jikai Wen)
MP6	upf2Δ	h+ his3D leu1-32 ura4D18 arg? upf2::KanMX6	Broгна lab (Jikai Wen)
MP7	upf1-flag	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E upf1:5flag::hphMX6	Broгна lab (Jianming Wang)
MP8	upf1-flag RpS2801-HA	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E UPF1:5FLAG::hphMX6 RpS2801:HA:kanMX6	This study
JM28	JM1-GFP	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E pREP41-GFP	Broгна lab (Jianming Wang)
JM27	JM1-PTC6	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E pREP41-GFP N6	Broгна lab (Jianming Wang)
JM30	JM2-GFP	h+ his3D leu1-32 ura4D18 arg? upf1::KanMX6 pREP41- GFP	Broгна lab (Jianming Wang)
JM29	JM2-PTC6	h+ his3D leu1-32 ura4D18 arg? upf1::KanMX6 pREP41- GFP N6	Broгна lab (Jianming Wang)
JM32	JM3-GFP	h+ his3D leu1-32 ura4D18 arg? upf2::KanMX6 pREP41- GFP	Broгна lab (Jianming Wang)
JM31	JM3-PTC6	h+ his3D leu1-32 ura4D18 arg? upf2::KanMX6 pREP41- GFP N6	Broгна lab (Jianming Wang)
MP9	wt GFP	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E pREP41-GFP	This study
MP10	wt PTC6	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E pREP41-GFP N6	This study
MP11	wt PTC27	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E pREP41-GFP N27	This study

MP12	wt PTC140	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E pREP41-GFP N140	This study
MP13	upf1Δ GFP	h+ his3D leu1-32 ura4D18 arg? upf1::KanMX6 pREP41- GFP	This study
MP14	upf1Δ PTC6	h+ his3D leu1-32 ura4D18 arg? upf1::KanMX6 pREP41- GFP N6	This study
MP15	upf1Δ PTC27	h+ his3D leu1-32 ura4D18 arg? upf1::KanMX6 pREP41- GFP N27	This study
MP16	upf1Δ PTC140	h+ his3D leu1-32 ura4D18 arg? upf1::KanMX6 pREP41- GFP N140	This study
MP17	upf3Δ GFP	h+ his3D leu1-32 ura4D18 arg? upf3::KanMX6 pREP41- GFP N140	This study
MP18	upf3Δ PTC6	h+ his3D leu1-32 ura4D18 arg? upf3::KanMX6 pREP41- GFP N140	This study
MP19	upf3Δ PTC27	h+ his3D leu1-32 ura4D18 arg? upf3::KanMX6 pREP41- GFP N140	This study
MP20	upf3Δ PTC140	h+ his3D leu1-32 ura4D18 arg? upf3::KanMX6 pREP41- GFP N140	This study
MP21	upf3Δ	h+ his3D leu1-32 ura4D18 arg? upf3::KanMX6	Brojna lab (Jikai Wen)
MP22	upf3Δ	h+ his3D leu1-32 ura4D18 arg? upf3::HphMX6	Brojna lab (Jikai Wen)
MP23	upf1-flag RpS2502-HA	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E UPF1:5FLAG::hphMX6 RpS2502:HA:kanMX6	This study
MP23	upf3-flag RpS2502-HA	h+ ade6-210 arg3D his3D leu1-32 ura4DS/E UPF3:5FLAG::hphMX6 RpS2502:HA:kanMX6	This study

8.2. Media

	Reagent	Quantity (for 1 L)	Final concentration
YES	glucose	30 g	3 %
	yeast extract	5 g	0.5 %
	adenine-HCl	225 mg	1.31 mM
	uracil	225 mg	2.01 mM
	L-histidine	225 mg	1.45 mM
	L-leucine	225 mg	1.71 mM
	L-lysine	225 mg	1.23 mM
	arginine	225 mg	1.3 mM
	agar (solely for solid media)	20 g	2 %
	ddH ₂ O	to 1 L	

	Reagent	Quantity (for 1 L)	Final concentration
EMMG	glucose	20 g	2 %
	potassium hydrogen phthalate	3 g	14.7 mM
	dibasic sodium phosphate	2.2 g	15.5 mM
	L-glutamic acid	3.38 g	2.3 mM
	adenine-HCl	225 mg	1.31 mM
	uracil	225 mg	2.01 mM
	L-histidine	225 mg	1.45 mM
	L-leucine	225 mg	1.71 mM
	L-lysine	225 mg	1.23 mM
	arginine	225 mg	1.3 mM
	50x salt stock (recipe below)	20 mL	1x
	1000x vitamin stock (recipe below)	1 mL	1x
	10000x mineral stock (recipe below)	0.1 mL	1x

	Reagent	Quantity (for 1 L)	Final concentration
MB	potassium dihydrogen phosphate	0.5 g	3.7 mM
	potassium acetate	0.36 g	3.7 mM
	magnesium sulphate heptahydrate	0.5 g	2 mM
	sodium chloride	0.1 g	1.7 mM
	calcium chloride dihydrate	0.1 g	0.7 mM
	di-ammonium sulphate	5 g	37.8 mM
	boric acid	500 µg	8.1 µM
	cupric sulfate pentahydrate	40 µg	0.16 µM
	potassium iodide	100 µg	0.6 µM
	ferric chloride hexahydrate	200 µg	0.7 µM
	manganese sulphate monohydrate	400 µg	2.4 µM
	sodium molybdate dihydrate	200 µg	0.8 µM
	zinc sulphate heptahydrate	400 µg	1.4 µM
	glucose	5 g	0.5 %
	biotin	10 µg	41 nM
	calcium pantothenate	1 mg	2.1 µM
	nicotinic acid	10 mg	81 µM
	myo-inositol	10 mg	55 µM
	uracil for ura4- strains or leucine for leu1- strains	150 mg	1.3 mM

	Reagent	Quantity (for 1 L)	Final concentration
NZY	yeast extract	5 g	0.5 %
	sodium chloride	5 g	85 mM
	magnesium chloride hexahydrate	2 g	10 mM
	NZ amine	10 g	1 %
	ddH ₂ O	To 1 L	

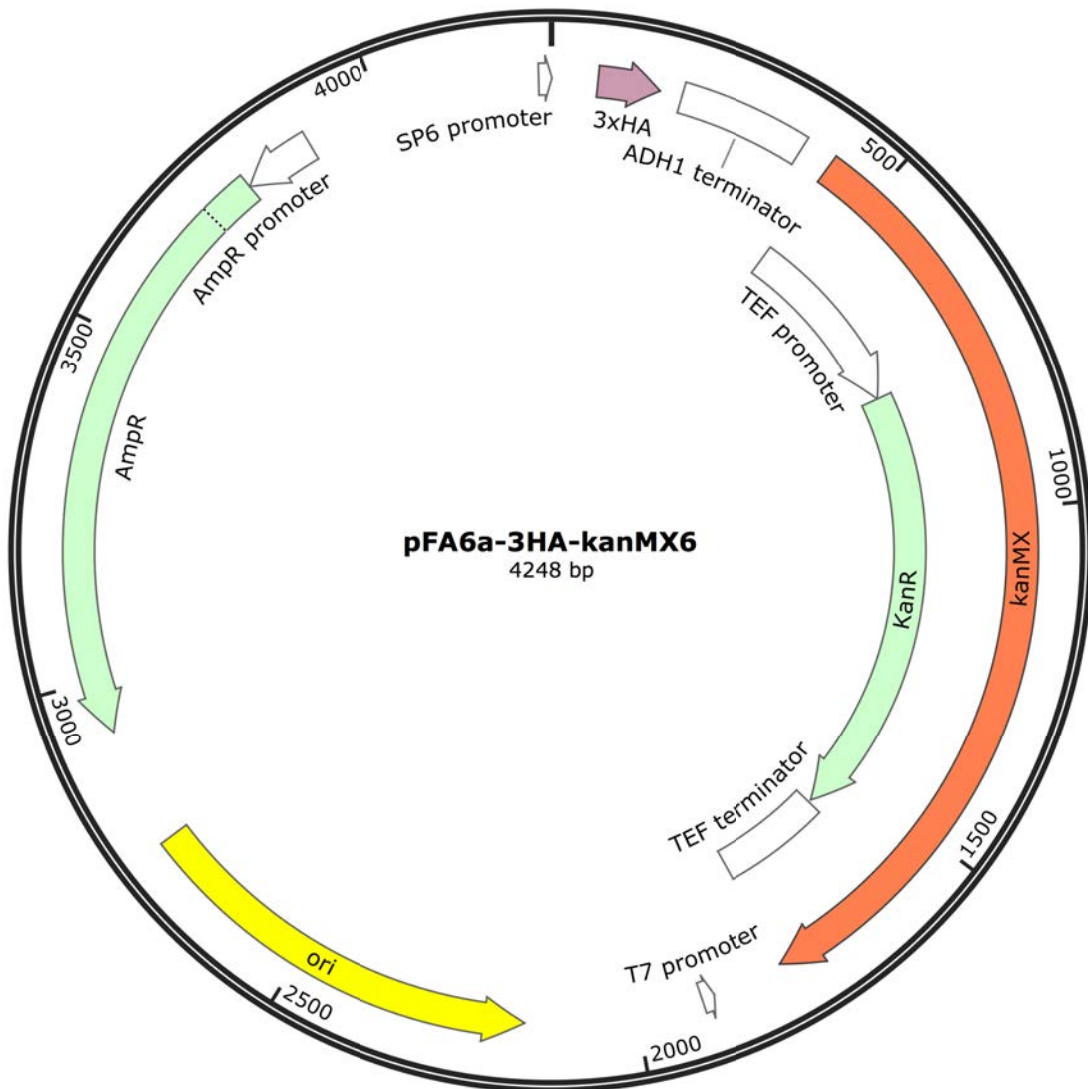
	Reagent	Quantity (for 1 L)	Final concentration
1000x vitamin stock	pantothenic acid	1 g	81.2 mM
	nicotinic acid	10 g	81.2 mM
	inositol	10 g	4.2 mM
	biotin	10 mg	40.9 μ M

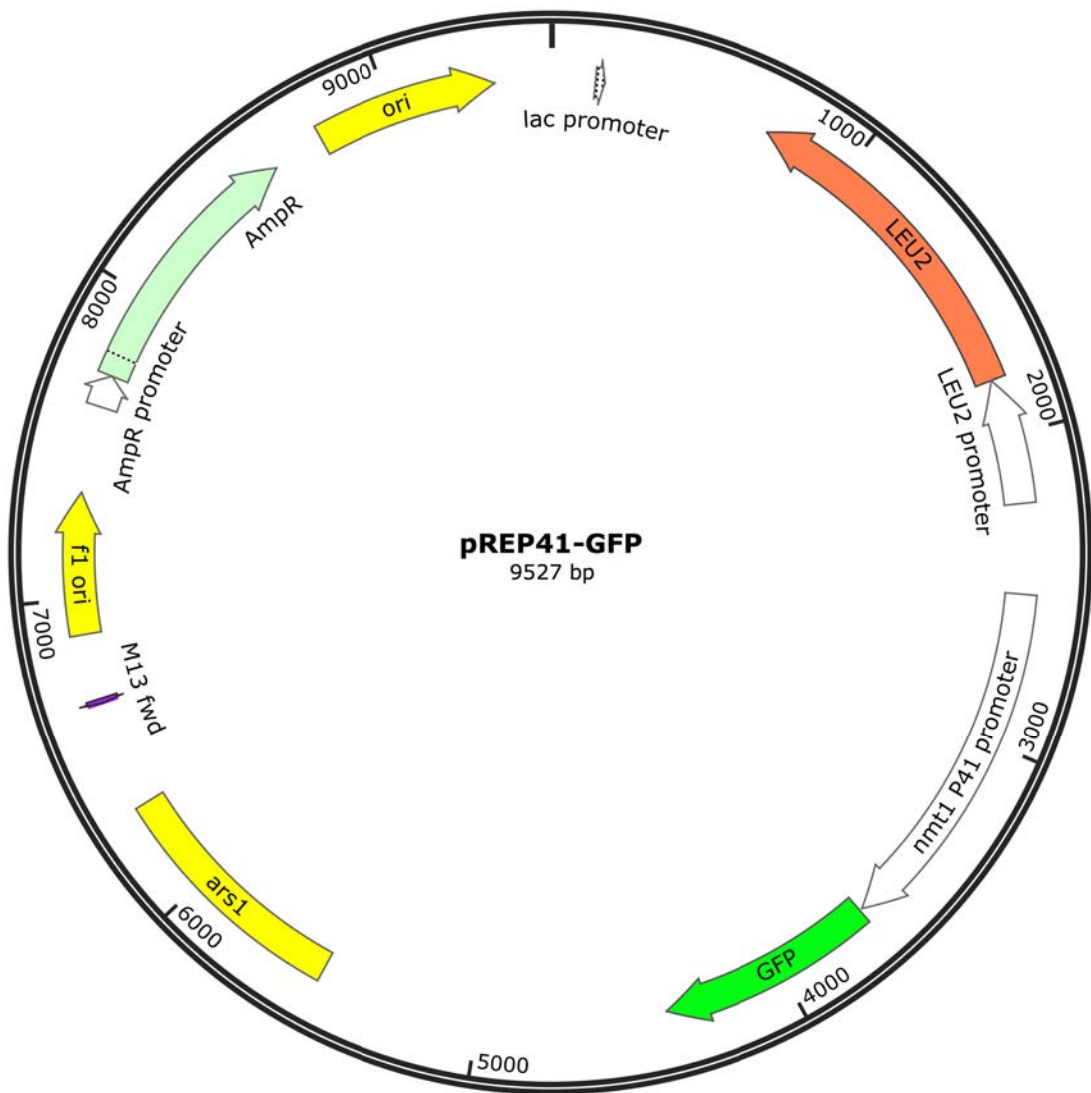
	Reagent	Quantity (for 1 L)	Final concentration
10000x mineral stock	boric acid	5 g	80.9 mM
	magnesium sulphate	4 g	33.2 mM
	zinc sulphate heptahydrate	4 g	13.9 mM
	ferric chloride hexahydrate	2 g	7.4 mM
	molybdic acid	0.4 g	0.32 mM
	potassium iodide	1 g	6.02 mM
	cupric sulfate pentahydrate	0.4 g	1.6 mM
	citric acid	10 g	47.6 mM

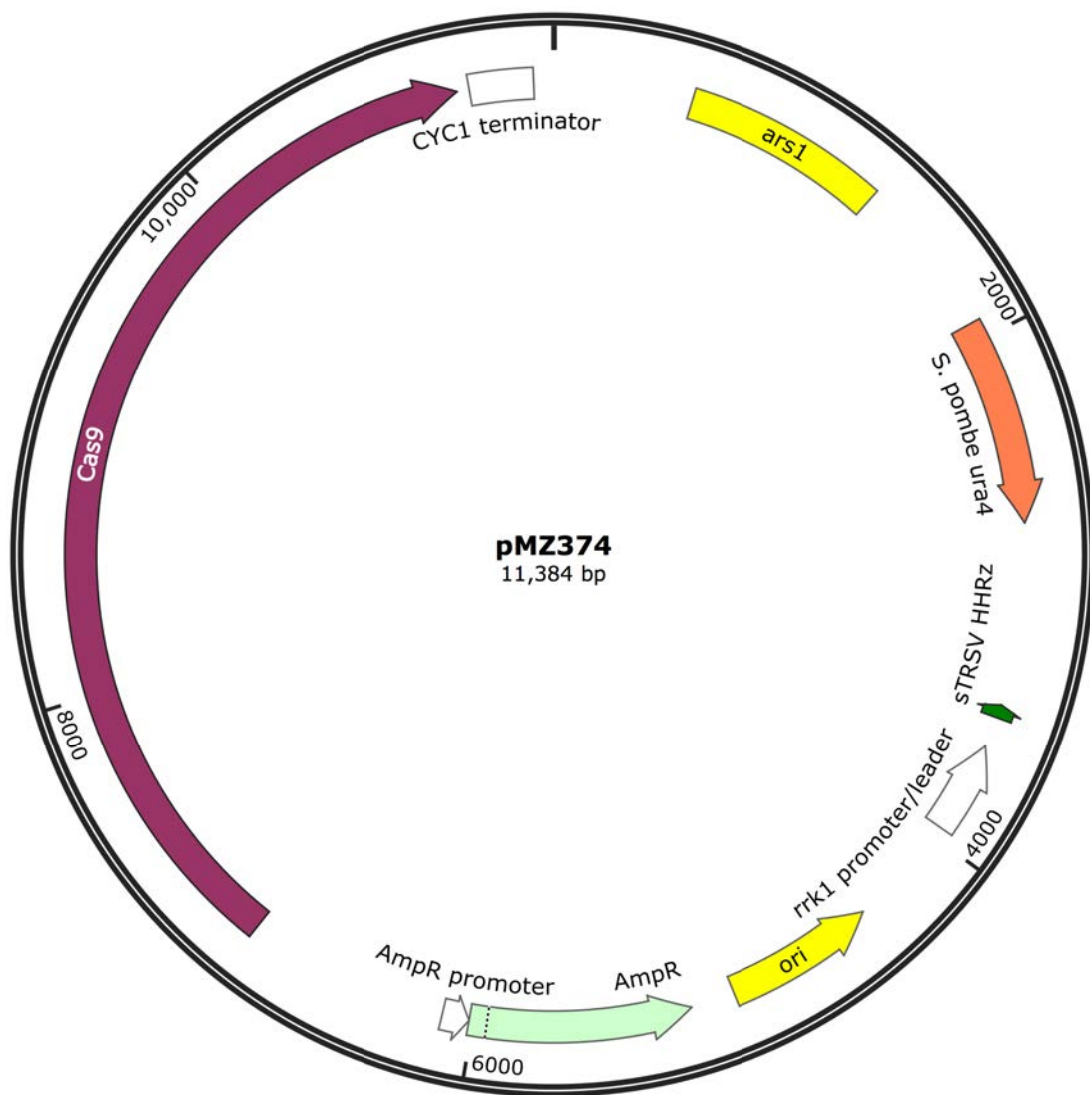
8.3. Plasmids

Plasmid name	Resistance in E. coli	Selectable marker in yeast	Comments
pFA6a-3xHA-KanMX6	amp	KanMX6	PCR based gene editing - C-terminal 3xHA tagging of a target gene
pREP41-GFP	amp	<i>LEU2</i>	Vector used for the episomal expression of GFP (from an <i>nmt41</i> promoter), obtained by cloning the GFP ORF into the multiple cloning site (Jikai Wen PhD Thesis)
pREP41-PTC6	amp	<i>LEU2</i>	Vector expressing an NMD reporter gene – GFP with a PTC at codon position 6 (Jikai Wen PhD Thesis)
pREP41-PTC27	amp	<i>LEU2</i>	Vector expressing an NMD reporter gene – GFP with a PTC at codon position 27 (Jikai Wen PhD Thesis)
pREP41-PTC140	amp	<i>LEU2</i>	Vector expressing an NMD reporter gene – GFP with a PTC at codon position 140 (Jikai Wen PhD Thesis)
pMZ374	amp	URA4	Vector containing a Cas9 gene and an <i>rrk1</i> placeholder for cloning gene-specific gRNA sequence and element for gRNA expression

8.4. Plasmid maps







8.5. Primers used for C-terminal protein tagging

Primer name	Description	Sequence
MP1	RpS2801_Ctl_F	ACCGGTTCTCGGTAAGTTTG
MP2	RpS2801_Ctl_R	GGGGATCCGTCGACCTGCAGCGTACGA ACGCAATCTTCTGGCCTCAC
MP3	RpS2801_Ctr_F	GTTTAAACGAGCTCGAATTCATCGAT AGCTGCTTTGGCCCATTTGTT
MP4	RpS2801_Ctr_F	TGTGTGTACATACTGCGTTTT
MP5	RpL2502_Ctl_F	CGTAAGAGTGTTCCCCATGC
MP6	RpL2502_Ctl_F	GGGGATCCGTCGACCTGCAGCGTACGA AAGGAAGCCAATGCGGTTGG
MP7	RpL2502_Ctl_F	GTTTAAACGAGCTCGAATTCATCGAT TAGCTTAAAACTAGTTTAATAATTCAC
MP8	RpL2502_Ctl_F	AATGGGGGTCCTCTAGTTCC

8.6. CRISPR/Cas9 gene editing

8.6.1. Ade6 gene primer design

Ade6 sequence with indicated designed PTCs:

```
CGACGACAAG  TATTCTTTCA  ACTTGGTGGT  GAGGTAAACG  AAATCCAGCC
TGGTGCAAGTA
61          TAAGGTATAA CGACAACAAA CGTTGCTTTA TATATGGATA ACGACATATA
CATTATTTTC
121          TTGAAAAAAT AGCAAAGCTA TATTGATTTT AGCGTATGCT GCTTTAAAGT
TGCATTTTAC
181          AATGCTTGGA AATGTAACGA TGACAGAATT AATACGATGC AAAACTCAAA
TAATAAACTG
241          CGCACTAACT CACTACAATA AACAACTTTG AAAAAAAGAT TCGTTTTTTC
AACATTTACC
301          ATCTCATTA  GCTGAGCTGC CAAGGTATAT ACATACTTCA TCGAATATGA
GCGAAAAACA
361          GGTGTAGGG ATCCTTGGAG GTGGTCAATT GGGCCGAATG ATGGTAGAGG
CAGCCCATCG
                                     N11
                                     GGA → TGA
421          CTAAACATC AAATGCATCA TCTTGGATGC AGCAAATTCT CCTGCCAAAC
AAATTGATGG
481          AGGACGTGAG CACATTGATG CATCATTTAC TGACCCCGAT GCAATTGTTG
AACTGTCTAA
                                     N45
                                     GGA → TGA
541          GAAGTGCACG TTATTAACAA CTGAAATTGA GCATATTAAC ACTGATGCCT
TGGCAGCCGT
601          TACGAAATCT GTTGCTGTTG AACCTCTCC TGCAACTCTG CGATGCATTC
AAGACAAATA
661          TCTTCAAAA CAGCATTTAC AGGTTTTTAA GATTGCACTT CCTGAATTTT
GCGATGCACC
721          TGACCAGGAA AGTGTGAAA AAGCAGGCCA AGAGTTTGGT TATCCTTTTG
TACTGAAAAG
781          TAAACATTG GCTTACGACG GTCGTGAAA TTACGTTGTT CATCAACCAT
CTGAGATTCC
841          TACTGCCATC AAAGCACTTG GTGATCGTCC GCTTTATGTT GAAAAGTTG
TTCCTTTCTC
901          CATGGAAATT GCAGTGATGG TAGTACGCAG TTTAGACGGA AAAGTTTATG
CTTATCCTAC
961          AACTGAGACC ATTCAAAGG ATAATGTTTG TCATTAGTA TATGCCCTG
CTCGTCTTCC
1021         CTTCTCAATT CAACAGCGTG CTCAAACCCT TGCCATGGAT GCGGTGCGCA
CTTTTGAAGG
                                     N244
                                     GAA → TAA
1081         CGCTGGTATA TATGGTGTAG AGATGTTTGT TTTGCAAGAT GGTGAGACCA
TTTTACTCAA
```

1141 CGAAATTGCT CTCGGCCTC ACAATTCAGG TCACTACACC ATTGAAGCTT
 GCCCAACTTC
 1201 TCAGTTTGAA GCTCACTTAC GGGCCATATG TGGTCTTCCT TTCAGCGAAA
 TCAACACCCA
 1261 ACTCTCGACT TCCACAACCTC ATGCGTTGAT GGTAATATT TTAGGTACTG
 ATGATCCTGA
 1321 TTATGTTTCA AAGATCGCTA AACGTTCTCT GTCCATTCCC GGTGCAACTT
 TGCATCTTTA
 1381 TGGTAAAGCT GAATCTAGAA AGGGTCGCAA GATGGGACAC GTTACCATCA
 TTTCTGATTC
 1441 ACCTCAAGAA TGTGAACGTA GGTATCAGAT GCTTCTTGAC GTCAAAGATC
 CTGTCGAATC
 1501 ACCTGTTGTT GGTATTATCA TGGGTTCCGA TTCTGATTGA AGCAAGATGA
 AAGATGCTGC
 1561 CGTCATTTA GATGAATTCA AGGTGCCTTA CGAACTTACT ATTGTTTCAG
 CTCACCGCAC
 1621 ACCAGATCGC ATGGTTACTT ATGCTCGTAC CGCAGCTTCA AGAGGGTTGC
 GTGTGATTAT
 1681 TGCTGGTGCT GGTGGTGCCG CTCATTTGCC TGGTATGGTT GCTGCAATGA
 CACCTCTTCC
 1741 AGTAATCGGC GTTCCTGTAA AAGGAAGCAC TCTTGACGGA GTTGA CTCTC
 TTCACTCTAT
 1801 TGTTGAGATG CCTCGAGG TG TCCCTGTCGC CACTGTTGCT ATCAATAATA
 GCCAAAATGC
 N490
 CGA → TGA
 1861 CGGTATTTTA GCCTGTCGTA TACTTGCTAC ATTTCAACCC TCCCTTTTGG
 CTGCTATGGA
 1921 GAGCTTTATG GACAATATGA AAGAGATTGT TTTAGAAAAG GCTGATAAAT
 TAGAAAAAGT
 1981 TGGTTGGAAA AATTATTCTG CATAGGCGAC CATAGACATA ACTGTAAAT
 GATTTTGCTT
 2041 GCTTTTGTGTT TTAAATAATA TTTTATTTTG AAATTCTAGA TTGTAAAATG
 TATTCTAAAA
 2101 AGTGATTAAA AATTAAAAAT AAAAACCCCC CGAATAATGT GCTGCGACGC
 AATTAAAAAA
 2161 AATAAGAATC CTGAATAATG TGCTGTGAAG CAGTTGAAAG AAAATTAAAG
 AACCCAATAT
 2221 AATATGCTAT AAAGCAATCA AAAAATAATA GTAAGAAACC CAAATAAAAC
 CCCGTTGACC
 2281 ATGTTCCCAA TTAAATTAAC GCATATTAAT GCAAAAAATG CAATATTAAG
 ACTAGTTATT
 2341 AAACC

PAM motifs coding sequence, UTR

Primers

MP125	sgRNA_ade6N11_F	CAGGTTGTAGGGATCCTTGG GTTTTAGAGCTAGAAATAGC
MP126	sgRNA_ade6N11_R	CCAAGGATCCCTACAACCTG TTCTTCGGTACAGGTTATG
MP127	ade6_N11_L_F	ACTTGGTGGTGAGGTAAACGA
MP128	ade6_N11_L_R	CGGCCCAATTGACCACCTCA AAGGATCCCTACAA CCTGTTT
MP129	ade6_N11_R_F	AAACAGGTTGTAGGGATCCTTTGAGGTGGTCAAT TGGGCCG
MP130	ade6_N11_R_R	AGAGTTGCAGGAGAGGGTTC
MP131	sgRNA_ade6N45_F	CCTGCCAAACAAATTGATGG GTTTTAGAGCTAGAAATAGC
MP132	sgRNA_ade6N45_R	CCATCAATTTGTTTGGCAGG TTCTTCGGTACAGGTTATG
MP133	ade6_N45_L_F	TGCTTGGAATGTAACGATGACA
MP134	ade6_N45_L_R	CATCAATGTGCTCACGTCCTCA ATCAATTTGTTT GCAGG
MP135	ade6_N45_altL_R	CATCAATGTGCTCACGTCCTCA ATCAATTT
MP136	ade6_N45_R_F	CCTGCCAAACAAATTGATTGAGGACGTGAGCACA TTGATG
MP137	ade6_N45_altR_F	AAATTGATTGAGGACGTGAGCACATTGATG
MP138	ade6_N45_R_R	GGCAGTAGGAATCTCAGATGGTT
MP139	sgRNA_ade6N244_	GATGCGGTGCGCACTTTTGA

	F	GTTTTAGAGCTAGAAATAGC
MP140	sgRNA_ade6N244_R	TCAAAAGTGCGCACCGCATC TTCTTCGGTACAGGTTATG
MP141	ade6_N244_L_F	AAACATTGGCTTACGACGGTC
MP142	ade6_N244_L_R	ACCATATATACCAGCGCCTT AAAAAGTGC
MP143	ade6_N244_altL_R	ACCATATATACCAGCGCCTT AAAAAGTGCGCACC GCATCCA
MP144	ade6_N244_R_F	TGGATGCGGTGCGCACTTTTT AAGGCGCTGGTAT ATATGGT
MP145	ade6_N244_altR_F	GCACTTTTT AAGGCGCTGGTATATATGGT
MP146	ade6_N244_R_R	GTAACGTGTCCCATCTTGCG
MP147	sgRNA_ade6N490_F	TCTATTGTTCAAGATGCCTCG GTTTTAGAGCTAGAAATAGC
MP148	sgRNA_ade6N490_R	CGAGGCATCTGAACAATAGA TTCTTCGGTACAGGTTATG
MP149	ade6_N490_L_F	GATGAAAGATGCTGCCGTCA
MP150	ade6_N490_L_R	ACAGTGGCGACAGGGACACCT CAAGGCATCTGA ACA
MP151	ade6_N490_R_F	TGTTCAAGATGCCTT GAGGTGTCCCTGTCGCCACT GT
MP152	ade6_N490_R_R	GTCGCAGCACATTATTCGGG

8.6.2. Ade2 gene primer design

Ade2 sequence with indicated designed PTCs:

CTAAGCATTT TATATTCTAT TTGACTATTG ATTGCCAATA AGCCTTTACG AGCGCATAAC
61 CATTATTTAC AAATAGCTTC AAAGTCAAAC TTTGAAGTCA ATAATAATTA
AAGAAATATG
121 ATCATATATT CTGATCTTTG TAATGGTGTT GGTCATAAAT TATACTGTTT
AACCGCTAAT
181 TTTGCTATAT CGTAACGTTA AGTGGAAAAA TAACCTTGAT CTTAGTGGTG
ATCAAGAATG
241 ACTACACAAG GCAGCGGCTA ACCGTACTTA CCGTTTTTCAG CATCACCAAA
ACGGAGATCT
301 AATAAAAACA TAGAAGAAAA AT**ATGGCTTC** **AGTGCGAGAA** **ACTGGTGTA**
ATGTTTCCAA
N12
TCC → TGA
361 **CGA****CGG**GTAA GAAAATTCTT TGAGAAGAAT TTACTATTCA CACAG**AATAA**
CCGTCGTCCT
421 **TGGCAGTCAA** **TGG**GTGACG AAGGAAAGGG TAAACTTGTT GATATCCTTT
GCGATAATGT
N24
TGG → TGA
481 **TGATGTCTGT** **GCTCGCTGTC** **AG**GTATGTTA CAAAAGCTTT ATATAGTGGT
GAACCGAACG
541 AGGTGATATT GTGTAAAATT TTTTTCCTACT CTACTTTGAC TCTTAAGTGC
GGGAGCTCGC
601 GGAATTACTG CCGATGAAAA ATAACATTTT TAATAAAAAA TTTAATTTTT
TGTTTAAATG
661 TTAATACAAA TCTGCTAAAT GACTTTTCCT ACTTGGTACA GTCAAGAAAA
GTATTTTAGA
721 GCTCACTGGA TAAGACGATC TGCATCAGTA ACAAATAAAT TTATACCAAT
TCTGAATGGT
781 TTCTTCTTTT TGATTTAATT ATGGAGGTTA CCTTTATGAT CTTGCCAGCT
TTTTTGTTG
841 CCCTTGATT GATTTTCAA ATATATAGCT AACTTTTATT TTTAG**GGTGG**
TAACAATGCA
901 **GGT**CATACCA TTGTTGCTAA TGGTGCTACT TACGATTTTC ACATTCTTCC
TTCAAGGACTC
N70 N53
TCA → TGA GGT → TGA
961 **GTAAATCCTA** **AATGCCAGAA** **CTTGATTGGT** **TCTGGTGTTG** **TCGTCTATTT**
ACCTGCTTTT
1021 **TTTAGTGAGT** **TAGAGAAACT** **TGAGCAAAAG** **GGTTTGAAAT** **GTAGAGATCG**
TATTTTCATT
1081 **TCTGATCGCG** **CTCATCTAGT** **ATTGACTAC** **CATCAACGTG** **CAGATGCTTT**
AAACGAGGCA
1141 **GAGCTCGGAA** **AACAAAGCAT** **CGGAACAACC** **GGAAAGGGTA** **TTGGTCCTGC**
TTATTCCACC

1201 AAGGCTACTC GCAGCGGTAT TCGTGTTTAT CACTTGTATC ATTGGGCTGA
 ATTTGAAGCT
 1261 CGTTACCGTA AGAACGTTGC TGAATTGCAG AAACGTTATG GTGCTTTTCA
 GTATGATGTT
 1321 GAAGCTGAAC TCATTCGTTA TAAGGAATTG GCTCAAAGAC TTAAGCCATT
 TGTCAATTGAT
 1381 GCTGTTGCGT TCATGTATGA AGCTTTACAA AGCAAGAAAC GTATCCTTGT
 CGAAGGTGCT
 1441 AACGCTTTGA TGCTTGATTT GGACTTTGGA ACTTATCCAT TCGTTACCAG
 CTCTAACACT
 1501 ACTGTTGGTG GTGTTTGCAC TGGTTTGGGT GTTCCTCCTC AGCGCATTGC
 TAACTCCATC
 1561 GGTGTAGTGA AAGCTTATAC AACTCGTGTA GCGGCTGGTC CTTTCCCAAC
 GGAACAGCTC
 1621 AACGAAATTG GTGACCATCT TCAAAGTGTT GGAAGAGAAG TTGGTGTGTTAC
 CACCGGTCGT

N303
 GGA → TGA

1681 AAACGCCGCT GCGGCTGGCT TGATTGGTTC GTTGTCAAAT ACTCTACTAT
 GATCAATGGT
 1741 TACACTTCTT TGAACCTTAC CAAGCTTGAC ATTTTGGATG CCTTGGATGA
 GATCAAAGTT
 1801 GCTGTTGCCT ACATTATTAA CGGTAAACGC ATTGAAACTT TTCCTGCTGA
 TCTTGATTCT
 1861 CTTGAAGAAG CCGAGATTGT ATACGAAACT TTCCCTGGTT GGAAAGTTCC
 TACAACAGGT
 1921 ATCACTCATT GGGATCAGAT GCCTGAGAAT GCCAAGAAAT ACATAGAGTT
 CATTGAAAAG
 1981 TTTGTTGGTG TTCCCATCAC TTTTCATCGGT GTTGGTCCTG GTCGTGATGA
 GATGTTGGTT
 2041 AAGGAGTAAG TTACGATTTA CAACTAGATA TCCTCACCAT TCTTCACTAC
 TTTCTTTATC
 2101 TAAATTTTGT TTATTATACA TCTTATTTTA GCTCTTTTCC ATTTATTTTT
 ATAGTGCTTT
 2161 TCTTTTTTTC GAATAGATTC ATAAAGCATA TAACTTTTTG ATTGTAATTT
 TATTTAAACG
 2221 TAACAATAAA GCTTGAAAGG CGAACAATAG CACAGAAATA ATTTTGATT
 AAATCCAACC
 2281 AGTATGGAAT TAGTATATGT AGTAAGAGAA AAGAGATTAT GCAATGTAGA
 ATAACTTAC
 2341 ACCGATCTTT AAAATAATTT AATCAAGGAT ATCCGTATCA AGGTACACAA
 TGCGA

PAM motifs coding sequence, UTR

Primers

MP153	sgRNA_ade2N12_F	GGTGTAATGTTTCCAATGA GTTTGTAGAGCTAGAAATAGC
MP154	sgRNA_ade2N12_R	TCATTGGAAACATTTACACC TTCTTCGGTACAGGTTATG
MP155	ade2N12_L_F	AGCCTTTACGAGCGCATAAC
MP156	ade2N12_L_R	CAAAGAATTTTCTTACCCATCATTT CAAA CATTTACACCAG
MP157	ade2N12_R_F	CTGGTGTAATGTTT GAA ATGATGGGTA AGAAAATTCTTTGAGA
MP158	ade2N12_R_R	TTCATCGGCAGTAATTCCGC
MP159	sgRNA_ade2N24_F	CCGTCGTCCTTGGCAGTCAA GTTTGTAGAGCTAGAAATAGC
MP160	sgRNA_ade2N24_R	TTGACTGCCAAGGACGACGG TTCTTCGGTACAGGTTATG
MP161	ade2N24_L_F	GCCTTTACGAGCGCATAACCA
MP162	ade2N24_L_R	GTTTACCCTTTCCTTCGTCACCT CATTG ACTGCCAAGG
MP163	ade2N24_R_F	TCGTCCTTGGCAGTCAATG AGGTGACG AAGGAAAGGGTA
MP164	ade2N24_R_R	TCGTCTTATCCAGTGAGCTCT
MP165	sgRNA_ade2N53_F	TTTGTAGGGTGGTAACAATGC GTTTGTAGAGCTAGAAATAGC
MP166	sgRNA_ade2N53_R	GCATTGTTACCACCCTAAAA TTCTTCGGTACAGGTTATG
MP167	ade2N53_L_F	GGGAGCTCGCGGAATTACT

MP168	ade2N53_L_R	CCATTAGCAACAATGGTATGT CAAGCAT TGTTACCACCCTA
MP169	ade2N53_R_F	TAGGGTGGTAACAATGCTTGACATACCA TTGTTGCTAATGG
MP170	ade2N53_R_R	TTTGTTTTCCGAGCTCTGCC
MP171	sgRNA_ade2N70_F	GATTTTCACATTCTTCCTTC GTTTATAGAGCTAGAAATAGC
MP172	sgRNA_ade2N70_R	GAAGGAAGAATGTGAAAATC TTCTTCGGTACAGGTTATG
MP173	ade2N70_L_F	GATGTCTGTGCTCGCTGTCAG
MP174	ade2N70_L_R	GGCATTTAGGATTACGAGTCCT CAAGG AAGAATGTGAAAATCGT
MP175	ade2N70_R_F	ACGATTTTCACATTCTTCCTT GAGGACT CGTAAATCCTAAATGCC
MP176	ade2N70_R_R	TGTTTTCCGAGCTCTGCCTC
MP177	sgRNA_ade2N303_F	CAAAGTGTTGGAAGAGAAGT GTTTATAGAGCTAGAAATAGC
MP178	sgRNA_ade2N303_R	ACTTCTCTTCCAACACTTTG TTCTTCGGTACAGGTTATG
MP179	ade2N303_L_F	CTTGTCGAAGGTGCTAACGC
MP180	ade2N303_L_R	TAACACCAACTTCTCTT CAACACTTTG AA
MP181	ade2N303_R_F	TTCAAAGTGTTTGAAGAGAAGTTGGTGT TA
MP182	ade2N303_R_R	CCAACCAGGGAAAGTTTCGT

8.6.3. RpS2801-HA protein tagging primer design

C-terminal tagging

1. Primers to incorporate sgRNA into the pMZ374 vector

sgRNA_RpS2801Cterm-F AGATTGCGTTAAAGCTGCTT gtttagagctagaaatagc

sgRNA_RpS2801Cterm-R AAGCAGCTTTAACGCAATCT ttctcggtacaggttatg

2. Primers to amplify 3-HA tag from pFA6a-3HA-kanMX6 plasmid with gene specific sequences:

RpS2801_3HA_hybrprimer_F GTGAGGCCAGAAGATTGCGT tacccatacgatgttcct

RpS2801_3HA_hybrprimer_R AATGGGCCGAAGCAGCTTTA gcactgagcagcgtaatc

3. Primers to amplify flanking regions homologous to the target site from genomic DNA:

RpS2801_CRISPR_C_term_L_F GGTTCCTCGGTAAGTTTGGGA

RpS2801_CRISPR_C_term_L_R ACGCAATCTTCTGGCCTCAC

RpS2801_CRISPR_C_term_R_F TAAAGCTGCTTGGCCCATTTGTTGTTAAGAGC

RpS2801_CRISPR_C_term_R_R TACAAGGACTCTGGAAAAGGGG

N-terminal tagging

1. Primers to incorporate sgRNA into the pMZ374 vector

sgRNA_RpS2801Nterm-F GAGGATAATGGATTCTTCTA gtttagagctagaaatagc

sgRNA_RpS2801Nterm-R TAGAAGAATCCATTATCCTC ttctcggtacaggttatg

2. Primers to amplify 3-HA tag from pFA6a-3HA-kanMX6 plasmid with gene specific sequences:

RpS2801_3HA_N_term_F TACATATCAAGAGGATAATG tacccatacgatgttcct

RpS2801_3HA_N_term_R ACAGGAACCTTAGAAGAATC gcactgagcagcgtaatc

3. Primers to amplify flanking regions homologous to the target site from genomic DNA:

RpS2801_CRISPR_N_term_L_F GGGAATCAAGACTGTGCTTCT

RpS2801_CRISPR_N_term_L_R CATTATCCTCTTGATATGTATAAGAGTGC

RpS2801_CRISPR_N_term_R_F GATTCTTCTAAAGTTCCTGTAAATTGG

RpS2801_CRISPR_N_term_R_R ACGAACAGGTCCTTAACGT

Desired sequence (final)

C-terminal tagging

Start codon; 3xHA; stop; point mutation

TCTTATACATATCAAGAGGATAATGGATTCTTCTAAGGTTCTGTAAATTGGCAAA
GGTCATTAAAGTGCTCGGTCGCACCGGTTCTCGGTAAGTTTGGGAACTTGTCAG
TTTATTGTTAAATTCAACATATTGTATTCACATAAATTGATTTGTGTTAAACGAATAA
ACATAAAAGATTCTAATTATTCCTCGCTTTTGTATATCCAGCAAGCAATGATTACTA
ACTTCTCTGTAGTGGTGGAGTCACCCAAGTCCGTGTCTGAATTTATGGATGACACTT
CTCGTTCTATCATTCGTAACGTTAAGGGACCTGTTTCGTGAAGATGACATCTTGTT
CTTCTTGAATCAGAGCGTGAGGCCAGAAGATTGCGTTACCCATACGATGTTCTG
ACTATGCGGGCTATCCGTATGACGTCCCGGACTATGCAGGATCCTATCCATATGA
CGTTCCAGATTACGCTGCTCAGTGTAAAGCTGCTTCGGCCCATTTGTTGTTAAGA
GCATCTTTGCAAAAAATATAATCTCGTTACACCTTTTACGATTCAA

N-terminal tagging

Start codon; 3xHA; stop; point mutation

TCTTATACATATCAAGAGGATAATGTACCCATACGATGTTCTGACTATGCGGGCT
ATCCGTATGACGTCCCGGACTATGCAGGATCCTATCCATATGACGTTCCAGATTAC
GCTGCTCAGTGCATTCTTCTAAAGTTCTGTAAATTGGCAAAGGTCATTAAAGT
GCTCGGTCGCACCGGTTCTCGGTAAGTTTGGGAACTTGTCAGTTTATTGTTAAAT
TCAACATATTGTATTCACATAAATTGATTTGTGTTAAACGAATAAACATAAAAGATTC
TAATTATTCCTCGCTTTTGTATATCCAGCAAGCAATGATTACTAACTTCTCTGTAGT
GGTGGAGTCACCCAAGTCCGTGTCTGAATTTATGGATGACACTTCTCGTTCTATCAT
TCGTAACGTTAAGGGACCTGTTTCGTGAAGATGACATCTTGTTTCTTCTTGAATCAG
AGCGTGAGGCCAGAAGATTGCGTTAAAGCTGCTTTGGCCCATTTGTTGTTAAGAGC
ATCTTTGCAAAAAATATAATCTCGTTACACCTTTTACGATTCAA