Characterization of Nonsense Mediated mRNA Decay in *Schizosaccharomyces pombe*

JIKAI WEN

Supervised by Dr. Saverio Brogna

September 2009

School of Biosciences
The University of Birmingham
Birmingham, B15 2TT

This dissertation is submitted for the degree of
DOCTOR OF PHILOSOPHY

UNIVERSITYOF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

SUMMARY

Nonsense mediated mRNA decay (NMD) is a translation-coupled process that preferentially destroys mRNAs harboring premature translation termination codons (PTCs). In mammalian cells, NMD is linked to pre-mRNA splicing. Typically, PTCs elicit strong NMD only if positioned upstream of at least one intron. The exon junction complex (EJC) is believed to mediate the link between splicing and NMD. However, recent studies have questioned the importance of splicing and the EJC in NMD and instead they have proposed that, from yeast to mammalian cells, NMD is mostly determined by the distance of the PTC from the 3' end.

In this study, to investigate the link between pre-mRNA splicing and NMD, I used the fission yeast *Schizosaccharomyces pombe*. Many genes carry introns in fission yeast and unlike in *Saccharomyces cerevisiae*, the genome encodes for proteins homologous to EJC components. *S. pombe* is a powerful model organism which is easily genetically manipulated and we envisaged that studying NMD in this organism should facilitate the understanding of the molecular mechanism and in particular the link between NMD and splicing.

During my PhD research, I have developed a versatile gene reporter system to study NMD; with it I discovered that splicing strongly enhances NMD in fission yeast and, surprisingly, that the EJC does not appear to be required. Unexpectedly, I found that splicing enhances NMD when the intron is either before or after the PTC, and furthermore, it does so only when the intron is close to the PTC. These observations suggest that the effect of splicing on NMD is direct and not a secondary consequence of splicing enhancing translation. Splicing is not an absolute requirement for NMD; I found that PTCs located early in the coding region could induce NMD even in the

absence of a nearby intron. However, against the prediction of current models, I found no strong correlation with the distance of the PTC from the 3' end. In summary, during my PhD a versatile system has been developed to study NMD in fission yeast and what I have observed challenges current NMD models and provides new mechanistic insights into NMD.

ACKNOWLEDGEMENTS

Thank you to my ex-colleagues in Wuhan Institute of Virology, China, who made it such a pleasurable experience working in the Xianen group. Special thanks are due to Prof. Xianen Zhang and Prof. Zhongxun Luo (deceased). Thank you for introducing me to biological research and for teaching me so much when I was an undergraduate and a master student.

Last but not least, I would like to thank all my friends from Birmingham and beyond, and my family. Without their plentiful supplies of love, support and help, my doctorate would have been neither half as productive nor half as fun.

Jikai Wen

September 2009

Birmingham

Table of Contents

ABBREVATIONS AND ANNOTATIONS

Abbreviations in the thesis

NMD	Nonsense-mediated mRNA decay
PTC	Premature termination condon
AS	Alternative splicing
PABPC	Poly(A) binding protein, cytoplasm
NPC	Nuclear pore complex
μORF	Small upstream open reading frame
UPF	Up-frameshift
eRF1/eRF3	Eukaryotic release factor 1 and 3
eIF3/eIF4AIII/eIF4G	Eukaryotic translation initiation factors
DSE	Downstream sequence
EJC	Exon junction complex
GFP	Green fluorescent protein
UTR	Untranslated region
ivs	Intervening sequence/intron
EST	Expressed sequence tag
SMG	Suppressor with morphological effect on genitalia
CBC	Cap binding complex
CBP	Cap binding protein
BSA	Bovine serum abulmin
HRP	Horseradish peroxidase
SDS	Sodium dodecyl sulfate

Tris	Tris(hydroxymethyl) aminomethane
EDTA	Ethylenediaminetetraacetic acid
PEG	Polyethylene glycol
FISH	Fluorescent *in situ* hybridization
DMSO	Dimethyl sulfoxide
HA	Hemagglutinin

Abbreviations names are usually given in full when they are first mentioned. This list only shows the most frequently used and those for which full names were not given in the text.

Annotations of the nomeclature in *S. pombe*

I have used a standardised style to refer to genes and proteins names. There are significant differences in the nomenclature among mammalian cells, *S. cerevisiae*, *D. melanogaster* and *S. pombe*. For example, *UPF1* is the gene which encodes the protein UPF1; UPF1 is the protein encoded by *UPF1*; *upf1Δ* means the *UPF1* gene was completely removed from the genome of *S. pombe*; *upf1* is a mutation of the gene in the genome. However, when the abbreviation refers to a gene or protein of a specific organism, that organism spefic nomeclature was used. For example, the *D. melanogaster* alcohol dehydrogenase gene is abbreviated *Adh* according to the fly nomenclature, not *ADH*. Similarly, if the abbreviation refers to a *S. cerevisiae* protein, it is, for example, Upf1p, not UPF1.

CHAPTER I. INTRODUCTION

Gene expression is the process that leads from transctription of DNA into mRNA and the translation of mRNA into protein. Although key steps are conserved between prokaryotes and eukaryotes, there are essential differences between the two groups of organisms. One of the key differences is that eukaryotic cells have a nucleus enclosed by a double membrane; this separates transcripion from translation, which occurs in the cytoplasm: the mRNA is exported to the cytoplasm through nucleor pore complexes (NPC) that span the nuclear membranes. In addition, in eukaryotes the primary transcripts (pre-mRNAs) undergo several covalent modifications before export of the mature mRNA to the cytoplasm. Nuclear pre-mRNAprocessing typically included including 3'-end processing and polyadenylation, 5'-capping and splicing. The 5'-cap and poly(A) tail are required stabilize the mRNA and also ply a key function in translation initiation and termination. Since pre-mRNA processing and translation are occurring in different compartments, it was believed that the should not be any link between the nucler history of the mRNA and its cytoplasmic fate. However, recent studies have indicated that pre-mRNA splicing can impinge on translation, mRNA localization and nonsense mediated mRNA decay (NMD) (Moore and Proudfoot et al., 2009). These findings suggests there are mechanisms that allow the ribosome to differentiate between spliced mRNA and transcript that have not undergone splicing; as reviewed below the link is probably mediated by proteins that are deposited on the transcript during pre-mRNA processing; these remain bound to the mRNA during export and then are recognized during translation of or by other cytoplasmic processes.

1.1 Nonsense-mediated mRNA decay and its substrates

Nonsense mediated mRNA decay (NMD) is a process served as a cellular mRNA qualiy control mechanism of gene expression. Here, I give a hystorical perspective of the phenomenon and an overiew of what is known, highlighting similarity and differences between model organisms. The first indications that NMD is a general feature of gene expression came from reports that nonsense mutations cause a rapid and drastic mRNA reduction in organism as different as bacteria, yeast and humans (Morse and Yanofsky, 1969; Peltz et al., 1993; Maquat, 1995). In eukaryotes, this phenomenon is thought to be a post-transcriptional process that probably protects cells from the potentially deleterious effect of truncated proteins. NMD has been extensively studied in *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana* and mammalian cells (He et al., 1993; Pulak and Anderson, 1993; Maquat, 1995; Brogna, 1999; Le Hir et al., 2001b; Gatfield et al., 2003; Amrani et al., 2004; Buhler et al., 2006; Kertesz et al., 2006).

Nonsense mutations generate premature termination codons (PTCs) and cause premature translation termination. Different biological processes can generate PTCs. For example, errors during DNA replication can cause point mutations, deletions, and insertions, which often generate PTCs. The expectation is that 3 out of 64 nucleotide substitutions (~5%) generate stop codons. Small insertions or deletions which cause frame-shifts, often also generate downstream PTCs in the coding region. In addition, PTCs can be generated by transcriptional errors and by abnormal pre-mRNA processing. Mutations that alter splicing signals generate nonsense mutations, frequently due to the retention of intronic sequences (Mendell and Dietz, 2001;

Holbrook et al., 2004). Abnormally spliced mRNAs are probably the most frequent NMD substrates in cells. Computational analysis of the human EST database showed that around 64% of human transcripts are alternatively spliced (AS), and about 35% of the transcripts generate at least one splice isoform encoding a PTC (Green et al., 2003; Lewis et al., 2003; McGlincy and Smith, 2008). In *S. cerevisiae* NMD inactivation causes an accumulation of many unspliced mRNAs, thus NMD might have an important role in discarding sub-optimally spliced transcripts (He et al., 1993; Sayani et al., 2008). On the basis of these findings, it has been proposed that NMD might contribute to splicing regulation (Green et al., 2003; Sayani et al., 2008). However, this proposal remains controversial because of recent studies that reported that only a small portion of the PTC-containing splice variants appears to be substantially regulated by NMD, which indicates that NMD might play a limited role in the regulation of PTC-containing AS transcripts (Baek and Green, 2005; Blencowe, 2006; Pan et al., 2006; McGlincy and Smith, 2008).

Studies in eukaryotes show that NMD substrates are not only PTC-containing transcripts. There is evidence that mRNAs containing small upstream open reading frames (μORFs) or bicistronic mRNAs are also NMD substrates (Oliveira and McCarthy, 1995). Additional NMD substrates are mRNAs with abnormally extended 3' untranslated region (3' UTR) (Muhlrad and Parker, 1999), and mRNAs with extra out-of-frame AUGs near the correct AUG (Welch and Jacobson, 1999) (table 1.1).

1.2 *Trans acting* factors and NMD mechanisms

NMD is not a passive mechanism and it requires specific *trans acting* factors. Some of these factors were initially identified in genetic screens aimed at the isolation of nonsense suppressor mutations. Others were identified in biochemical studies or

Table 1.1 Sources of NMD substrates

Aberrant mRNA generated by DNA alterations

Nonsense mutations	Nucleotide substitutions; 5% of all mutations generate PTCs (Chang and Kan, 1979)
Insertions and deletions	Random nucleotide insertions or deletions leading to a frameshift and PTCs (Culbertson et al., 1980)
Splice signals mutations	Mutations in splicing signals can cause splicing inhibition or aberrant splicing, leading to a frameshift or to intron-encoded PTCs (Mendell and Dietz, 2001; Holbrook et al., 2004)
Somatic DNA rearrangement	During programmed somatic maturation of the immunoglobulin and T-cell receptor genes, two thirds of the rearrangements generate frame shifts, often producing PTCs (Carter et al., 1995; Li and Wilkinson, 1998; Weischenfeldt et al., 2008)

Aberrant mRNA generated by transcription or posttranscriptional alterations

Transcription errors	The transcription errors at low frequency can produce PTCs.
Alternative splicing	Alternative splicing often produces frame shifts and PTCs (Green et al., 2003; Lewis et al., 2003; Sayani et al., 2008)

Physiological NMD substrates

Programmed translational frameshift	+1 or -1 frameshift can generate PTCs (Culbertson et al., 1980; Lee et al., 1995)
mRNAs with introns in 3' UTR	Observed in mammalian cells (Mendell et al., 2004)
Transcribed pseudogenes	*C. elegans, S. cerevisiae* and *S. pombe* (He et al., 2003; Mitrovich and Anderson, 2005; Rodriguez-Gabriel et al., 2006)
Transposons and retroviruses	Observed in *S. cerevisiae* and mammalian cells (He et al., 2003; LeBlanc and Beemon, 2004)
mRNAs with uORFs	The termination codon of the uORF is likely to be recognized as a PTC (He et al., 2003; Mitrovich and Anderson, 2005)
mRNAs with long 3' UTRs	Observed in yeast, fly, plant and human (He et al., 2003; Mitrovich and Anderson, 2005; Schwartz et al., 2006; Kebaara and Atkin, 2009)
mRNAs encoding selenoproteins	UGA is recognized as a PTC at low selenium concentrations (Moriarty et al., 1998; Sun and Maquat, 2002)
Bicistronic mRNAs	Observed in *S. cerevisiae* (He et al., 2003)

This table is modified from Muhlemann et al. (2008)

cropped up in various genetic screens aimed at studying unrelated processes. These proteins (listed in Table 1.2) are described below.

1.2.1 Genetic identifications of *trans acting* factors functioning in NMD

Trans acting factors involved in NMD were first identified in genetic screens in *S. cerevisiae* and *C. elegans* (Culbertson et al., 1980; Hodgkin et al., 1989). In *S. cerevisiae,* these factors were identified in a screen for suppressors of *his4-38* frameshift mutations, in a genetic background with a tRNA frameshift suppressor (Culbertson et al., 1980). The tRNA suppressor caused a low level of translation read-through that was sufficient to inhibit the His- phenotype at 30 °C, but not at 37 °C. Three mutations that suppress the His- phenotype at 37 °C were isolated and termed Up-frameshift (*upf*), *UPF1, UPF2,* and *UPF3* (Leeds et al., 1992)*. UPF1* was also hit in other screens in budding yeast. For example, in screens for mutations that increase by -1 the ribosomal frameshift efficiency, the mutations are named *mof* (**m**aintenance **o**f **f**rame) (Dinman and Wickner, 1994) and *ifs* (**i**ncreased **f**rame**s**hifting) (Lee et al., 1995). *UPF2* was also identified between mutations that affect translation initiation, named *sua* (**s**uppressor **o**f upstream **A**TG) (Hampsey et al., 1991; Pinto et al., 1992). As a consequence *UPF1* is also named *MOF4/IFS2* (Cui et al., 1996), *UPF2* (Cui et al., 1996) also *SUA1/IFS1* (Cui et al., 1996), and *UPF3* also *SUA6* (Cui et al., 1995). Additionally, *UPF1* is known as *NAM7*, since it was also isolated in a screen for mutations that suppress mitochondrial intronic mutations defective in RNA splicing (Altamura et al., 1992). Interestingly, mutations in other genes - *MOF2/ SUI1, MOF5, MOF8,* and *PRT1* - also stabilize PTC-containing mRNAs (Cui et al., 1999; Welch and Jacobson, 1999). *MOF2/SUI1* encodes for the translation initiation factor eIF1, which is essential for cap mediated initiation of translation (Pestova et al., 1998). The *PTR1* gene encodes for the p90 subunit of eIF3 (eIF3b), which is another essential

translation initiation factor (Isken et al., 2008; Martineau et al., 2008). Furthermore, one subunit of translation initiation factor eIF3 (eIF3e), named as integration site 6 protein (INT6) in mammalian cells, was co-immunoprecipitated with CBP80 and UPF2; interestingly, INT6 is not, unlike other eIF3 subunits, essential for global translation, but depletion of INT6 strongly inhibits NMD in human cells (Morris et al., 2007).

In *C. elegans,* NMD genes were originally isolated in a suppressor screen and grouped as *smg* (suppressor with morphogenetic effect on genitalia) mutations, which were classified as allele-specific but not gene-specific informational suppressors (Hodgkin et al., 1989; Pulak and Anderson, 1993; Morrison et al., 1997; Cali and Anderson, 1998; Cali et al., 1999). Genetic analysis of *smg 1-7* mutants and *smg*-suppressible alleles suggested that *SMG* genes functioned in all tissues of *C. elegans* at all times of development and were suggested to be involved in the same mRNA process or pathway (Hodgkin et al., 1989; Pulak and Anderson, 1993; Cali et al., 1999). Among these seven *SMG* genes, *SMG2*, *SMG3* and *SMG4* correspond to *UPF1*, *UPF2* and *UPF3*, respectively in *S. cerevisiae*. Recently, two additional proteins, named SMGL-1 and SMGL-2 have been identified in *C. elegans,* which are required for NMD in both nematodes and human cells (Longman et al., 2007).

1.2.2 The UPF and SMG proteins

Upf1p (also known as RENT1 in human) is a RNA binding protein with ATPase dependent helicase activity (Czaplinski et al., 1995). In *S. cerevisiae* and human cells, Upf1p associates with the ribosome via the interaction with the eukaryotic translation release factors eRF1 and eRF3 during premature translation termination (Weng et al., 1996a, b; Czaplinski et al., 1998; Weng et al., 1998; Kashima et al., 2006; Ivanov et

al., 2008). UPF1 has two separate functional domains: a cysteine-histidine rich (CH) domain at the amino terminus and an ATPase/helicase domain at the carboxyl terminus. A subset of mutations in the helicase domain inactivates NMD, but the protein is still able to enhance translation termination (Weng et al., 1996a). Another subset of mutations in the CH domain reduces the ability of UPF1 to enhance translation termination at a stop codon (Weng et al., 1996b). Some studies reported that certain mutations in *UPF1* gene result in a high frequency of ribosomal frameshifting and suggest that UPF1 may be directly involved in translation termination in addition to mRNA decay (Cui et al., 1996; Cui et al., 1998; Ruiz-Echevarria et al., 1998b). However, the role of UPF1 in translation termination is still controversial. Some other studies in *S. cerevisiae* indicate that UPF1 is not directly involved in translation termination and the nonsense suppression phenotype is probably a secondary effect of mRNA stabilization (Bidou et al., 2000; Harger and Dinman, 2004).

A number of observations indicate that UPF1, UPF2 and UPF3 form a complex (Weng et al., 1996b; He et al., 1997). Yeast two hybrid studies identified two interactions between UPF proteins, UPF1-UPF2 and UPF2-UPF3 (He et al., 1997). The interaction between UPF1 and UPF3 is bridged by UPF2. Deletion of UPF2 completely abolished the UPF1-UPF3 interaction and high expression of the full-length UPF2 or a fragmented UPF2 keeping the UPF1-UPF3 interaction domain, strongly enhanced the association of UPF1-3 in the complex. Surprisingly, mutations in the UPF1 domain which interacts with UPF2 do not abolish NMD (He et al., 1996). In human cells, UPF2 also interacts with UPF1 and UPF3 (Lykke-Andersen et al., 2000; Mendell et al., 2000; Serin et al., 2001; Schell et al., 2003; Chamieh et al., 2008). UPF2 contains three continuous middle domains which resemble eukaryotic

initiation factor 4G motifs (MIF4G) (Letunic et al., 2002). MIF4G domains are found in several proteins involved in mRNA processing and translation, such as eIF4G and CBP80 (Marcotrigiano et al., 2001; Mazza et al., 2001).

In humans there are two UPF3 homologues, UPF3A and UPF3B (Lykke-Andersen et al., 2000).(Lykke-Andersen et al., 2000). Both proteins contain a putative RNA recognition motif (RRM) at the amino terminus, which is normally involved in RNA binding. Another recent study has reported the structure of the complex between the interaction domains of UPF2 (residues 761-1054) and UPF3B (residues 42-143) (Kadlec et al., 2004). Against expectations, this.structure revealed that UPF3 (UPF3B) interacts with UPF2 via the RRM domain (Kadlec et al., 2004). In fact, the RRM domain of UPF3 does not bind RNA, instead the third MIF4G domain of UPF2 binds to RNA in the UPF2/UPF3 complex (Kadlec et al., 2004). Another study reported the structure of the CH domain of UPF1 and showed that this domain that interacts with UPF2, and has remarkable structural similarities to the RING domain found in ubiquitin ligases (Kadlec et al., 2006). A more recent study suggested that in the UPF trimetric complex, the association of UPF2 and UPF3B to UPF1, stimulates the ATPase and RNA helicase activities of UPF1, activating NMD (Chamieh et al., 2008).

Additional data also indicate that these proteins should form a trimetric complex, however, the abundance of the proteins in the cell varies greatly. In *S. cerevisiae* the copy number of UPF proteins was measured: 1600, 160 and 80 copies of UPF1, UPF2 and UPF3 per cell were estimated by Western blot of GST or HA fused Proteins and 6091, 1280 and 1250 copies per cell using GFP fusions (Maderazo et al., 2000; Ghaemmaghami et al., 2003). In mammalian cells, UPF proteins are very abundant,

tens of thousands to millions copies per cell, much more than the protein levels in budding yeast (Maquat and Serin, 2001; Culbertson and Neeno-Eckwall, 2005). However, UPF1 is almost five to ten times more abundant than UPF2 and UPF3. Possibly, as suggested by many studies, UPF1 has additional non-NMD roles, independent of UPF2 and UPF3, (Kaygun and Marzluff, 2005; Kim et al., 2005; Azzalin and Lingner, 2006; Isken et al., 2008).

A number of studies indicate that UPF1 can be phosphorylated and that regulation of the dephosphorylation /phosphorylation state of UPF1 is important for NMD in metazoa and human cells (Ohnishi et al., 2003; Grimson et al., 2004; Isken et al., 2008). SMG1 is a phoshatidylinositol kinase-related kinase that phosphorylates UPF1 (Yamashita et al., 2001; Grimson et al., 2004). SMG5 and SMG7 are involved in the dephosphorylation of UPF1 (Anders et al., 2003; Chiu et al., 2003; Ohnishi et al., 2003; Fukuhara et al., 2005). The crystal structure of the amino-terminal domain of SMG7 revealed a 14-3-3-like domain (also found in SMG5 and SMG6), which mainly binds to the protein motifs containing phosphorylated threonine or serine (Rittinger et al., 1999; Fukuhara et al., 2005). Mutations in the 14-3-3 domain of SMG7 impair its interaction with UPF1 *in vitro* and recruitment of UPF1 to processing bodies (P bodies are cytoplasmic sites which accumulate enzymes involved in mRNA turnover and mRNAs storage (Sheth and Parker, 2003, 2006)) (Fukuhara et al., 2005). SMG5 and SMG6 contain a PilT N-terminus domain (PIN domain) at the carboxyl terminal (Glavan et al., 2006). PIN domains often exhibit nuclease activity, and they are found in many proteins, in bacteria and eukaryotes (Wall and Kaiser, 1999; Clissold and Ponting, 2000). The crystal structure of the PIN domains of SMG5 and SMG6 revealed that both have a similar fold structure but have different amino acids in the active sites – SMG5 lacks two key residues, which are

believed to be essential for catalysis (Glavan et al., 2006). In agreement with the prediction that only the PIN domain of SMG6 has ribonuclease activity, recent studies in *D. melanogaster* and human cells have shown that SMG6 functions as an endonuclease via its carboxyl terminal PIN domain, and that NMD is probably started by the endonucleolytic cleavage of mRNA in both *Drosophila* S2 cells and human cells (Huntzinger et al., 2008; Eberle et al., 2009). Two additional proteins, SMG-8 and SMG-9, have also been shown to bind SMG1 (Yamashita et al., 2009). Both of them suppress SMG1 kinase activity and are components of the NMD inducing complex- a complex containing the NMD factors, SMG1, UPF1, and the release factor eRF1 and eRF3 (SURF) (Kashima et al., 2006; Yamashita et al., 2009). SMG-8 also seems to play an important role in the interaction between the complex and EJC (Yamashita et al., 2009).

There are no apparent SMG1 and SMG5-7 orthologs in *S. cerevisiae*, therefore it was expected that UPF1 should not be phosphorylated in this organism. However, it was recently reported that both UPF1 and UPF2 are phosphorylated in budding yeast (Wang et al., 2006). Surprisingly, the phosphorylation of UPF2 but not UPF1 plays an important role in NMD, probably by modulating the interaction UPF2 with Hrp1p (Wang et al., 2006). However, further studies will need to clarify whether the phosphorylation and dephosphorylation cycle of UPF1 is important to NMD in budding yeast as it is in human cells.

1.2.3 Translation termination mechanism

In eukaryotes, translation termination requires two release factors, eRF1 and eRF3. eRF1 recognizes all three termination codons in the A site of the stalling ribosomes and interacts with eRF3 (the carboxyl terminal domain) to form a complex. This

complex then promotes translation termination and releases the polypeptide chain from the ribosome; this reaction requires activation of the GTPase function of eRF3 when the eRF1/3 complex associates with ribosome (Frolova et al., 1994; Zhouravleva et al., 1995; Frolova et al., 1996). Through its amino terminal domain, eRF3 interacts directly with the carboxyl terminus of cytoplasm poly(A) binding protein (PABPC, Pab1p in *S. cerevisiae* and PABPC1 in human) in both human cells and *S. cerevisiae* (Hoshino et al., 1999; Kozlov et al., 2001; Hosoda et al., 2003).

PABPC is a multiple functional protein and has been implicated in several processes in *S. cerevisiae*: polyadenylation, mRNA export, stability and translation (Zhao et al., 1999; Brune et al., 2005). PABPC contains four highly conserved RRM-like domains (RBD) at the amino terminus and a helical domain at the carboxyl terminus (containing 4 helices in *S. cerevisiae* but 5 in human), and two domains are connected by a unstructured proline-rich linker (Mangus et al., 2003).

The studies in *S. cerevisiae* and human cells revealed that UPF1 directly interacts with the eRF1/3 complex, suggesting that the association of UPF1 might cause aberrant translation termination (Weng et al., 1996a, b; Czaplinski et al., 1998; Weng et al., 1998). It is still unknown how UPF1 is recruited to premature termination sites and whether UPF1 is also recruited to normal termination sites. A recent study reported that efficient translation termination requires interaction between eRF3 and PABPC, and that UPF1 can inhibit this interaction (Ivanov et al., 2008). These and other observations indicate that translation termination is intrinsically different at PTC (see below).

1.2.4 PTC recognition and NMD mechanisms

1.2.4.1 *cis-acting* elements

A key question in the NMD field is how premature translation termination is distinguished from normal termination. In *S. cerevisiae*, it was proposed that the presence of a downstream sequence element (DSE) distinguishes the NMD-inducing PTC from a normal stop codon (Peltz et al., 1993; Zhang et al., 1995). The consensus sequence motif for the DSE was proposed to be TGYYGATGYYYYY (Y stands for either T or C). Two copies of this DSE motif were identified in the *GCN4*, *PGK1*, *ADE3* and *HIS4* genes and were required for strong NMD (Zhang et al., 1995; Ruiz-Echevarria and Peltz, 1996). It has been proposed that the DSE triggers NMD only if located no more than 200 nt downstream of the PTC (Ruiz-Echevarria and Peltz, 1996; Ruiz-Echevarria et al., 1998a).

It has been proposed that the DSE recruits some RNA binding proteins involved in the PTC recognition. In agreement with this suggestion, it was reported that the hnRNP-like RNA binding protein Hrp1p binds the DSE and activates NMD (Gonzalez et al., 2000). Purified Hrp1p binds specifically to the RNA substrates harboring a DSE and interacts with UPF1. Mutations in the *HRP1* gene caused a 5 to 10 fold stabilization of nonsense containing mRNAs without affecting the stability of wide type mRNAs, and one of these mutations abolished both its binding to the DSE and the interaction with UPF1 (Gonzalez et al., 2000). As referred above, Hrp1p also interacts with phosporylated UPF2 (Wang et al., 2006). These results suggest that the DSE is critical for eliciting NMD because it recruits specific RNA binding proteins that mediates the recruitment of NMD factors. Hrp1p is also involved in mRNA 3' end processing in yeast (Kessler et al., 1997; Gross and Moore, 2001). Contrary to

earlier studies, the structure of the Hrp1p-RNA complex revealed that Hrp1p recognizes a six base motif, AUAUAU, which is very different from the DSE consensus sequence (Perez-Canadillas, 2006).

1.2.4.2 The *faux* 3' UTR model

Since the very first study of NMD in eukaryotes, it was apparent that not all nonsense mutations reduce the mRNA level: NMD is most readily apparent in the case of 5' proximal mutations, but not with mutations located further downstream. In the first study of NMD, the authors characterized a number of nonsense mutations in the *URA3* gene of *S. cerevisiae* and found that amber mutations (UAG) in the first half of *URA3* gene caused a drastic reduction of the *URA3* mRNA, while downstream mutations only slightly affected the *URA3* mRNA level (Losson and Lacroute, 1979). This polarity phenomenon has been confirmed in the *PGK1* and *CYC1* genes and the general conclusion is that nonsense mutations in the 5' half cause strong and rapid mRNA, whilst mutations in the second half of the gene behave like a normal stop codon (Losson and Lacroute, 1979; Peltz et al., 1993; Yun and Sherman, 1995).

As referred above, the polarity effect was initially explained with the DSE model: the more 3' located is the PTC, the less likely is to find a DSE. However, it was reported that mRNAs with extended 3' UTRs are also NMD substrates in *S. cerevisiae* and *C. elegans* (Pulak and Anderson, 1993; Muhlrad and Parker, 1999). This discovery suggested that the important distinction between normal and premature terminations might simply rely on the distance between the stop codon and the 3' end of mRNA. More recent studies indicate that premature translation termination is intrinsically aberrant (Amrani et al., 2004). According to this model – the *faux* 3' UTR model - efficient termination requires an interaction with factors associated with the 3' UTR.

The initial *faux* 3' UTR model proposed that inefficient termination is due to lack of interaction between the terminating ribosome and PABPC (Pab1p in *S. cerevisiae*) (Amrani et al., 2004). This model is mainly based on the observations that the ribosome leaves a very different "toe-print" near PTCs than that left near normal stop codons. In addition, the abnormal ribosome toe-print is not detected in the absence of UPF1 or if the PTC was flanked by a normal 3' UTR. The artificial tethering of PABPC at a downstream position close to the PTC also strongly inhibited NMD. A recent study reported that *GFP* mRNA, whose 3' UTR was removed by a self-cleaving hammerhead ribozyme, was susceptible to NMD only if an artificial poly(A) track (75 nt) is added at its 3' end (Baker and Parker, 2006). This study therefore further indicated that the distinction of a PTC from a normal stop codon simply depends on the distance from the poly(A) tail. The *faux* 3' UTR model thus provides a simple explanation for the polarity effect on nonsense mutations in *S. cerevisiae* (Fig. 1A).

In higher organisms, this model could explain the NMD polarity observed in the alcohol dehydrogenase (*Adh*) gene in *D. melanogaster*, chloramphenicol acetyltransferase (*Cat*) gene in *Drosophila* S2 cells, and immunoglobulin μ gene (Ig-μ) in human cells (Brogna, 1999; Gatfield and Izaurralde, 2004; Buhler et al., 2006; Behm-Ansmant et al., 2007). In agreement with the *faux* 3' UTR model, extension or shortening of the distance between PTCs and the poly(A) tail affected NMD in several systems (Buhler et al., 2006; Behm-Ansmant et al., 2007; Eberle et al., 2008). Moreover, tethering of PABPC at a downstream site close to the PTC suppressed NMD; the mRNA lacking the poly(A) tail escapes NMD and depletion of PABPC strongly inhibits NMD in *Drosophila* S2 cells, suggesting that the poly(A) tail and PABPC are crucial in NMD (Behm-Ansmant et al., 2007).

1.2.4.3 Splicing-dependent NMD - The EJC model in pioneer round translation

Unlike in *S. cerevisiae*, in many instances NMD is splicing dependent in mammalian cells. Early studies of NMD of the β-glogin mRNA and triosephosphate isomerase (TPI) mRNA in human cells indicated that NMD might target nucleus-associated mRNAs, and reported that there was no change in the stability of cytoplasmic mRNAs with or without nonsense mutations (Baserga and Benz, 1992; Cheng and Maquat, 1993). The further studies with human TPI mRNA also reported that nucleus associated NMD took place after splicing, (Belgrader et al., 1994; Cheng et al., 1994). In another study, using the T-cell receptor beta (TCR-β) reporter genes, it was discovered that a PTC triggered NMD only if it was upstream of a functional and spliced intron (Carter et al., 1996). Similar observations were reported in the human TPI gene: nonsense mutations located less than 50 to 55 nt upstream of 3'-most intron were not subjected to NMD (Zhang and Maquat, 1996; Zhang et al., 1998a). It was concluded that splicing of the 3' most intron marks to the exon junction position.

This hypothesis was substantiated by the discovery that pre-mRNA splicing deposits multiple proteins about 20-24 nt upstream of the exon-exon junction site (Le Hir et al., 2000a). This set of proteins is the signature of exon exon junctions and forms a well-defined complex called the exon junction complex (EJC) (Le Hir et al., 2000b; Le Hir et al., 2001b). *In vitro* assembled EJC contains five components, SRm160, DEK, RNPS1, Y14, and REF (Le Hir et al., 2000a). To date, more than a dozen proteins has been identified as components of the EJC, including splicing co-activators and alternative splicing factors (SRm160, RNPS1 and Pinin), the mRNA export proteins (UAP56, REF/Aly and NXF1/TAP: p15), mRNA localization factors (Y14, Magoh, MLN51 (known as Barentsz, BTZ, in *Drosophila*) and eIFIII4A) (Kataoka et al., 2000; Le Hir et al., 2000a; Lykke-Andersen et al., 2000; Zhou et al.,

2000; Kataoka et al., 2001; Kim et al., 2001a; Kim et al., 2001b; Le Hir et al., 2001a;

Le Hir et al., 2001b; Luo et al., 2001; Lykke-Andersen et al., 2001; Chan et al., 2004;

Degot et al., 2004; Ferraiuolo et al., 2004; Palacios et al., 2004; Shibuya et al., 2004).

The EJC provides a direct link between splicing and NMD by serving as the anchor

site for UPF2 and UPF3 (Le Hir et al., 2001b). A conserved domain in UPF3A and

UPF3B associate with Y14 and RNPS1, and this interaction is essential for NMD:

NMD was strongly suppressed in the cells where the specific functional domains of

Y14 and RNPS1 were deleted or where the proteins were depleted by RNAi (Lykke-

Andersen et al., 2001; Gehring et al., 2003; Singh et al., 2007). RNAi depletion of

MAGO, MLN51 and eIF4AIII in mammalian cells also leads to defects in NMD

(Ferraiuolo et al., 2004; Palacios et al., 2004; Shibuya et al., 2004). Recently, two

functionally distinguishable EJC subgroups were defined, RNPS1-type EJC, which

requires normal level of UPF2 to trigger efficient NMD, and Y14-Magoh-eIF4AIII-

BTZ type, which activates NMD in the cells with a low UPF2 level or in UPF2-

depleted cells (Gehring et al., 2005). All of these reports indicated that the EJC is

essential for NMD in mammalian cells (Fig. 1B). These recent studies also indicated

that the association between EJC and UPF1 activates the ATPase-dependent helicase

activity of UPF1 and induces the formation of the SURF complex (Kashima et al.,

2006; Chamieh et al., 2008).

Some studies in human cells indicate that NMD occurs while the mRNA is still

associated with the nuclear cap binding complex (CBC, formed by CBP80 and

CBP20 - Cbc1p and Cbc2p in *S. cerevisiae*), but not with eIF4E (cytoplasmic cap
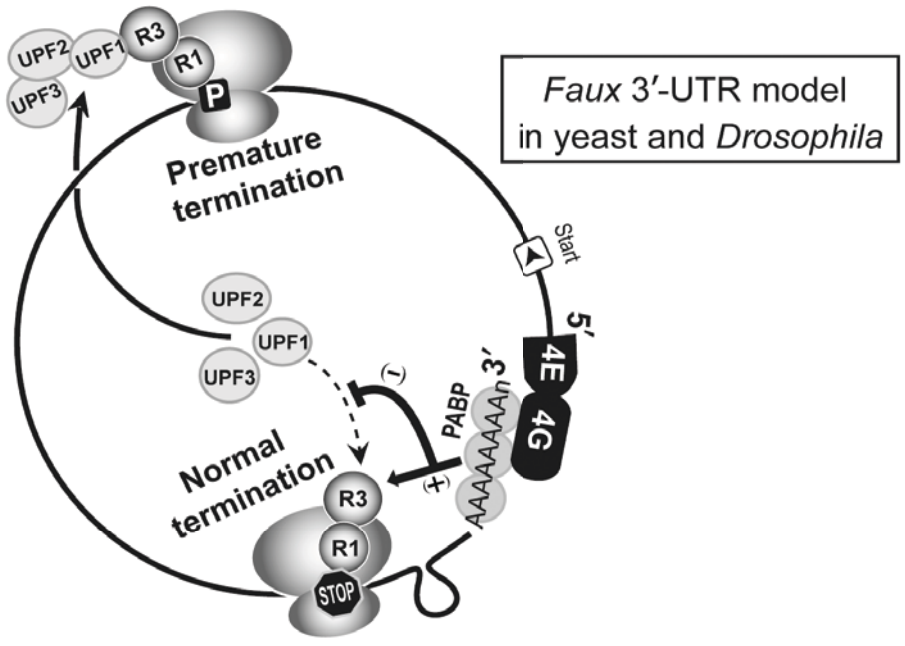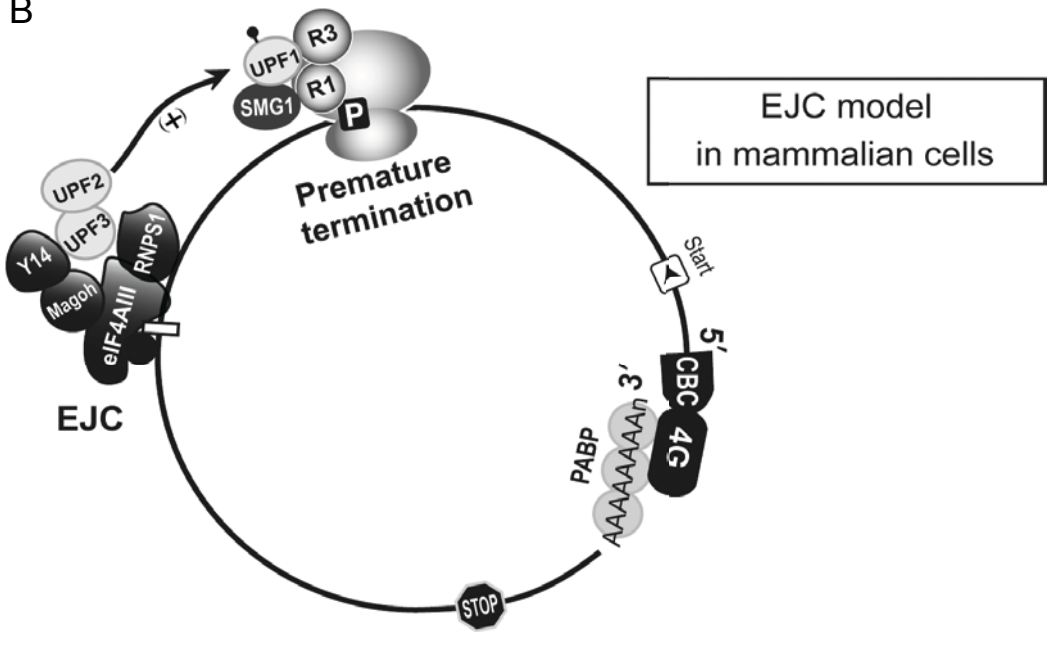
Figure 1.1 The *faux* 3' UTR and the EJC models of NMD

(A) *Faux* 3' UTR model in *S. cerevisiae* and *D. melanogaster.* The ribosome is stalled at termination codon. If the termination is far away from the 3' end, poly(A) tail PABPC (Pab1p in yeast) is unable to interact with the terminating ribosome and termination is deemed aberrant and triggers a strong mRNA decay (NMD). Termination is normal when the stop codon is close to the poly(A) tail because PABPC can interact with the terminating ribosome.

(B) EJC model in mammalian cells. NMD mainly occurs in the pioneer round of translation. If the PTC is upstream of an intron, the EJC accelerates the recruitment of UPF proteins. The association of UPF proteins causes triggers NMD. The main proteins and complexes in NMD are shown: UPF proteins, translation release factor 1 and 3 (R1 and R3), ribosomes, EJC complex, cap binding complex (CBC), translation initiation factor 4G and 4E (4G and 4E), cytoplasmic poly(A) binding protein (PABP).

binding protein) bound transcripts (Ishigaki et al., 2001). The EJC appears to associate with CBP80 bound mRNA but not with eIF4E bound mRNA (Lejeune et al., 2002). These studies therefore indicate that CBP80 is required for NMD in mammalian cells. RNAi depletion of CBP80 stabilizes NMD substrates (Hosoda et al., 2005). In the same study, there is also evidence that CBP80 directly associates with UPF1, and as I referred above, that UPF1 is activated by an association with the EJC.

1.2.5 Problems with current NMD models

EJC model is not generally applicable in mammalian cells

Many studies in mammalian cells support the EJC model of PTC recognition, but several observations contradict its predictions. Firstly, there is a report that PTCs located less than 50 nt from the last exon-exon junction can also cause strong NMD in TCR-β mRNA (Wang et al., 2002a). In addition, there is a report that strong NMD can also occur without downstream introns in the β-globin mRNA, but instead needs an upstream intron (Neu-Yilik et al., 2001; Matsuda et al., 2007). NMD without the requirement for a downstream intron has also been observed in the immunoglobulin-μ (Ig-μ) mini-gene system (Buhler et al., 2006). All of these observations suggest that there is also an EJC-independent NMD mechanism in mammalian cells. In addition, the finding that NMD does not require introns or EJC proteins in *Drosophila* S2 cells or *C. elegans* questioned the applicability of the EJC model in other organisms. As mentioned above (1.2.4.2), the *faux* 3' UTR model explains NMD better than the EJC model in S2 cells. Furthermore, a number of recent studies have questioned the importance of downstream introns and the EJC in NMD in mammalian systems which were previously believed to require a downstream intron (Buhler et al., 2006; Eberle et al., 2008; Singh et al., 2008). These studies have shown that shortening or

lengthening of the distance between the PTC and the 3' UTR respectively suppressed or enhanced NMD (Buhler et al., 2006; Eberle et al., 2008; Singh et al., 2008). It was also shown that extending the distance between the normal stop codon and the 3' end caused an mRNA reduction that required the presence of UPF proteins, suggesting that the normal stop codon behaved like a PTC when placed further away from the 3' end (Singh et al., 2008). NMD could be also suppressed by folding the poly(A) tail back to the vicinity of the PTC, or tethering PABPC1 near to the PTC (Eberle et al., 2008). In summary, the emerging view is that the *faux* 3' UTR model might explain NMD better than the EJC model even in mammalian cells (Fig. 2).

Inconsistencies of the faux 3' UTR model

As mentioned above, many recent studies proposed that the *faux* 3' UTR model explains NMD better than the EJC model. However, I will argue below that not all available data are consistent with the *faux* 3' UTR model. The *faux* 3' UTR model provides a good explanation for the 'polarity effect'. However, the polarity effect is not gradual or linear with the distance from a PTC to the 3' end in *S. cerevisiae, D. melanogaster* and mammalian cells (Yun and Sherman, 1995; Behm-Ansmant et al., 2007; Eberle et al., 2008). In addition, the polarity effect is not typical in mammalian NMD: most of PTCs at reasonable positions cause mRNA reduction to a similar degree.

The *faux* 3' UTR model also does not take into account that mRNA is probably translated in closed-loop conformation (Tarun et al., 1997; Wells et al., 1998). In a circular mRNP, the 3'UTR is in fact close to both ends of the mRNA; therefore 5' proximal NMD-inducing PTCs may be as close to the 3' UTR as NMD-insensitive PTCs in the 3' half of the coding region. Several observations support the view that

the mRNA is kept in a circular conformation during translation. The conformation is stabilized by the interaction between the 5' end and the 3' end, which is mediated by translation initiation factor eIF4E, eIF4G and PABP (Tarun and Sachs, 1996; Wells et al., 1998; Sachs and Varani, 2000; Amrani et al., 2008). This interaction between eIF4G and PABP is evolutionary conserved and the double rows structure of the ribosome resembling closed hairpins has been observed by electron microscopy (Christensen et al., 1987; Le et al., 1997; Imataka et al., 1998; Christensen and Bourne, 1999). The observation that in the human β-globin transcript, PTCs at codon 5 and 17 do not trigger strong NMD whereas PTCs at codon 26 and 39 do, was taken as a validation of the *faux* 3' UTR model prediction that 5' proximal PTCs should not induce NMD (Romao et al., 2000). A similar phenomenon was also observed when the PTC was located at codon 32 in the β-globin mRNA (Eberle et al., 2008). However, we have previously argued that PTCs very close to the AUG do not induce NMD because translation is likely to reinitiate downstream of the PTC (Brogna and Wen, 2009). In addition, there are many examples of PTCs very close to the AUG which induce strong NMD; for example, PTCs at codon position 6 or 7 in *S. cerevisiae* (GFP reporter and *CYC1* mRNA) and codon 23 in human TPI mRNA still cause strong NMD (Belgrader et al., 1994; Yun and Sherman, 1995; Kuperwasser et al., 2004).

In support of the *faux* 3' UTR model, many studies have reported that the tethering of PABP nearby the PTC suppresses NMD (Amrani et al., 2004; Behm-Ansmant et al., 2007; Eberle et al., 2008; Singh et al., 2008). However, we have argued that the tethering of PABPC might stabilize the NMD substrates simply because it prevents
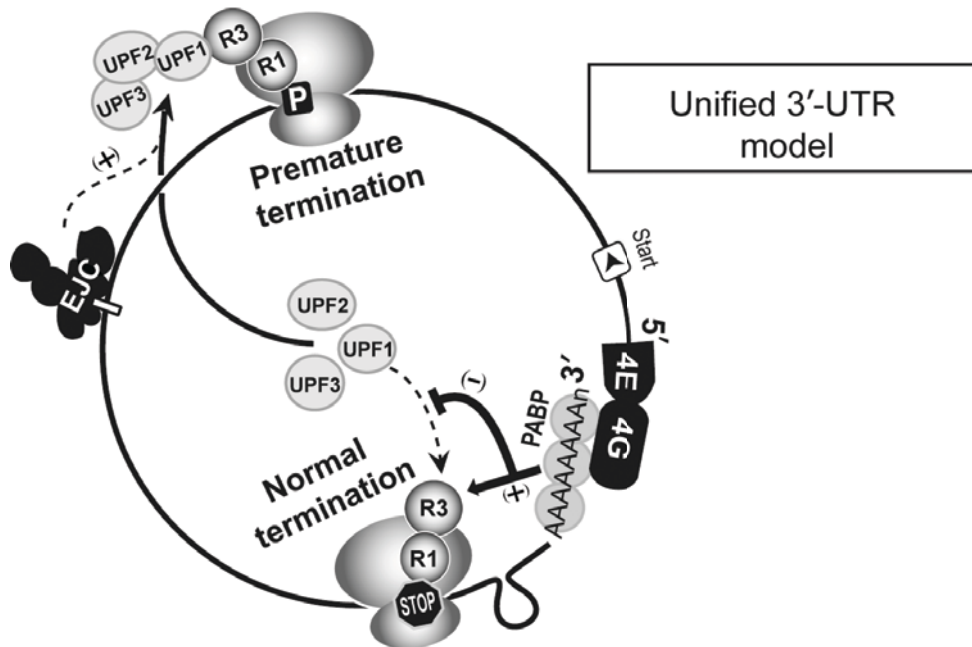
Figure 1.2 The unified 3' UTR model for NMD.

NMD is mainly determined by the flanking 3 UTR. PABP plays an antagonistic role in NMD. NMD is elicited if the PTC is far way from the poly(A) tail. When termination occurs near the poly(A) tail, PABP prevents the recruitment of UPF1, preventing NMD. The presence of a downstream intron deposits an EJC which can further stimulate the recruitment of UPF proteins and NMD.

standard mRNA decay rather than inhibits the NMD processes (Brogna and Wen, 2009). The *faux* 3' UTR model predicts a key role of PABPC in NMD. However, contrary to this prediction, it was reported that NMD is not affected by deletion of PABPC's carboxyl terminal region -which is thought to mediate the interaction with terminating ribosomes to trigger the deadenylation of mRNAs (Simon and Seraphin, 2007). The *faux* 3' UTR model was further questioned by a recent study in *S. cerevisiae* (Meaux et al., 2008). This study reported that PTCs elicit strong NMD even in poly(A)-less transcripts with ribozyme-generated 3' ends, or in the cells with the deletion of *pab1,* the gene encoding PABPC in *S. cerevisiae* . This study suggested that neither the poly(A) tail nor PABPC is essential for NMD in *S. cerevisiae*.

In *Arabidopsis thaliana* both intron and intronless genes are subjected to NMD (Kertesz et al., 2006). In this study, it was reported that either the insertion of long DNA fragments in the 3' UTR or introducing an intron into the 3' UTR both trigger strong NMD.

1.3 Cellular mRNA decay pathways and NMD

There are many observations indicating that once it has been recognized as an NMD substrate, the mRNA is degraded by the canonical mRNA decay pathway that destroys "normal" mRNAs (He and Jacobson, 2001; Chen and Shyu, 2003; Lejeune et al., 2003; Sheth and Parker, 2003). Here, I present a brief review of mRNA decay in eukaryotes and the fate of NMD substrates following PTC recognition.

Eukaryotic mRNAs have two internal stability elements, the 7–methylguanosine cap at the 5' end and poly(A) tail at the 3' end. Deadenylation typically precedes the breakdown of the mRNA. In *S. cerevisiae* deadenylation is catalyzed by

Table 1.2 Proteins involved in human NMD and the homologs in *S. cerevisiae*

Protein	Localization (human)	Characteristics/Functions	Homolog in *S. c*
UPF1	Shuttle protein; mainly in cytoplasm	ATPase, ATP dependent helicase; SMD; Histone mRNA decay; telomere maintenance; DNA repair (Weng et al., 1996a)	Upf1p
UPF2	Shuttle; perinuclear and cytoplasm	Telomere maintenance (Dahlseid et al., 2003; Azzalin et al., 2007;)	Upf2p
UPF3A	Shuttle; mainly in nucleus	Weak NMD (Kunz et al., 2006)	Upf3p
UPF3B	Shuttle; mainly in nucleus	EJC component (Le Hir et al., 2001b; Gehring et al., 2003)	Upf3p
SMG1	Cytoplasm	PPI 3-kinase; maintenance of genome stability, stress response and DNA repair (Yamashita et al., 2001; Brumbaugh et al., 2004)	No
SMG5	Shuttle; mainly in cytoplasm	PP2A interaction and dephosphorylation of UPF1 (Chiu et al., 2003)	No
SMG6	Shuttle; mainly in cytoplasm	Endonuclease; endonucleolytic cleavage (Huntzinger et al., 2008; Eberle et al., 2009)	No
SMG7	Shuttle; mainly in cytoplasm	PP2A interaction and dephosphorylation of UPF1 (Chiu et al., 2003)	Est1p *
NAG (SMGL-1)	Unclear	Function currently unclear (Longman et al., 2007)	No
DHX34 (SMGL-2)	Unclear	DEAD box protein 34, helicase (Longman et al., 2007)	No
SMG8	Unclear	Subunit of SMG1 complex (Yamashita et al., 2009)	No
SMG9	Unclear	Subunit of SMG1 complex (Yamashita et al., 2009)	No
eIF4AIII	Shuttle; mainly in nucleus	EJC scaffold protein (Palacios et al., 2004; Gehring et al., 2005)	Fal1p
MAGO	Shuttle; mainly in nucleus	EJC core component, MAGO/Y14 heterodimer (Le Hir et al., 2001a; Gehring et al., 2005)	No
Y14	Shuttle; mainly in nucleus	EJC core component, MAGO/Y14 heterodimer (Le Hir et al., 2001a; Gehring et al., 2005)	No

MLN51	Shuttle; mainly in cytoplasm	EJC core component (Palacios et al., 2004; Gehring et al., 2005)	No
RNPS1	Shuttle; mainly in nucleus	EJC component; splicing coactivator (Le Hir et al., 2000a; Lykke-Andersen et al., 2001; Gehring et al., 2005)	No
PYM	Shuttle; mainly in cytoplasm	EJC component; interacts with MAGO/Y14 (Bono et al., 2004)	No
eIF3b	Mainly in cytoplasm	Subunit of translation initiation factor 3 (Welch and Jacobson, 1999; Martineau et al., 2008)	Prt1p
INT6	Mainly in nucleus	Subunit of translation initiation factor 3 (Morris et al., 2007)	No
PNRC2	Cytoplasm	Energy balance control; estrogen synthesis; UPF1 binding; DCP1 binding (Cho et al., 2009)	No

* Contradictory reports of its function in NMD

A

P bodies

Dcp2   Xrn1

Dcp1

LSm 1-7

eRF1

eRF3

Pab1

Upf1

Upf3   Upf2

Ski complex

exosome

Hrp1

S. cerevisiae

B

eRF3   eRF1

PABPC

SMG5

PP2A

SMG7

SMG6

SMG8   SMG1

SMG9

UPF1

UPF3   UPF2

P bodies

Dcp2   Xrn1

Dcp1

PNRC2

LSm 1-7

Ski complex

exosome

Y14   eIF4A3

RNPS1

PYM   MAGO   MLN51

EJC
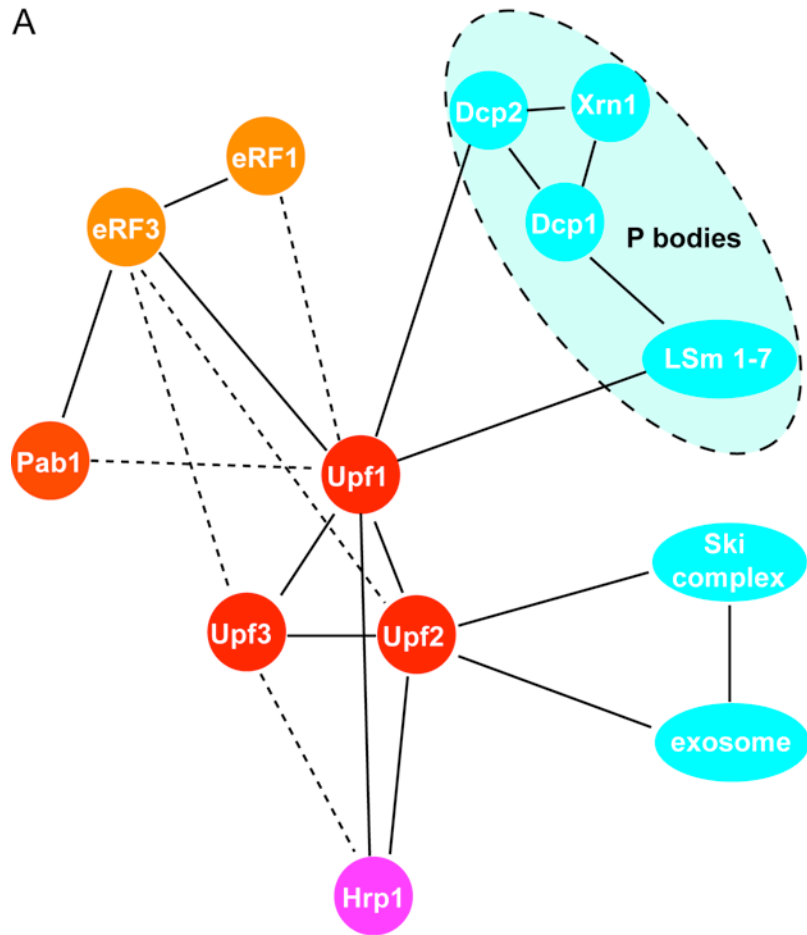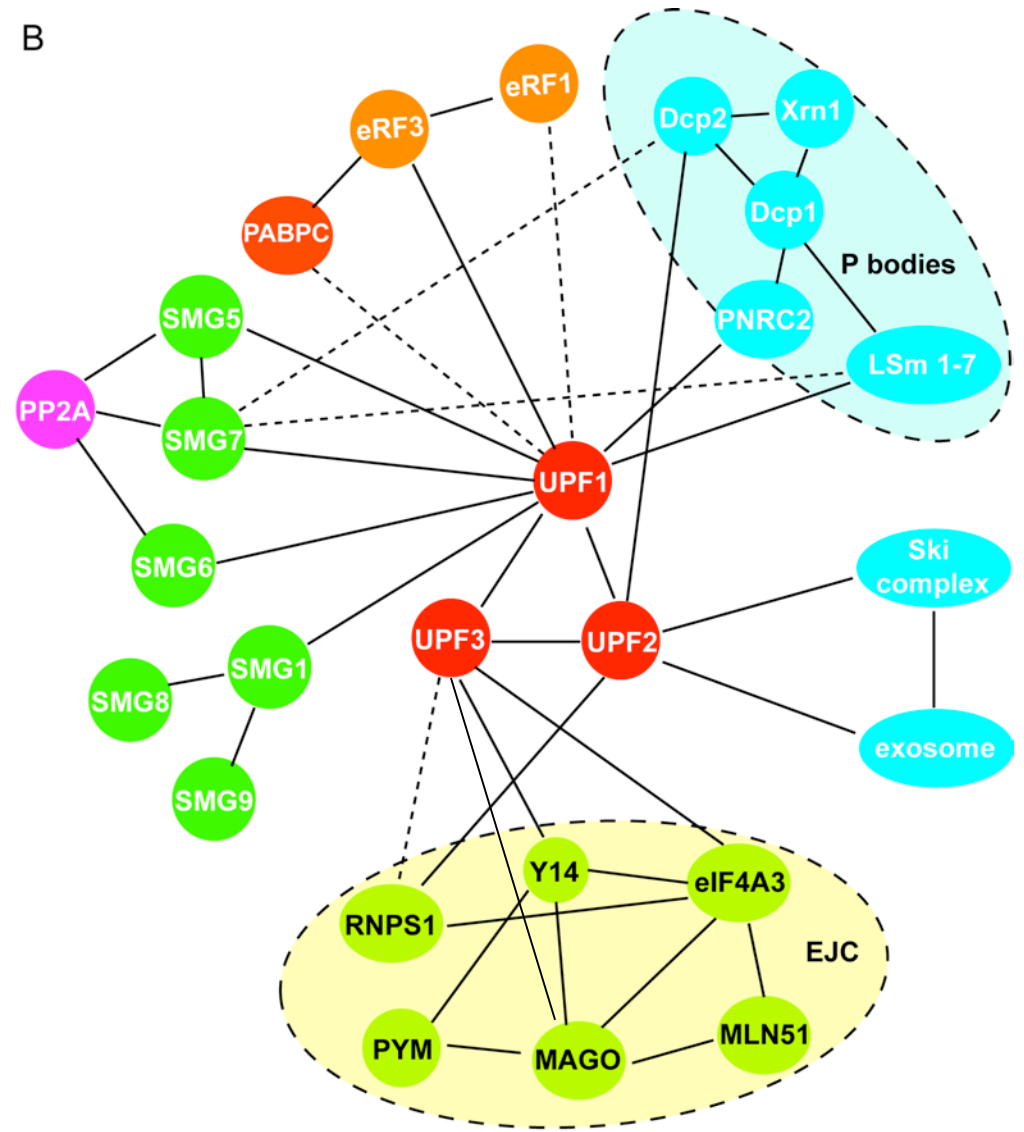
Mammals

Figure 1.3 The inter-connection between NMD factors and other proteins involved in global mRNA decay.

(A and B) Protein interaction network in *S. cerevisiae* and human cells. The specific proteins which are involved in NMD in *S. cerevisiae* and human are shown in purple (Hrp1p and PP2A) and green (SMG proteins and EJC components), which are summarized in Table 1.2. The proteins involved in normal mRNA decay process are shown in turquoise and described in Section 1.3. The EJC and P body are indicated by the ovals with dash lines. The drawing is according to Figure 5 in Conti and Izaurralde, (2005).

PAN2-PAN3, a PABP-dependent poly(A) nuclease, which trims the mRNA's poly

(A) tail to about 60-80 nt length, and then CCR4-NOT, the main deadenylase, shorten

the poly(A) tail further to 10-12 bases (Brown and Sachs, 1998; Funakoshi et al.,

2007). In higher eukaryotes, such as *X. laevis* and mammalian cells, PARN (poly(A)

ribonuclease) is the main poly(A) nuclease of 5' capped mRNA (Garneau et al.,

2007). Following the shortening of the poly(A) tail, mRNAs are attacked by the

cytoplasmic exosome, a 3'-5' exonucleolytic complex composed of 10-12 subunits

and associated with the Ski complex (consist of Ski2, Ski3, Ski7 and Ski8) (Garneau

et al., 2007; Ibrahim et al., 2008; Schmid and Jensen, 2008). At the same time, the 5'

ends are decapped by Dcp1 and Dcp2 enzymes. Decapped mRNAs are subjected to

5'-3' degradation by the exoribonuclease XRN1 (Anderson and Parker, 1998; Decker

and Parker, 2002; Lejeune et al., 2003; Mangus et al., 2003). The rapid decapping of

PTC-containing mRNAs is probably mediated by the interaction of UPF1 with

Dcp1:Dcp2 (Lykke-Andersen, 2002; Wagner and Lykke-Andersen, 2002). Staufen

mediated mRNA decay (SMD) and stem loop mediated histone mRNA decay might

cause rapid decay also as a consequence of recruiting UPF1 (Kaygun and Marzluff,

2005; Kim et al., 2005). The importance of decapping in NMD was further

emphasized by the recent report that human proline-rich nuclear receptor coregulatory

protein 2 (PNRC2) enhances NMD by stimulating the recruitment of DCP enzymes to

UPF1 and the NMD machinery (Cho et al., 2009).

As mentioned above, it has been reported that mRNA substrates are degraded by the

standard 5'-3' and 3'-5' enzymes that destroy the bulk of mRNA. In *S. cerevisiae*,

XRN1 was believed as the main exonuclease involved in NMD (He and Jacobson,

2001; He et al., 2003). However, more recent studies have reported that in *D.*
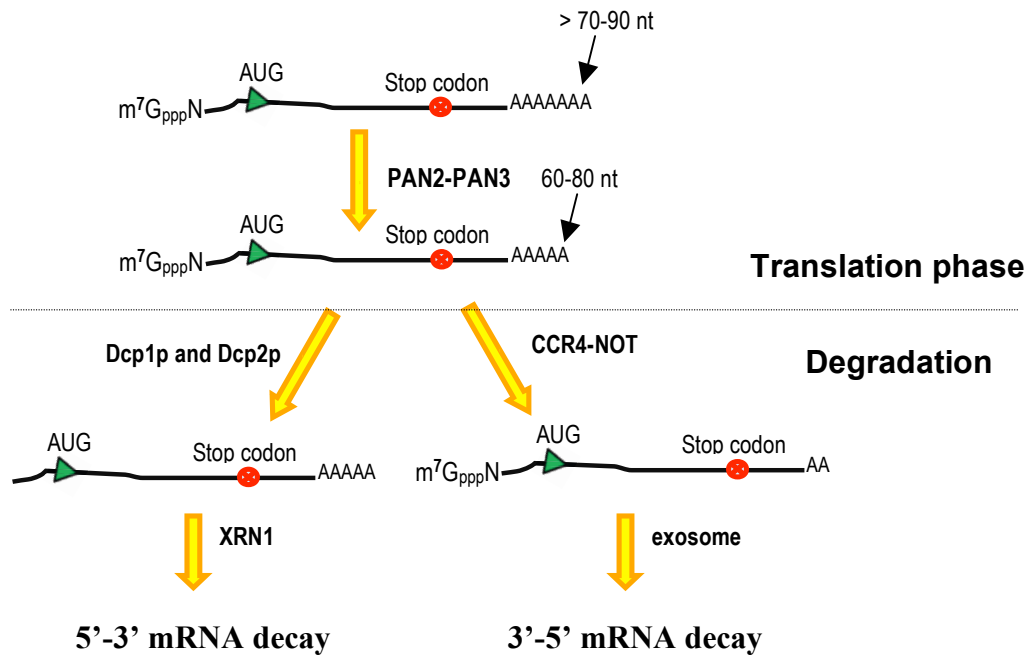
Figure 1.4 Global mRNA decay in *S. cerevisiae.*

The mRNA is initially deadenylated by PAN2-PAN3 during translation, and then is either further deadenylated to an oligo tail by CCR4 –NOT and degraded by the exosome, or decapped by Dcp1 and Dcp2 and degraded from 5' to 3' by XRN1.
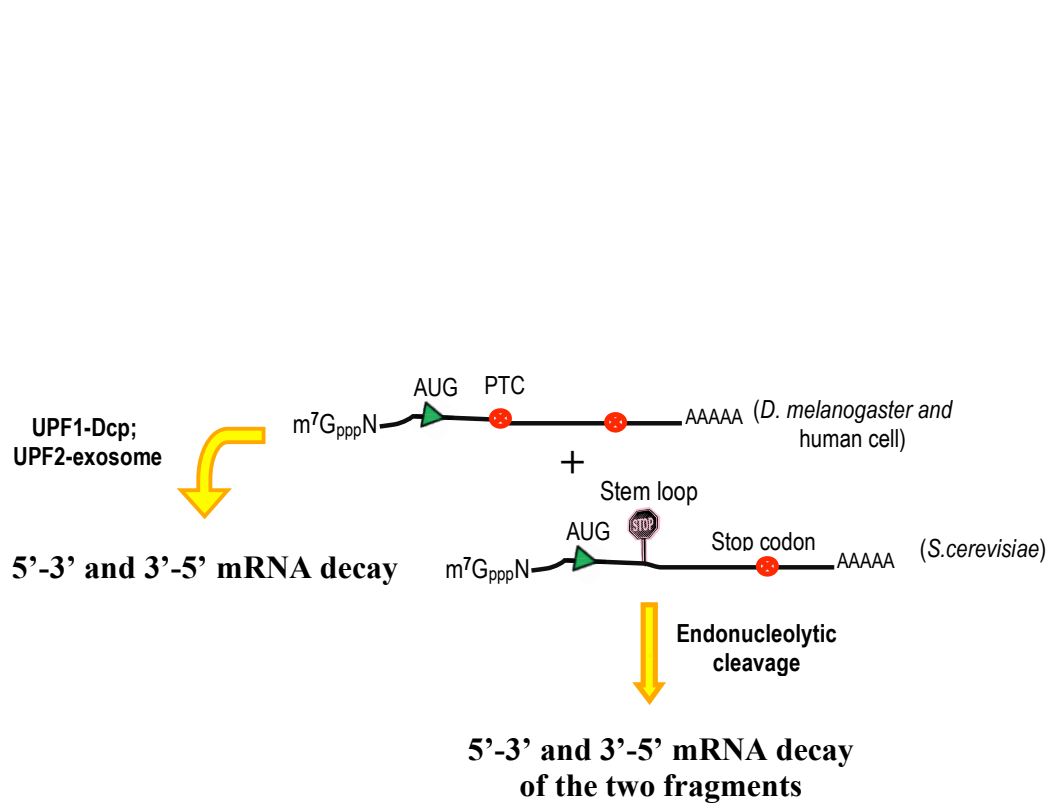
Figure 1.5 Aberrant mRNA decay and endonucleolytic cleavage

Two different mRNA decay processes degrade aberrantly translated mRNAs. In one pathway (left diagram), the mRNA is rapidly decapped and deadenylated, and then degraded by exonucleases from both ends. In the other pathway (right diagram), the degradation is initiated by the endonucleolytic cleavage followed by 5' to 3' and 3' to 5' decay of the two RNA fragments.

melanogaster and human cells, NMD substrates appear to be degraded by an alternative pathway, involving an endonucleolytic cleavage (Gatfield and Izaurralde, 2004; Behm-Ansmant et al., 2007; Huntzinger et al., 2008; Eberle et al., 2009). This endonucleolytic cleavage occurs at the multiple sites near the PTC, surprisingly both before and after the stop codon. In both *D. melanogaster* and human cells, this cleavage depends on a functional SMG6 protein (Huntzinger et al., 2008; Eberle et al., 2009). Evidence of the involvement of the endonucleolytic cleavage was reported also in the no-go mRNA decay (NGD) in *S. cerevisiae:* a stem loop that prevents translation elongation, also triggers mRNA decay (Doma and Parker, 2006). These observations suggest a general link between the ribosome stalling and endonucleolytic cleavage. In summary, NMD substrates are likely to be degraded by a number of pathways, requiring both endonucleases and exonucleases.

1.4 Does NMD occur in the nucleus?

1.4.1 Nucleus associated NMD in human cells

In an early study of NMD in mammalian cells it was observed that the β-globin mRNA was drastically reduced in beta-thalassemias caused by nonsense mutations (Chang and Kan, 1979). Surprisingly, it was also observed that mRNA level was also reduced in the nuclear fraction (Chang and Kan, 1979). This finding led to the speculation that NMD may be occurring while the mRNA is still associated with the nucleus. The reduction of the nuclear-associated mRNAs has been observed in transcripts coding for beta-globin (Baserga and Benz, 1992), dihydrofolate reductase (Urlaub et al., 1989) and triosephosphate isomerase (Belgrader et al., 1994; Cheng et al., 1994; Zhang and Maquat, 1996). It appears that the level of nucleus-associated

mRNA was always reduced to almost the same amount as the cytoplasmic mRNA with premature terminal codon (PTC) in human cells (Maquat, 1995). Furthermore, several studies reported that there is no obvious difference in stability between mRNA carrying PTCs and the corresponding wild-type mRNA, in spite of the low steady-state level of the PTC-containing mRNAs (Baserga and Benz, 1992; Cheng and Maquat, 1993). These observations were the first suggestion that PTC-containing transcripts are degraded in the pioneer round translation while they are still associated with the nucleus(Maquat, 1995; Hentze and Kulozik, 1999). It was also proposed that mRNA escaping this first round of translation escape NMD as well. Nucleus associated NMD in mammalian cells presents a paradox, because it is believed that only cytoplasmic ribosomes recognize and scan the open reading frames.

1.4.2 Can PTCs be recognized in the nucleus?

Studies in immunoglobulin genes indicated that PTCs could inhibit splicing of the upstream intron in *vivo* and in nuclear extracts (Lozano et al., 1994; Aoufouchi et al., 1996). Meanwhile, studies of the fibrillin (FBN1) RNA and the RNAs of parvovirus minute virus of mice suggested that the selection of splice sites could be influenced by the translating reading frames because nonsense mutations led to exon skipping but missense mutations did not (Dietz et al., 1993; Dietz and Kendzior, 1994; Gersappe and Pintel, 1999). However, the recent studies proposed a more reasonable explanation: that the typical pre-mRNA contains exonic splicing enhancers (ESE) and the effect of nonsense mutations in exon skipping may simply cause the dysfunctional ESEs in splice sites choice (Valentine, 1998; Cartegni and Krainer, 2002). Indeed, a recent study using FBN1 RNA indicated that nonsense mutations activate exon skipping by the disruption of an ESE, rather than a frame-dependent nuclear scanning mechanism (Caputi et al., 2002).

Nevertheless, additional studies suggest that PTCs can affect pre-mRNA splicing in the nucleus (Muhlemann et al., 2001; Li et al., 2002; Wang et al., 2002b). Furthermore, it was discovered that nonsense mutations also negatively affect pre-mRNA 3' end processing and lead to longer poly (A) tails in both mature mRNA and the unspliced pre-mRNA (Brogna, 1999). Therefore, there are still data that suggest the existence of an ORF scanning mechanism in nucleus.

This view is supported by several studies that have reported that transcription and translation are coupled in *Drosophila* and mammalian cells (Iborra et al., 2001; Brogna et al., 2002). In addition, it was reported that blocking mRNA export does not inhibit NMD of TCR-β in human cell (Buhler et al., 2002).

Other studies, however, have reached opposite conclusions. One study reported that the domain-negative polypeptides of UPF proteins inhibited NMD only when they were overexpressed in the cytoplasm, but not when they were retained in the nucleus by introducing a nuclear localization signal (NLS) (Singh et al., 2007). This later discovery is in agreement with the discovery that block of mRNA export prevents NMD in *S. cerevisiae* (Kuperwasser et al., 2004).

Therefore, there are many inconsistencies in the available data. My view is that only additional studies can resolve the issue of whether there is mechanism that scans the ORF in the nucleus.

1.5 Different phenotypes of NMD mutants in eukaryotes

Deletion of all three *upf* genes in *S. cerevisiae* does not have any obvious phenotype (He et al., 1997). However, mutations in NMD factors cause obvious phenotypes in *A. thaliana*, nematodes, fly, zebra fish and mice. In *A. thaliana*, mutations of *upf1* result

in longer and heavier seeds (Yoine et al., 2006). Moreover, in this study it was found that knocking out *upf1* causes seedling lethality. In *C. elegans*, mutant alleles of *smg1-6* are viable but have morphogenetic defects of genitalia and one double mutant (*smg1;smg3*) grows distinctly slower than either of the parental mutants (Hodgkin et al., 1989). In *D. melanogaster*, mutations in *upf1* and *upf2* cause lethality during larval development (Metzstein and Krasnow, 2006). Interestingly, mutations in NMD factors also enhance the expression of transgenes with no obvious NMD features – this mutations were named 'photoshop' because they were isolated for their ability to enhance the expression of GFP transgenes (Metzstein and Krasnow, 2006). In mice Rent1 (UPF1) is essential in embryonic viability (Medghalchi et al., 2001). In a very recent study in zebrafish, it was shown that mutations or deletions of *upf1*, *upf2, smg1, smg5* or *smg7* have a severe impact on embryonic development, early patterning and viability (Wittkopp et al., 2009).

However, it is not yet clear whether it is the lack of NMD that causes these phenotypes, or whether instead these proteins have some other important functions outside NMD. Some studies indicated that NMD controls the telomere length and telomeric silencing in budding yeast, probably by regulating the mRNA level of telomere-related genes (Lew et al., 1998; Dahlseid et al., 2003; Enomoto et al., 2004). Some NMD factors, such as SMG1, UPF1, UPF2 and SMG6 also seem to be involved in telomere maintenance in mammalian cells (Reichenbach et al., 2003; Azzalin et al., 2007). Furthermore, SMG1 and UPF1 are involved in genome stability and cell cycle regulation (S phase arrest) (Brumbaugh et al., 2004; Kaygun and Marzluff, 2005; Azzalin and Lingner, 2006).

1.6 NMD in *S. pombe*

NMD in *S. pombe* has not been substantially investigated before this study. Two previous studies had shown that UPF1 and UPF2 are required for NMD also in *S. pombe* but did not investigate the key issue of how PTCs are recognized (Mendell et al., 2000; Rodriguez-Gabriel et al., 2006). One of these studies reported that early PTCs trigger stronger NMD than a PTC located later in the *ADE6* coding region. In a microarray study aimed at the identification of mRNAs controlled by UPF1 in *S. pombe,* the authors revealed that there are only 27 genes (about 0.6% in the whole genome) whose mRNA levels were upregulated more than 2 fold in a *upf1Δ* mutant (Rodriguez-Gabriel et al., 2006). Surprisingly, UPF1 appears to affect the expression of very few genes in comparison to *S. cerevisiae*, where more than 580 genes (about 12% of the whole genome) are upregulated at least 2 fold in *upf1Δ* (He et al., 2003). Among these NMD substrates in *S. pombe*, the authors found that half of the genes contain at least one out of frame AUG within the 80 nt upstream of the annotated initiation codon. The reason that the proportions of NMD substrates are so different between two yeasts *S. cerevisiae* and *S. pombe* is still unclear. Possibly, the study in *S. pombe* using cDNA microarray does not show the complete category of NMD sensitive mRNAs or 'junk' RNAs due to the deficiency of pre-mRNA detections by cDNA microarray. The new microarray investigation with the information of intron sequences would be great helpful to further clarify the native NMD substrates in *S. pombe*. In this study it was discovered that both UPF1 and UPF2 deletion leads to increased sensitivity to oxidative stress (Rodriguez-Gabriel et al., 2006).

CHAPTER II MATERIALS AND METHODS

2.1 A note on solutions and buffers

Solutions were prepared from analytical grade reagents supplied by Sigma-Aldrich,
VWR or Fluka. All of the solutions and buffers were made in deionised water (Elix 5,
Millipore) and sterilized by either autoclaving or filtration (0.22 μm, Millipore). All
of the solutions used in RNA experiments were prepared in sterilized glassware and
treated overnight with 0.1% (v/v) diethyl pyrocarbonate (DEPC), left overnight in a
laminar flow hood and then autoclaved. Tris buffer solutions were not DEPC treated,
but were prepared with DEPC treated water.

2.2 DNA cloning in *Escherichia coli*

Most standard protocols were as described in Molecular Cloning 2nd edition
(Sambrook et al., 1989).

2.2.1 *E. coli* strains

DH5α

Φ80 *lacZΔM15 Δ*(*lacZYA -argF*) *U169 recA1 endA1 hsdR17* (r$_k$-, m$_{k+}$) *phoA supE44*
λ– *thi-1 gyrA96 relA1*

XL1 blue

recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ *proAB lacI*^{q}*ZΔM15* Tn*10*
(Tet^r)]

2.2.2 Bacterial growth media

LB broth and LB agar recipes are shown in Appendix I.

2.2.3 Growth of bacteria cultures

E.coli transformants were routinely grown overnight at 37 °C. For liquid culture, the flasks were kept in a shaking incubator at 220 rpm. Otherwise, *E. coli* were grown on inverted 9-cm LB agar plates.

2.2.4 Ligation and *E. coli* transformation

Ligation of DNA fragments was typically performed in a 20 μL reaction containing 100 ng of linearized plasmid and a four fold molar excess of the insert DNA, typically with 10 units of T4 DNA ligase (New England Biolabs, NEB). The ligation reaction was kept at 18 °C overnight or at room temperature for 2 hours. 100 μL of *E. coli* competent cells were typically transformed with 5 μL of ligation mixture as follow. Ligation mixture was mixed with competent cells and kept on ice for 20 minutes; the cells were then heat shocked at 42 °C for 45 seconds and cooled on ice for 2 minutes; the competent cells were mixed with 0.5 mL of SOC media and incubated at 37 °C for 1 hour, with gentle shaking. The cells were briefly centrifuged and spread on an LB plate containing 100 μg/mL ampincillin.

The competent cells were typically made by the Rubidium method: DH5α or XL1 blue were grown in 100 mL LB with 10 mM $MgCl_2$ and 10 mM $MgSO_4$ at 37 °C until the cell density was about to 0.5 at OD_{650}; the cells were then harvested at 5000 rpm for 15 minutes at 4 °C, resuspended in 30 mL filter sterilized Rb buffer 1(100 mM RbCl, 50 mM $MnCl_2 \cdot 4H_2O$, 80 mM KAc, 10 mM $CaCl_2 \cdot 2H_2O$, 15% glycerol and adjusting pH to 5.8 with 0.2 M acetic acid), and incubated on ice for 1 hour; the cells were then pelleted and resuspended in 8 mL filter sterilized Rb buffer 2 (10 mM RbCl, 10 mM MOPS, 75 nM $CaCl_2 \cdot 2H_2O$, 15% glycerol and adjusting pH to 6.8 by

NaOH); the competent cells were aliquoted with 200 µL per tube, and kept at -80 °C for future use.

2.2.5 Small-scale preparation of plasmids

A single colony was inoculated into 5 mL of LB broth containing 100 µg/mL ampicillin and grown overnight. Plasmid DNA was typically purified from a 1 mL aliquot of this culture using the following boiling-prep method:

1. 1 ml of the cell culture was transferred into a fresh 1.5 mL tube and spun briefly at 13000 rpm and the supernatant discarded.

2. 110 µL of ice cold STET buffer (8% sucrose, 50 mM Tris pH 8.0, 50 mM EDTA pH 8.0, 5% Triton X-100) containing 5 µL of 20 mg/mL lyzozyme was added into each sample and the pellet was then completely resuspended by pipetting up and down.

3. The samples were placed in boiling water for 20 seconds and then centrifuged at 13000 rpm for 10 minutes. The pellets were removed by using sterile toothpicks.

4. 110 µL of isopropanol was added to the supernatant, mixed and centrifuged at 13000 rpm for 15 minutes.

5. The supernatant was discarded. The pellet was washed with 70% ethanol, air-dried and resuspended in 40 µL TE (10 mM Tris·Cl, pH 8.0, 1 mM EDTA, pH 8.0) containing 1 µL of 1 mg/mL RNase A stock. The DNA samples were incubated at 65 °C for 20 minutes to remove the RNA and stored at -20 °C if required for future use.

When needed the extra 4 mL of cell culture was used to prepare pure plasmids preps for sequencing and yeast transformations (typically using Fermentas GeneJET plasmid Miniprep kit).

2.2.6 Large-scale preparation of plasmid DNA

Typically a single colony was inoculated into 1 mL of LB broth containing 100 μg/mL ampincillin and grown overnight, then 200 μL of the overnight culture was incubated into 100 mL of LB broth containing 100 μg/mL ampincillin for 4-5 hours until OD_{650} 0.8-1.0. Plasmid DNA was then prepared from the culture by using commercial kits (typically QIAfilter Plasmid Midi Kit, Qiagen). The extracted plasmid DNA was resuspended in 500 μL TE, pH 8.0, and the concentration of plasmid DNA was measured with a spectrometer (ND-1000, NanoDrop).

2.2.7 Restriction enzyme digestion

Restriction enzyme digestions were carried out in a 10-50 μL reaction All restriction enzymes used in the study were obtained from New England Biolabs (NEB). The conditions of the single-enzyme or double-enzyme digestion were followed according to the NEB enzyme instructions. For a sequential digestion, the initial reaction contained the enzyme that is active in the buffer with the lowest salt concentration. After the reaction had proceeded for 2 hours, the second enzyme and the buffer with the higher salt concentration were added and the reaction continued to proceed for further 1 hour.

2.2.8 Dephosphorylation of DNA

Antarctic phosphatase (NEB) was used to remove the 5' terminal phosphates of the DNA. This procedure was generally applied to prevent self-ligation of digested

plasmid DNA. Following the restriction enzyme digestion, 1 μL of antarctic phosphatase (5 units/μL) was added into the reaction and incubated at 37 °C for 1 hour. The DNA samples was then inactivated at 65 °C for 15 minutes or purified by gel electrophoresis and gel extraction using a silica powder based technique (see below).

2.2.9 DNA Purification

In my study, two methods were used to perform DNA purifications following PCR and restriction digestion. One is the polyethylene glycol (PEG) method, and the other is gel extraction.

2.2.9.1 PEG purification

1. An equal volume of the PEG solution (13% PEG8000 (w/v), 0.6 M NaAc, and 6mM $MgCl_2 \cdot 6H_2O$) was added to each DNA sample, mixed by vigorous vortexing, and kept at room temperature for 20 minutes. If the DNA fragment size was less than 300 bp, three volumes of the PEG solution were used instead.

2. The sample was centrifuged at 13000 rpm for 20 minutes and the supernatant was completely removed by a Pasteur pipette without touching the pellet.

3. The DNA pellet was washed by 1 mL 96% ethanol and centrifuged at 13000 rpm for 3 minutes, and then washed again by 70% ethanol.

4. The pellet was air-dried and dissolved in 20-30 μL TE buffer.

2.2.9.2 Gel purification

The DNA fragment was sliced-out of the gel and placed into a 1.5 mL eppendorf tube. The DNA was then purified by silica powder as described in the manufacturer's instructions (Silica Bead DNA Gel Extraction Kit, Fermentas).

2.2.10 Standard PCR

All primers used in my study are shown in Appendix II. The primers were purchased from either Sigma or MWG. The PCR conditions varied, depending on the DNA polymerase, the melting temperature (T_m) of the primer and the length of DNA to be amplified. PCR was run in a thermal cycler (PTC-200, DNA Engine) and the product analyzed by agarose gel electrophoresis.

2.2.10.1 PCR for colony screening

For a bacteria colony PCR, the fresh colonies were mixed with 10 μL PCR solutions which contained 1× PCR buffer, dNTP mixture (0.2 mM for each), 1.5 mM MgCl2, 2 μM primers and 0.25 U Taq DNA polymerase (typically GoTaq, Promega), and amplified using standard cycling parameters.

The procedure for a yeast colony PCR is a slightly modification of a previously described protocol (Danilevich and Grishin, 2002). A fresh *S. pombe* colony was treated with 50 μL of lysis buffer D2 (4 M guanidine hydrothiocyanate, 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, 0.1 M β-mercaptoethanol, and 0.5% laurylsarcosine). The samples were kept in boiling water for 5 minutes, and then centrifuged at 12000 rpm for 1 minute. The pellets were harvested, washed twice with 1 ml deionized water. Finally, the pellets were dissolved in 20 μL water and 2-4 μL of the cell samples were amplified in 25 μL PCR reaction for 35 cycles.

2.2.10.2 PCR for cloning

Pyrobest DNA polymerase (TaKaRa) was used to amplify the DNA fragments from yeast genomic DNA. 20 ng of *S. pombe* genomic DNA was used template and amplified in 50 μL PCR reactions which contained 1× pyrobest buffer, dNTP mixture (0.2 mM for each), 2 μM primers and 0.5 U pyrobest DNA polymerase. The PCR was run as follow: 95 °C initial denaturation for 5 minutes; 95 °C for 30 seconds, (T_m-5) °C as the annealing temperature for 30 seconds, 72 °C extension for 1 minute per kb of the expected DNA length and running for 30 or 35 cycles; 72 °C extension for 5 minutes.

Phusion DNA polymerase (NEB) was used to amplify DNA fragments from plasmids. ~1 ng of plasmid DNA was used as template and amplified in 50 μL reactions which contained 1× HB buffer, dNTP mixture (0.2 mM for each), 2 μM primers and 1 U Phusion DNA polymerase. The PCR amplification was as follow: 98 °C denaturation for 1 minutes; 98 °C for 5 seconds, Tm °C as the annealing temperature for 20 seconds, 72 °C extension for 0.5 minute/kb of the expected DNA length and running for 25 or 30 cycles; 72 °C extension for 5 minutes.

2.2.11 Reverse transcription of RNA

Reverse transcription was performed on 1 μg of total-RNA using an oligo-dT primer with a 5' end anchor sequence and SuperScript II reverse transcriptase (Invitrogen). The reverse transcription process was followed as described in the manufacturer's instructions. 1μL of reverse transcription product (20 μL) was used in PCR amplification by a combination of gene specific primers and an anchor primer.

2.2.12 Agarose gel electrophoresis of DNA

PCR or restriction enzyme digestion of DNA samples was resolved in agarose gels to confirm and separate the correct bands by molecular weight. DNA samples and the loading control were mixed with DNA loading buffer (10× stock, 20% glycerol, 0.1 M EDTA, pH8.0, 1.0% SDS, 0.25% bromphenol blue and 0.25% xylene cyanol), loaded onto the 0.8% -2% (w/v) horizontal agarose gel and run in TAE buffer (40 mM Tris base, 40 mM acetic acid and 2 mM EDTA) with 0.5 μg/mL ethidium bromide at a constant voltage of 90 V. The 1kb DNA ladder and the 100 bp DNA ladder were used as the loading control (NEB). The DNA sizes in descending order correspond to 10000 bp, 8000 bp, 6000bp, 5000bp, 4000 bp, 3000 bp, 2000 bp, 1500 bp, 1000 bp and 500 bp for 1 kb ladder; 1500 bp, 1200 bp, 1000 bp, 900 bp, 800 bp, 700 bp, 600 bp, 500 bp, 400 bp, 300 bp, 200 bp and 100 bp for 100 bp DNA ladder respectively.

2.2.13 DNA sequencing

Sequencing of some DNA samples was carried out by the Functional Genomics and Proteomics Unit of the School of Biosciences. Some samples were sequenced and analyzed by GATC biotech (Germany).

2.3 Plasmids construction

The plasmids used in this study are derived from pREP41 (Map 1, Appendix III) (Maundrell, 1993). pREP41 is a *S. pombe E. coli* shuttle vector. The plasmid contains the pUC19 backbone, and is easily engineered and manipulated in *E. coli*. The *LEU2* gene from *S. cerevisiae* is incorporated as the selection marker, which was used to screen the transformants by complementation of the leucine auxotroph of *leu1Δ32* or *leu1Δ* strains. In *S. pombe* cells, the expression is driven by the *nmt41* promoter (no

message thiamine, moderate strength) with a high level of transcription when thiamine is absent (Basi et al., 1993; Zurlinden and Schweingruber, 1997). The plasmid constructs used for chromosomal integration into *S. pombe* genome were derived from pDUAL-nmt41 (Map 2, Appendix III). pDUAL-*nmt41* is a shuttle vector, like pREP41 but containing *URA4* as the selective marker in episomal expression (Matsuyama et al., 2004). The marker for integration is *LEU1* 5' half fragment which rescues the *leu1d32* mutation by homologous recombination only if correctly integrated. All of the primers used in the cloning are listed in Appendix II (Table 1); and all of plasmids in the study are listed and shown in Appendix II (Table 3) and Appendix III (Map 1-11).

2.3.1 GFP reporters (intronless NMD reporters)

The GFP reporters described in this study derive from a plasmid clone encoding GFP S65T (gift from Professor M. Rosbash), (Kuperwasser et al., 2004). The GFP coding region was amplified by PCR using primers: gfp-ks and gfp-rev, digested with *BamH I* and inserted into the *BamH I* site of pREP41. Then the colonies were confirmed by small scale DNA preparation and *Nde I* digestion (pREP41 has a unique *Nde I* site, 17 bp upstream of the *BamH I* site, and the *GFP* coding region has a *Nde I* site, positioned at 244 bp from the beginning of the ORF).

The nonsense mutation (TAA) at triplet position 6 (PTC6) was introduced directly with the primer, N6-sens (combined with gfp-rev); the other two nonsense mutations (triplets positions 27 and 140) were introduced by overlapping PCR. The principle of overlap PCR is shown in Figure 2.1. The primer pairs used to introduce the TAA
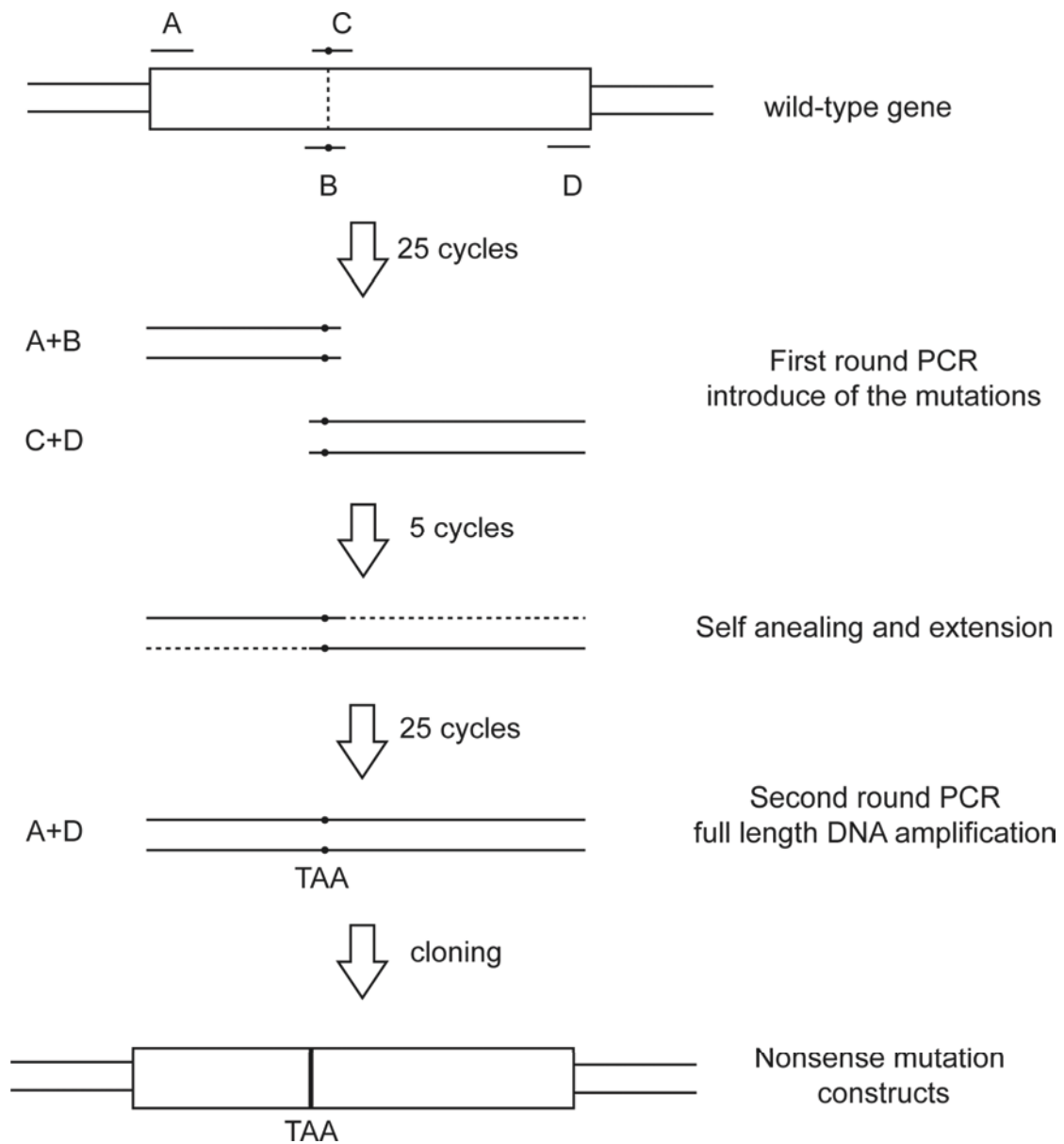
Figure 2.1 The principle of overlap PCR to introduce nonsense mutations.

This strategy was used to introduce the nonsense mutations into *GFP* or the endogenous gene *YPT3* and mutate the 5' splice site in the intron-containing NMD reporters. A-D, four primers used in overlap PCR.

mutations at triplets positions 27 and 140 are sens-N27 /rev-N27 and sens-N140 /rev-N140, respectively; the primers for the full length amplification of *GFP* ORF are gfp-ks and gfp-rev. Primer sequences are shown in Table 1, Appendix II. All of the intronless GFP constructs were initially screened by proof reading PCR with the primers N6-ch-sens, N27-ch-sens and N140-ch-sens, restriction digestion (*Nde I*), and further confirmed by sequencing with primer SSP41 flanking the cloning site. The schematic maps are shown in Fig. 4.1A and Map 3, Appendix III.

By using the same cloning and screening strategy, I made several batches of NMD reporters based on the intronless GFP reporter.

2.3.2 GFP reporters with longer AUG-PTC distance

To make the constructs with extended AUG-PTC distance, an extra *Avr II* site, between AUG and PTC6, was introduced into the GFP coding region with the primers, Gavr-2-sens, PTC6avr-2-sens and gfp-rev. After the desired colonies were screened by colony PCR, as described in 2.3.1, a small scale DNA preparation was carried out for the further confirmation with *Nde I* and *Avr II* digestions. The intermediate plasmids were named pGFP-AvrII and pPTC6-AvrII.

Then, two in-frame DNA fragments, 147 bp and 297 bp, both from the EGFP coding region, were amplified from pC3-GFP (lab stock) with the primers InHAeG-sens, InHAeGS-rev (147 bp) and InHAeGL-rev (291 bp). The PCR products were digested by *Avr II* and cloned into the intermediate plasmids. All of the constructs with the longer AUG-PTC distance were screened by colony PCR (the primers InHAeG-sens and gfp-rev) and *Nde I* digestion, and confirmed by sequencing with SSP41. The schematic maps are shown in Fig. 4.7A and Map 4, Appendix III.

2.3.3 GFP reporters with long 3' UTRs

Two DNA fragments, 102 bp and 418 bp, were amplified from the coding region of the firefly luciferase gene with the primers, luc102-sens, luc102-rev and luc418-rev. The PCR products digested by *Xma I* and cloned into the intronless GFP reporter constructs at the *Xma I* site, which is located 6 bp downstream of the GFP coding region, immediate upstream of the *nmt1* 3' UTR. The constructs with the long 3' UTR were screened by colony PCR (the primers luc102-sens and RSP41), and further confirmed by *Xma I* digestion. The schematic maps are shown in Fig. 4.2A and Map 5, Appendix III.

2.3.4 GFP reporters with longer PTC-3' end distances

An *Avr II* site was introduced directly into the GFP coding region immediately upstream of the normal stop codon with the primers Gavr2-243-rev and gfp-ks/N6-sens. The PCR fragments containing an *Avr II* site at the end of the GFP coding region were cloned into pREP41 as described in 2.3.1. The intermediate plasmids were named pGFP-Avr243, pPTC6-Avr243, pPTC27-Avr243 and pPTC140-Avr243.

An in-frame 402 bp DNA fragment was then amplified from the EGFP coding region with the primers InHAeG-sens and HAeG402-rev, and cloned into the plasmids at *Avr II* site as described in 2.3.2. All the constructs with longer PTC-3' end were screened by colony PCR with the primers InHAeG-sens and RSP41 and *BamH I* digestion, and further confirmed by sequencing with the primer RSP41. The schematic maps are shown in Fig. 4.3A and Map 6, Appendix III.

2.3.5 GFP reporters with *ade6* 3' UTR

The *ade6* 3' UTR (1.1 kb) was amplified with the primers, ade6-3Us and ade6-3Ur, from *S. pombe* genomic DNA, and digested by *Xma I* and *Sac I*. Then, *nmt1* 3' UTR which is located between the two enzyme sites was replaced by *ade6* 3' UTR by *Xma I* and *Sac I* digestion. The replacement was confirmed by colony PCR with the primers, Sgfp-sens and ade6-3Ur and *Xma I* and *Sac I* digestion. The schematic maps are shown in Fig. 4.4A.

2.3.6 GFP reporters with short PTC-3' end distance

The 5' half of the GFP coding region (327 bp from ATG) was amplified with the primers, gfp-ks and GFPsh-rev, and cloned into pREP41 at *Bam HI* site as described in 2.3.1. The nonsense mutation at triplet position 6 was introduced directly by the primer N6-sens. The constructs were screened by colony PCR (the primers, SSP41 and GFPsh-rev) and *Nde I* digestion, and confirmed by sequencing. The schematic maps of the constructs are shown in Fig. 4.5A.

2.3.7 GFP$_{ivs}$ reporters (intron-containing NMD reporters)

A unique *Pml I* site is located at 326 bp in the GFP coding region. The second intron of *UBC4* was amplified from *S. pombe* genomic DNA by the primers, INubc4-sens and INubc4-rev, phosphorylated by T4 DNA polynucleotide kinase (NEB), and cloned in GFP coding region at *Pml I* located at 326 bp in the GFP coding region. All the intron-containing GFP reporters were screened by colony PCR with the primers SSP41 and INubc4 rev, and further confirmed by sequencing with the primers SSP41 and RSP41 respectively. The intron insertion splits the GFP coding region at triplet position 110. The schematic maps are shown in Fig. 5.1A and Map 7, Appendix III.

2.3.8 GFP$_{ivs}$ reporters with 5' splice-site mutation

Derived from the GFP$_{ivs}$ reporters, the 5' splice-site was mutated by overlap PCR with the primers, 5SP-SENS and 5SP-REV. The cloning and screening strategy is the same as described in 2.3.1, with the exception that the primer 5sp-ch-sens is used in proof reading PCR. The 5'splicing site GTAT was mutated as CTAA, which caused an in frame nonsense mutation at triplets position 111. The 5' splice-site mutation constructs were confirmed by sequencing with the primers SSP41 and RSP41. The schematic maps are shown in Fig. 5.2A.

2.3.9 GFP reporters with a 3' UTR intron

The intron from *UBC4* was amplified with the primers IN3UTR-sens and IN3UTR-rev, digested by *Xma I*, and cloned into the GFP reporters at the *Xma I* site, which is located 6 bp downstream of the GFP normal stop codon. All the constructs were screened by colony PCR with primers, IN3UTR-sens and RSP41, and confirmed by sequencing with the primer RSP41. The schematic maps are shown in Fig. 5.7A and Map 8, Appendix III.

2.3.10 GFP$_{ivs}$ reporters with long PTC-intron distance

Derived from GFPivs reporters, an *Avr II* site was introduced into the GFP coding region at triplet position 120 by overlap PCR with the primers, Gavr-120-sens, Gavr-120-rev, gfp-ks and gfp-rev. The plasmids were screened by colony PCR with primer SSP41 and gfp-rev and *Avr II* digestion, and confirmed by sequencing with the primer RSP41. The two intermediate plasmids were named p*GFP*ivs-Avr120 and pPTC140ivs-Avr120.

Then, two different lengths of DNA fragments, 147 bp and 291 bp, as described in 2.3.2, were cloned into the intermediate plasmids. The constructs were screened by colony PCR with the primers InHAeG-sens and RSP41, and confirmed by sequencing with the primers SSP41 and RSP41. The schematic maps are shown in Fig. 5.8A and Map 9, Appendix III.

2.3.11 Integrated GFP reporters

All of the integration constructs are derived from pDUAL-HFF41 (Matsuyama et al., 2004). pDUAL-HFF41 was cleaved by *BamH I* to remove the DNA fragment of C terminal fused His-tag and Flag-tag. The digested vector was self-ligated and harvested as pDUAL-*nmt41* (Map 2, Appendix III).

The GFP fragments with/without introns were digested by *BamH I* and cloned into pDUAL-*nmt41* at *BamH I* site. The integrated GFP reporters were screened by colony PCR with the primers SSP41 and gfp-rev and confirmed by sequencing with the primers, SSP41 and Sgfp. The schematic maps are shown in Map 10, Appendix III.

2.3.12 *YPT3* and *YPT3* cDNA NMD constructs

For the intron-containing *YPT3* constructs, the *YPT3* gene, which contains two introns, was directly amplified from *S. pombe* genomic DNA by the primers ypt3-sens and ypt3-rev, and inserted into pREP41 at *BamH I*. The nonsense mutation at triplets position 6 was introduced directly with the primer ypt3N6-sens; the other nonsense mutation at triplets position 131 was mutated by overlap PCR with the primers ypt3N130-sens, ypt3N130-rev, ypt3-sens and ypt3-rev.

For making the *YPT3* cDNA constructs, the *YPT3* cDNA was obtained by reverse transcription from total RNA extracts from *S. pombe*. The *YPT3* cDNA cloning use

the same procedure as indicated above. The *YPT3* and its cDNA NMD constructs were screened by colony PCR with the primers SSP41 and ypt3-rev and confirmed by sequencing with the primers SSP41 and RSP41. The schematic maps are shown in Fig. 5.3A and Map 11, Appendix III.

2.4 *S. pombe* methods

2.4.1 Strains

The *S. pombe* strains used in the project are shown in Appendix II, Table 3. All strains were stored at -80 °C in YES or EMM media plus 40% sterile glycerol. Deletion strains were validated by PCR of genomic DNA, using the appropriate primer pair (Appendix II Table 1), A schematic of how knock-out and knock-in cells were verified by PCR is shown in Fig. 2.2.

2.4.2 *S. pombe* media and culturing

S. pombe cells were cultured in YES or EMM medium. The recipes of these media are shown in Appendix I. For cultures on Petri dishes, cells were streaked on YES or EMM plates and incubated at 30 °C or 25 °C for temperature sensitive strains (ts strains) for 2-4 days until colonies reached at least 1 mm diameter.

Liquid cultures were typically inoculated from single colonies from streaked plate. First in small start-up cultures, 1-2 mL culture was grown overnight at 30 °C or 25 °C (ts strains) in shaking incubator. Fresh small cultures were then used to inoculate 10-50 mL (1 to 50 dilution) and grown overnight.

Fig. 2.2

A

B

C
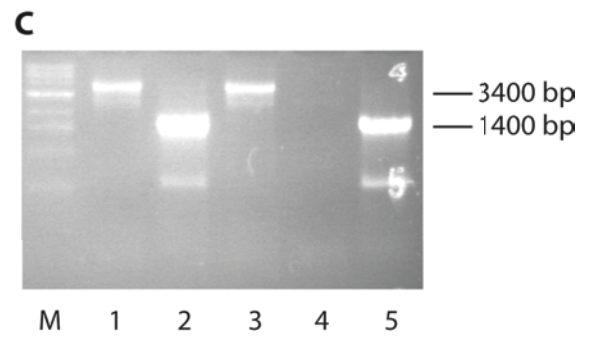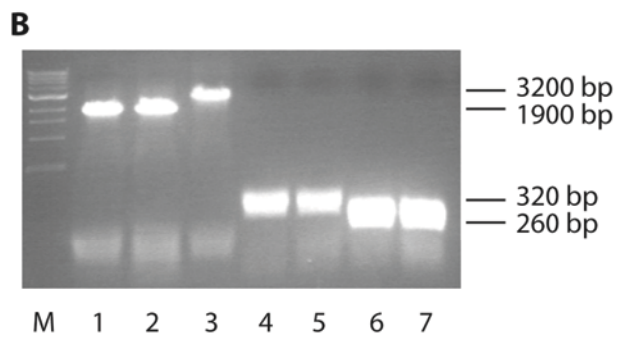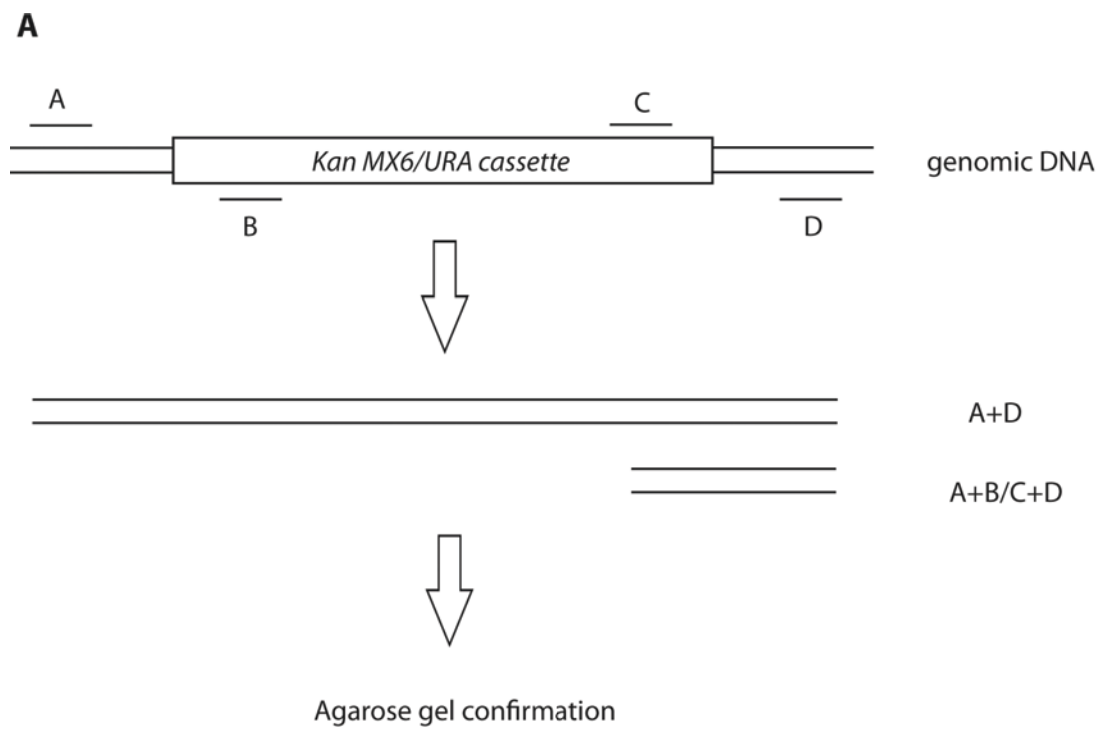
Figure 2.2 *S. pombe* strain confirmations by PCR of genomic DNA.

(A) Schematic of the PCR procedure to test if the marker cassette was inserted in the right location.

(B-C) Confirmation of *upf1Δ* and *exo2Δ* strains in which the Kan MX6 cassette and URA cassette were used to knock out the genes respectively. (B) Genomic DNA from *upf1Δ* was prepared and assayed by PCR with the primers upf1ch-sens1 and upf1ch-rev1 (lanes 1 and 2, *upf1Δ* and lane 3, wild-type as the control); and the combined primers of upf1ch-sens/kan-inp1r (lane 4 and 5) and kan-inp1s/upf1ch-rev (lane 6 and 7). (C) Genomic DNA from *exo2Δ* was assayed with primers exo2-sens and exo2-rev (lanes 1 and 3, *exo2Δ* and lane 4, wild-type); and the combined primers exo2-sens and ura4ch-rev (lane 2 and 5). The expected DNA fragments were produced in the deleted strains, 1.9 kb in *upf1Δ* (B, lane 1) but 3.2 kb in wild-type (B, lane 3); 3.4 kb in *exo2Δ* (C, lane 1), but no DNA present in wild-type (C, lane 4, 5.4 kb DNA expected)

2.4.3 *S. pombe* DNA transformation

Three slightly different procedures were used, depending on the aims for episomal

expression, DNA integration or gene replacement. The strain to be transformed was

cultured in 5-10 mL YES until the cell density reached 0.8-1.0 at OD_{650} ($1\times10^{7,}$

measured by haemocytometer). Cells were made competent for transformation using

the LiAc method described below (Forsburg and Rhind, 2006).

1. The cell culture was centrifuged at 3000 rpm for 5 minutes and the pellet was
   washed once with 10 mL of sterile water.

2. The cells were centrifuged again in the same condition and resuspended in
   100-200 μL of sterilized water. An equal volume of LiAc buffer (10 mM Tris-
   HCl pH 8.0, 1 mM EDTA, 0.1 M Lithium acetate) was added and thoroughly
   mixed with the resuspended cells (cell density should be above to 5×10^{8}).

3. For each transformation, 100 μL of LiAc-resuspended cells were transferred
   into a 1.5 mL sterile microtube and kept at room temperature (lower than 28
   °C) for 15-30 minutes.

4. Then 1 μg of plasmid DNA (episomal expression) or 10 μg linerized DNA for
   homologous recombination (plasmid integration or gene knock-out), and 2 μL
   of 10 mg/mL ssDNA were mixed with 100 μL of LiAc-treated cells. The
   mixture was then kept at room temperature for 20-30 minute.
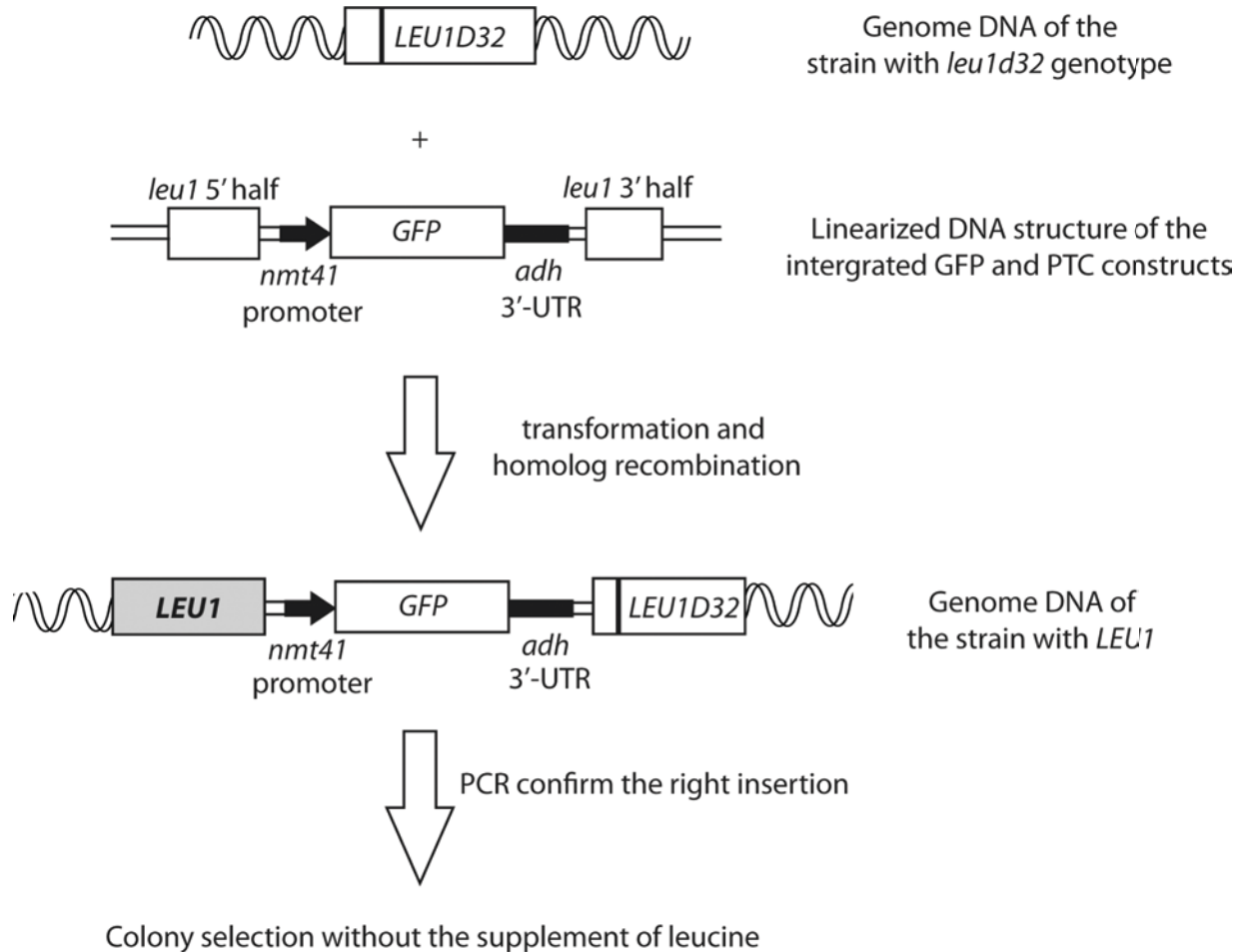
Figure 2.3 Plasmid integrated into the genome of *S. pombe*.

The *LEU1* 5' half and 3' half DNA sequences in pDUAL-recombine with the homologous sequences of the chromosomal *leu1* locus. The recombination reaction removes the mutation of the *leu1d32* allele and leads to a fully functional *LEU1* gene and leucine prototrophy, giving in cells that had a successful recombination. Between the two *LEU1* recombination sequences there is the *nmt41* promotor, multiple cloning sites and the *adh* terminator (Map 2, Appendix III).

5. After, 220 μL of 50% PEG solution and 40 μL of LiAc buffer were added to the transformation mixture and mixed gently. The sample was kept at 30 °C (25 °C for ts strain) for 1 hour.

6. Then the sample was heat shocked at 42 °C for 20 minutes; briefly centrifuged at 3000 rpm and the pellet washed with 1 mL of sterile water

7. At the end, the cells were pelleted, and resuspended in 100 μL of sterilized deionized water and spread on the appropriate selection plate

2.4.4 Genomic DNA extraction

Single colonies were grown in 20 ml cultures to stationary phase for 1.5-2 days, and spun down at 3000 rpm for 5 minutes. The pellets were then briefly washed in sterile water and the genomic DNA was extracted using the CPES method described below (Alfa et al., 1993).

1. The cell pellet was resuspended in 1ml of Citrate/ Phosphate buffer (50 mM citric acid and phosphate pH5.6, 40 mM EDTA pH 8.0, and 1 M sorbitol).

2. 50 μl of Zymolyase 20T (Europa Bioproducts Ltd) was added and then the mixture was incubated at 37 °C for 60 minutes. Cells are ready for next step if more than 50% cells show a loss in refractivity under the microscope, (the change in refraction is apparent by adding 1% SDS to the aliquot of culture to be examined)

3. Cells were pelleted at 8000 rpm and resuspended in 0.55 ml TE and 1 % SDS and incubated at 65°C for 1 hour for lysis. Then 300 μl of 3 M potassium acetate (pH 5.0) was added and the tubes were kept on ice for 15 minutes.

4. After the samples were centrifuged at 13000 rpm for 15 minutes at 4 °C, 600 μl of supernatant was transferred into a fresh tube, and an equal volume of ice-cold isopropanol was mixed completely with the supernatants. The mixture was incubated at -20°C for 10 minutes and centrifuged at 13000 rpm for 15 minutes at 4 °C.

5. After washed by 70 % cold ethanol, the pellet was resuspended in 300 μl TE buffer with 50 μg/ml RNase, and incubated at 65 °C for 15 minutes.

6. The sample was mixed with 300 μl of saturated phenol (pH 7.9):chloroform:isoamyl alcohol (25:24:1), vortexed vigorously and centrifuged. The supernatant was then transferred into a 1.5 mL sterile tube. The phenol: chloroform extraction was repeated several times (typical 2 times) until the interface turned clear.

7. The water phase was transferred into a fresh tube, and then 30 μl of 3 M sodium acetate (pH 7.5) and 700 μl of cold absolute ethanol were added to precipitate the DNA. The tubes were kept at -20 °C for 5-10 minutes.

8. After the tubes were centrifuged at 13000 rpm for 10 minutes at 4 °C, the pellet was washed with 70 % ethanol and air-dried. The dry pellet was dissolved in 50 μl TE and 1 μl of the solution was run in agarose gel to estimate the quality and quantity of the genomic DNA (there should be a band at proximate 20 kb).

2.4.5 RNA extraction

Total-RNA was purified from either 10 mL (plasmid transformed cells) or 20 mL cultures (strains with integrated constructs). The RNA was extracted with the hot-acid phenol method, with slight modifications (Ausubel et al., 1996).

1. Cells were centrifuged at 3000 rpm for 3 minutes at 4 °C, washed once with 10 mL of sterile water.

2. Cells were centrifuged again and the pellet was resuspended in 600 μL of TES buffer (10 mM Tris pH 7.5; 10 mM EDTA pH 8; 0.5% SDS). Then the resuspended cells were mixed with 600 μL of acid phenol pre-heated at 65 °C.

3. The sample was incubated at 65 °C for 1 hour and vortexed for 10 seconds every 10 minutes. At the end the sample was cooled on ice briefly.

4. The sample was centrifuged at 13000 rpm for 20 minutes at 4 °C; 550 μL of the aqueous phase was transferred into a clean sterile tube and mixed (by vortexing) with 500 μL acid phenol:chloroform:isoamyl alcohol (25:24:1) .

5. The mixture was centrifuged at 13000 rpm for 15 minutes at 4 °C and 500 μL of the aqueous phase transferred into a clean sterile tube. The aqueous phase was re-extracted with 500 μL chloroform: isoamyl alcohol (24:1) as above.

6. The final aqueous phase was transferred into a nuclease free sterilized tube and the RNA precipitated by adding 50 μL of 3M NaAc (pH 5.2) and 1.2 mL of pre-chilled 100% ethanol.

7. The tube was kept at -20 °C for 20 minutes, and then centrifuged at 13000 rpm for 15 minutes at 4 °C; the pellet was washed once by 70% ethanol, dried and dissolved in 50 μL DEPC-treated water.

2.4.6 Northern blot analysis of RNA samples

2.4.6.1 Gel electrophoresis

5 to 20 μg of RNA samples were mixed with 5.5 μl of 35% formaldehyde, 15 μl of 100% formamide, 3 μl 10× MOPS (0.2 M MOPS, 80 mM NaAc and 10 mM EDTA) buffer and DEPC water to the final volume of 30 μl. The sample mixtures were incubated at 65 °C for 20 minutes and then immediately put on ice for at least 5 minutes. The 1.2 % agarose/formaldehyde gel (1.4 g agarose, 12 ml 10× MOPS, 21.45 ml formaldehyde, and 86.25 ml DEPC water) was pre-run for 5 minutes. 1 μl of RNA loading buffer was added into each sample and the gel was run at 80 V for 4-5 hours in 1× MOPS buffer.

2.4.6.2 Capillary blot, crosslinking and staining

The gel was then washed in DEPC water for 20 minutes and in 20× SSC (3M NaCl, 300 mM Sodium citrate pH 7.0) buffer for 20 minutes. During gel washing, the membrane (HyBond, Amersham Biosciences), 3 MM papers and paper towels were prepared. The RNA was transferred onto HYBOND membrane by the blot and left overnight as described in Molecular cloning II (Sambrook et al., 1989).

The membrane was removed from the blot and put on wet filter paper to be crosslinked by 254 nm ultraviolet light (UV) at 120 J. After UV crosslinking, the membrane was washed by DEPC water for 5 minutes, stained with 50 mL methylene blue (0.5 M Na acetate, pH 5.2, 0.04% methylene blue) by shaking for 5 minutes and

destained by DEPC water for 10 minutes until the 18S and 25S rRNA bands were visible on the membrane.

2.4.6.3 Probe labeling and hybridization

30 ml hybsol buffer (1.5× SSPE (0.27 M NaCl, 15 mM sodium phosphate pH 7.4, 1.5 mM EDTA), 7% SDS and 10% PEG8000) containing 100 μg/ml boiled ssDNA and 250 μg/ml heparin was added into a hybridization tube and the prehybridization was carried out in the hybridization oven at 68 °C for 4 hours at 30 rpm.

The radioactive probe was synthesized by random priming as following. Firstly, 40 ng of template DNA was mixed with 10 μl of 5× reaction buffer (promega, random primer already premixed in the buffer) and 30 μl of sterile water. The mixture was boiled at 100 °C for 5 minutes and immediately cooled on ice for 5 minutes. After the sample was denatured, 2 μl of dNTPmix (15 mM for each) without dCTP, 1 μl of Klenow DNA polymerase and 3-5 μl of P35-dCTP (3000cci) were added into the sample and kept at room temperature for 1-2 hours. The G50 sepharose column was used to purify the radioactive labeled probes. The purified probes were heated at 100 °C for 5 minutes and immediately cooled on ice for 5 minutes. Meanwhile, the prehybridization buffer was discarded, and then the labeled probes and 20 ml of Hybsol containing 100 μg/ml boiled ssDNA and 250 μg/ml heparin were added into the tube and mixed well together. The hybridization was carried out at 68 °C overnight.

2.4.6.4 Membrane washing, signal development and quantitative analysis

The membrane was washed with 60 ml of 2× SSC and 0.1% SDS with four intervals, 2, 5, 30 and 30 minutes respectively and then washed with 0.2× SSC and 0.1% SDS

for an extra 30 minutes at 68 °C. After washing, the membrane was wrapped in Saran film. The membrane was either exposed to X-ray film at -70 °C in a cassette with intensifying screens and developed by X-graph or exposed to a Kodak phosphor-imaging screen and developed by the phosphorimager *FX* (*Bio-Rad*). The signal intensities were analyzed using the Quantity-one program (*Bio-Rad*).

2.4.7 Counting the number of cells

Cell cultures were diluted in water (1:10), and 10 μL of the dilution were loaded onto a haemocytometer and counted with a phase-contrast microscope equipped with a 40× objective. The calculation of cell numbers is based on the manufacturer's instructions.

2.4.8 Fluorescent *in situ* hybridization (FISH)

FISH was performed as previously described for *S. cerevisiae* with minor modifications (Kuperwasser et al., 2004). All the procedures that are related to the fluorescent dye and labeled probes were performed in a dark environment. The tubes or containers were also wrapped in foil.

Cell cultures (2 mL, $OD_{600} \sim 0.5$) were fixed with 0.4% formaldehyde and 10% glacial acetic acid at room temperature for 15 minutes. The cells were pelleted and washed with 1 mL of SPPi buffer (1.2 M sorbitol, 0.1 M potassium phosphate, pH 6.5) three times. Cells were resuspended in 50 μl of SPPi. Before FISH, the cell wall was removed using the following spheroplasting procedure. Cells were incubated with Novozyme 234 (200 μg/mL, Interspex Products Inc) and zymolyase 20T (500 μg/mL, Europa Bioproducts Ltd) for 15-30 min at 37 °C in SPPi buffer with 0.2 % β-mercaptoethanol, 20 mM Vanadyl-ribonucleoside, until approximate 80% cells were spheroplasted (confirmed by microscope observation). The spheroplasts were

centrifuged at 3500 rpm for 2 minutes, washed by 1 mL of SPPi buffer once and resuspended in 40 μL SPPi buffer. Then, 20 μL of the cell suspensions was spread onto one of the circles of a coated coverslip (see below) and incubated in a humid chamber for 1-2 hour. The coverslip was washed as follow: 5 minutes in SPPi buffer twice; 5 minutes in 0.1 M potassium phosphate pH 6.5; 5 minutes in 0.1 M potassium phosphate pH 6.5 with 0.1% NP40; 5 minutes in 0.1 M potassium phosphate pH 6.5. Lastly the coverslip was washed in 70% ethanol at - 20 ^{o}C for 20 minutes. Before hybridization, the coverslip was briefly washed twice with 2× SSC, each for 5 minutes, and then treated with 40% formaldehyde in 2× SSC for 20 minutes. After washing and discarding the SSC, The coverslip was covered with 20 μL per circle of hybridization solution containing Cy3-labbeled oligonucleotide probes (see below) and incubated in a dark humid chamber at 37 ^{o}C for at least 10 hours (typically overnight). Post-hybridization the coverslip was washed as follow: twice with 40% formamide in 2× SSC at 37 ^{o}C for 15 minutes; with 2× SSC with 0.1% Triton X-100 for 15 minutes at room temperature; twice with 1× SSC for 15 minutes and twice with 1× PBS for 5 minutes. The wells were then dried for 5-10 min (avoid over drying the coverslip prevent cell shrinkage). Drop 6-7 μL of mounting solution (1×PBS, pH 8.0, 1 mg/mL phenylenediamine, 90% glycerol, optional for 0.1 mg/mL DAPI) was dropped onto the coverslip, and the coverslip picked-up with a glass slide. The coverslip was immobilized with nail polish. The slides were viewed with a fluorescent microscope equipped with Ds-Red filter-set using a 100× oil lens (Nikon ECLIPSE E600). Images were captured with a CCD camera (HAMAMATSU C4743-95) using the Simple PCI 4.0.6.1605 software (HAMAMATSU).

For probe labeling, the oligos used for FISH were a gift from Dr. T. H. Jensen (University of Aarhus) and listed in Table 1, Appendix II. The GFP probes consist of

four amino modified oligonucleotides complementary to the GFP coding region, as previously described (Dower and Rosbash, 2002). The oligos were labeled with Cy3 using the Cy^{TM} 3-mono-Reactive Dye Pack (GE Healthcare). 20 μg of amino-modified oligonucleotide (4 oligos, 5 μg for each sample dissolved in water) was dried in a speed vacuum concentrator and then dissolved in 25 μL of 0.1M sodium bicarbonate pH 9.0. 25 μL of the Cy3 dye solution (made by 75 μL DMSO per vial as described in manufacturer's instructions) was mixed with the oligos; the tube was wrapped in foil and incubated 3 hours at room temperature in the dark. After the conjugation, the sample was mixed with 50 μL of water and purified by a G50 column (Sambrook et al., 1989).(Sambrook et al., 1989). The labeled oligos were kept at a concentration of 200 ng/μL; the tubes were wrapped in foil and stored at -20 °C.

For attaching the spheroplasted cells, on the glass coverslip (20 mm×70 mm), 6 wells (~6 mm square each) were made by using a hydrophobic pen. 20 μL 0.1% poly-L-lysine (Sigma, P8920) was added into each well and incubated in a humid chamber for 2 hours. The poly-L-lysine solution was discarded and the wells were briefly washed twice by 20 μL water and then air dried to be used.

For hybridization, 5 μL of labeled oligos were mixed with 20 μL of 10 mg/mL salmon sperm DNA and 20 μL of 10 mg/mL yeast tRNA. The mixture was lyophilized in a speed vacuum and resuspended in 80 μL of 100% formamide (Sigma Aldrich). The solution was heated at 100˚C for 3 minutes, cooled on ice, and then mixed with 80 μL of ice-cooled hybridization buffer (4 mg/mL BSA, 4 mM Vanadyl-ribonucleoside (200 mM stock, Sigma), 20% Dextran Sulfate, 4× SSC).

2.4.9 Protein extraction from *S. pombe* cells

Two different methods were used to extract proteins from yeast cells. One method makes use of glass beads. A 5-15 mL aliquot of a fresh cell culture (cell number > 10^8) was spun at 3000 rpm for 5 minutes, and the pellet was washed with 1.2 mL of cold distilled water, and then transferred into a 1.5 mL fresh screw cap tube and spun again. The pellets were resuspended in 200 µL of protein lysis buffer (50 mM Tris, pH 7.5,150 mM NaCl, 5 mM EDTA, 10% glycerol, 1 mM phenylmethylsulphonyl-fluoride (PMSF),) and about 400 µL of 0.5 mm acid-washed glass beads (Sigma) were added into each tube. Each sample was vortexed at 5500 rpm for 15 seconds in Precellys 24 homogenizer (Bertin Technologies) and kept at 4 °C for 2 minutes. The process was repeated 2-3 times until more than 70% of cells were broken as observed under the microscope. To recover the cell lysate, the bottom of the tube was pierced with a needle (the needle can be heated over a flame to facilitate piercing). The punctured tube was placed in a fresh 1.5 mL tube and centrifuged at 13000 rpm for 10 minutes at 4 °C. 30 µL of supernatant was mixed with 30 µL of 2× SDS loading buffer (120 mM Tris-HCl, pH 6.8, 8% SDS, 8% β-mercapto-ethanol, 25 mM EDTA, 0.04 % bromophenol blue, 20% glycerol). The extract was boiled for 5 minutes and either loaded on a SDS-PAGE gel or stored at -20 °C for future use.

The second protein extraction method involves sodium hydroide mediated lysis of the cells (Matsuo et al., 2006). A 2-5 mL aliquot of cell culture was spun down and washed once with 0.5 mL of water. The pellet was completely resuspended in 300 µl water, and then 130 µL of 1 M NaOH was added and into the sample and mixed gently by flipping the tube. The tube was incubated at room temperature for 10 minutes. The sample was then spun at 8000 rpm for 3 minutes at 4 °C and the supernatant was fully removed. 50-100 µL of 1× SDS loading buffer was used to

resuspend the cells completely, and then boiled for 5 minutes and centrifuged at 13000 rpm for 2 minutes. Typically 20 μL of the sample were loaded on the gel or stored at -20 ^{o}C for future use.

2.4.10 SDS-PAGE and Western blot

SDS-PAGE gels (10%) were prepared as described in Molecular Cloning 2nd edition (Sambrook et al., 1989). SDS-PAGE gels were run at 120 V for 60-100 minutes at room temperature by using 1× running buffer (10× stock buffer: 0.25 M Tris-HCl, 1.92 M Glycine, 1.0% SDS, pH 8.3). The proteins were transferred onto nitrocellulose membranes (Schleicher & Schuell) using a standard wet electroblotter apparatus at 100 V for 60-90 minutes in cold room in transfer buffer (25 mM Tris-HCl, 190 mM Glycine, 20% Methanol). Membranes were blocked with 30 ml of 5% skimmed milk-TBS (25 mM Tris-HCl, 137 mM NaCl) at room temperature for 1-2 hours and then washed with TBS 0.1% Tween 20 (TBST) three times, each for 10 minutes. The primary antibody was diluted 1:3000 directly in TBST and incubated with the membrane overnight in the cold room. The following day, the membrane was washed with TBST three times, each for 10 minutes. The secondary antibody, horseradish peroxidase (HRP) label anti-mouse (Sigma), was diluted 1:10000 in TBST and incubated with the membrane for 40-60 minutes. The membrane was washed with TBST three times, each for 10 minutes. The signal was developed by Supersignal West Pico Chemiluminescent Substrate kit (PIERCE). 0.5 mL of Supersignal WestPico Luminol/Enhance solution was mixed with 0.5 mL Stable Peroxide solution and the mix spread uniformly onto the membrane and incubated for 5 minutes. The solution was drained away and the membrane wrapped in cling film. The signal was detected with X-ray films or visualized with a ChemiDoc imaging system (*Bio-Rad*).

2.4.11 Polysome profile

2.4.11.1 Preparation of sucrose gradients

Tubes with 10-50% (w/v) sucrose gradients were prepared with DEPC-treated water and 10× polysome buffer (100 mM Tris acetate pH7.4, 700 mM ammonium acetate, 40 mM magnesium acetate). Sucrose solutions were prepared as detailed in Table 2.1 and dispensed, starting with 50% solution, into SW41 polyallomer centrifuge tube (Beckman Coulter) in 2.25mL fractions and quick frozen in liquid nitrogen between each layer. The dispensing of sucrose solutions started with 50% sucrose, following with 40%, 30%, 20% and 10% sucrose solution sequentially. Gradients were stored at -80 $^{\circ}$C, defrosted and equilibrated at 4 $^{\circ}$C for 12 hours prior to loading the samples.

Table 2.1 Sucrose solutions for gradient preparation

	50%	40%	30%	20%	10%
10× polysome buffer (mL)	3	3	3	3	3
Sucrose (g)	15	12	9	6	3
DEPC water	Adjust the final volume to 30 mL				

2.4.11.2 Cell extract preparation

A 50 mL *S. pombe* cell culture was grown at 30 $^{\circ}$C until OD_{650} 0.3-0.4, mixed with 500 µL of 10 mg/mL cycloheximide and incubated for further 15 minutes. The following steps were all performed on ice or in a cold room. After the cycloheximide treatment, the cells were harvested by centrifuging at 3000 rpm and washed by 10 mL of lysis buffer (20 mM HEPES pH7.4, 2 mM magnesium acetate, 100 mM potassium acetate, 100 µg/mL cycloheximide, 0.5 mM dithiothreitol). The cells were pelleted

again and resuspended in 300 µL of lysis buffer with RNase inhibitor (40 U/mL, Invitrogen) and protease inhibitor cocktail, EDTA free (Roche). The suspension was transferred into a fresh 2 mL screw-cap tube and filled with acid-washed glass beads (0.5 mm) until the beads reached the meniscus of the solution. Cells were lysed by vortexing for three rounds of 15 seconds at 5500 rpm in homogenizer with 2 minutes cooling in between. Typically more than 70% of cells were broken after 2-3 rounds. An additional 300 µL of lysis buffer with 40 U/mL RNase inhibitor and protease inhibitor cocktail was added to dilute the sample. The 2 mL screw-cap tube was pierced at the bottom with a needle, placed trough a hole cut through the cap of a 15 mL screw cap tube, and centrifuged at 5000 rpm for 5 minutes at 4 °C to recover the cell lysate. The lysate was transferred into a fresh 1.5 mL tube and cleared by centrifugation at 13000 rpm for 15 minutes at 4 °C. The supernatant was transferred into a fresh 1.5 mL tube and kept on ice .A 1:10 dilution of the extract was used to measure the absorbance at 260 nm with a NanoDrop spectrometer (ND-1000 NanoDrop). The absorbance at 260 nm is proportional to the RNA concentration and therefore can be used to normalize polysome extracts loading.

2.4.11.3 Sedimentation of translation complexes

The polysome extracts were adjusted to 10 Abs_{260} units/mL and carefully loaded onto the top of the 10-50% sucrose gradients. The tubes were carefully balanced, loaded into the buckets of a Beckman SW41 rotor and centrifuged at 38000 rpm for 160 minutes. The gradients were pumped (from the bottom, using a steel capillary) through a flow-through UV spectrophotometer (Pharmacia LKB-Optical Unit UV-1) by using a peristaltic pump (P-1, Pharmacia) with the speed at 1.2 mL/min. The absorbance was measured at 254nm and recorded on crart recorder (Pharmacia LKB

REC 102, speed setting, 1 cm/min). Fractions (0.8 mL per tube, 15 fractions) were collected with a fraction collector (FRAC100, Pharmacia).

2.4.11.4 RNA and protein purification from sucrose fractions

To extract the RNA from the fractions, the phenol:chloroform method was adopted. A 400 μL aliquot of every fraction was mixed with 600 μL of phenol:chloroform and centrifuged at 13000 rpm at 4 °C for 10 minutes. The aqueous phase was transferred into a fresh tube and extracted with 400 μL of chloroform:isoamyl alcohol (24:1) and centrifuged at 13000 rpm at 4°C for 10 minutes. The aqueous phase was transferred into a nuclease free sterilized tube, and then mixed with 40 μL of 3M NaAc (pH 5.2) plus 1 mL of 100% ethanol (pre-chilled) and kept at -20 °C for 30 minutes. The sample was centrifuged at 13000 rpm for 15 minutes and washed once by 70% ethanol (RNase free). The RNA sample was then dried and dissolved in 20 μL of DEPC water with 60% of formamide for future use.

To precipitate the proteins from the fractions, a 400 μL aliquot was mixed with 70 μL of 100% (w/v) trichloroacetic acid and kept at 4 °C overnight. The sample was centrifuged at 13000 rpm for 20 minutes to precipitate proteins from the sucrose solution. The pellet was washed with 300 μL of acetone three times, dried in a heat block at 95 °C, resuspended in 30 μL of 1× SDS loading buffer and boiled for 5 minutes. Protein samples were run in 10% SDS-PAGE and analyzed by Western blot.

RESULTS

Aims of the study

As described in Introduction (Chapter I), the studies of NMD mechanism how the PTC has been differentiated from the normal termination codon mainly studied in *S. cerevisiae*, *D. melanogaster* and mammalian cells. Both of the proposed dominated merchanisms, *faux* 3' UTR model and EJC model, which explain NMD in eukaryotes have the contradictory observations, suggesting that NMD might be a complexed consequence of mRNA reductions due to PTCs. To further investigate the explanations, the fission yeast, *S. pombe*, was chosen as the model organism in my PhD study to investigate NMD mechanism, mainly including the following questions:

1. Whether the current explanation of NMD mechanism, faux 3' UTR, are applied in S. pombe?

2. Since S. pombe have much more introns in its genome, in comparison with S. cerevisiae, whether intron or splicing play a certain role in NMD in S. pombe?

3. As suggested in fly and human cells, whether the endonucleolytic cleavage occurred in NMD in S. pombe?

CHAPTER III. BIOINFORMATIC ANALYSIS OF NMD PROTEINS IN *S. POMBE*

3.1 Introduction

It has already been reported that the genome of *S. pombe* contains orthologs of NMD factors such as UPF1 and UPF2 (Mendell et al., 2000; Rodriguez-Gabriel et al., 2006). However, a systematic search for putative homologs of all known NMD factors has not yet been performed. In this chapter, I present the results of such a bioinformatic search. The study aimed to identify all the possible orthologs of NMD proteins in *S. pombe*, and also to analyze the domain organization of the proteins and their molecular evolution by using multiple sequence alignments, from single cell organisms, such as yeast, to metazoa.

3.2 Results

3.2.1 The orthologs of UPF proteins in *S. pombe*

At the start of the project, I searched for homologs of the *S. cerevisiae* proteins, UPF1, UPF2 and UPF3 in the *S. pombe* complete protein database (http://www.genedb.org/) by omniBlast; The other protein sequences were retrieved from at the NCBI website (http://www.ncbi.nlm.nih.gov/sites/entrez?db =protein) and shown in Table 3.1.

Table 3.1 The homologs of NMD factors.

	Human	X. laevis	D. melanogaster	A. thaliana	C. elegans	S. pombe	S. cerevisiae
UPF1	Q92900	AAH73441	NP_572767	AAL92018	NP_490829	NP_593080	NP_013797
UPF2	AAG60689	NP_001089661	NP_572434	NP_181459	NP_500974	NP_593784	NP_011944
UPF3	AAG48511 (3B) NP_075387 (3A)	ENSXETG000 00007440	NP_726375	NP_001117406	CAD18872	NP_593705	CAA97074
SMG1	NP_055907	156595	NP_727132	No.	AAD48773	No	No
SMG5	NP_056142	AAI23019*	NP_609685	No.	NP_491929	No	No
SMG6	NP_060045	356705*	NP_651321	No.	NP_497566	No	No
SMG7	NP_775179	NP_989258	No.	NP_197441	NP_501033	N.A.	N.A.
eIF4AIII	NP_055555	AAH84859	NP_649788	NP_188610	AAB96704	NP_592863	NP_010304
MAGO	NP_002361	NP_001079724	NP_476636	NP_171716	P49029	NP_596666	No
Y14	NP_005096	NP_001083872	NP_610454	NP_564591	CAA83626	NP_594439	No
RNPS1	NP_006702	Q5XG24	NP_649903	NP_173107	AAK21429	NP_596549	No
MLN51	O15234	NP_001089026	NP_733229	NP_850944	NP_493346	No	No

Underlined accession numbers are derived from http://www.xenbase.org/ by *blastp*. Asterisk means that the sequences are not identified by *blastp* of SMG5/6 from *C. elegans*, but identified by using SMG5/6 from human. N.A. means that the sequences are not identified by *blastp*, although studies suggested the homolog in *S. cerevisiae*.

I identified orthologs of the three genes: *SPAC16C9.06c* (*UPF1*)*, SPAC19A8.08*
(*UPF2*) and *SPAC13G7.03* (*UPF3*). Protein sequence alignment with the
corresponding orthologs in *A. thaliana, C. elegans, D. melanogaster*, Human*, S.
cerevisiae, S. pombe* and *X. laevis*, showed the UPF1 and UPF2, are highly conserved
from yeast to human (Fig 3.1 and 3.2), while UPF3 is much less conserved (Fig 3.3).

S. pombe UPF1 shares 55.3% amino acid identity with human UPF1and 53.5% with
S. cerevisiae UPF1. The fact that *S. pombe* UPF1 is slightly closer related to the
human ortholog than *S. cerevisiae* UPF1. High sequence identity is apparent in two
domains of UPF1, the cysteine-histidine rich domain (CH domain) and the RNA
dependent ATPase/helicase domain. As described in the Introduction, the CH domain
binds zinc and is involved in the interaction with UPF2 and eRF3 (Weng et al., 1996a;
He et al., 1997; Serin et al., 2001; Kadlec et al., 2006; Ivanov et al., 2008; Clerici et
al., 2009). The helicase domain in the middle part of the protein is essential for NMD
(Czaplinski et al., 1995; Weng et al., 1996a; Bhattacharya et al., 2000; Cheng et al.,
2007). As mentioned previously, the carboxyl termini of *C. elegans* and human UPF1
have serine/threonine-glutamine rich (SQ-rich) motifs, which act as the
phosphorylation sites of UPF1 and then regulate NMD via a phosphorylation/
dephosphorylation cycle of UPF1 (Page et al., 1999; Pal et al., 2001; Yamashita et al.,
2001). The alignment clearly shows that the SQ-rich motifs are present only in the
UPF1 homologs derived from multicellular species. The lack of SQ motifs suggests
that the phosphorylation of UPF1 might not be required for NMD in both yeasts or
that at least the phosphorylation sites of UPF1 is very different between yeast and
metazoa, although a study has reported the UPF1 and UPF2 are also phospholylated
in *S. cerevisiae* (Wang et al., 2006).

Xl-UPF1
At-UPF1
Ce-UPF1
Dm-UPF1
H-UPF1
Sc-UPF1
Sp-UPF1

CH domain

ATP binding

To next page
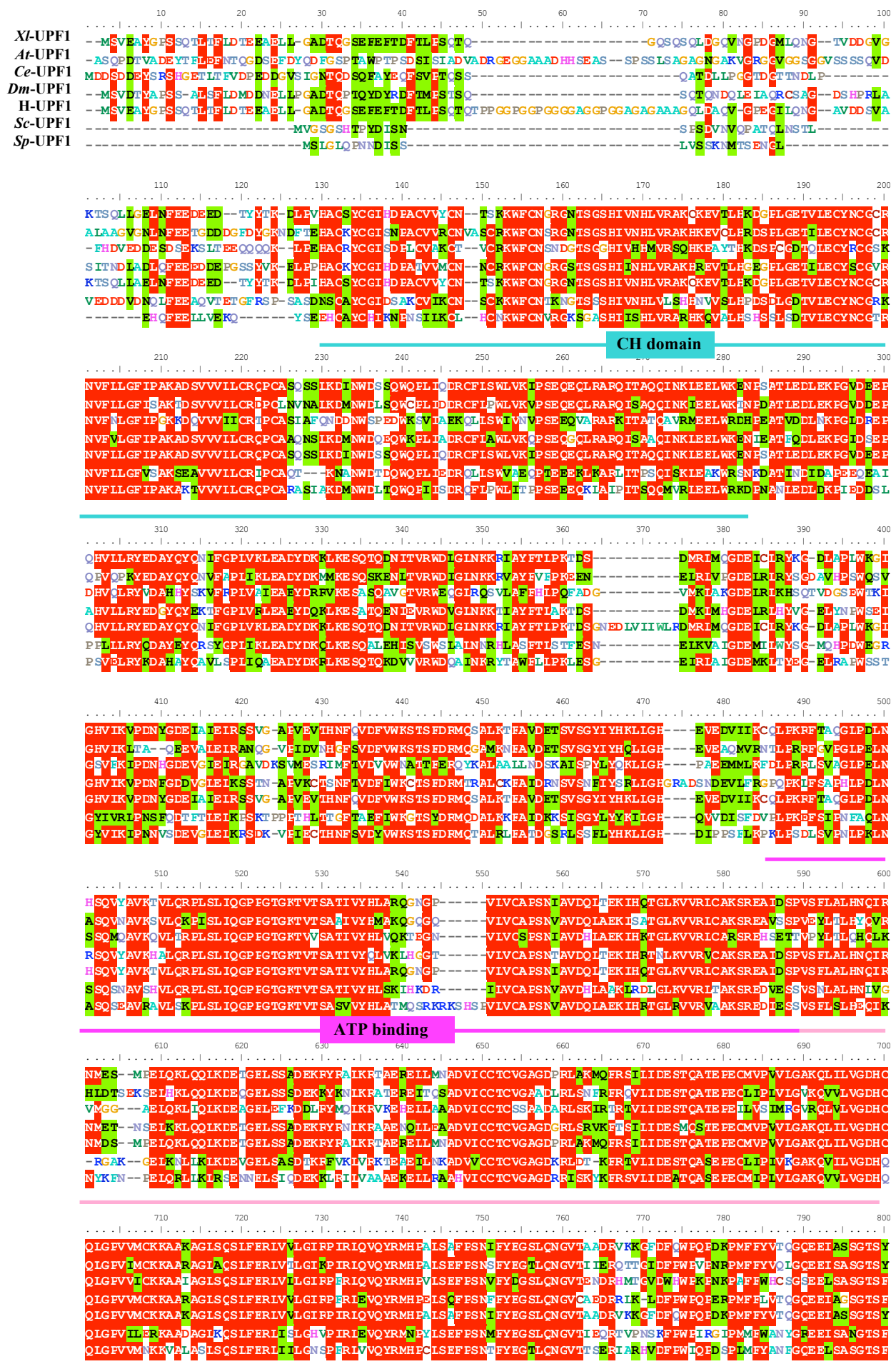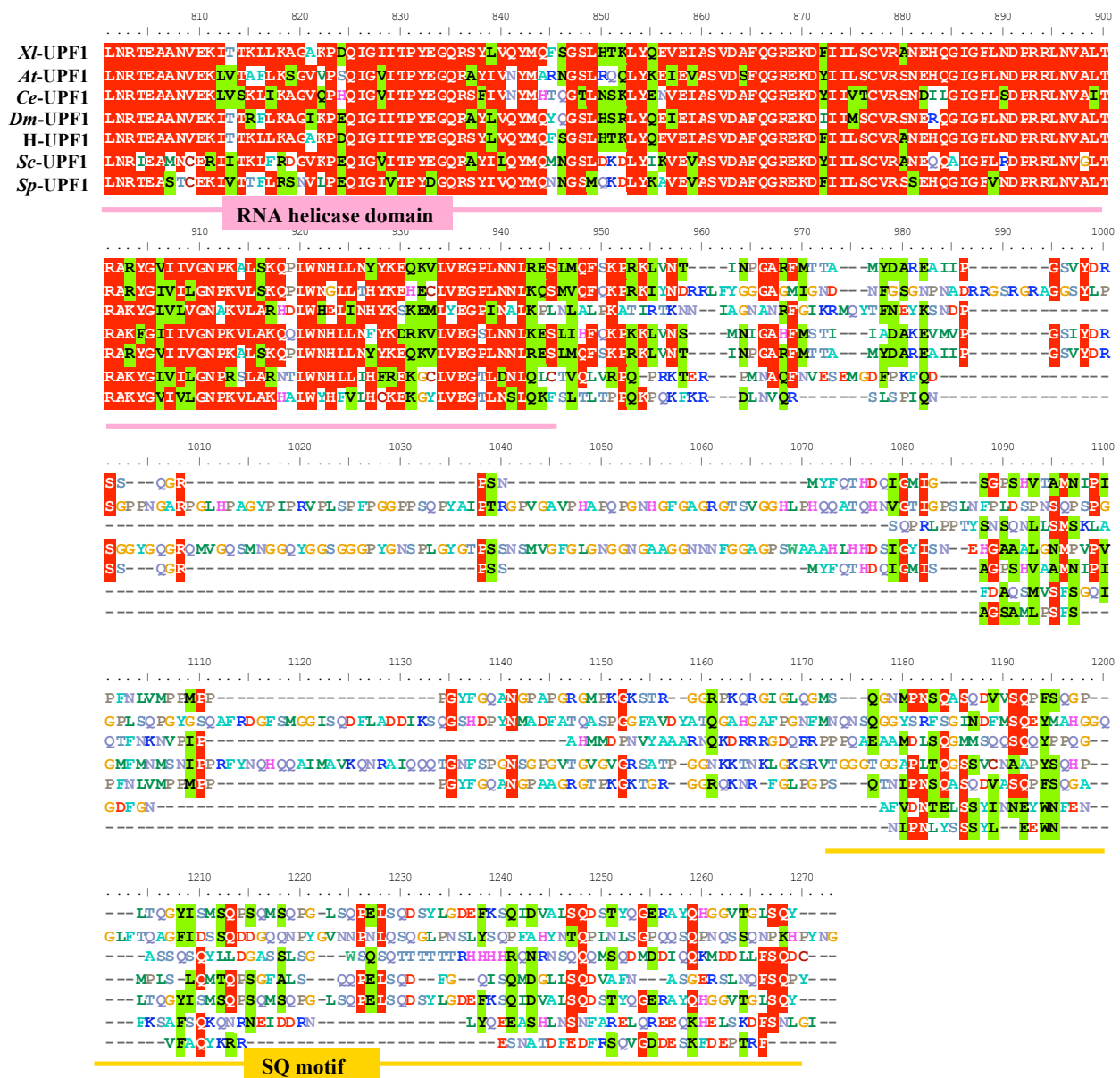
Figure 3.1 UPF1 protein analysis.

Sequence alignment of UPF1 in *S. cerevisiae* (*S.c.*)*, S. pombe* (*S.p.*)*, C. elegans* (*C.e.*)*, A. thaliana* (*A.t.*)*, D. melanogaster* (*D.m.*), *X. laevis* (*X.l.*) and human (H). Protein sequences are listed in Table 3.1. The alignment was performed with ClustalW, BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical residues are highlighted in red; similar residues in green.

Multiple sequence alignment of UPF2 showed 22.5% identity of amino acids between *S. pombe* and human, 19.8% between *S. cerevisiae* and human, and 16.5% between *S. cerevisiae* and *S. pombe*. UPF2 contains three consecutive well-conserved MIF4G domains (middle domain of eukaryote initiation factor 4G) and the last MIF4G mediates the interaction with the RRM domain of UPF3 (He et al., 1997; Serin et al., 2001; Kadlec et al., 2004). In UPF2, downstream of the MIF4G domains, there is the conserved carboxyl terminus that interacts with UPF1 (He et al., 1997; Serin et al., 2001). Additionally, the amino terminus of human UPF2 contains two functional domains, the nuclear localization signals which are not conserved in *S. cerevisiae* and *S. pombe*, and an extra UPF1 interaction domain (residues 93 -133) which is not present in both yeasts (Fig. 3.2) (Serin et al., 2001). In contrast, *S. cerevisiae* UPF2 is phosphorylated at an amino-terminal serine residue, and the residues in the amino-terminus mediate the interaction with Hrp1p and modulate NMD (Wang et al., 2006). The region between the third MIF4G and UPF1 binding domain, an extraordinary long Glu and Asp rich sequence, termed as acidic amino-acid rich domain (Ac domain), has been reported as an essential motif to interact with eRF3 in *S. cerevisiae* and human cells (He et al., 1997; Serin et al., 2001). As for the UPF1 interaction and eRF3 interaction regions, the Ac domain is not well conserved between *S. cerevisiae* and the other organisms, especially *S. pombe* (Fig. 3.2).

Unlike UPF1, UPF3 is not well conserved between yeast and metazoa. The protein is poorly conserved even between *S. cerevisiae* and *S. pombe*. Protein sequence alignment showed the 20.7% amino acid identity between *S. pombe* and human, 9.5% between *S. cerevisiae* and human, and 12.9% between *S. cerevisiae* and *S. pombe*. The UPF2 interaction region is well conserved between the two yeast species and the other organisms. The carboxyl terminus of human UPF3 has been previously shown

Xl-UPF2
At-UPF2
Ce-UPF2
Dm-UPF2
H-UPF2
Sc-UPF2
Sp-UPF2

Hrp1p binding domain

MIF4G

MIF4G
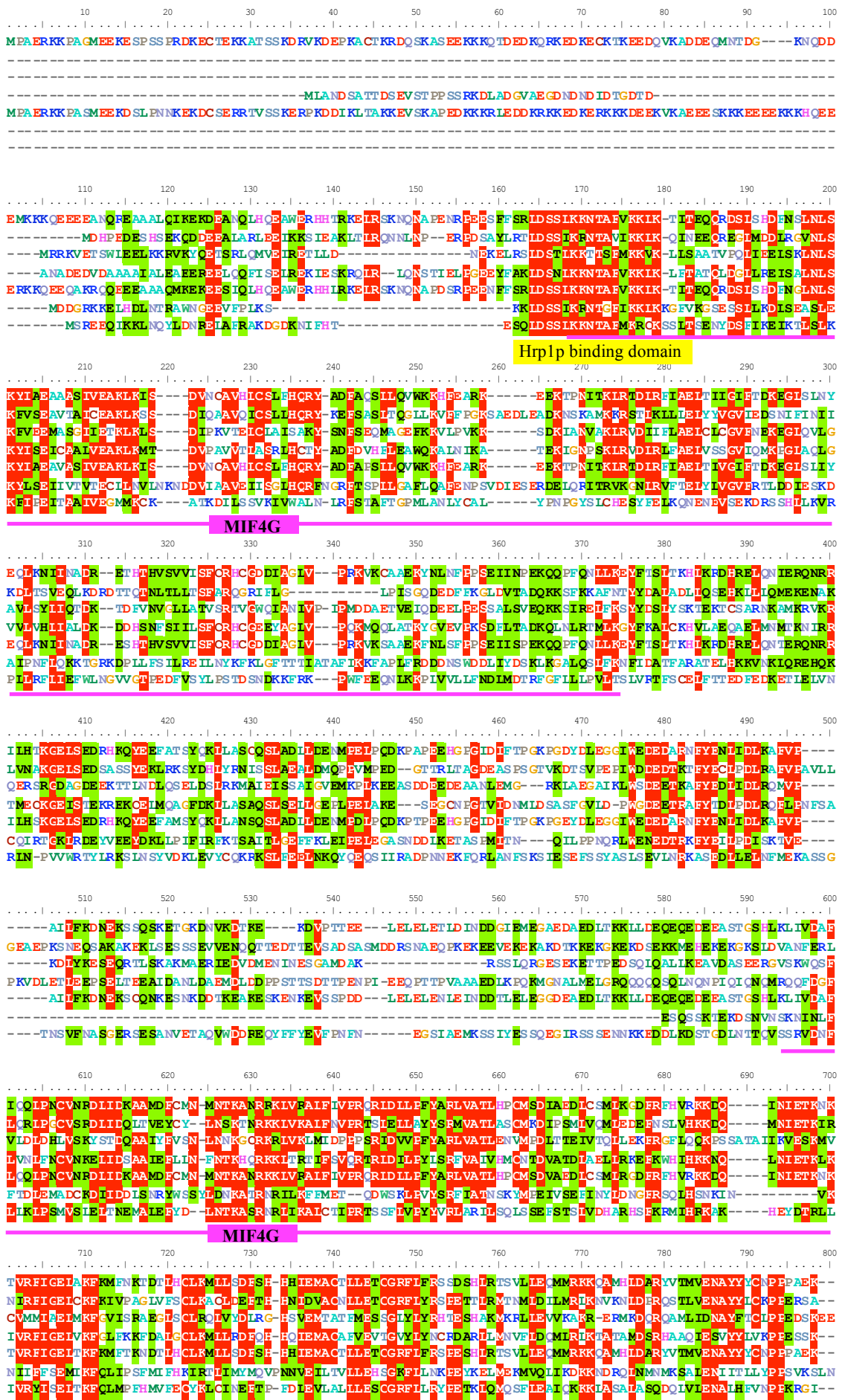
To next page

Figure 3.2 UPF2 protein analysis.

Sequence alignment of UPF2 in *S.c., S.p., C.e., A.t., D.m., X.l.* and human. Protein sequences are listed in Table 3.1. The alignment was performed with ClustalW, BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical residues are highlighted in red; similar residues in green.

Figure 3.1 UPF3 protein analysis.

Sequence alignment of UPF3 in *S.c., S.p., C.e., A.t., D.m., X.l.* and human. Protein sequences are listed in Table 3.1. The alignment was performed with ClustalW, BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical residues are highlighted in red; similar residues in green.

to interact with Y14. This region is highly conserved in human, fly and *Xenopus* (Kim

et al., 2001a). However it is not presented in *S. pombe* and *S. cerevisiae* UPF3 (Fig.

3.3).

3.2.2 SMG1, and SMG 5-7 homologs

As mentioned in the introduction, SMG1 is the protein kinase that phosphorylates

UPF1 in *C. elegans, D. melanogaster* and humans. SMG1 has not been reported in the

rest of the model organisms, especially in yeast. In *X. laevis*, the putative protein

(Table 3.1) is identified as the conserved SMG1 ortholog by *Blastp* in Xenobase

(http://www.xenbase.org). The homolog of SMG1 seems also to be identified as TOR

(target of rapamycin, Phosphatidylinositol-3 kinase family) in *A. thaliana, S. pombe*

and *S. cerevisiae*, but they are questioned (Yamashita et al., 2001; Brumbaugh et al.,

2004).

The multiple sequences alignment indicated that only the kinase catalytic domain is

highly conserved, while the carboxyl terminal domain, and amino terminus are not

well conserved (data not shown). The catalytic domain is similar to that of other

protein kinases such yeast, TOR1 and TOR2, which are the subunits of TORC1 (TOR

complex 1), a main regulator for cell growth in response to environmental signals

(Heitman et al., 1991; Helliwell et al., 1994). However, there is no evidence that TOR1/2

or TORC1 are involved in mRNA stability in yeast – most probably TOR1 and TOR2

are not the homologs of SMG1 in yeast. Therefore, there is no indication of which

could be the kinases that phosphorylate UPF1 and UPF2 in *S. cerevisiae* and *S.*

pombe.

I could not find apparent homologs of SMG5, SMG6, and SMG7 in *S.cerevisiae* and

S. pombe. However some studies in *S. cerevisiae*, indicate that Ebs1p, as the homolog

of SMG6, SMG6 and SMG7. The shared region with Ebs1p is the14-3-3 domain. It is not clear whether Ebs1p is important for NMD in *S. cerevisiae:* while one study concluded Ebs1p is involved in NMD, the other suggested no effect (Ford et al., 2006; Luke et al., 2007). I also identified the Ebs1p homolog in *S. pombe,* but it is not well conserved relative to SMG7 (data not shown).

3.2.3 EJC components

By searching the *S. pombe* protein database for putative orthologs of EJC components, I identified orthologs of MAGO (Fig. 3.6), Y14 (Fig. 3.7) and RNPS1 (Fig. 3.5). However the putative *S. pombe* Y14 only shares the similarity with the carboxyl terminus of other Y14 proteins, and appears to lost the amino terminal domain (Fig. 3.7). The lack of the amino terminal region in *S. pombe* Y14 has been previously reported (Lau et al., 2003; Shi and Xu, 2003). I could not identify any protein similar to MLN51 (barentsz in *D. melanogaster*). Compared to other EJC components MLN51 is also not very well conserved, even amongst multicellular organisms (Fig. 3.8). As in *S. cerevisiae*, there is also a clear ortholog of eIF4AIII (Fig. 3.4). All of them show the high conservations across organisms. The presence of clear orthologs of EJC components in *S. pombe* suggested that splicing might be required for NMD also in this organism, similar to mammalian cells. Alternatively, it is possible that these proteins might have a non-NMD function as reported in *C. elegans* and fly (see discussion below).

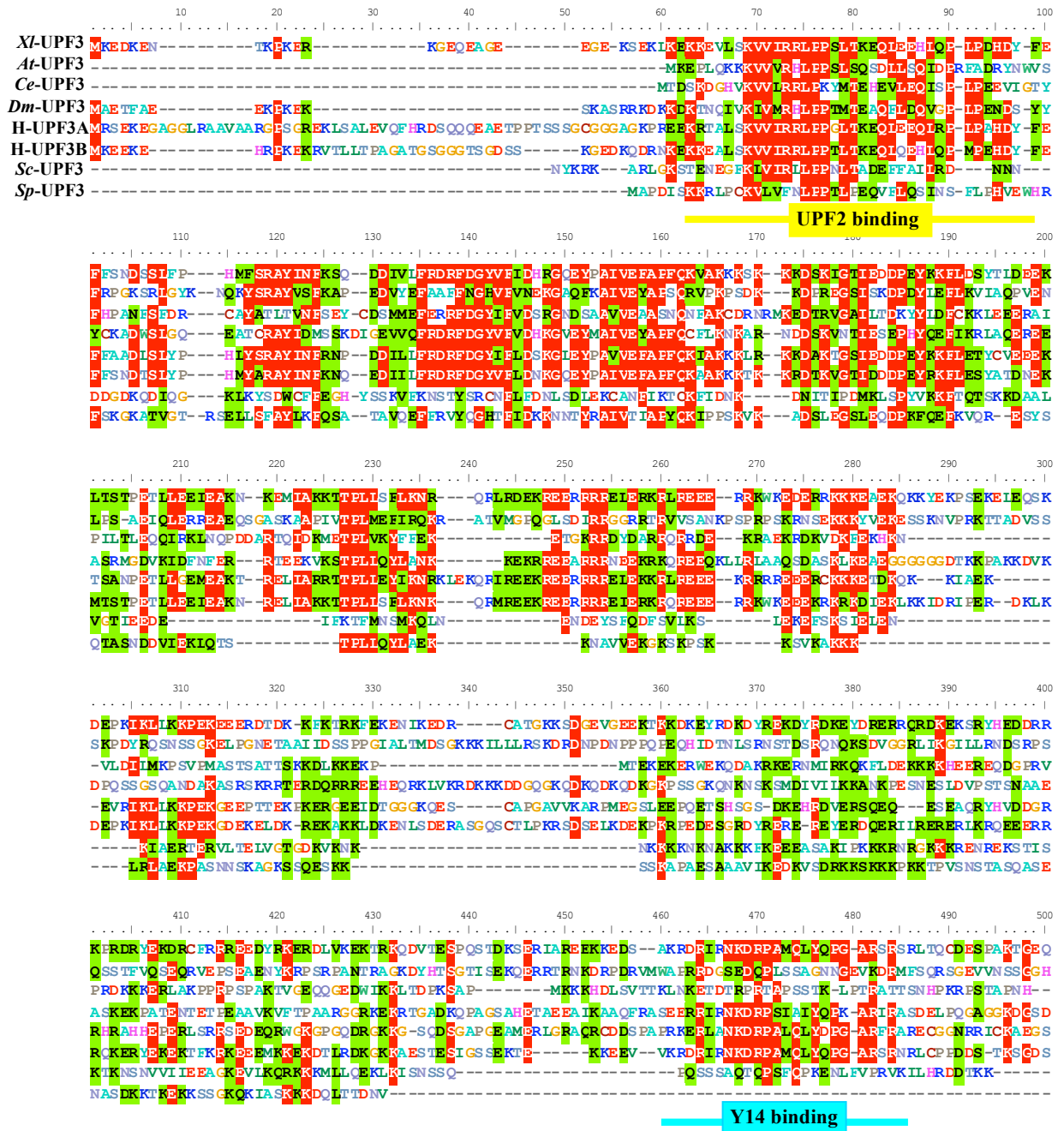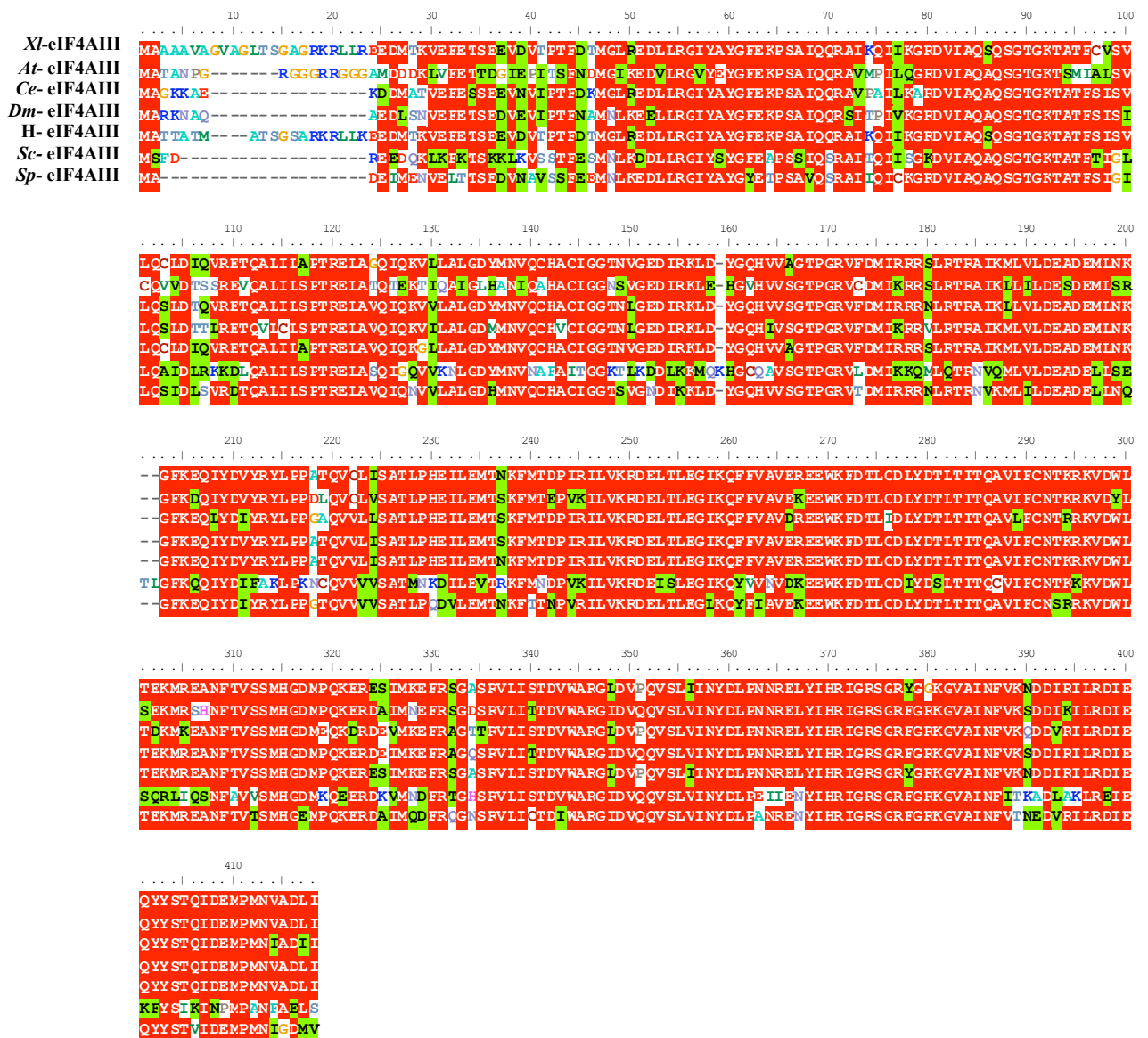**Figure 3.4 eIF4AIII protein analysis.**

Sequence alignment of eIF4AIII in *S.p., C.e., A.t., D.m., X.l.* and human. Protein
sequences are listed in Table 3.1. The alignment was performed with ClustalW,
BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical
residues are highlighted in red; similar residues in green. No such homolog is
identified in *S. cerevisiae*.
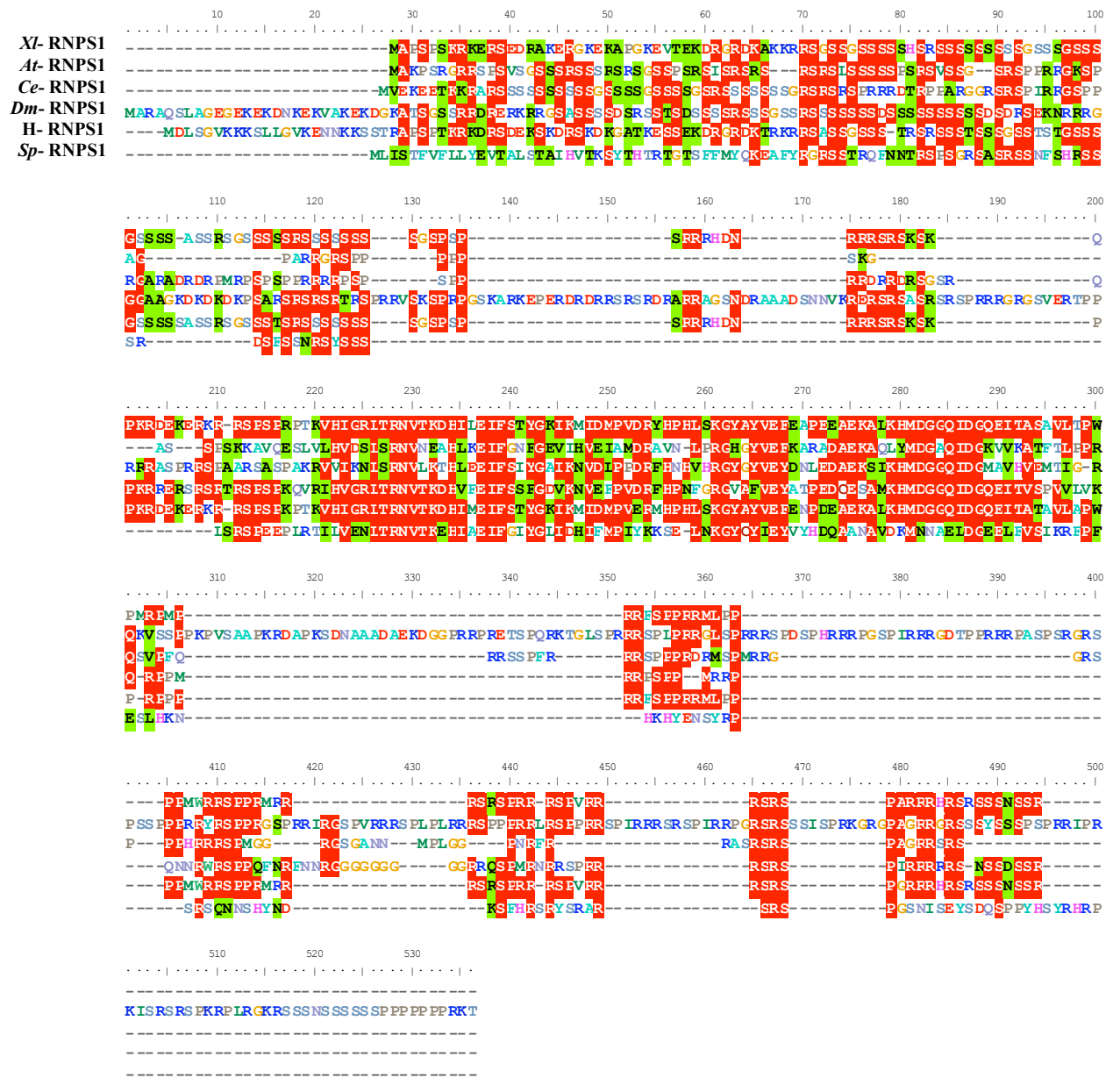
Figure 3.5 RNPS1 protein analysis

Sequence alignment of RNPS1 in *S.p., C.e., A.t., D.m.*, *X.l.* and human. Protein
sequences are listed in Table 3.1. The alignment was performed with ClustalW,
BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical
residues are highlighted in red; similar residues in green. No such homolog is
identified in *S. cerevisiae*.

Figure 3.6 MAGO protein analysis

Sequence alignment of MAGO in *S.p., C.e., A.t., D.m.*, *X.l.* and human. Protein sequences are listed in Table 3.1. The alignment was performed with ClustalW, BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical residues are highlighted in red; similar residues in green. No such homolog is identified in *S. cerevisiae.*

Figure 3.7 Y14 protein analysis

Sequence alignment of Y14 in *S.p., C.e., A.t., D.m., X.l.* and human. Protein sequences are listed in Table 3.1. The alignment was performed with ClustalW, BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical residues are highlighted in red; similar residues in green. No such homolog is identified in *S. cerevisiae*.
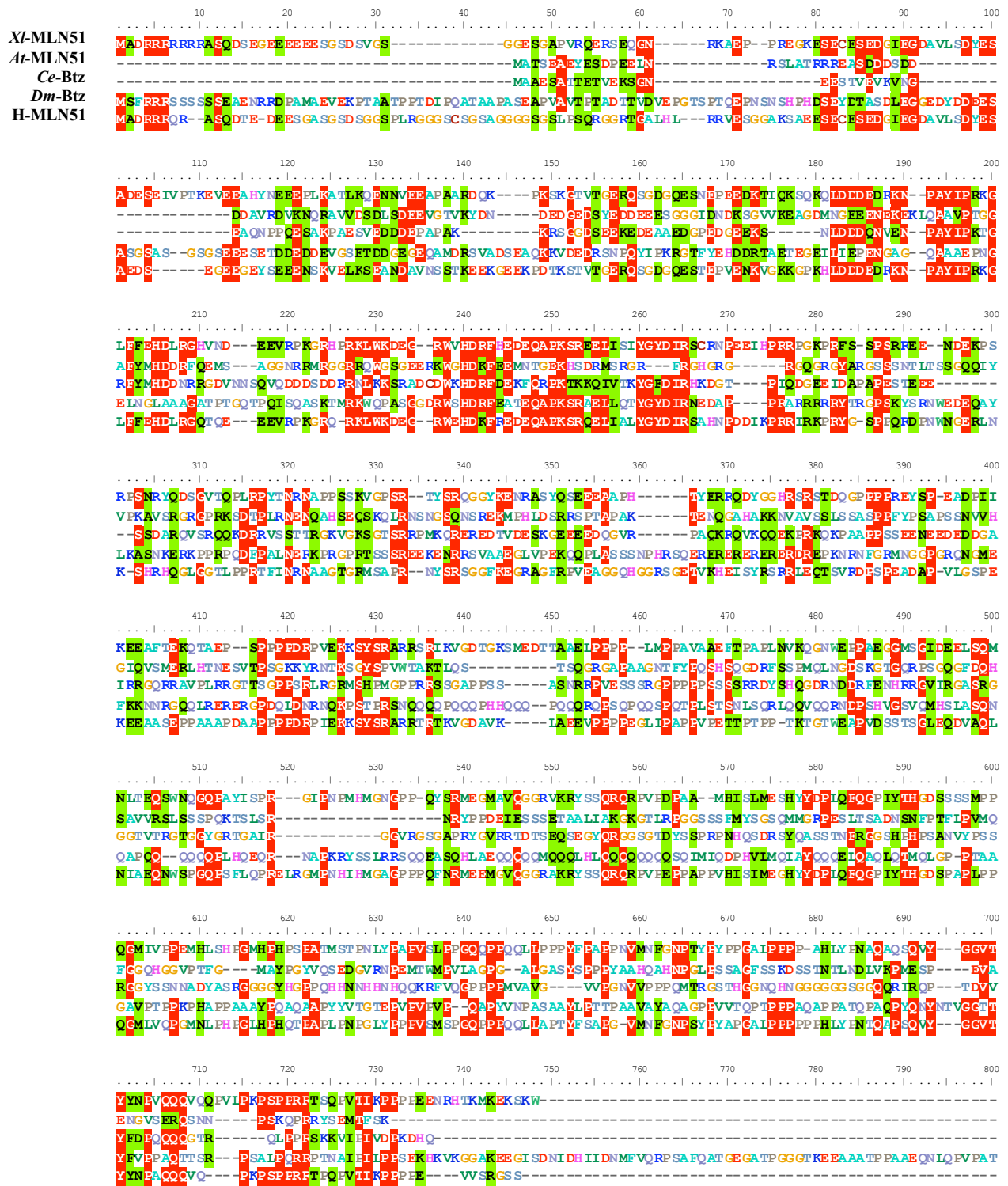
Figure 3.8 MLN51 protein analysis

Sequence alignment of MLN51 in *C.e., A.t., D.m., X.l.* and human. Protein sequences are listed in Table 3.1. The alignment was performed with ClustalW, BioEdit (described in http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Identical residues are highlighted in red; similar residues in green. No homologs of MLN51 are identified in *S. pombe* and *S. cerevisiae.*

3.3 Discussion and implications for this research project

The fact that UPF1, UPF2 and to a less extent of UPF3, are conserved in *S. pombe* implies that the mechanism of NMD is probably similar to *S. cerevisiae*. However, the fact that many EJC components are also present in fission yeast, and that 43% of genes typically have introns in this organism raises the question of whether splicing is required for NMD in *S. pombe* (Wood et al., 2002)? The presence of EJC components alone is not a sufficient indication of a link between splicing and NMD. In fact, several studies have reported that NMD does not require splicing in *C. elegans* and *Drosophila* S2 cells, despite EJC components being present in these two organisms (Gatfield et al., 2003; Longman et al., 2007). In addition in *S. cerevisiae*, a clear eIF4AIII homolog is present but still here is no evidence of a link between NMD and splicing. Instead *S. cerevisiae* eIF4AIII has an essential function in the biogenesis and maturation of ribosomal RNA (Kressler et al., 1997). Based on the fact that most of EJC components are present and given that about half of the mRNAs are spliced, I decided to investigate whether splicing and the EJC are involved in NMD in fission yeast.

The fact that *S. pombe* UPF1 lacks the SQ motifs at the C-terminal and absence of SMG5, SMG6 and SMG7 homologs indicate that in *S. pombe*, UPF1, is not regulated by phosphorylation and dephosphorylation, unlike in mammalian cells, *C. elegans* and *D. melanogaster*. But the observation that UPF1 and UPF2 are the phospoproteins in *S. cerevisiae* suggest that that this issue should be still investigated (Wang et al., 2006). In addition, the absence of an obvious SMG6 homolog suggests that there might not have endonucleolytic cleavage in yeast NMD, unlike in *D. melanogaster* and human cells (Huntzinger et al., 2008; Eberle et al., 2009).

CHAPTER IV. THE *FAUX* 3' UTR MODEL DOES NOT EXPLAIN NMD IN *S. POMBE*

4.1 Introduction

PTCs located within the 5' half of the coding region usually trigger stronger NMD than PTCs within the 3' half. This polar effect of PTCs is typical of NMD in *S. cerevisiae* (Losson and Lacroute, 1979; Peltz et al., 1993; Yun and Sherman, 1995). A similar polar effect has also been observed in *D. melanogaster* in the *Adh* gene, *Cat* gene and in *GFP* (Brogna, 1999; Gatfield et al., 2003; Behm-Ansmant et al., 2007). In addition, immunoglobulin µ gene reporters also show that 3' proximal PTCs do not trigger strong NMD in human cells (Buhler et al., 2006; Eberle et al., 2008)

The polar effect is currently explained by the *faux* 3' UTR model (Amrani et al., 2004; Eberle et al., 2008). This model proposes that translation termination at NMD-inducing PTCs is biochemically abnormal and inefficient because it occurs away from the normal 3' UTR; while the termination at normal stop codons is more efficient due to the interaction with certain *trans acting* factors located in the 3' UTR. The original *faux* 3' UTR model proposed that poly(A) tail binding protein (PABP) is the most important among these termination-enhancing factors. It has been suggested that cytoplasmic poly(A) binding protein (PABPC) plays a pivotal role in NMD as some studies have shown that artificial tethering of PABPC downstream of a PTC strongly inhibits mRNA reductions in budding yeast and S2 cells in *D. melanogaster* (Amrani et al., 2004; Behm-Ansmant et al., 2007; Eberle et al., 2008). However, other studies have reported that transcripts without the poly(A) tail can still be subject to NMD, both in mammalian cells and in budding yeast (Maquat and Li, 2001; Meaux et al.,

2008). It is then proposed other unknown factors recruited to the 3' UTR might be involved in NMD (Singh et al., 2008).

In this chapter, I have undertaken to investigate whether the *faux* 3' UTR model could be applied in fission yeast and whether PABPC is required for NMD. In particular, I have investigated whether NMD is affected by positioning of PTCs along the gene, whether the length and nature of the 3' UTR is a key NMD determinant and whether the distance between the 5' end and a PTC is also important for NMD.

4.2 Results

4.2.1 NMD reporters used in the study

To study NMD, I generated the plasmid reporters expressing green fluorescent protein (GFP). The reporters are based on the pREP41 plasmid, a widely used fission yeast shuttle vector regulated by a derivative of the *nmt1* (no message thiamine) promoter, which allows episomal expression (Basi et al., 1993; Maundrell, 1993). Based on this GFP construct, three further constructs were made where the nonsense mutation (TAA) was introduced in the coding region at three different positions, replacing codons 6, 27 and 140 respectively (named as pPTC6, pPTC27 and pPTC140) (Fig. 4.1A).

4.2.2 NMD depends on the position of the PTC along the coding region

For the assessment of NMD, the three PTC-containing reporters were transformed along with the PTC-less control into yeast, either a wild-type strain (SPJK039) or *upf1Δ* (SPJK033) and *upf2Δ* (SPJK029) mutants. Strong GFP fluorescence was observed both in the cytoplasm and in the nucleus in the control, but not in the PTC-

Fig. 4.1

A

nmt41 prom 1   243 nmt1 term   pGFP
                                                                         pPTC6
                                                                         pPTC27
                       GFP   pPTC140

PTC6   PTC27   PTC140

B ——— *wt* ———

GFP PTC6 PTC27 PTC140

1   2   3   4

% 100 16±7 43±4 82±20

C ——— *upf2Δ* ———

GFP PTC6 PTC27 PTC140

1   2   3   4

% 100 73±6 85±8 114±13

D ——— *upf1Δ* ———

GFP PTC6 PTC27 PTC140

1   2   3   4

% 100 67±9 65±12 103±7

— GFP

— Rpl32
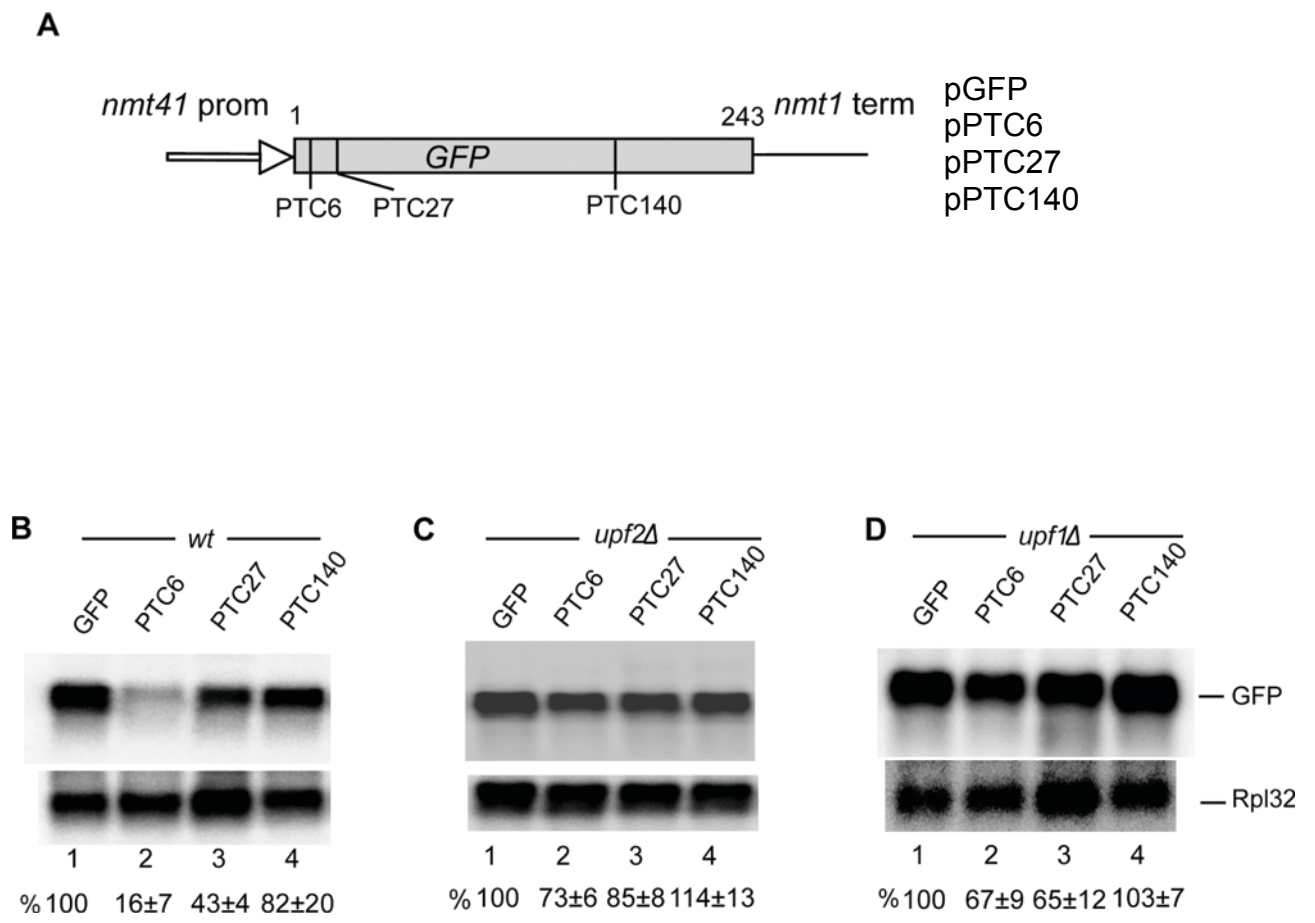
Figure 4.1 NMD polarity effect in *S. pombe*

(A) Diagram of the GFP reporters. Nonsense mutations (TAA) were introduced at codon positions 6, 27 and 140. Detailed information is in Materials and Methods (Chapter II).

(B-D) Northern blot analysis of total-RNA from wild-type (B), *upf2Δ* (C) and *upf1Δ* (D) cells, transformed with the plasmids indicated. Top panels show hybridization with a GFP specific probe; bottom panels show hybridization with a probe specific for the large ribosomal protein 32 (*Rpl32*) mRNA, as a loading control. The GFP probe was PCR amplified with primers gfp-ks and gfp-rev, and the rpl32 probe with rpl32-sens and rpl32-rev (Appendix II Table 1). The values below each lane are percentages with standard deviation of the GFP mRNA present in the PTC-less control, lane 1. Bands intensities were quantified with a phosphor-imager and standardized to the RpL32 signal (described in Materials and Methods, Chapter II). Calculations are based on three independent experiments.

containing constructs (Fig. 4.8D). Total RNA was extracted and analyzed by Northern blot, using ^{32}P labeled probes. The PTC6 mutation triggered strong mRNA reduction in wild-type cells: the mRNA was less than 20% of that of the PTC-less control (Fig. 4.B, lane 2). The PTC27 mutation triggered a moderate mRNA reduction (43%) and PTC140 triggered only a minor mRNA reduction (82%) (Fig. 4.1B, lanes 3-4).

As expected, the mRNA levels were restored in both *upf1Δ* and *upf2Δ* cells: the mRNA levels were from 80% to 110% of that of the PTC-less control (Fig. 4.1C and 1D). These results are consistent with the expectation that this mRNA reduction is the result of the canonical NMD, requiring UPF1 and UPF2 as in other organisms. The observation that early PTCs trigger stronger mRNA reduction also indicates that NMD in fission yeast is likely to be polar as reported in budding yeast and fly.

4.2.3 Lengthening of the 3' UTR only partially enhances NMD.

The *faux* 3' UTR model proposes that the distance between a PTC and the 3' UTR is the key determinant of NMD. The model is supported by studies in *S. cerevisiae, D. melanogaster* and human cells (Amrani et al., 2004; Eberle et al., 2008; Singh et al., 2008). To investigate whether the distance of the PTC from the 3' end is also important for NMD in *S. pombe*, I inserted either a 102 bp or 418 bp DNA fragment (derived from the firefly luciferase coding region) 6 nt downstream of GFP normal stop codon, immediately upstream of the *nmt1* 3' UTR sequence (Fig 4.2A). The constructs were named pGFP-102, pPTC6-102, pPTC27-102, pPTC140-102, pGFP-418, pPTC6-418, pPTC27-418, and pPTC140-418. According to the *faux* 3' UTR model, it was expected that both insertions would enhance NMD as the distance between the PTC and the poly(A) tail is longer. For example, in PTC27-102 the distance between the PTC and the poly(A) tail is 890 nt, while in the original PTC6

reporter is 813 nt; the distances between the PTC and the poly(A) tail in PTC27-418 and PTC140-418 are 1106 nt and 867 nt, respectively: these three constructs were expected to trigger strong NMD, similar to PTC6 in the original reporter.

RNA analysis showed that the 102-nt insertion significantly enhanced NMD of the mRNA with PTC27 (27% vs. 43%, Fig. 4.2B lane 7 vs. lane 3). However, the 102-nt insertion did not enhance NMD in the reporter with PTC6 (Fig. 4.2B lane 2 and lane 6). Furthermore, the level of the PTC140-102 mRNA was reproducibly higher than the one with the original 3' UTR (Fig. 4.2B lane 4 vs. lane 8). In the reporters with the 418-nt insertion, the 418-nt insertion significantly decreased the level of the mRNAs with PTC27 and PTC140, 24% and 53% respectively of the mRNA produced by the GFP reporter with the original 3' UTR (Fig. 4.2C lane 4 and lane 5); but the insertion did not affect the mRNA level of PTC6 (Fig. 4.2C lane 3). Notably, there was no relationship between the distance of the PTC from the 3' end and mRNA reduction. For example, it was expected that the 418-nt insertion would make PTC140 behave like PTC6 in the original reporter, triggering strong NMD; instead the PTC140 mRNA level was hardly reduced compare to the PTC-less control (Fig. 4.2C). Consistent with the view that mRNAs with an artificially long 3' UTR turn to be NMD substrates, the mRNA without PTCs was also reduced to about 50% of that with the normal 3' UTR (Fig. 4.2C lane 2) and the mRNA level was restored in the *upf2Δ* strain (Fig. 4.2D lane 2).

To further investigate whether NMD was affected by the distance between the PTC and poly(A) tail, a new set of constructs were made by inserting a 402 bp DNA fragment into the GFP coding region, 6 bp upstream of the normal stop codon

Fig. 4.2
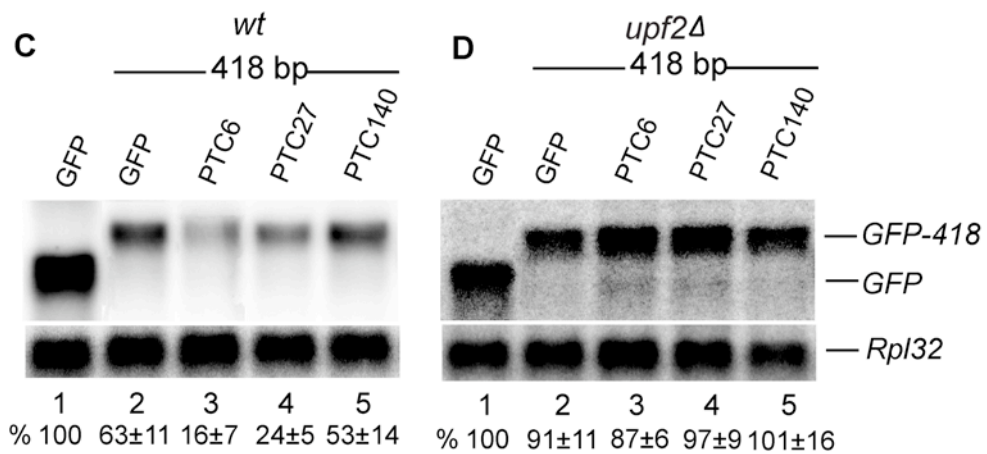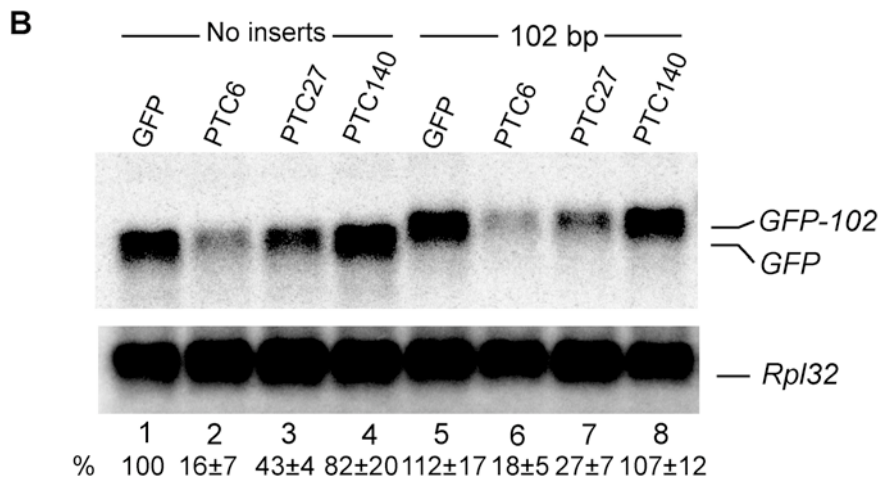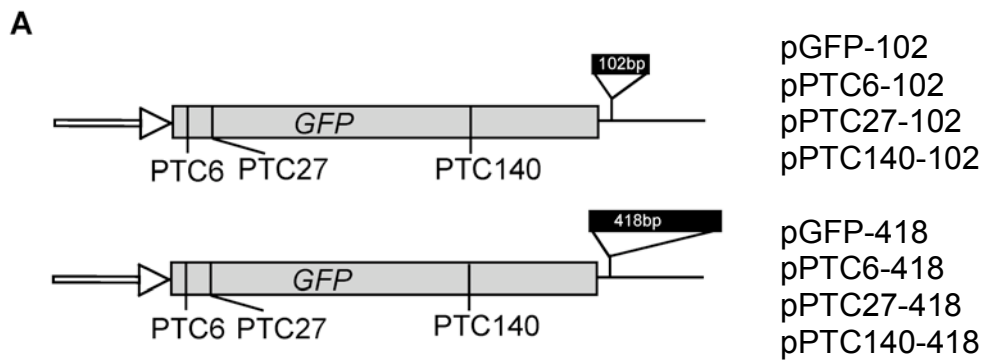
Figure 4.2 NMD does not strongly correlate to the length of the 3' UTR

(A) Diagram of reporters with different length of 3' UTRs. The pGFP-102 reporter carries a 102 bp insert at 9 bp downstream of the normal stop codon, and pGFP-418, a 418 bp DNA fragment at the same position. The inserted DNA fragments were cloned from firefly luciferase coding region, described in Material and Methods (Chapter II).

(B-D) Northern blot analysis of total-RNA from wild-type (B and C) and *upf2Δ* (D) cells, transformed with the indicated plasmids. Top panels show hybridization with a GFP specific probe as in Figure 4.1B; bottom panels with *Rpl32* mRNA as a loading control as in Figure 4.1B. The values below each lane are percentages with standard deviation from three independent experiments.

(Fig. 4.3A). The constructs were named as pGFP-402, pPTC6-402, pPTC27-402, and pPTC140-402.

The PTC140-402 and PTC27-402 were expected to show strong NMD: in the PTC140-402 reporter, the PTC is at the same distance from the poly(A) tail as PTC6 in the original reporter (851 nt); in the PTC27-402 reporter the PTC is further away from the poly(A) tail, 1190 nt. RNA analysis showed that PTC6 still induced strong NMD (13%, Fig. 4.3B lane 2), like in the original reporter. However, NMD was only partially enhanced in the PTC27-402 (27% vs. 43%, Fig. 4.3B lane 3) and in PTC140-402 reporters, compare to the original reporters (63% vs. 84%, Fig 4.3B lane 4). The mRNA levels were restored in the *upf2Δ* strain, consistent with the expectation that these mRNA reductions were the result of NMD (Fig 4.3C).

In summary these data indicate that the distance between a PTC and the poly(A) tail is only a minor determinant of NMD in *S. pombe*. These data seems not to be consistent with the key prediction of the *faux* 3' UTR model, which states that the strong NMD is mainly depended on an abnormally long distance between a PTC and the 3' end.

4.2.4 The nature of the 3' UTR does not affect NMD

To investigate whether the nature of the 3' UTR is important for NMD, we exchanged the *nmt1* 3' UTR of my NMD reporters with that of two other genes. One was the 3' UTR of the *ade6* gene. This 3' UTR was chosen because it has been previously reported that nonsense mutations in this gene triggers strong NMD in *S. pombe* (Mendell et al., 2000). The other 3' UTR from the *adh* gene was chosen because this gene is constitutively expressed at high level (data shown in Chapter V, Fig. 5.4A, lane 1-4).
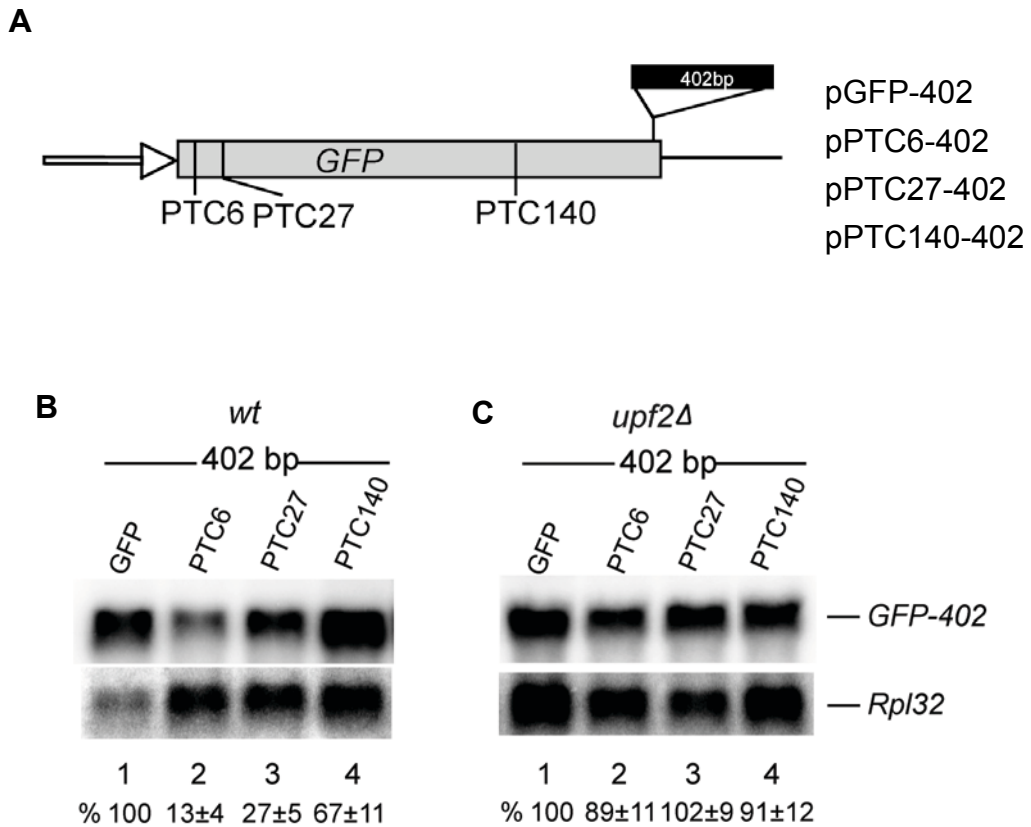
Figure 4.3 Lengthening the distance between a PTC and the 3' UTR does not greatly affect NMD.

(A) Diagram of GFP reporters for long distance. A 402 bp DNA fragment from firefly luciferase gene was cloned 6 bp upstream of the normal stop codon.

(B and C) Northern blot analysis of total-RNA from wild-type (B) and *upf2Δ* (C) cells, transformed with the plasmids indicated. Top panels show hybridization with a GFP specific probe as in Figure 4. 1B; bottom panels with a probe specific for *Rpl32* mRNA, a loading control as in Figure 4.1B. The values below each lane are percentages with standard deviation from three independent experiments.
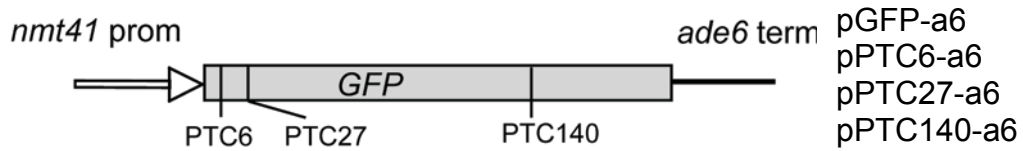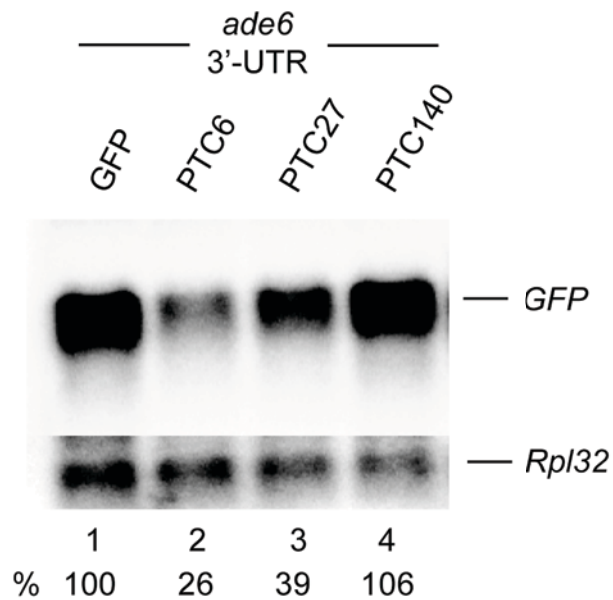
Figure 4.4 The nature of 3' UTR does not affect NMD in *S. pombe.*

(A) Diagram of GFP reporters with *ade6* 3' UTR from *S. pombe ade6* gene. Three nonsense mutations (TAA) are the same as described in Figure 4.2A.

(B) Northern blot analysis of total-RNA from wild-type cells, transformed with the plasmids indicated. Top panels show hybridization with a GFP specific probe as in Figure 4.1B; bottom panels with *Rpl32* as a loading control as in Figure 4.1B.

RNA analysis of these reporters with the ade6 3' UTR shows a similar NMD pattern: the mRNA levels of PTC6, 27 and 140, relative to the PTC-less control, are comparable to that of the reporters with the *nmt1* 3' UTR (Fig. 4.4B). These results suggest that the nature of the 3' UTR sequences might not have a major role in NMD.

4.2.5 Shortening of the distance between a PTC and the poly(A) tail only modestly affects NMD

The data reported above indicate that lengthening distance between a PTC and the poly(A) tail only slightly affects NMD in *S. pombe*. Here I tested the second prediction of the *faux* 3' UTR model: shortening of the distance between a NMD-inducing PTC and the 3' end should prevent NMD. In particular, I reasoned that the drastic mRNA reduction associated with PTC6 should be suppressed by shortening of the 3' UTR. To test this prediction, new constructs were made with a shorter GFP coding region (327 nt instead of the normal 729 nt) (Fig. 4.5A). In the shorter GFP reporter, named as pGFP-sh and pPTC6-sh, the distance between PTC6 and the poly(A) tail is the same between PTC140 and the poly(A) tail in the initial reporter with the full-length GFP (449 nt). RNA analysis showed a strong mRNA reduction of the PTC6-sh reporter (11% of the PTC-less control, Fig. 4.5B lane 2) Unexpectedly, the PTC6-sh mRNA level was only partially stabilized in *upf2Δ*, showing 42% mRNA level normalized by GFP-sh (Fig. 4.5B, lane 4).

The northern blot showed an additional transcript with an approximate length of 1300 nt. This transcript was probably generated due to the inefficient termination at the canonical poly(A) site; the polymerase probably read-through the first poly(A) site and allows processing at an alternative poly(A) site located further downstream of the stop codon. To evaluate this possibility, the 1020 nt *nmt1* 3' UTR sequence was

analyzed bioinformatically with two poly(A) prediction programs: "*S. cerevisiae* 3'

end processing predictor" and human "PolyApred" (Graber et al., 2002; Ahmed et

al., 2009). The *S. cerevisiae* 3' end processing predictor program (but not the

PolyApred) predicted a poly(A) signal at about 140 nt downstream of normal stop

codon; this is the experimentally verified main polyadenylation of the *nmt1* gene

(Aranda and Proudfoot, 1999). Whereas both the prediction programs identified a

strong poly(A) signal about 850 nt further downstream; this polyadenylation site

probably accounts for the long transcripts in the northern blot (Fig. 4.5B and 5C). The

S. cerevisiae 3' end processing predictor program identified a weak poly(A) signal in

the middle of the *nmt1* 3' UTR sequence, 460-500 nt from the start of 3' UTR; but in

the northern blot there is no evidence that this site is used (Fig. 4.5). Quantification of

the short and long GFP-sh transcripts in *upf2Δ* indicates that the two poly(A) signals

are used with equal frequency, which strongly suggested that shortening of the ORF

activates downstream poly(A) processing. Furthermore, The long transcripts are

strongly stabilized in the *upf2Δ* strain regardless of whether it carried the PTC6

mutation (Fig. 4.5B, lane 3 and lane 4). This is consistent with the observations that

mRNAs with the native long 3' UTR were destabilized by NMD in *S. cerevisiae*, but

this finding also contradicts the prediction according to the *faux* 3' UTR model where

PTC6 would be expected to cause a more intense NMD in the transcripts with the

longer 3' UTR; instead of what I found: that the level of the longer transcript was not

affected by the PTC6.

Fig. 4.5

A

pGFP-sh

pPTC6-sh

B

C
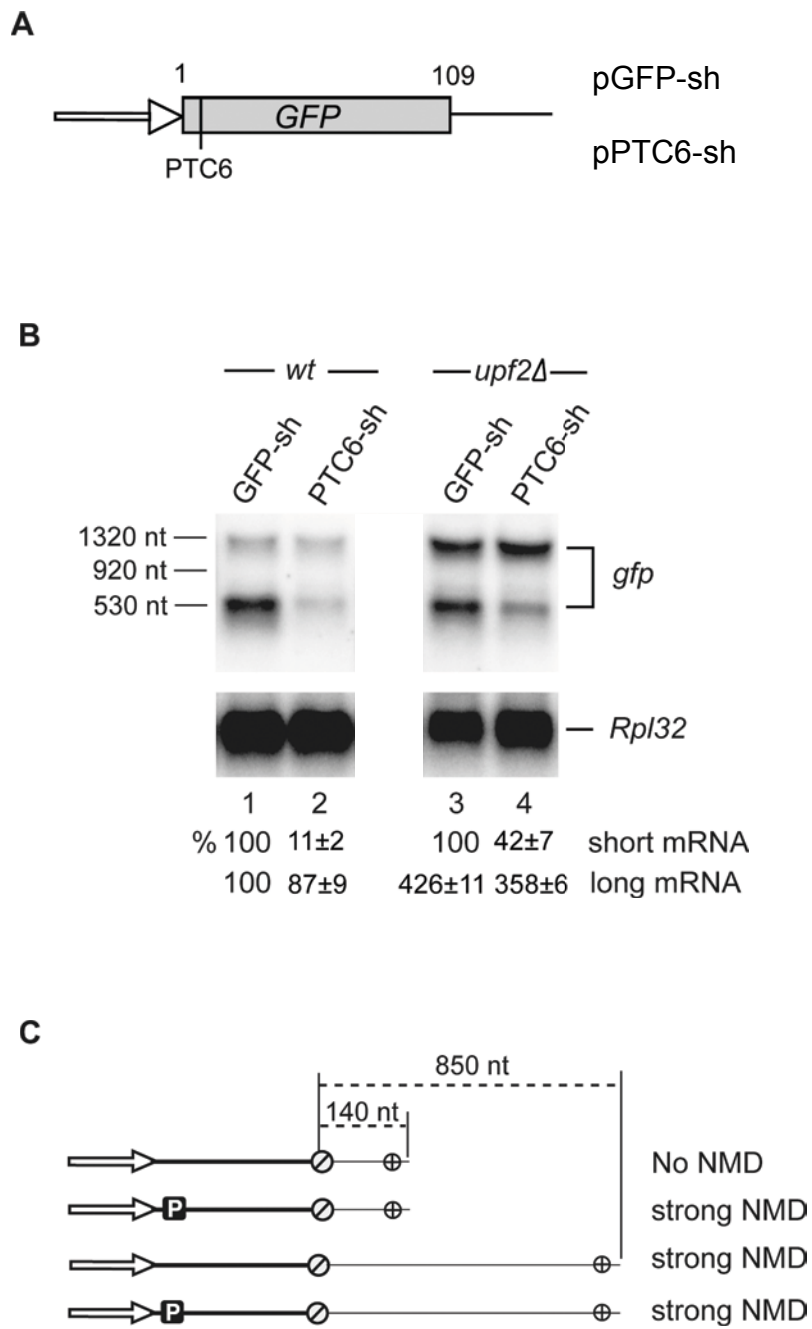
Figure 4.5 Shortening of the distance between a PTC and the poly(A) tail does not prevent NMD.

(A) Diagram of reporters with a deletion derivative of GFP. The 3' half of GFP was deleted until the normal stop codon (402 bp deletion).

(B) Northern blot analysis of total-RNA from wild type and *upf2Δ* cells. Top panels show hybridization with a GFP specific probe as in Figure 4.1B; bottom panels with a probe specific for *Rpl32* mRNA, as a loading control as in Figure 4.1B. The values below each lane are percentages with standard deviation from two independent experiments.

(C) Diagram of the transcripts with the short or the long 3' UTR, generated from the two polyadenylation sites.
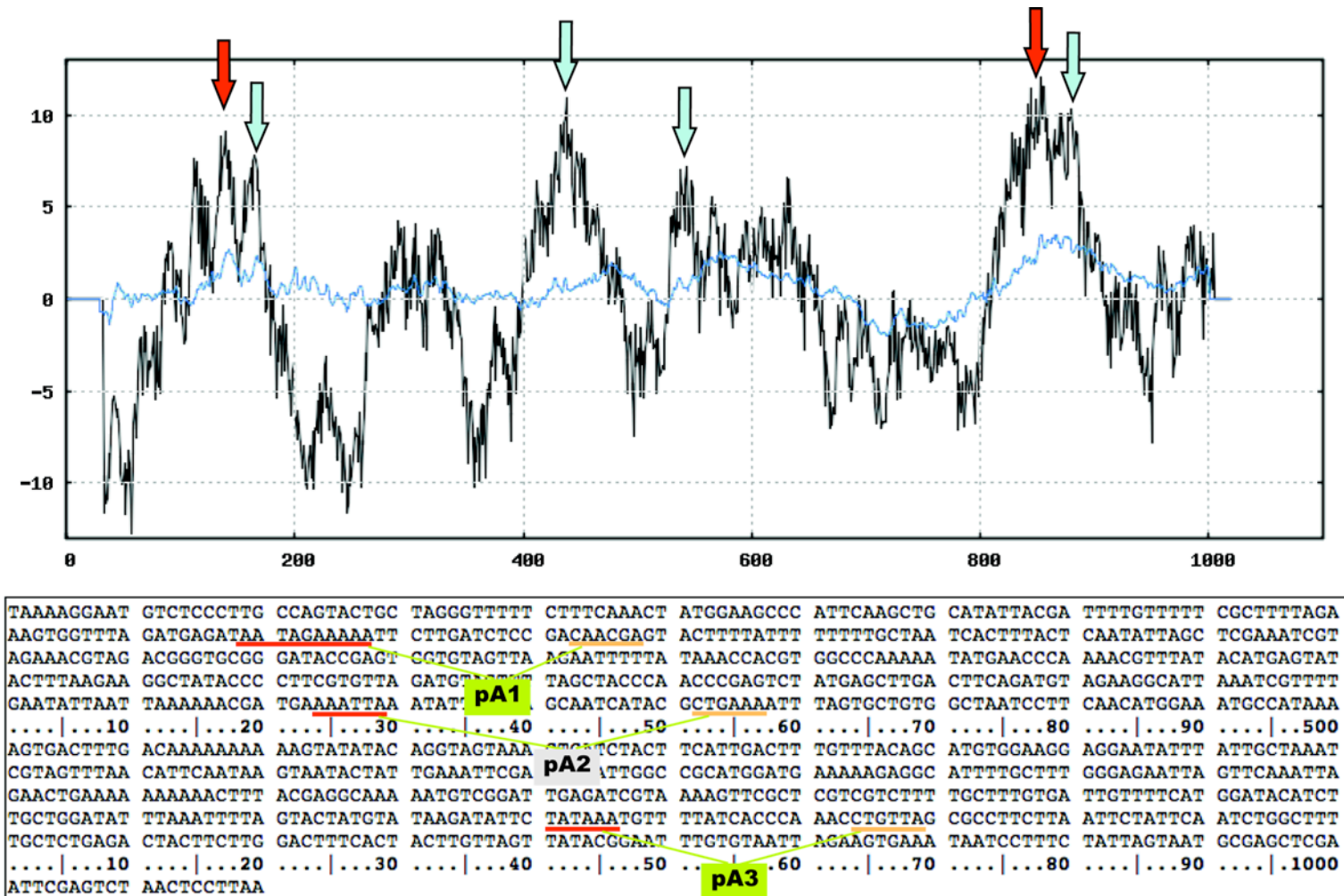
Figure 4.6 Prediction of polyadenylation signal sequence and cleavage site in *nmt1* 3' UTR.

The *nmt1* 3' UTR was analyzed with the "*S. cerevisiae* 3' end processing predictor" program. The most probable poly(A) sites are indicated by arrows; the red arrows indicate the sites that appear to be used in the NMD reporter (pA1 and pA3 shown in sequences); blue arrows indicate predicted sites that do not appear to be used in the reporters. Underlined in red are the poly(A) signal sequences, in orange putative cleavage sites.

4.2.6 Cytoplasmic PABP is not essential for NMD in *S. pombe*

The third prediction of the original *faux* 3' UTR model is that cytoplasmic PABP (PABPC) is important for NMD. To elucidate the function of PABPC in *S. pombe* NMD, the NMD reporters were transformed in a knock out strain, *pab1Δ* (Thakurta et al., 2002). *PAB1* is the only gene encoding the PABPC homolog in *S. pombe*. Unlike in *S. cerevisiae*, *S. pombe* cells with the deletion of *PAB1* are viable; although the deletion causes slow growth of the cells (Thakurta et al., 2002). RNA analysis showed strong NMD of the PTC6 reporter, moderate NMD with PTC27 and a weak NMD with PTC140; the results are very similar to the NMD pattern observed in the wild-type strain (Fig. 4.7). The data indicate that in *S. pombe*, PABPC does not play an important role in NMD, suggesting that NMD might be independent of PABPC.

4.2.7 The half-life of mRNA with a PTC is extended in NMD mutant strain

The primary signature of NMD is the reduction of steady-state of mRNA caused by a reduced mRNA stability of the mature PTC-containing mRNA. To assess whether NMD is also due to a reduced mRNA stability in my system, I measured the half-life of the most affected NMD reporter, the PTC6 mRNA, and compared it with a PTC-less control. The half-life of the mRNA was estimated following transcription inhibition with 1, 10 phenanthroline, which is a metal chelating compound that deplete Zn^{2+} from RNA polymerase II (Falchuk et al., 1976). GFP mRNA levels were assayed at six regular time intervals, 0, 2, 10, 20, 40, and 60 min, after 1,10-phenanthroline was added to inhibit transcription. This drug has been previously used to inhibit transcription and to measure global mRNA stability in *S. cerevisiae* and *S. pombe* cells, and the effects on mRNA stability were similar to it in *rpb1-1*, a temperature-sensitive allele of RNA polymerase II (Rodriguez-Gabriel et al., 2003;

Grigull et al., 2004; Lackner et al., 2007). I found that the PTC6 mRNA decays faster, however the decay rate is not gradual like that of the PTC-less mRNA: it is initially fast but then becomes slow, compared to that of the control (Fig. 4.8A and 8C). This typical biphasic decay profile has been reported in mammalian cells and *S. cerevisiae* (Leeds et al., 1991; Lykke-Andersen et al., 2000; Maquat, 2004; Guan et al., 2006). In *upf2Δ* cells, the PTC6 mRNA was stabilized and the decay kinetics were similar to that of the PTC-less mRNA in the wild-type strain (Fig. 4.8B and 8C). Notably, in the *upf2Δ*, the PTC-less mRNA was also stabilized, indicating that the depletion of UPF2 might also stabilize mRNAs without PTCs. It supported the observation that the fluorescence of GFP in *upf2Δ* was more intense than it seen in wild-type, which has been previously reported as the "photoshop" phenotype in *D. melanogaster* NMD mutants (Fig. 4.8D)
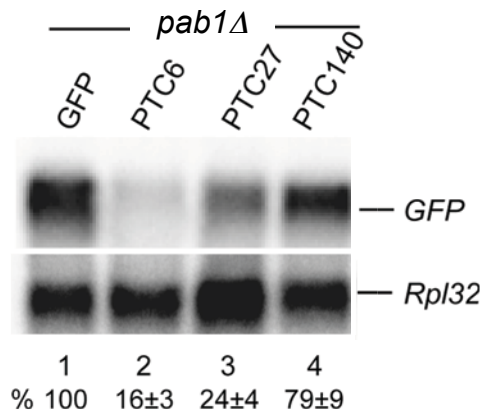
Fig 4. 7 Pab1p is not required for NMD.

Northern blot analysis of total-RNA from *pab1Δ* cells. Top panels show hybridization with a GFP specific probe as in Figure 4. 1B; Bottom panel with a probe specific for *Rpl32* mRNA-the loading control as in Figure 4. 1B. The values below each lane are percentages with standard deviation from three independent experiments.

Fig. 4.8
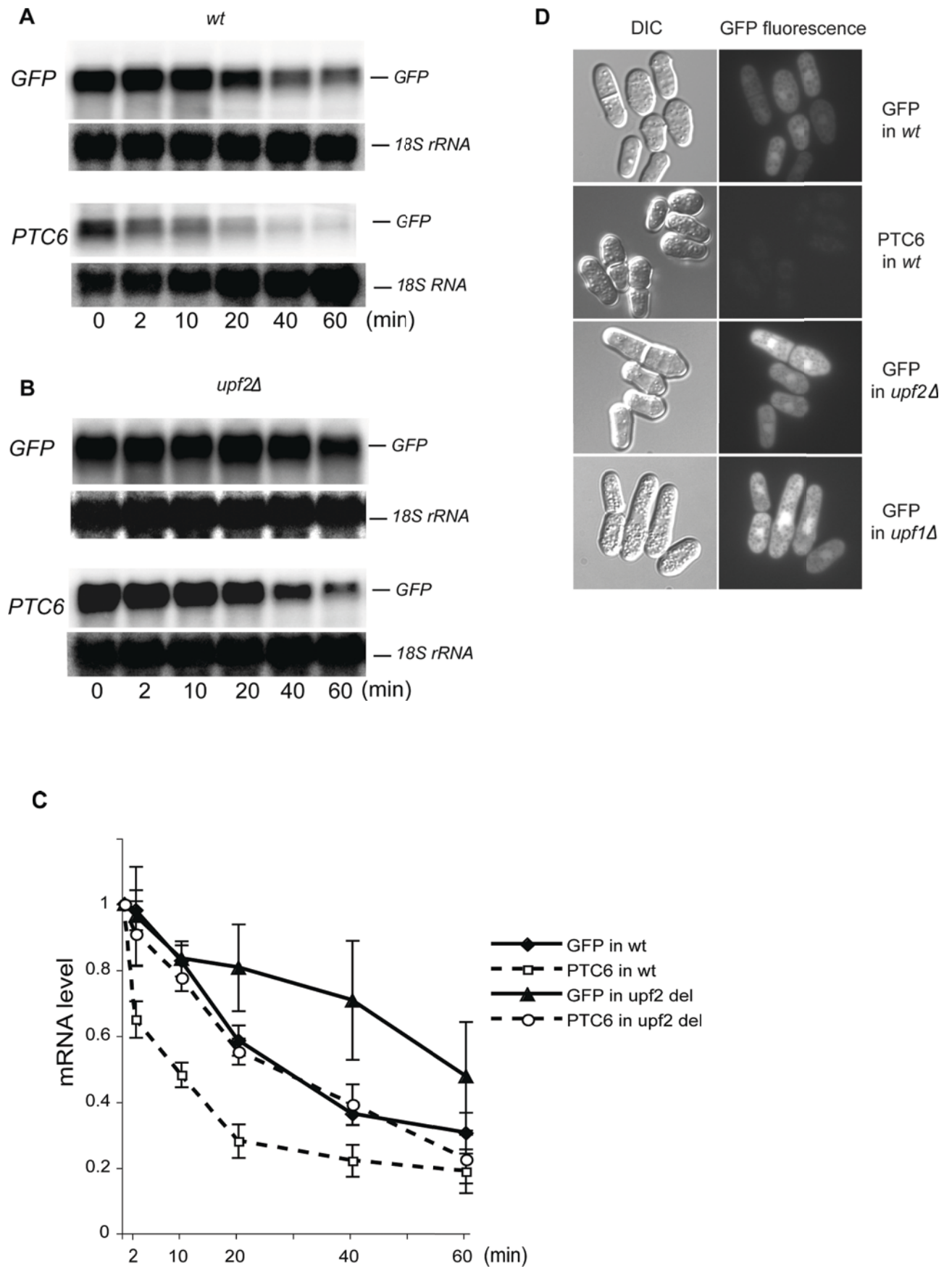
Figure 4. 8 mRNA half-life analysis and GFP fluorescence.

(A-C) Stability of the GFP mRNA in wild type (A) and *upf2Δ* (B) following transcription inhibition with 1, 10-phenanthroline. The PTC-less control (top panel) and PTC6 mRNA (bottom panel) were independently monitored and (C) quantified by Northern blot as mentioned above, except that values were standardized to the 18S rRNA signal (this RNA does not show visible decay over the time period of the experiment). Total-RNA was extracted at 6 different time points after adding the drug: 0, 2, 10, 20, 40 and 60 min. Error bars are standard deviations from three different experiments.

(D) GFP fluorescence analysis by fluorescent microscope. The GFP and PTC in wild type, *upf1Δ* and *upf2Δ*, were observed under the 100X oil lens, and an image captured of GFP fluorescent by CCD camera with an exposure fixed at 0.1112 seconds, as describe in Materials and Methods (Chapter II).

4.3 Discussion

In agreement with what has been reported in *S. cerevisiae* and in *D. melanogaster*, the data presented in this chapter clearly indicate that in my reporter system, NMD is also polar in *S. pombe*: only early PTCs in GFP coding region lead to a drastic mRNA reduction. The more apparent NMD when a PTC is located at the beginning of the gene would suggest that a key NMD determinant is the length of the distance from the 3' end, as indicated by the *faux* 3' UTR model. However, contrary to the *faux* 3' UTR model, I found that NMD does not clearly correlate with the distance between a PTC and the 3' end according to my NMD reporters. For example, neither lengthening of the 3' UTR nor lengthening of the distance between a PTC and the 3' end did strongly enhance NMD of the PTC140 reporter, even though in these new reporters the PTC was as far from the 3' end as PTC6 in the original reporter showing strong NMD. Moreover, shortening of the distance between PTC6 and the poly(A) tail did not suppress NMD. In the reporter carrying the deletion of the 3' half of GFP (pPTC6-sh), PTC6 still triggered strong NMD as much as it did in the original reporter, while it was expected that PTC6 should cause the minimal NMD when close of the 3' UTR. These data challenge the notion that the distance between a PTC and the poly(A) tail is the key NMD determinant.

The stabilization of the extra long transcripts of GFP-sh and PTC6-sh in *upf2Δ* mutants is in agreement with the observation that some transcripts with a long 3' UTR are substrates for NMD. However, as mentioned above, it still contradicts to the *faux* 3' UTR model since the presence of PTC6 does not change the mRNA level, despite the longer distance between the termination codon and the 3' end. This further

challenges the notion that the distance between a PTC and the poly(A) tail is a key determinant of NMD.

In addition, I also found that lengthening of the distance between the start codon and the PTC only minimally affected mRNA levels (data not shown). I did not investigate whether the length of the 5'-UTR affects NMD because it was anticipated that such changes would also change the translation efficiency and therefore complicate the interpretation of the results as reported in a similar study in *S. cerevisiae* (Yun and Sherman, 1995).

The NMD polarity is clearly seen in the NMD reporters carrying different 3' UTRs. Thus NMD efficiency appears not to be related to the nature or length of the 3' UTR. These observations are in agreement with a recent study in *S. cerevisiae*, which reported that the GFP and endogenous mRNA with a ribozyme-generated 3' end is also subjected to NMD, even without the poly(A) tail (Meaux et al., 2008). Further challenging the key role of PABPC in NMD, this study also concluded that in *S. cerevisiae,* PTC-containing mRNAs without poly(A) tails are also subject to NMD even in a strain deleted of Pab1p (Meaux et al., 2008).

However, it is feasible that other proteins associated with the 3' end could affect NMD. Interestingly, a recent study has suggested that Pab2p, the homolog of the nuclear PABP in *S. pombe*, associates to ribosomes, and is not replaced by PABPC in the pioneer round translation (Lemieux and Bachand, 2009). Pab2p might play a role similar to Pab1p in the early stage of translation. Future studies should address whether Pab2p is involved in *S. pombe* NMD.

In summary, these results indicate that the distance of the PTC from the both 5' and 3' end is not a key determinant for the efficient NMD in *S. pombe*. Instead, it is possible

that a downstream sequence element (DSE) plays an important role in the polarity

effect of NMD, as suggested by early NMD models (Zhang et al., 1995; Ruiz-

Echevarria et al., 1998a). This will be further discussed in the Conclusion and

Discussion (Chapter VII).

CHAPTER V. SPLICING ENHANCES NMD IN *S. POMBE*

5.1 Introduction

A number of studies have reported that in mammalian cells, NMD is a nucleus associated event and in some instances occurs only when a PTC is located upstream of at least one intron (Cheng et al., 1994; Carter et al., 1996; Thermann et al., 1998; Zhang et al., 1998a; Zhang et al., 1998b). In addition, it has been reported that NMD affects most of CBP80-associated mRNAs, and that UPF2, UPF3 and EJC associate with CBP80-associated mRNA but not with eIF4E-associated mRNA in human cells (Ishigaki et al., 2001; Lejeune et al., 2002). These studies strongly suggest that newly synthesized transcripts are subjected to NMD more than steady state mRNA associated with eIF4E. The linkage between introns and NMD is the exon-exon junction complex (EJC) – a multiprotein complex deposited on spliced mRNAs about 20-24 nt before the splice junctions, which remains associated with the mRNAs during export to the cytoplasm (Le Hir et al., 2000a; Le Hir et al., 2000b). The EJC probably serves as a binding platform for the NMD factors (Le Hir et al., 2001b; Chamieh et al., 2008). It has been reported that UPF3B directly associates with Magoh/Y14 and RNPS1, three core protein components of the EJC (Kim et al., 2001a; Le Hir et al., 2001b; Lykke-Andersen et al., 2001; Gehring et al., 2003; Singh et al., 2007). The other core components eIF4AIII and MLN51 are also required for efficient NMD in mammalian cells; RNAi depletion and mutation of these proteins strongly inhibit NMD (Shibuya et al., 2004; Gehring et al., 2005).

However, the requirement for a downstream intron in mammalian NMD was recently challenged by many other studies in human cells. First, the PTCs that are closer than 50-55 nt to the last exon-exon junction do still induce strong NMD in TCR-β and Ig-μ

transcripts (Carter et al., 1996; Wang et al., 2002a; Buhler et al., 2004). Second, at least one type of NMD appears to be independent of downstream introns; this alternative pathway is termed "failsafe NMD" and has been observed in human β-globin transcripts (Neu-Yilik et al., 2001; Matsuda et al., 2007). More recent studies have reported that extending or shortening of the distance of a PTC from the 3' end also strongly affects NMD in mammalian cells. And in agreement with the *faux* 3' UTR model, tethering of PABPC downstream position close to the PTC inhibits NMD, (Buhler et al., 2006; Eberle et al., 2008). These investigators concluded that the *faux* 3' UTR model explains NMD better than the EJC model in mammalian cells. Although all core EJC components are conserved in *D. melanogaster* and *C. elegans*, studies in both organisms have revealed that neither splicing nor EJC components are required for NMD (Gatfield et al., 2003; Longman et al., 2007). In *S. cerevisiae* there is no report of splicing enhancing NMD. In addition, most genes don't have introns in budding yeast and the genome does not encode the homologs of EJC components with the exception of eIF4AIII.

Study of the link between splicing and NMD is key to understand the mechanism by which pre-mRNA processing affects downstream events such as translation. To unveil this mechanism it is preferable to use a model organism that is amenable to genetic manipulations. My PhD project aimed to explore whether I could investigate these issues in the fission yeast, *S. pombe*. The genome of this organism encodes most of the EJC core components (described in Chapter III) and introns are commonly distributed in genes; therefore I wanted to investigate whether splicing affects NMD in *S. pombe*. In this chapter, I mainly focus on the relationship between introns and NMD, and whether the EJC is required for the intron-dependent NMD.

5.2 Results

5.2.1 Intron-containing NMD reporters

To investigate the role of splicing in NMD, I inserted an intron in the coding region of the GFP reporter described in Chapter IV – the intron derives from the endogenous S. *pombe* gene *ubc4*. The insertion of the intron splits the GFP coding region at codon 110. To assay NMD, the PTC6, PTC27, PTC140 mutations were introduced into the GFP coding region - the constructs were named pGFP$_{ivs}$, pPTC6$_{ivs}$, pPTC27$_{ivs}$ and pPTC140$_{ivs}$ (Fig. 5.1A).

Among these four constructs, PTC6 and PTC27 are located upstream of the intron, 313 nt and 250 nt away from the intron, respectively; this is more than the 55 nt boundary beyond which introns do not appear to enhance NMD in mammalian cells (Zhang et al., 1998a). Instead PTC140 is 89 nt downstream of the intron. Strong GFP fluorescence was observed in the PTC-less control, suggesting that the intron is correctly and efficiently spliced (Fig. 5.1B). Also, the RT-PCR was performed to confirm the same length of the DNA bands between GFP without intron and GFP with intron, which suggests the intron is spliced correctly in the intron-containing constructs (Fig. 5.1C).

5.2.2 An intron strongly enhances NMD and abolished the polar effect of NMD.

To assess the intron's function in NMD, these constructs were transformed into wild-type and *upf2Δ* strains, the mRNA levels were analyzed by Northern blot. RNA analysis showed that the mRNA level was strongly reduced in all PTC-containing reporters: PTC6 induced a strong mRNA reduction (17% mRNA level of the control),

Fig. 5.1

A

nmt41 prom   1   Intron   GFP   243   nmt1 term   pGFPivs
PTC6 PTC27   PTC140   pPTC6ivs
                                                          pPTC27ivs
                                                          pPTC140ivs

B

wt   upf2Δ

DIC

Fluorescence

GFPivs   ivsPTC6   GFPivs   ivsPTC6

C

GFP   GFP-ivs   PTC27-ivs

2 kb
1 kb
0.5 kb

M   1   2   3

D ———— wt ————

GFPivs   ivsPTC6   ivsPTC27   ivsPTC140

pre-mRNA
GFP

Rpl32

1   2   3   4
% 100   17±5   25±11   20±7

E ———— upf2Δ ————

GFPivs   ivsPTC6   ivsPTC27   ivsPTC140

pre-mRNA
GFP

Rpl32

1   2   3   4
% 100   80±12   105±3   115±7
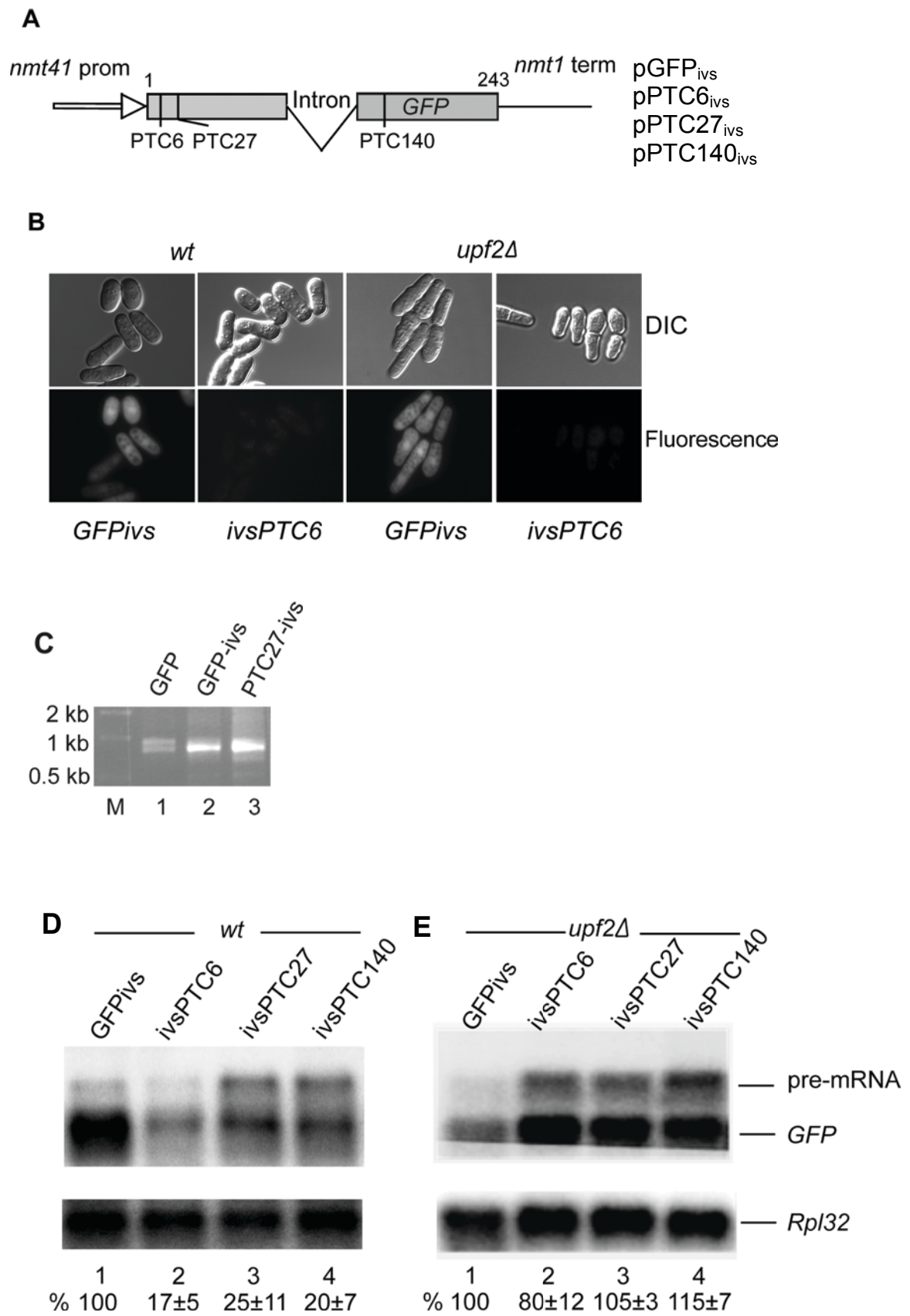
Figure 5.1 The presence of an intron enhances NMD in *S. pombe*.

(A) Diagram of the intron-containing GFP reporters. Nonsense mutations (TAA) were introduced at codon positions 6, 27 and 140. The intron was PCR amplified with primers INubc4-sens and INubc4-rev, see Appendix II Table 1; and inserted between 328th nt and 329th nt of codon 110.

(B) GFP fluorescence of cells transformed with the indicated intron-containing reporters. No GFP fluorescence was observed in cells transformed with PTC6, PTC27 and PTC140 reporters (data not shown for PTC 27 and PTC140).

(C) RT-PCR confirmation of accurate splicing of the intron-containing reporters. Reverse transcription was performed as described in Chapter II. (D and E) Northern blot analysis of total-RNA from wild-type (C) and *upf2Δ* (D) cells transformed with the indicated reporters. The top panels show hybridization with the GFP probe (the slower migrating band corresponds to unspliced pre-mRNA), the bottom panels with the Rpl32 probe. Values are standardized percentages of the GFP mRNA in the PTC-less control, lane 1. Quantifications are based on three independent experiments.

as in the corresponding intronless reporter (Fig. 5.1D, lane 2). However, PTC27 and PTC140 also triggered strong mRNA reductions, 25% and 20% mRNA levels of GFP with intron, respectively, while these mutations did not cause strong NMD in the intronless reporters (Fig. 5.1D, lane 3 and lane 4). The presence of the intron clearly enhanced NMD: the intron caused the additional 22% mRNA reductions in PTC27-intron construct, corresponding to approximate 1 fold increase in NMD, and a 60% further reduction in the PTC140-intron reporter (3 fold increase in NMD), even though PTC140 is located 90 nt downstream of the intron. This suggested that the intron in the GFP coding region strongly enhances NMD (Table 5.1). The presence of the intron also abolished the NMD polarity effect that has been observed in the intron-less reporters. More importantly, in contrast to the prediction of the EJC model, this splicing-dependent NMD enhancement is independent of whether the intron is located downstream or upstream of the PTC. In all cases the mRNA levels were restored in *upf2Δ cells*, indicating that the degradation is elicited by NMD, not from another RNA decay mechanism.

5.2.3 Splicing also enhances NMD in an endogenous gene.

To confirm whether introns enhance NMD in general, a new NMD reporter was generated based on the endogenous gene *SPAC18G6.03*, which encodes a GTPase, named as Ypt3p in *S. pombe*. The gene has two introns in the coding region. I cloned the wild-type gene with the introns and a cDNA derivative without any intron into pREP41. Two nonsense mutations were introduced into the coding region at residue 6 and residue 131 (Fig. 5.2A).

RNA analysis showed that PTC6 and PTC135 cause a strong mRNA reduction in the endogenous gene constructs, 19% and 24% of *ypt3* mRNA levels respectively, and

Fig. 5.2

A

B

C
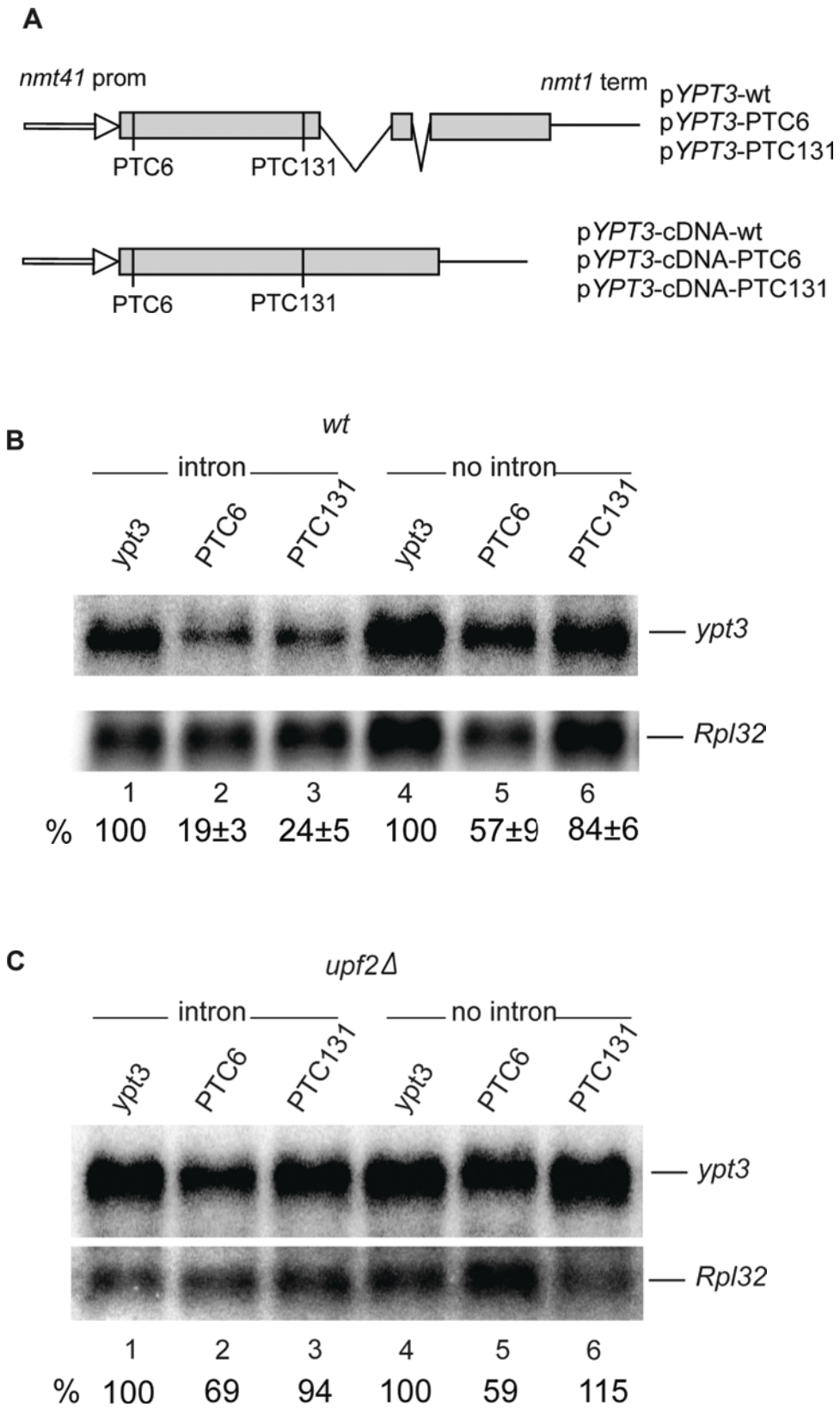
Figure 5.2 Introns are necessary for strong NMD in the *ypt3* gene.

(A) Diagram of the reporters expressing either the wild-type endogenous *ypt3* gene or a cDNA derivative lacking both introns; PTCs are at codon positions 6 and 131, detailed information described in Materials and Methods, Chapter II.

(B and C) Northern blot analysis of total-RNA from wild-type (B) and *upf2Δ* (C) cells transformed with the indicated reporters. Top panels show hybridization with a *ypt3* specific probe which was amplified with primers, ypt3-sens and ypt3-rev (listed in Table 2.1, Materials and Methods); bottom panels with the *RpL32* probe as in Fig. 4. 1B. Values are standardized percentages of the *ypt3* mRNA signal in the PTC-less control, lane 1; standard deviations are from three independent experiments.

the mRNA levels were increased in the *upf2Δ* strain (Fig. 5.2B and 2C). Instead,

minimal mRNA reduction occurred in the cDNA constructs, 57 % of mRNA levels

for PTC6 normalized by the wild-type control and 84% for PTC131 (Fig. 5.2B, lane 5

and lane 6). These data indicate that the intron-dependent enhancement might be

general feature of NMD in *S. pombe*.

5.2.4 Intron-dependent NMD depends on splicing

To further assess whether the intron effect on NMD is splicing dependent, I generated

four additional constructs carrying a mutation at the 5' splice site, <u>GT</u>AG to <u>AT</u>AA

(Fig. 5.3A). Additional information is mentioned in Materials and Methods (Chapter

II). The constructs were named as pGFP$_{ivs-mut}$, pPTC6$_{ivs-mut}$, pPTC27$_{ivs-mut}$ and

pPTC140$_{ivs-mut}$, respectively.

As expected, the 5' splice site mutation almost abolished most of splicing, resulting in

the accumulation of a longer size transcript because the intron failed to be removed. A

PTC leads to an obvious reduction in mRNA level but, like in the original intronless

report, only when the PTC is located at the beginning of the gene: PTC6-5' splice site

mutant (PTC6-ss) still triggered strong mRNA reduction; PTC27-ss triggered

moderated NMD; and PTC140-ss caused minimal mRNA reduction, (Fig. 5.3B). The

absence of strong mRNA reduction in PTC27-intron and PTC140-intron further

indicate that splicing is able to enhance NMD in *S. pombe*. As mentioned above, the

retention of the intron generates additional intron-encoded PTCs; therefore in the

GFP-wt and PTC140 reporter the earliest PTC is actually at the triplets position 110.

It is possible that the unspliced transcripts escape NMD because they are retained

Fig. 5.3

A

B
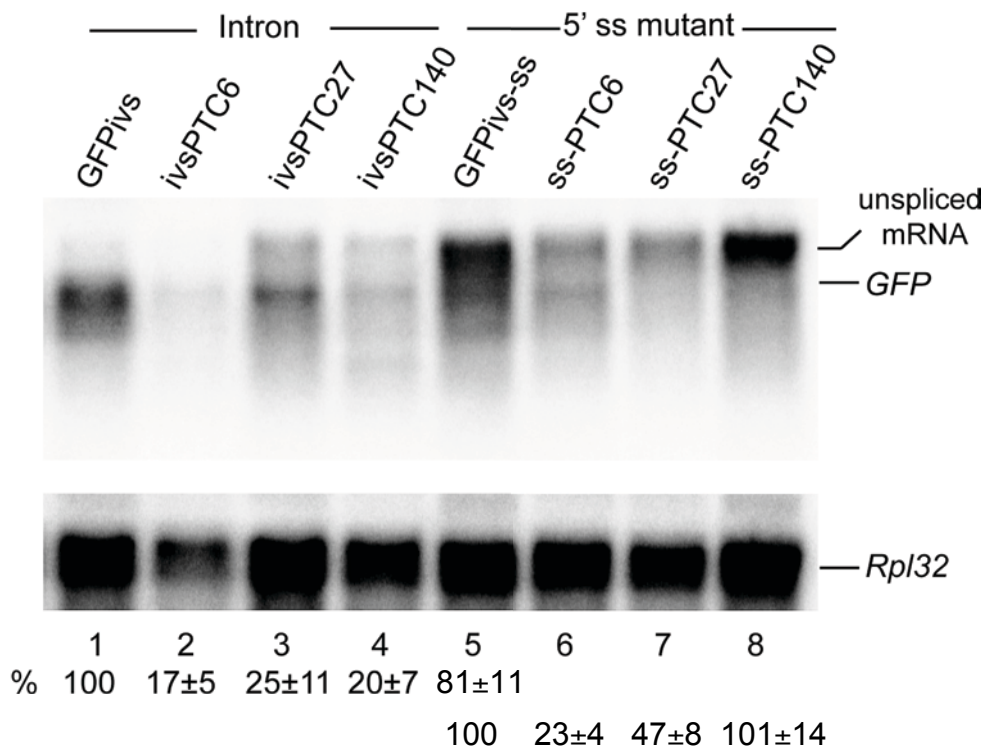
C
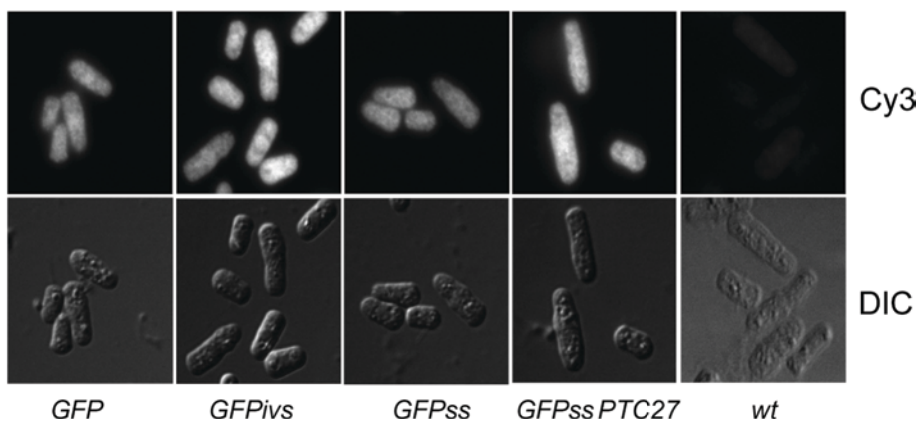
Figure 5.3 Inhibition of splicing prevents NMD.

(A) Diagram of the GFP reporters carrying a mutation in the intron 5' splice site. The mutagenesis of 5' splice site and generation of constructs are described in Materials and Methods, Chapter II.

(B) Northern blot analysis of total RNA from wild-type cells transformed with the indicated reporters. Top panel shows hybridization with the *GFP* probe as describe in Fig. 4.1B; bottom panel hybridization with the *Rpl32* probe as Fig. 4.1B. Values are percentages of the standardized intensity of the GFP mRNA in lane1 (lanes 1-4) and lane 5 (lanes 5-8). Calculations are based on three independent experiments.

(C) FISH analysis of cells transformed with reporters carrying either the original intronless reporter (GFP), or the intron-containing construct (GFPivs), 5' splice site mutant constructs with or without a PTC as indicated (GFPss and GFPssPTC27). The mRNA appears to be distributed throughout the cell in all strains, including the one carrying a PTC and the mutated intron (GFPssPTC27). The rightmost column shows FISH of cells not expressing GFP.

in the nucleus, as suggested by previous study (Legrain and Rosbash, 1989). To investigate the issue further, I assayed the sub-cellular distribution of GFP mRNAs by fluorescent in situ hybridization (FISH). However, the FISH data confirmed that the unspliced RNA with the mutated 5' ss accumulates throughout the cell, similar to the fully spliced mRNA (Fig. 5.3C). These results suggest that transcripts that fail to be spliced are not necessarily retained in the nucleus in *S. pombe*. Meanwhile, these data indicated that the unspliced mRNAs become NMD substrates

Furthermore, in the unspliced transcript, PTC111 (the earliest intron-encoded PTC) did not cause apparent NMD in GFP-ss and PTC140-ss, despite the fact that the distance between PTC110 and the 3' UTR (~760 nt) is longer than that between PTC6 and the 3' UTR (711 nt) in the intronless reporter. Therefore, this piece of data further supports the results in Chapter IV and indicates that a long distance between a PTC and the 3' UTR is not sufficient to trigger NMD (Fig. 5.3B).

5.2.5 Splicing enhances NMD also in chromosomal reporters.

To assess whether NMD is affected by the reporters expressed from a plasmid or from a chromosomal location, the GFP and PTC constructs, with or without introns, were cloned into the integration vector, pDUAL. pDUAL can be integrated by homologous recombination into the *leu1* locus and rescue the phenotype of leucine auxotroph by the *leu1D32* null allele.

The integration strains were generated as described in Materials and Methods, Chapter II. The mRNA levels were analyzed in the integration strains and showed a similar mRNA reduction as in the episomal reporters. PTC6 still triggered strong mRNA reduction; but PTC27 and PTC140 did not. Instead in the intron-containing reporters the PTC triggered strong mRNA reduction at all three positions (19% to

Fig. 5.4
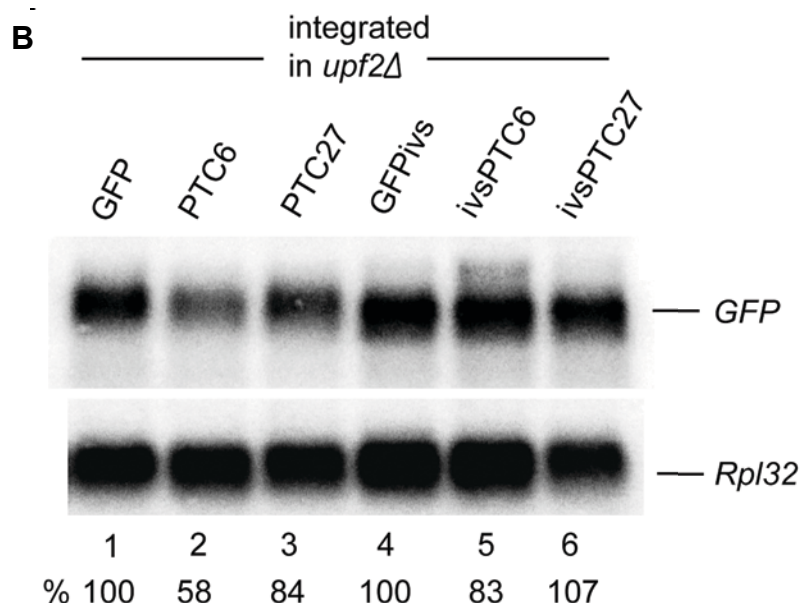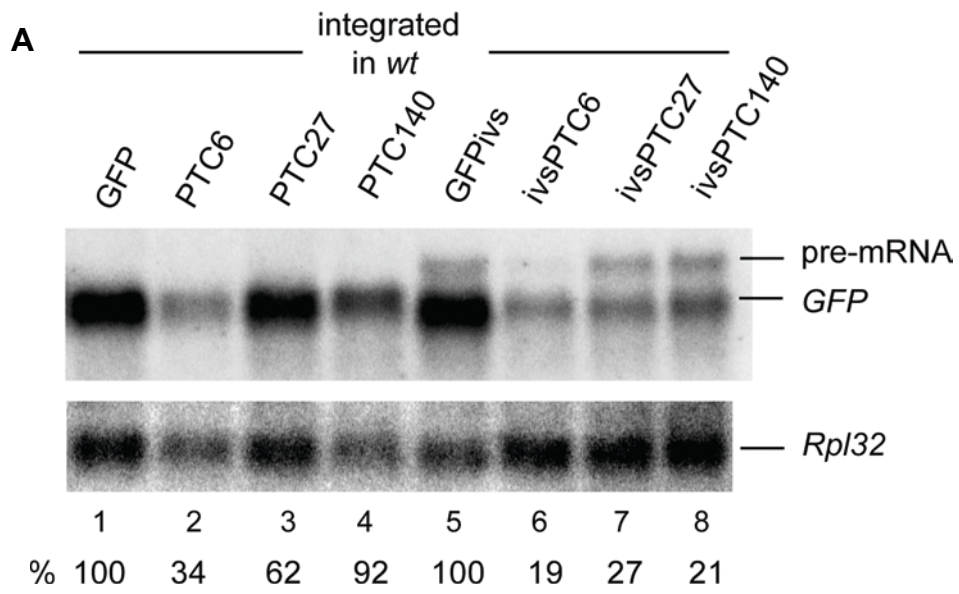
Figure 5.4 Intron enhanced NMD in the integration strains.

(A and B) Northern blot analysis of total RNA from the integration strains (derived from the wild-type or the *upf2Δ* strain) carrying the reporters indicated. Top panels show hybridization with the *GFP* probe as Fig. 4.1B; bottom panels with the *Rpl32* probe as Fig. 4.1B. (A) Values are percentages of the standardized intensity of the GFP mRNA in lane1 (lanes 1-4) or lane 5 (lanes 5-8); (B) standardization relative to lane 1 and lane 4.
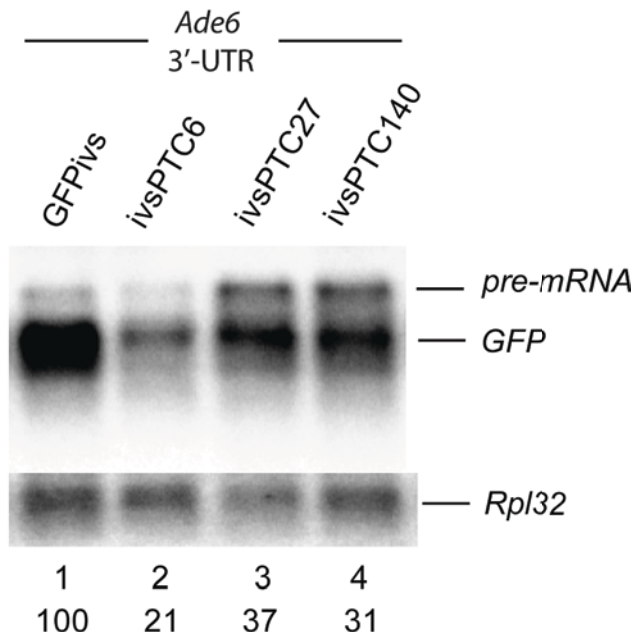
Figure 5.5 An Intron enhanced NMD independently of the nature of the 3' UTR.

Northern blot analysis of total RNA from cells transformed with the indicated reporters. Top panel shows hybridization with the *GFP* probe as Fig. 4.1B; bottom panel with the *Rpl32* probe as Fig. 4.1B. Values are percentages of the standardized intensity of the GFP mRNA in lane 1.

27% of GFP control), as in the episomal intron-containing reporters (Fig. 5.4A). In all cases the mRNA levels were increased in the *upf2Δ* strain, indicating that the mRNA reduction is caused by the canonical NMD (Fig. 5.4B). Four similar constructs with different 3' UTRs also showed a similar NMD pattern. The constructs which have the *ade6* instead on the *nmt1* 3' UTR (described in Materials and Methods, Chapter II), were transformed into *S. pombe* cells (Fig. 5.5). PTC6, PTC27 and PTC140 still triggered strong mRNA reductions similar to the constructs with the *nmt1* 3' UTR. These results indicate that splicing enhances NMD regardless of the nature of the 3' UTR.

5.2.6 Splicing-dependent NMD does not require EJC components.

As reviewed in the Introduction, studies from mammalian cells indicated that eIF4AIII, MAGO/Y14, MLN51 and RNPS1 are essential for NMD, and depletion of any of these components by RNAi in human cells strongly inhibited NMD (Gehring et al., 2005; Gehring et al., 2003; Lykke-Andersen et al., 2001; Shibuya et al., 2004). To address the EJC function in *S. pombe* NMD, I assayed for NMD in *S. pombe* strains deleted for known orthologs of EJC components. The fission yeast genome carries proteins highly similar to these EJC components with the exception that there is no homolog of MLN51 (Chapter III). The eIF4AIII ortholog could not be deleted, probably because it is essential like the homologous protein in *S. cerevisiae*, which functions in ribosome biogenesis and 40s ribosome subunit maturation (Kressler et al., 1997). In *S. pombe* Y14 does not have the conserved N-terminal domain, which has been reported as the additional interaction site for MAGO binding (Kataoka et al., 2001; Lau et al., 2003; Shi and Xu, 2003). Therefore, two knock out strains, *magoΔ* and *rnps1Δ*, were selected to study whether NMD depends on the presence of these proteins - *rnps1Δ* was generated by homologous recombination as described in
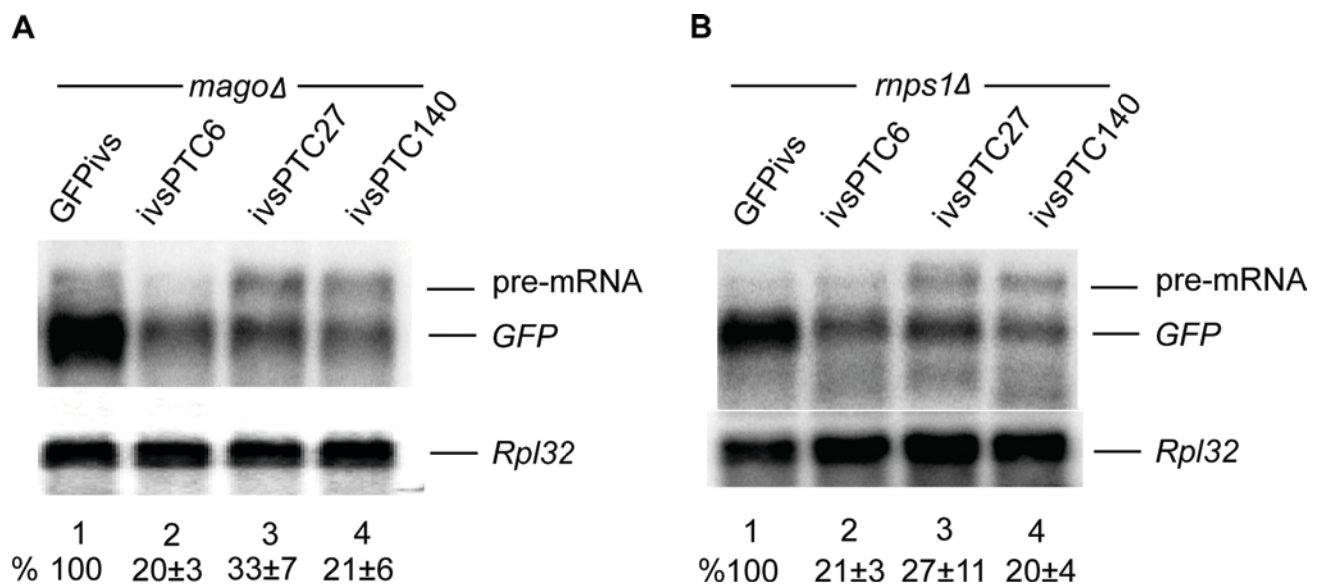
Figure 5.6 MAGO and RNPS1 are not required for NMD

(A and B) Northern blot analysis of total-RNA extracted from *mago* or *rnps1* deletion strains transformed with the indicated reporters. The top panels show hybridizations with the GFP probe as described in Fig. 4.1B, the bottom panel with the *Rpl32* probe in Fig. 4.1B. Values are percentages of the intensity of the GFP mRNA in the PTC-less control in lane 1; standard deviations are from three independent experiments.

Materials and Methods (Chapter II); the *magoΔ* is viable and has been previously described (Decottignies et al., 2003).

The four intron-containing NMD constructs were transformed into *magoΔ* and *rnps1Δ* cells and the mRNA levels were analyzed by Northern blot. The results showed the similar phenomenon of NMD as in the wild-type strain- the three intron-containing constructs have the strongest NMD. In particular, PTC27 and PTC140 caused a strong mRNA reduction (27% and 21% in *magoΔ* and 24% and 20% in *rnps1Δ*). These data indicate that splicing-dependent NMD does not require the EJC components, MAGO and RNPS1 (Fig. 5.6A and 6B).

5.2.7 Introns in 3' UTR trigger strong mRNA reduction

Some studies revealed that an intron in the 3' UTR caused strong mRNA reductions in mammalian cells (Mendell et al., 2004; Wittmann et al., 2006; Giorgi et al., 2007; Singh et al., 2008). To assess whether a 3' UTR intron has a similar effect in fission yeast, an intron was inserted 9 nt downstream of the normal stop codon. RNA analysis of strains transformed with these reporters, revealed a strong mRNA reduction only in the PTC-less reporter and in PTC6 reporter (~20% mRNA levels of GFP control), but not in PTC27 or PTC140 reporter. The fact that strong mRNA reduction occurred in the PTC-less reporter is consistent with the view that an intron downstream of a termination codon can make a normal stop codon behave like a PTC and trigger strong NMD. The finding is in agreement with the prediction of the EJC model and a key distinction between normal termination and NMD-inducing termination is the presence of a downstream intron. However, against the prediction of the EJC model, PTC27 and PTC140 transcripts were not affected, perhaps

Fig. 5.7

A

B

C
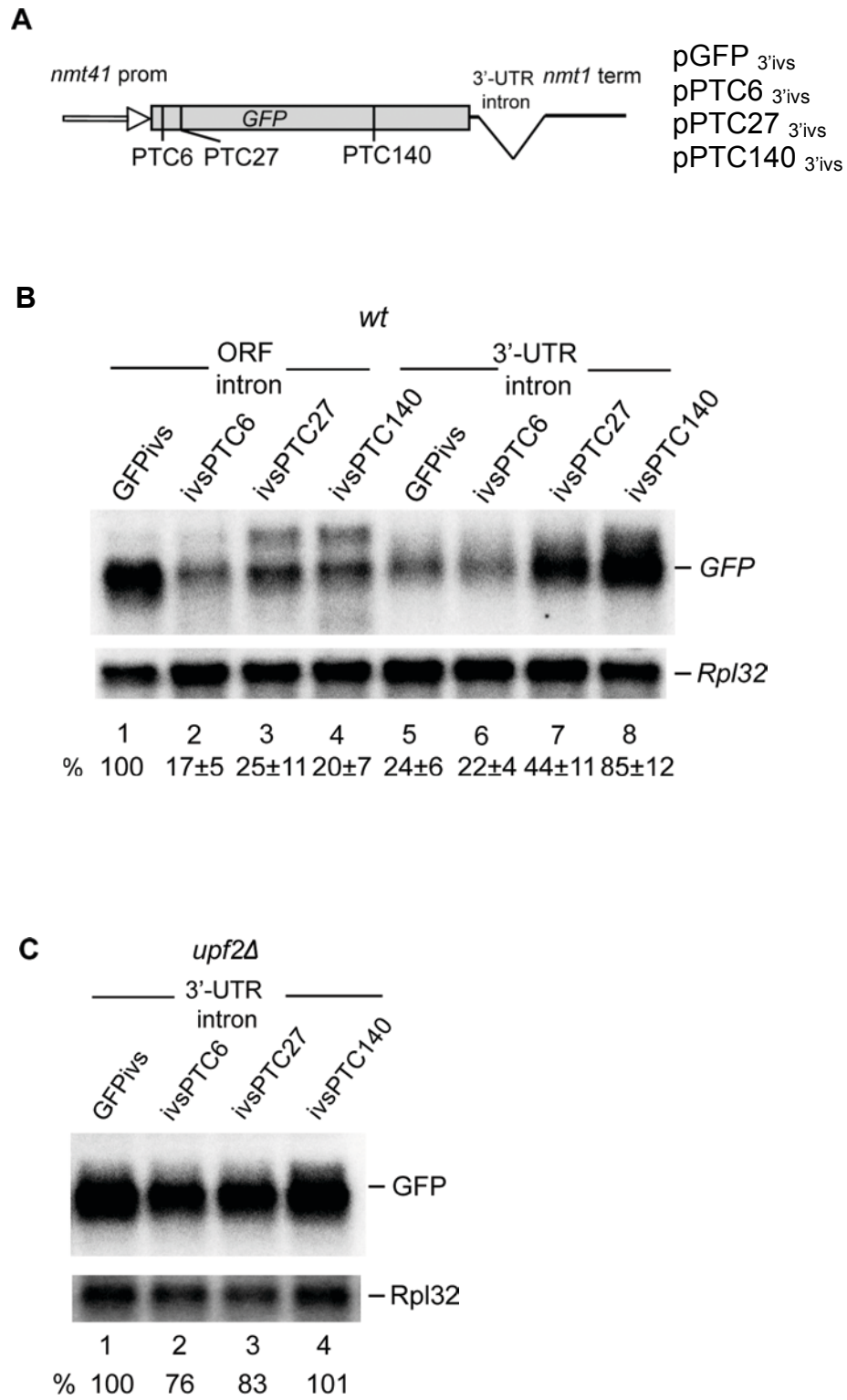
Figure 5.7 intron in 3' UTR caused *GFP* as the NMD substrate.

(A) Diagram of the reporters with a 3' UTR intron. The generation of these constructs is described in Materials and Methods, Chapter II. In the construct the intron is located 9 nt after GFP ORF.

(B and C) Northern blot analysis of total-RNA extracted from wild-type (B) or *upf2Δ* (C) strains transformed with the reporters with the 3' UTR intron. Top panels show hybridization with the GFP probe as described in Fig. 4.1B, bottom panels with the *Rpl32* probe as Fig. 4.1B. The values in B are percentages of the GFP mRNA signal of the PTC-less control; standard deviations are from three independent experiments.

suggesting that splicing can affect NMD only when termination occurs close to the splice site (Fig. 5.7B). In agreement with this prediction, in the reporters with the intron in the coding region, NMD was enhanced in PTC140-intron reporter much more than that in PTC27- intron (60% vs. 20%), while NMD in PTC6-intron was not affected by the intron at all. In these reporters the intron is located at codon position 110, which is far away from PTC6 (313 nt) and PTC27 (250 nt), but close to PTC140 (89 nt). However, in the reporter containing the 3' UTR intron, the intron is close to normal strop codon (12 nt), but far way from PTC140 and PTC27 (324 and 657 nt, respectively). It seems that the intron enhances NMD, only if the PTC is closer than ~240 nt to the splice site (Table 5.1 and Fig. 5.9).

5.2.8 An Intron enhances NMD only when is close to the PTC

To further investigate whether the effect of an intron on NMD depends on the distance from the PTC, I have made new constructs in which the distance between PTC140 and the intron is lengthened by either 147 bp or 291 bp - the two fragments derive from the EGFP coding region and lead to longer ORFs, which give rise to larger proteins in the PTC-less reporters. In the reporter with the 147 bp insertion (PTC140ivs-147), the distance between the PTC and the intron was extended to 249 nt (83 triplets) which is exactly the same as the distance between PTC27 and the intron in pGFPivs-PTC27; Meanwhile, in the PTC140ivs-291 construct, the distance between the PTC and the intron was extended to 390 nt (130 triplets) which is longer than the distance between PTC6 and the intron (313 nt) or the distance between PTC140 and the 3' UTR intron (323 nt) which do not have the enhancement of NMD even in the presence of a downstream intron. RNA analysis of cells transformed with

Fig. 5.8
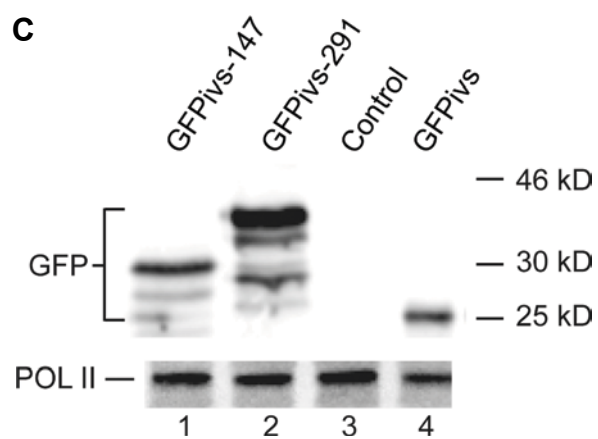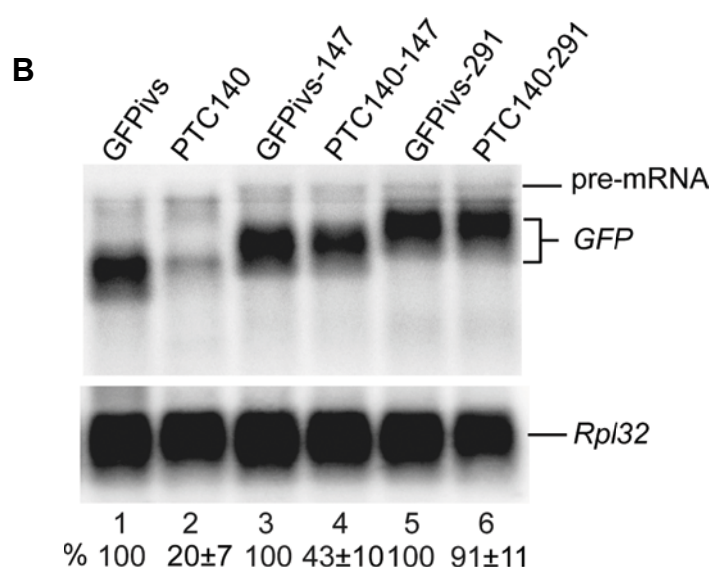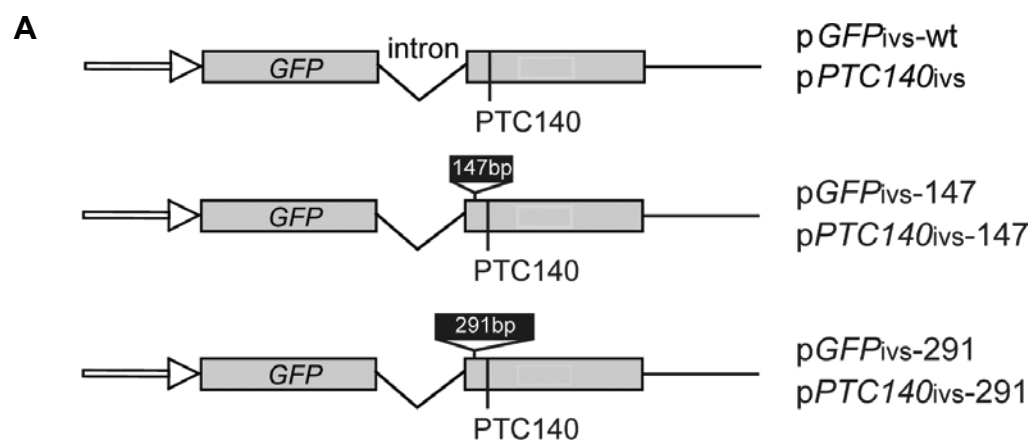
Figure 5.8 NMD depends on the distance between PTC and intron

(A) Diagram of the reporters in which the distance between the intron and the PTC was expanded. In the constructs, two insertions, 147 bp and 291 bp, are located between codons 120 and 121 (described in Materials and Methods, Chapter II).

(B) Northern blot analysis of total- RNA extracted from wild-type cells transformed with the reporters with the indicated reporters. Top panels show hybridization with the GFP probe as described in Fig. 4.1B, bottom panels with the *Rpl32* probe as Fig. 4.1B. Values are percentages with standard deviations (from three experiments) of the intensity of the PTC-less GFP mRNA control.

(C) Western blot analysis of whole-cell protein extracts of cells transformed with the indicated reporters. The top panel shows a blot probed with a goat polyclonal anti-GFP (AbD Serotec), bottom blot re-probed with the monoclonal 8WG16 (Covance) anti-Pol II antibody, as loading control. The secondary antibody was a rabbit HRP-conjugated anti-goat IgG (probe the anti-GFP) and anti-mouse IgG (probe the anti-Pol II) which was detected with chemiluminescent HRP substrate.

these reporters showed that both insertions suppressed the mRNA reduction: the reporter with the 147 bp insertion produced 43% of the mRNA level of the related control, a clear increase of the PTC140-mRNA level compared to 20% of the initial reporter (pGFPivs-PTC140); and the reporter with the 291 bp insertion yielded 91% of the level of the related control mRNA, a clear increase of relative levels (20% to 91%)- the enhancement of mRNA reduction by the intron was fully suppressed, as the extent of mRNA reduction is similar to the PTC140 intronless construct. The experimental data suggest that both insertions suppress NMD and confirm that the enhancement of NMD in the intron-containing reporters depends on the distance between the PTC and the intron.

The insertion of two spacers in the PTC-less reporter did not noticeably change the mRNA level, arguing against the inserts per-se having major effects on transcription or processing. Moreover, the protein yield was either comparable or increased, suggesting that the insertions also do not inhibit translation. Thus the reduced NMD is not caused by less efficient translation of the mRNA (Fig. 5.8C). In summary, it appears that the extent of intron-enhancement NMD depends on the distance from the intron to the PTC: the shorter the distance between the splice site and the PTC, the stronger NMD enhancement occurred (Fig. 5.9).

5.2.9 PABPC is not involved in NMD in *S. pombe*

As referred in the Introduction (Chapter IV), several studies suggest that PABPC play an important role in NMD. To address the function of PABPC in splicing-dependent NMD in *S. pombe*, the intron-containing reporter was transformed into the *pab1Δ* strain. RNA analysis showed that strong mRNA reductions still occurred in all three

Table 5.1 The extent of NMD and the distance of PTC to splice junction

Constructs	Splice junction (nt position)	PTC (nt position)	NMD enhancement (fold increase)	Distance from splice junction to PTC (nt)
pGFP-3'ivs	745	730-732	3.54	12
pGFPivs-PTC140	329	418-420	3.08	89
pPTC140ivs-147	329	577-579	0.98	248
pGFPivs-PTC27	329	79-81	0.74	250
pPTC140-3'ivs	745	418-420	0.1	321
pGFPivs-PTC6	329	16-18	0	316
pPTC140ivs-291	329	721-723	0	394
pPTC27-3'ivs	745	79-81	0	664
pPTC6-3'ivs	745	16-18	0	727

The grey-scaled data are used to make the Figure 5.9 for the relationship between NMD enhancement and the PTC-intron distance.
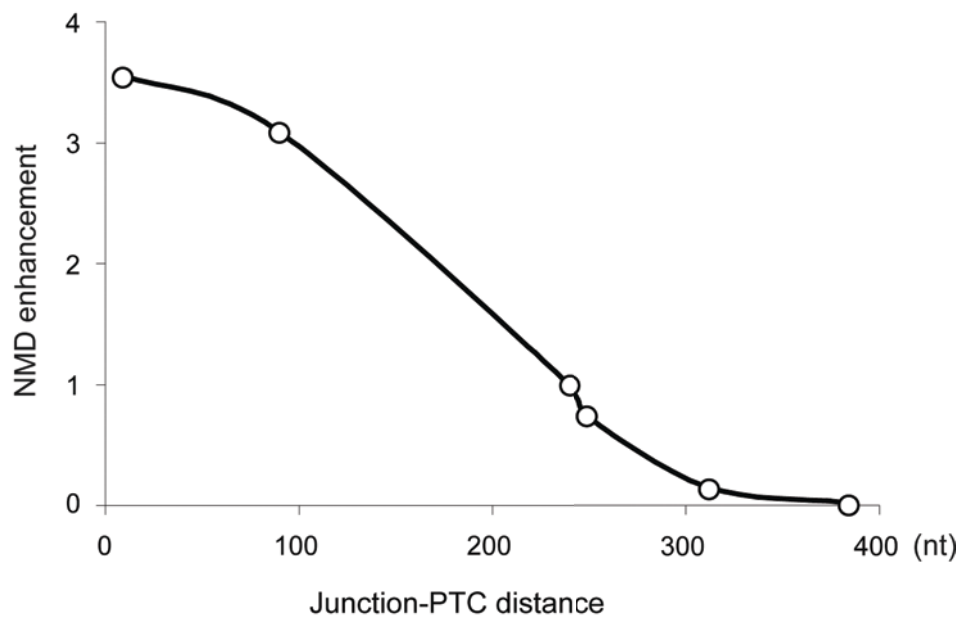
Figure 5.9 Splicing enhancement of NMD is depended on the PTC-intron distance.

An intron enhances NMD only when it is close to the PTC. Plot shows the relationship between the distance (in nucleotides) of the intron from the PTC and NMD enhancement. The figures are based on the data reported in this paper; from left to right (close to distant): p*GFP*-3'ivs, p*GFP*ivs-PTC140, pPTC140ivs-147, p*GFP*ivs-PTC27, pPTC140-3'ivs, and pPTC140ivs-291 (Table 5.1). NMD level correspond to the folds of mRNA reduction standardized by the reduction of the corresponding intronless reporter.

PTC-intron containing constructs, very similar to the intron enhancement in wild-type strain (Fig. 5.10). This suggests that PABPC is not involved in splicing-dependent NMD, similar to what was previously shown for intronless reporters (Chapter IV, Fig. 4.8).

5.2.10 Splicing-dependent NMD does not correlate with a reduced stability of the mature mRNA

The data I have presented in this chapter clearly show that splicing enhances NMD when the intron is close to the PTC. To address whether the reduction is caused by a lower mRNA stability, the mRNA half-lives were estimated after transcription inhibition with 1, 10-phenathroline (described in Chapter II and Chapter IV Section 2.10). RNA samples were harvested at six time points and assayed by Northern blot. Surprisingly, RNA analysis showed that in the intron-containing reporter, the decay of spliced PTC140 mRNA is very similar to that of the PTC-less control, despite that the fact that mRNA levels of PTC140 are very low compared to the control mRNA (Fig. 5.11A). The decay rate of spliced PTC140 mRNA was also not greatly affected in *upf2Δ* cell (Fig. 5.11C). These data strongly suggest that only newly made and spliced PTC140 mRNAs undergo accelerated decay in the pioneer round of translation as mentioned in the Introduction.
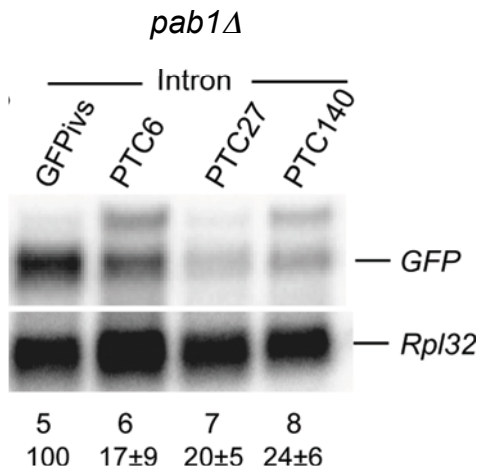
Figure 5.10 Intron enhancement in NMD does not require PABPC in *S. pombe*.

Northern blot analysis of total RNA extracted from wild-type cells transformed with the indicated reporters. Top panel shows hybridization with the GFP probe, bottom panel with the *Rpl32* probe as described in Fig. 4.1B. Values are percentages with standard deviations (from three experiments) of the intensity of the PTC-less GFP mRNA control.

Fig. 5.11

A wt

GFPivs

ivsPTC140

Time(min)   0   2   10   20   40   60

B upf2Δ

pre-mRNA
GFP
18S rRNA

pre-mRNA
GFP
18S rRNA

0   2   10   20   40   60

C

GFPivs in wt
PTC140ivs in wt
GFPivs in upf2del
PTC140ivs in upf2del

mRNA levels

2   10   20   40   60   (min)
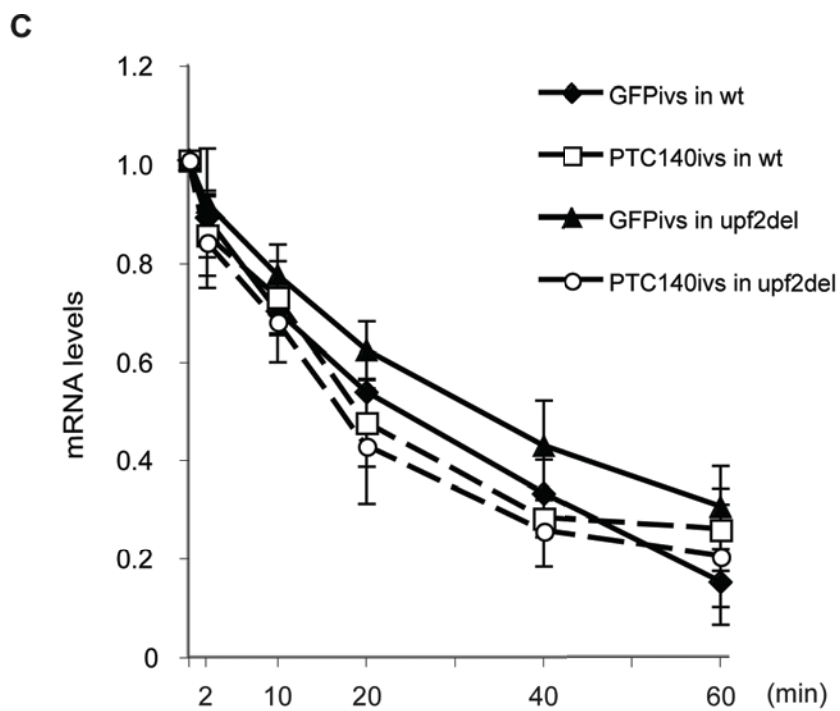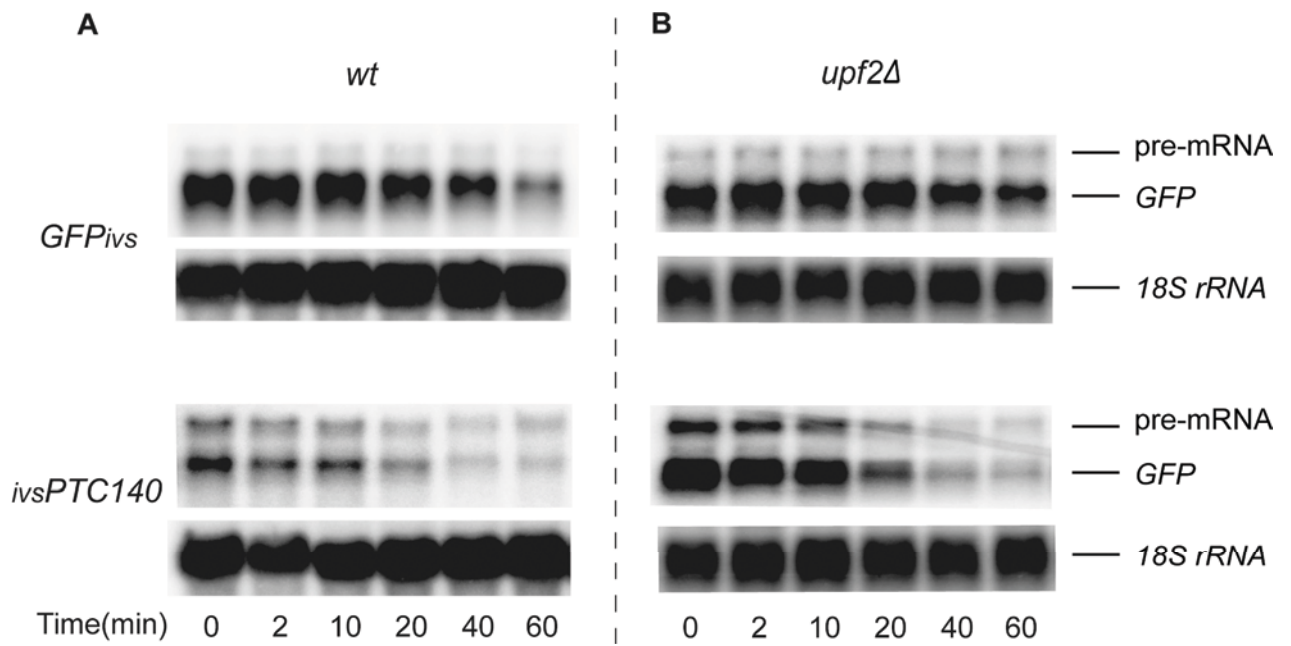
Figure 5.11 mRNA half-life analysis of the intron-containing constructs.

(A-C) Stability of the GFP mRNA in wild type (A) and *upf2Δ* (B) following transcription inhibition with 1, 10-phenanthroline. The PTC-less control (top panel) and PTC6 mRNA (bottom panel) were independently monitored and (C) quantified by Northern blot as described in Fig. 4.10, except that values were standardized to the 18S rRNA signal (this RNA does not show visible decay over the time period of the experiment). Total-RNA was extracted at 6 different time points after adding the drug: 0, 2, 10, 20, 40 and 60 min. Error bars are standard deviations from three different experiments.

5.3 Discussion

An open issue in the field was whether the link between splicing and NMD is a new acquisition of multicellular organisms or instead whether it has been lost in budding yeast (in which all the available experimental evidence indicates that introns are not involved in NMD). Fission yeast is an ideal model organism to address this issue, because 43 % of its genes (~2200 genes) carry introns (Wood et al., 2002). It was estimated that fission yeast diverged from budding yeasts 330-420 Myr ago (Wood et al., 2002). In particular, as reported above (Chapter III), the genome of this organism contains orthologs of EJC components.

Here, I found that introns play a very important role in fission yeast NMD by using NMD reporters. The presence of an intron close to the PTCs enhanced NMD in both GFP and the reporter derived from the endogenous gene *ypt3*. This intron-enhancement is directly coupled to splicing because mutations of the 5' splice site that blocks splicing, strongly inhibits NMD. In addition, an intron located in the 3' UTR induced strong NMD of the wild-type mRNA. It appears, that an intron in the 3' UTR can make a normal stop codon behave like a PTC; this observation is in agreement with a number of reports in mammalian cells (Carter et al., 1996; Thermann et al., 1998). However, in contrast to what has been reported in mammalian system, an intron enhances NMD even when it is very close to the stop codon, and there is no evidence of a 50-55 nt boundary in this system, unlike what was reported by several studies in mammalian cells (Cheng et al., 1994; Zhang and Maquat, 1996; Zhang et al., 1998a; Zhang et al., 1998b).

A second intriguing finding is that an intron can enhance NMD regardless of whether it is positioned upstream or downstream of the PTC in my system. The results contradict the main prediction of the EJC model: that an intron can induce NMD only if located downstream of the termination codon. The fact that an intron can also enhance NMD when positioned before the PTC seems to exclude that an EJC is mediating the influence in NMD. To recognize the PTC the ribosome has to scan through the splice site region and strip off any associated proteins. In addition, I found that NMD is not affected by deletion of RNPS1 and MAGO, which are two core components of EJC that appear to be essential for NMD in human cells. Therefore, the EJC seems not appear to be involved in *S. pombe* NMD (Discussed in Chapter VII).

Another interesting discovery is that the extent to which an intron can enhance NMD depends on the distance between the PTC and the intron. My data suggest that in splicing dependent NMD, the intron must be within ~240 nt from the PTC, regardless of whether it's downstream or upstream.

CHAPTER VI. ADDITIONAL DATA IN *S. POMBE* NMD.

As described in the previous chapters, the main focus of my research was to study the link between pre-mRNA splicing and NMD in *S. pombe*. However, I have also investigated other issues in this project. In one of these spin-off projects, I tested whether NMD is initiated by endonucleolytic cleavage in *S. pombe*, as reported in *Drosophila* and mammalian cells (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008; Eberle et al., 2009). In addition, in this chapter I present some initial data that indicate UPF1 may associate with both large and small ribosome subunits.

6.1 No evidence of endonucleolytic cleavage in *S. pombe* NMD

6.1.1 Introduction

Initial studies in *Drosophila* S2 cells have shown that low molecular weight RNA degradation intermediates, deriving from PTC-containing transcripts, were accumulated in cells depleted of XRN1(Gatfield and Izaurralde, 2004). This strongly suggested that endonucleolytic cleavage initiates NMD in *Drosophila* S2 cells. After the endonucleolytic cleavage, the 5' and 3' fragments are vulnerable to exosome mediated 3' to 5' degradation and to XRN1 mediated 5' to 3' decay, thus they are not normally detected in wild-type cells. Recent studies suggested that the endonucleolytic cleavage also initiates NMD in mammalian cells and that SMG6 is the main endonuclease in both mammalian cells and *Drosophila* S2 cells (Huntzinger et al., 2008; Eberle et al., 2009).

It is not known whether endonucleolytic cleavage occurs in *S. pombe* NMD. A previous study in *S. cerevisiae* showed no evidence for endonucleolytic cleavage (Doma and Parker, 2006). It was reported that the endonuleolytic cleavage can occur

in the proximity of stem loops that cause ribosome stalling, but this endonucleolytic cleavage occurred only in mRNAs with a stem loop, but not in rare codon or PTC-containing mRNAs (Doma and Parker, 2006). As I described in Chapter III, there is no SMG6 homolog in *S. pombe* and in *S. cerevisiae*; thus endonucleolytic cleavage is not expected in *S. pombe* or *S. cerevisiae* NMD. In agreement with this prediction, the data I report below confirm that the endonucleolytic cleavage is not involved in *S. pombe* NMD.

6.1.2 Results

6.1.2.1 No evidence of endonucleolytic cleavage in *exo2Δ cells*

To address whether PTCs trigger endonucleolytic cleavage of the mRNA in *S. pombe*, I assayed the available NMD reporters in a strain deleted for *exo2* (SPAC17A5.14) – this codes for the *S. pombe* otholog of XRN1 (Szankasi and Smith, 1996). Similar to previous studies, it was expected that if an endonucleolytic cleavage occurred in *S. pombe*, then 3' degradation fragments should be readily detectable in *exo2Δ* cells. In addition, the length of these 3' half fragments should correlate with the PTC position – earlier PTCs should produce longer 3' fragments than later ones. Contrary to this expectation, I could detect only the full-length mRNA with all of the NMD reporters (Fig. 6.1). Furthermore, the overall NMD was not affected in the *exo2Δ* strain, with none of the reporters (Fig. 6.1). These results indicated that the endonucleolytic cleavage is unlikely to occur during NMD in *S. pombe*.
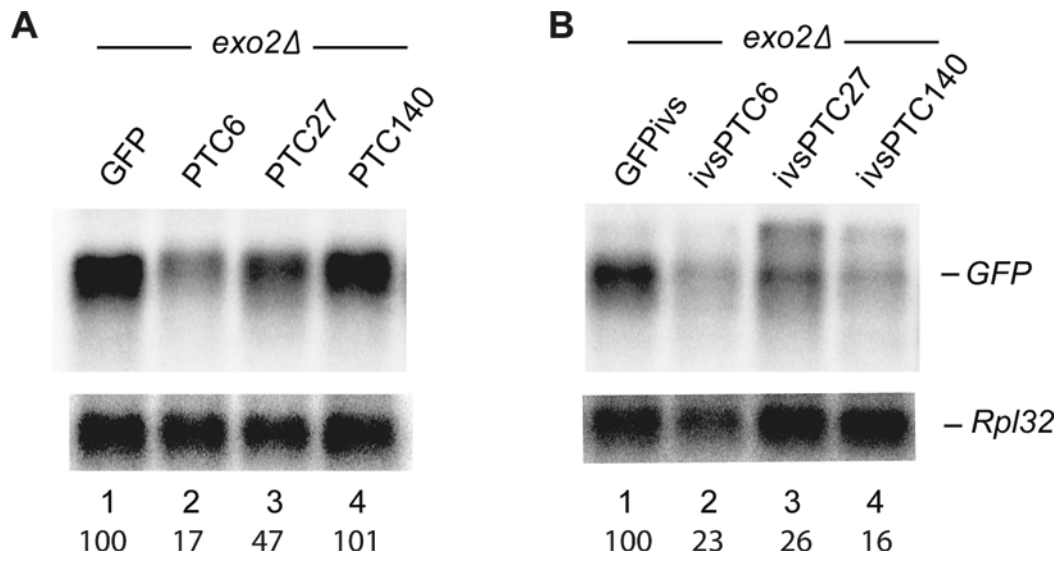
Figure 6.1 NMD in *exo2Δ*-XRN1 mutant.

Northern blot analysis of total RNA from *exo2Δ* cells transformed with the NMD reporters indicated, (A) intronless reporters and (B) intron-containing reporters. As in Fig. 4.1B, top panels shows hybridization with the *GFP* probe and bottom panels with the *Rpl32* probe.

6.2 UPF1 is associated with both polyribosome and separated ribosomal subunits

6.2.1 Introduction

As reported in the Introduction, UPF1 plays a pivotal role in NMD and understanding its molecular function (or functions) is key to a full understanding of the NMD phenomenon. It is believed that UPF1 is recruited to the terminating ribosome by an interaction with eRF1 and eRF3; this is mainly based on observations that eRF1 and eRF3 interact with UPF1 directly (Czaplinski et al., 1998; Wang et al., 2001; Kobayashi et al., 2004; Kashima et al., 2006; Ivanov et al., 2008). Polysome profiling in *S. cerevisiae* has shown that UPF1 mainly associates with monosome and polyribosomes and is unlikely to interact with split subunits (Atkin et al., 1995; Atkin et al., 1997). I was interested in investigating the association of UPF1 with ribosomes in *S. pombe* further. Below, I report some initial data indicating that UPF1 probably associates with the dissociated ribosome subunits as well as polyribosomes. The data suggest that UPF1 might be also involved in translation initiation.

6.2.2.1 UPF1 associates with polyribosomes

To investigate the association of UPF1 with ribosomes, I used a strain expressing a human influenza hemagglutinin (HA) tagged UPF1 (UPF1-HA3) (Rodriguez-Gabriel et al., 2006). The HA tagged UPF1 can be easily detected with commercially available anti-HA antibodies. To test whether UPF1 associates with functional ribosomes, I assayed whether they are associated with polysomes: cell extracts were separated by sucrose gradient (10% -50%) centrifugation and the fractions - corresponding to free proteins, 40S, 60S, 80S and polysomes - analyzed by

Fig. 6.2

A

polysomes

80s

60s

40s

1 2 3 4 5 6 7 8 9 10 11 12 13

28s
18s
5s, 5.8s

B

C   1 2 3 4 5 6 7   8 9 10 11 12 ——UPF1-3HA

C

PTC6   polysome   80s+60s

GFP

1 2 3 4 5 6 7 8 10 13 ——GFP

GFP   polysome   80s+60s

GFP

1 2 3 4 5 6 7 8 9 10 ——GFP
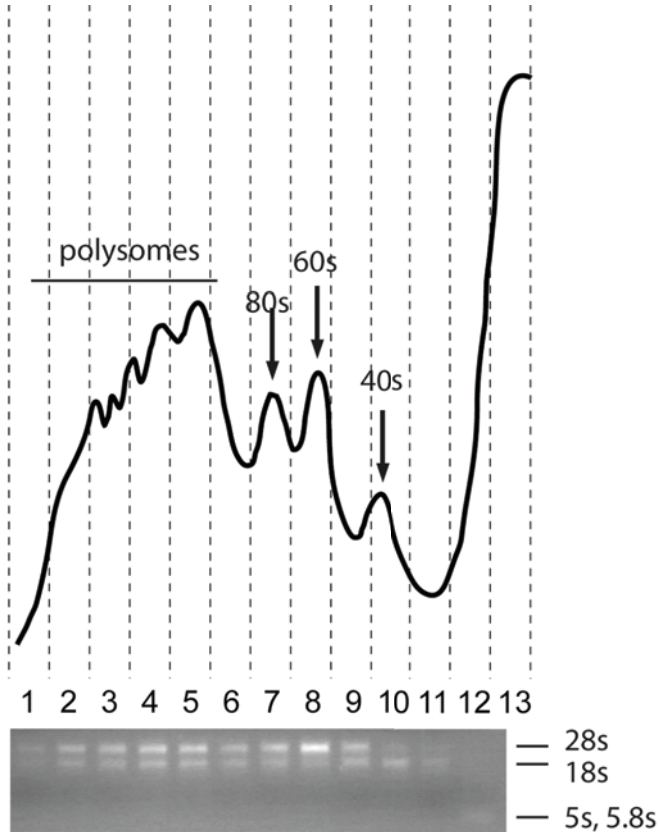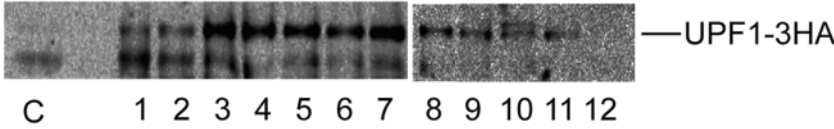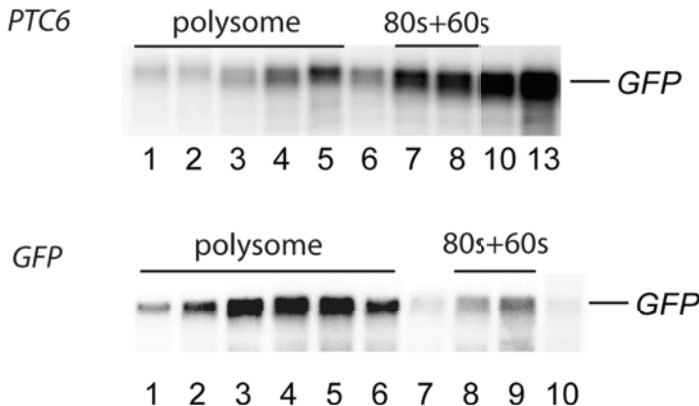
Figure 6.2 UPF1 and PTC-mRNA distribution analysis by polysome profile

(A) Sucrose gradient analysis of ribosomes. The gradient was monitored for AU_{258} 254 by pumping the content through a UV detector connected to a charter recorder. 16 fractions were collected but the first two and the last one fraction were discarded. The fractions were analyzed by native agarose gel electrophoresis to visualize 25S and 18S rRNAs (panel below polysome profile).

(B) Western blot of the polysome fractions, to visualize the distribution of UPF1. UPF1 was detected with a monoclonal anti-HA antibody (12CA5, Cancer Research UK). Lanes 1-12 correspond to fractions 1 to 12; C is an aliquot of the unfractionated protein extract.

(C) Northern blot analysis of the RNA isolated from the fractions to visualize the distribution of the GFP mRNA using a specific probe as in Fig 4.1B. Top one shows the GFP mRNA distribution in cells expressing the PTC6 transcript, bottom panel the distribution in cells expressing the PTC-less control. The polysome profile in A was obtained with cells expressing the PTC6 reporter.

immunoblotting with tag-specific antibodies (Fig. 6.2) (procedures described in Chapter II, 2.4.10). I found that UPF1 is associated not only with polyribosomes but also with 80S monosome, 40S and 60S ribosome subunits (Fig. 6.2B). I also investigated the distribution of the GFP reporter mRNA in these fractions. As expected, the PTC-less mRNA is mostly associated with polyribosome fractions (Fig. 6.2C). Instead the mRNA with PTC6 was mostly associated with lighter fractions corresponding to monosome or separate subunits (Fig. 6.2C). The association of the PTC6 mRNA with light polyribosome and monosome fractions is in agreement with the prediction that no more than one translating ribosome should associate with this mRNA (Fig. 6.2C), because the open reading region (18 nt) is shorter than the region protected by a single ribosome (30 nt) (Ingolia et al., 2009)

6.2.2.2 UPF1 interacted with separated ribosome subunits

The association between ribosome subunits depends on the presence of $Mg2^+$ ions. Therefore, incubation with EDTA dissociates the ribosomes into 40S and 60S subunits; this step makes it possible to distinguish polyribosome association from the association with RNP or other large complexes.

The western blot revealed that UPF1 mainly associates with 60S subunits, and weakly with 40S subunits (Fig. 6.3). Therefore, in translating ribosomes UPF1 might associate only with the 60S, while it might associate with both subunits prior to subunits joining and initiating translation (Fig. 6.2B) (see discussion below). For unknown reasons in my assays the mRNAs were strongly degraded in the presence of EDTA so that I could not study the mRNAs distribution in EDTA-treated polysome fractions as done above.
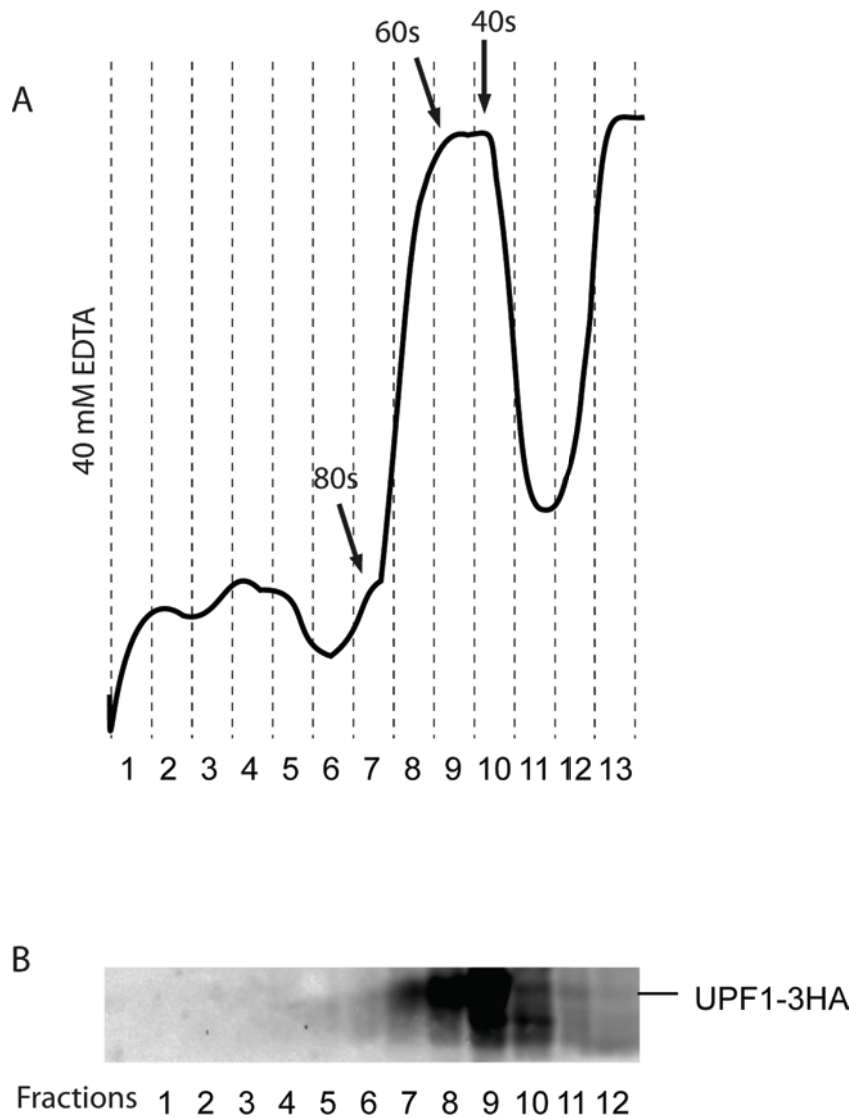
Figure 6.3. UPF1 distributions upon EDTA-induced polysome breakdown.

(A) Sucrose gradient analysis of ribosomes. The gradient was monitored for AU_{258} 254 by pumping the content through a UV detector connected to a charter recorder. 16 fractions were collected but the first two and the last one fraction were discarded.

(B) Western blot of the polysome fractions, to visualize the distribution of UPF1. UPF1 was detected with a monoclonal anti-HA antibody. Lanes 1-12 correspond to fractions 1 to 12.

6.3 Discussion

6.3.1 Is an endonucleolytic cleavage required for NMD in *S. pombe*?

In contrast to the studies in *Drosophila* and mammalian cells, here I did not find the supported evidence of the PTC-induced endonucleolytic cleavage. The results also indicate that NMD might not require XRN1 and is not affected by XRN1 deletion. The fact that there is no SMG6 homolog in this organism further supports this conclusion. SMG6 and the endonucleolytic cleavage might be specific to NMD in multicellular organisms.

However, the possibility of an endonucleolytic cleavage for some kind of mRNA decay in mammalian cells and *Drosophila* S2 cells has not been discounted yet. As referred above, there is evidence that a stem loop structure can elicit the endonucleolytic cleavage in *S. cerevisiae* (Doma and Parker, 2006); therefore, it is possible that other forms of ribosome stalling during translation elongation can induce the endonucleolytic cleavage in yeast. This indicates that either the endonucleolytic cleavage is not specific in PTC-borne mRNAs or that it might have different mechanisms between yeasts and metazoa. The endonucleolytic cleavage still requires to be further confirmed in *S. pombe*.

6.3.2 UPF1 function in NMD

On the basis of published data, indicating that UPF1 interacts with eRF1 and eRF3, it is widely expected that UPF1 should be found associated only with terminating 80S ribosomes. Here, however, I found that UPF1 seems to associate with separate 40S and 60S subunits. This observation suggests that UPF1 might function in stages of

translation other than termination, for example translation initiation or re-initiation. In agreement with my findings, it has been recently reported that in mammalian cells phosphorylated UPF1 suppresses translation initiation by interacting with eIF3 to inhibit the forming of 80S/Met-tRNAi^{Met}/mRNA initiation complexes (Isken et al., 2008). Additionally, another study indicates that UPF1 is required for suppressing translation of specific mRNAs but not required for their mRNA decay (You et al., 2007).

The fact that most of the UPF1 in the cells is found associated with polyribosomes, suggest that this protein has a general role in translation and it is not just recruited to NMD substrates. If UPF1 was only recruited to NMD substrates, it should only be found in light polyribosomes and 80S fractions, because NMD degrades most of mRNAs with early PTCs. Further studies will need to be performed to clarify the function of UPF1 in translation.

CHAPTER VII. CONCLUSION AND DISCUSSION

7.1 Development of a new experimental system to study NMD in *S. pombe*

As mentioned in the introduction, the NMD mechanism in *S. pombe* has not been well characterized in previous studies; we only knew that mRNA derived from some nonsense alleles of the *ade6* gene is subjected to NMD and that UPF1 and UPF2 are required for this to occur (Mendell et al., 2000; Rodriguez-Gabriel et al., 2006). The aim of my project was to develop a system in which I could investigate the NMD mechanism more in detail. To achieve this I used a *gfp* reporter gene which was known to be subject to NMD in *S. cerevisiae* (Kuperwasser et al., 2004). In the initial constructs I introduced PTCs at three different triplets positions: 6, 27 and 140. The GFP constructs were cloned in the pREP41 vector, which allows GFP expression from the *nmt41* promoter (Basi et al., 1993; Maundrell, 1993). The *nmt41* promoter is less active than the original *nmt1* promoter and chosen mainly to prevent saturation of the NMD machinery. There is no evidence that the NMD machinery can be saturated by a high level of mRNA substrates, however I reasoned that the level of the UPF proteins is low in *S. cerevisiae* (especially UPF2 and UPF3), and it would therefore be safer to keep the expression of the reporters at a low level (Maderazo et al., 2000; Ghaemmaghami et al., 2003).

To assay NMD, I introduced these reporters into the cells either by transformation (for episomal expression) or by homologous recombination-mediated chromosomal integration. As described in the Results and in Material and Methods, I have produced several NMD reporters. All the constructs and the corresponding NMD phenotypes are listed in Table 7.1.

Table 7.1 the extent of mRNA reductions in NMD constructs.

Constructs (pREP-)	PTC position (nt)	mRNA reduction (%)	Extent of NMD
GFP-418	730-732	37	Moderate
GFP-102	730-732	-12	No
GFP-ss	331-333	19	Weak
GFP-3'ivs	730-732	76	Strong
ivsPTC140-291	721-723	9	Weak
ivsPTC140-147	577-579	57	Moderate
PTC140	418-420	18	Weak
PTC140-ss	331-333	18	Weak
PTC140-102	418-420	-7	No
PTC140-418	418-420	47	Moderate
PTC140-402	418-420	33	Moderate
ivsPTC140	418-420	80	Strong
PTC140-3'ivs	418-420	15	Weak
PTC27	79-81	57	Moderate
PTC27-ss	79-81	69	Strong
PTC27-102	79-81	73	Strong
PTC27-418	79-81	76	Strong
PTC27-402	79-81	73	Strong
ivsPTC27	79-81	75	Strong
PTC27-3'ivs	79-81	56	Moderate
PTC6	16-18	84	Strong
PTC6-ss	16-18	83	Strong
PTC6-102	16-18	82	Strong
PTC6-418	16-18	84	Strong
PTC6-402	16-18	87	Strong
PTC6-sh	16-18	89	Strong
ivsPTC6	16-18	83	Strong
PTC6-3'ivs	16-18	78	Strong
PTC6-291	319-321	71	Strong
PTC6-147	175-177	77	Strong

Gray columns mean intron-containing NMD reporters; the PTC-less control constructs are not listed.

7.2 NMD is polar in *S. pombe*

As mentioned above, in many systems it has been reported that NMD is polar: the PTC at the beginning of the gene induced stronger NMD than the PTC located further downstream. Previous studies have reported that nonsense mutations located at the beginning of the *S. pombe ADE6* gene also induce stronger NMD than mutations further downstream (Mendell et al., 2000). In this study, I have also found that PTCs induce strong NMD only when they are located at the beginning of the coding region, while PTCs in the second half of the coding region do not induce strong NMD. Surprisingly, I found that this strong NMD does not require Pab1p, the unique cytoplasmic PABP in *S. pombe*, and the extent of NMD does always correlate with the distance between the PTC and either the poly(A) tail or 5' end. These observations argue that the *faux* 3' UTR model cannot explain the polarity of NMD in *S. pombe*.

7.3 Splicing dependent NMD

A key discovery of this study is that in the presence of a nearby intron, PTCs located in the second half of the coding region also cause strong NMD. It appears that splicing abolishes the PTC polar effect that was seen with intronless reporters. Additionally, inserting an intron in the 3' UTR makes the normal stop codon behave like a PTC and causes a strong mRNA reduction, as previously reported for mammalian systems (Carter et al., 1996; Thermann et al., 1998). Similarly, I found that the removal of the introns from an endogenous gene also drastically reduces NMD, arguing that introns have a general role in *S. pombe* NMD. Collectively, these results suggest the existence of splicing dependent NMD in *S. pombe*. The observation that the half-life of the PTC-containing mRNA (PTC140-ivs) does not

differ significantly from the corresponding PTC-less control, suggests that splicing - dependent NMD might only occur in the pioneer round of translation. These observations are in agreement with studies in mammalian cells, where it was found that only newly made transcripts are subject to NMD, while mRNAs that survive the initial stages are stable at the steady state (Urlaub et al., 1989; Baserga and Benz, 1992; Cheng and Maquat, 1993).

As referred in the introduction, many recent studies have questioned the role of splicing in NMD (Gatfield et al., 2003; Amrani et al., 2004; Buhler et al., 2006; Behm-Ansmant et al., 2007; Eberle et al., 2008; Singh et al., 2008). However my data clearly indicate that splicing is an important NMD determinant in fission yeast and strongly support the view that pre-mRNA splicing can affect the stability of the PTC-containing mRNA (Cheng and Maquat, 1993; Carter et al., 1996; Zhang et al., 1998a; Le Hir et al., 2000a; Nott et al., 2004; Gudikote et al., 2005). On the basis of these findings, one would have expected intron-containing genes to be over-represented with in the set of endogenous NMD substrates identified in a microarray study of NMD mutant strains in *S. pombe* (Rodriguez-Gabriel et al., 2006). However, only 22% these putative NMD substrates contain introns (6 out of 27 genes) (Rodriguez-Gabriel et al., 2006). Given that 43% of the genes have introns in *S. pombe*, the findings suggest that intron-containing genes are not preferred NMD substrates.

In addition, splicing of several genes is alternatively regulated during the cell cycle in *S. pombe* (Nakamura et al., 1997; Malapeira et al., 2005; Moldon et al., 2008). However, none of these genes is affected by NMD, in contrast to the view that unspliced transcripts are the most abundant NMD substrates in *S. cerevisiae* (He et al., 1993; He et al., 2003).

7.3.1 Splicing-dependent NMD does not require EJC components

Initially, the discovery that splicing enhances NMD, suggested the involvement of the EJC as in mammalian cells. However, I found that contrary to the prediction of the EJC model, the intron enhanced NMD, regardless of whether it was positioned before or after a PTC in NMD reporters in *S. pombe*. It is difficult to explain how the ribosome could recognize the PTC without scanning through preceding exon junctions -the ribosome tunnel through which the mRNA moves is just large enough to accommodate naked single stranded RNA, suggesting that the mRNP is stripped off during ribosome scanning (Yusupova et al., 2001).

Furthermore, in contrast with the EJC model, I found that NMD was not affected by the deletion of orthologs of EJC components MAGO and RNPS1, of which MAGO appears to be a core component (Chamieh et al., 2008). These results are also in disagreement with the report that in mammalian cells, MAGO/Y14 is required for standard NMD and RNPS1 for UPF2-dependent NMD (Gehring et al., 2005). Among the other core EJC components which are essential in mammalian NMD, MLN51 is not present in *S. pombe*; and Y14 lacks the amino terminus that is important but dispensable for the interaction with MAGO (Kataoka et al., 2001; Lau et al., 2003; Shi and Xu, 2003); and eIF4AIII, the putative core of the EJC, the gene could not be deleted in *S. pombe,* probably because like in *S. cerevisiae* the protein might have an essential function in rRNA biogenesis and maturation (Kressler et al., 1997). My results clearly indicate that MAGO and RNPS1 are not involved in either splicing-dependent or splicing-independent NMD in *S. pombe*. I can't exclude that eIF4AIII can assemble an EJC complex alone in *S. pombe*, but this is unlikely. For example, there is a protein very similar to eIF4AIII in *S. cerevisiae* yet no evidence of the

existence of an EJC (Chapter III, Fig. 3.4). Conclusively, if an EJC is present in *S. pombe,* most probably is not involved in NMD.

Observations that an intron can also enhance NMD when positioned before the PTC, have been reported for the β-globin gene in mammalian cells: it was found that an upstream intron is required for NMD when downstream introns are absent (Matsuda et al., 2007). However, this mode of NMD appears to require some of the EJC proteins, although RNPS1 was also dispensable (Matsuda et al., 2007).

The homologs of eIF4AIII, MAGO, RNPS1 and Y14 are present and well conserved in *S. pombe.* It would be interesting to further investigate whether they can assemble an EJC-like complex and whether they have a function in linking pre-mRNA splicing with downstream events such as mRNA export, transport and translation. Perhaps in *S. pombe* the EJC is involved in mRNA localization, as it has been observed of *oskar* mRNA in *D. melanogaster* (Le Hir et al., 2001a; Hachet and Ephrussi, 2004; Palacios et al., 2004). Some studies suggest that some proteins specially associate to spliced mRNAs in the nucleus and are involved in mRNA export in *S. pombe* (Carnahan et al., 2005; Thakurta et al., 2007).

7.3.2 Local mRNP anisotropy can affect NMD: a new model

An alternative explanation of why introns enhance NMD might be that spliced mRNAs are translated more efficiently and as a result PTCs are recognized more frequently, enhancing NMD. Several studies reporting that splicing can enhance translation support this view (Nott et al., 2003; Wiegand et al., 2003; Nott et al., 2004; Sanford et al., 2004; Ma et al., 2008; Michlewski et al., 2008). In particular, it has been reported that strengthening of splice signals enhances both translation and NMD, while mutations that weakened splicing have the opposite effects (Gudikote et al.,

2005). It has been suggested that the EJC enhances the initiation of the pioneer round of translation directly, but the efficiency of the first round might positively affect subsequent rounds and steady-state translation (Ishigaki et al., 2001; Ma et al., 2008). Splicing might also enhance translation in *S. cerevisiae* and in other yeast species including *S. pombe* (Juneau et al., 2006; Man and Pilpel, 2007). However, I found that splicing strongly enhances NMD only if the intron is positioned close to the PTC; indeed this enhancement disappears if the PTC is more than 240 nt away from the intron (Chapter V, Fig. 5.9 and Table 5.1). These observations argue that it is the proximity of an intron to PTCs, which influences NMD, not just whether the mRNA has undergone splicing.

In agreement with the above view, I found no correlation between changes in translation yields attributable to the position of the intron along the coding region and NMD. Thus, my data indicate that pre-mRNA splicing impinges on NMD directly, yet exclude that an EJC-like protein complex mediates the effect. While I propose that proteins remaining associated with the mRNA after splicing mediate the interaction, the effect is not caused by a specific factor deposited at a particular distance before the splice junction (Le Hir et al., 2000b; Le Hir et al., 2001b); instead, I postulate that there are splicing-associated proteins (SAP) that bind at either sides of the splice junction, and the recruitments of SAPs change the local mRNP configuration. Such splicing-specific changes in mRNP composition have been reported in recent studies: it has been found that in both in human cells and *S. cerevisiae* there are difference in mRNP composition around the exon junction between spliced and unspliced mRNA (Merz te al., 2007; Fabrizio et al., 2009). The region near the splicing joint, termed as the splicing affected region, where splicing could affect NMD. I postulate that translation termination in this region is not efficient during the pioneer round of

translation. If the termination codon were close to SAPs, the transient interaction would trigger strong NMD by accelerating the recruitment of UPF proteins. When termination occurs away before the splicing affected region, the mRNP will not influence termination.

I propose that these proteins deposited at either side of the junction change the local mRNP arrangement compared to the same sequence in the intronless reporter. This model postulates that during translation of newly synthesized transcripts, the mRNP local arrangement affects either translation termination or the fate of the post termination ribosome, leading to NMD (Fig. 7.1). It is also feasible that there is a connection between the splice region and the mRNP components outside the translated sequence, which could affect termination or post-termination events. Some of the physical interactions of splicing with 5' capping and 3' end formation, which occur during pre-mRNA processing, may be maintained in the mature mRNP (Proudfoot et al., 2002). In particular, interactions between cap binding complex (CBC) and splicing have been reported in *S. cerevisiae* (Colot et al., 1996; Fortes et al., 1999); notably the yeast CBC can interact with the translation initiation factor eIF4G plus the CBP80 subunit of CBC has a structural similarity to eIF4G (Fortes et al., 2000; Marintchev and Wagner, 2005); furthermore, NMD substrates are preferentially associated with CBC rather than the translation initiation factor eIF4E in mammalian cells (Ishigaki et al., 2001). A key assumption of the model is that the anisotropy in mRNP composition and structure is only transient; the prediction is that splicing only enhances the NMD of newly synthesized transcripts before the disassociation of the "processing" factors, in agreement with some earlier data indicating that steady-state mRNA is immune to splicing-dependent NMD in
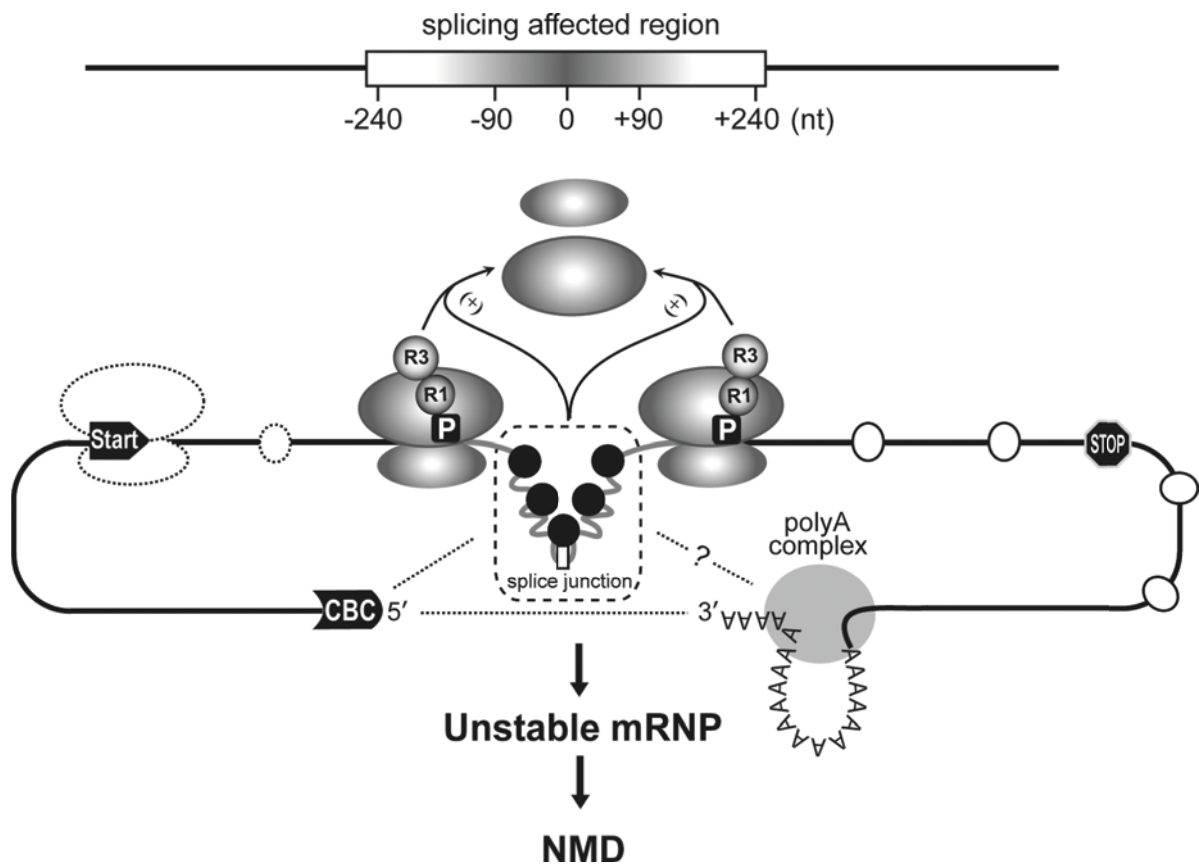
Figure 7.1 Anisotropy mRNP model of splicing dependent NMD

mRNP anisotropy around the splice junction enhances NMD. Pre-mRNA splicing deposits proteins around the splice junction. These proteins remain associated with the mRNA and change the local structure of the mRNP in nucleus and mRNA export to cytoplasm. The proteins remain associated with the mRNA transiently, possibly until the first round of translation. The cap binding complex (CBC) might be associated with the splicing region and together with other proteins affects either translation termination of the fate of the post-termination ribosome (e.g. ribosome release). White ovals are non-specific RNA binding proteins; black circles are proteins deposited by splicing. R1, translation release factor 1; R3, translation release factor 3.

mammalian cells (reviewed in Maquat, 2004). Instead, provided that the PTC is in a region not subjected to splicing-independent NMD-the transcripts are stable afterward.

Finally, there might be other mechanism that could explain the link between splicing and NMD. For example, the spliced mRNA that carries covalent modifications near the splicing junction which are absent in the mRNA generated by intronless reporters. In agreement with this speculation, it has been reported that adenosine methylation might be involved in pre-mRNA processing and mRNA translation (Carroll et al., 1990; Tuck et al., 1999).

7.4 How to explain the polarity effect in intronless genes?

7.4.1 The *faux* 3' UTR model does not explain NMD in *S. pombe*

In agreement with observations in other organisms, I observed PTCs induced NMD also in intronless reporters. However, NMD was apparent only with PTCs located at the beginning of the GFP coding region. Instead, PTCs in the second half of the coding region did not induce apparent NMD. Initially, these observations seem in agreement with the *faux* 3' UTR model (Chapter I, Section 1.2.4.2 and Fig. 1.1A). However, while this model predicts that the distance between PTC and poly(A) tail determine the extent of mRNA reduction, I found that extending the distance only slightly enhanced NMD (from the 18% mRNA reduction to 33%, PTC140). As seen in Table 7.2 (a lists the NMD phenotype in relation to 3' UTR length), the cumulative result of my study is that there is no apparent correlation between NMD and the distance of the PTC from the 3' end (Fig. 7.2). I also found no correlation with the distance from the 5' end. Furthermore, I found that PABPC is not required for NMD

Table 7.2 Relationship between NMD and the distance of PTC from the 3' end.

Constructs (pREP-)	PTC position (nt)	mRNA reduction (%)	Distance from 3' end to PTC (nt)
GFPwt	N	0	84
GFP-402	N	0	84
GFP-sh	N	0	84
GFP-102	730-732	-12	186
PTC140	418-420	18	411
PTC6-sh	16-18	89	411
GFP-418	730-732	37	502
PTC140-102	418-420	-7	513
PTC27	79-81	57	750
PTC140-ss	331-333	18	769
GFP-ss	331-333	19	769
PTC6	16-18	84	813
PTC140-402	418-420	33	813
PTC140-418	418-420	47	829
PTC27-102	79-81	73	852
PTC6-102	16-18	82	915
PTC27-ss	79-81	69	1120
PTC27-402	79-81	73	1152
PTC27-418	79-81	76	1168
PTC6-ss	16-18	83	1183
PTC6-402	16-18	87	1215
PTC6-418	16-18	84	1231

The gray columns indicate the variance of mRNA reductions in the PTC constructs with the similar distance between the PTC and the poly(A) tail.
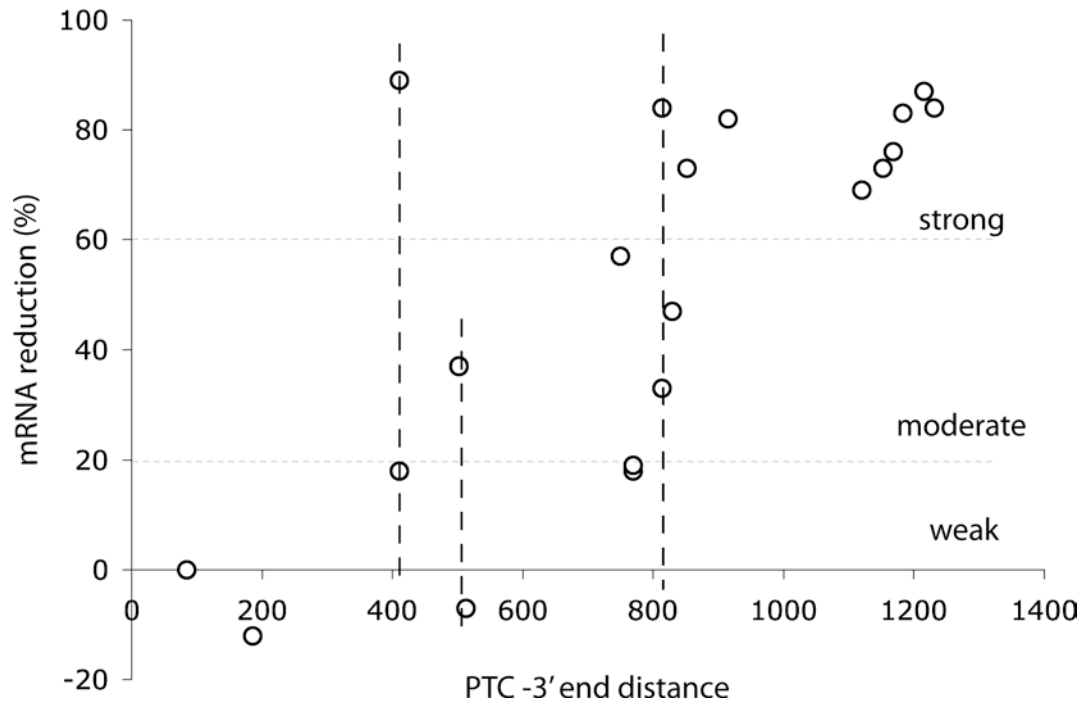
Figure 7.2 No clear correlation exists between NMD and distance of the PTC from the 3' end.

Circles show the extent of NMD in reporters with different 3'UTR lengths (constructs listed in Table 7.2). The X axis indicates the distance between PTCs and 3' end, the Y axis show mRNA reduction. The mRNA levels are normalized by the level of the mRNA in the PTC-less control (0% reduction).

of either spliced or unspliced mRNAs; this observation is in agreement with the recent report that neither the poly(A) tail nor PABPC are required for NMD in *S. cerevisiae* (Meaux et al., 2008). In conclusion, these results argue against the notion that it is primarily a long distance from either the 3' end or the 5' end which triggers splicing-independent NMD; instead the data suggest that some sequence elements flanking the PTC might affect NMD; this view is in agreement with earlier studies in *S. cerevisiae* which concluded that downstream sequence elements (DSE) located after the PTC modulate NMD (Peltz et al., 1993; Zhang et al., 1995).

7.4.2 Is NMD affected by a DSE sequences in *S. pombe*?

One of the observations of this study is that PTC6 always triggers strong NMD regardless of the distance between PTC6 and the 5' or 3' end (Table 7.1). Instead PTC140 is always weakly affected in all intronless reporters, again regardless of the distance from the ends (Table 7.1). Given that neither the distance of the PTC from the 3' end nor the distance from the 5' end is the essential NMD determinant , how can the polarity effect be explained?

Based on these experiments, I propose that there might be a sequence downstream of PTC6, which enhances NMD. The DSE should reside in the coding region between nt 19 and 327 (Fig. 7.3), corresponding to PTC6-downstream sequence of the short-PTC6 reporter (PTC6-sh) that showed very strong NMD. It is likely that the core DSE sequence is located between nt 19 and 81, because PTC6 triggered strong NMD but PTC27 only caused moderate NMD. Therefore, in the GFP coding region, there might be at least one DSE-like sequence enhancing NMD. This DSE is probably different
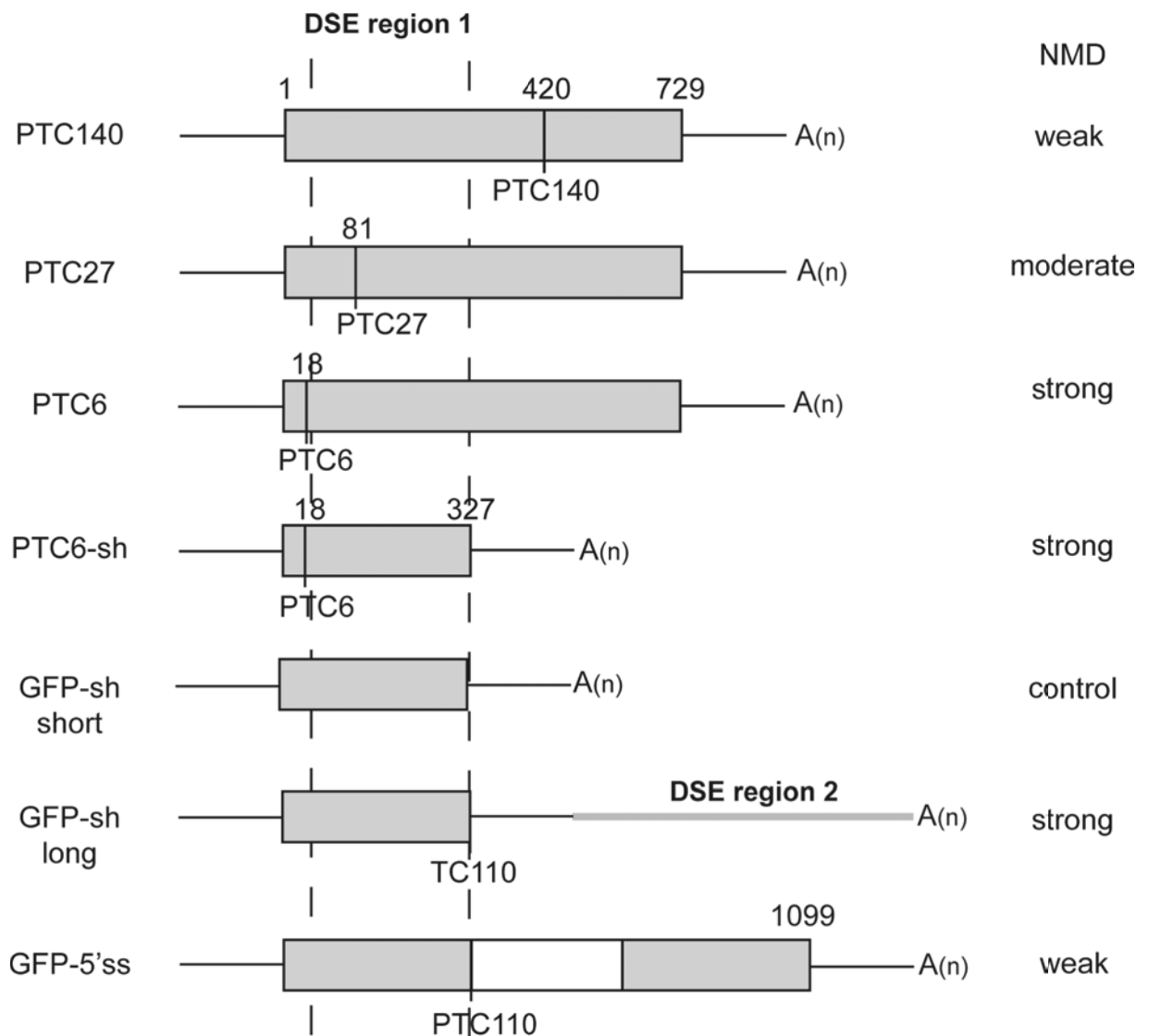
Figure 7.3 Multiple DSEs might be involved in NMD.

The schematic maps of the NMD constructs and northern blot are shown in Figure 4.1 Figure 4.5 and Figure 5.3. The presence of a DSE in Region 1 is suggested by the strong NMD in all reporters with PTC6 and by the weak NMD in PTC110 and PTC140. The presence of a DSE in Region 2 is suggested by the strong NMD of the long transcript with a further downstream poly(A) site (produced by the GFP-sh reporter). Probably there a sequence downstream of the normal poly(A) site that function as a DSE when included in the transcript.

from DSE consensus proposed in *S. cerevisiae* (Zhang et al., 1995; Ruiz-Echevarria et al., 1998a). In *S. cerevisiae* the DSE consensus is TGYYGATGYYYYY (Y stands for either T or C), but I could not find it in the GFP coding region. Perhaps the DSE might not be well-conserved sequence in different organisms and alternative sequence motifs can act as DSE in *S. pombe*. Perhaps in all cases the DSE enhances recruitment of UPF proteins.

7.5 NMD cannot be explained by a single mechanism

None of the available NMD models explains all published data. We have proposed that this is probably due to the fact that several mechanisms might act together to destabilize PTC-containing mRNAs (Brogna and Wen, 2009).

My data suggests that there are at least two different NMD mechanisms in *S. pombe;* splicing-dependent NMD and splicing-independent NMD. Given that deletion of *UPF1* and *UPF2* suppresses NMD in both cases, the mechanisms are probably not entirely distinct. However, the genetic analysis alone may be misleading because NMD mutants might have a non-specific mRNA stabilization phenotype: In *upf1Δ* and *upf2Δ* strains PTC-less mRNAs are also stabilized. This is not an idiosyncrasy of my experimental system, similar observations have been previously reported in *Drosophila* both in cell culture and in flies (Gatfield et al., 2003; Metzstein and Krasnow, 2006); examples of over-stabilization of control PTC-less mRNAs upon knock-down of NMD factors were also reported in mammalian cells yet not explained (Fig. 1 and 4, Eberle et al., 2008). Furthermore, early studies in budding yeast also reported the cases of over-stabilization of wild-type mRNAs in *UPF1* mutants (table 1 in Peltz et al., 1994). Additionally, microarray expression studies have clearly indicated that mRNAs with no obvious NMD features are also stabilized in NMD

mutants (Rehwinkel et al., 2005). Furthermore, I cannot conclude that the two mRNA reduction mechanisms are similar just because lack of UPF1 or UPF2 suppresses them both. Generally, NMD is viewed as a single well-defined biochemical process, which, similar to other cellular mechanisms, should be conserved across all groups of eukaryotic organisms. Against this view, there are cases in which UPF2 and UPF3 are not required for NMD (Gehring et al., 2005; Chan et al., 2007). On the basis of these conflicting observations it is plausible that the nonsense-mediated mRNA reduction (NMMR), typically attributed to NMD, is actually the compound effect of diverse mechanisms that destabilize translated mRNPs and destroys the released and unprotected mRNAs which are not efficiently translated (Wen and Brogna, 2008; Brogna and Wen, 2009)

Although various mechanisms may contribute to NMD, it is clear that there are features like the PTCs polar effect that are common to many experimental systems. We have recently proposed a model to explain the polar effect in general (Fig. 7.4) (Brogna and Wen, 2009). This model proposes that ribosome subunits are released efficiently after translation termination and the release is not affected by the position of PTCs. However, if releasing occurs at the beginning of the coding region, the long mRNA region downstream of the PTC will not be traversed by ribosome subunits, destabilizing the mRNP. Once the mRNP is destabilized, it will become accessible to nucleases, either exonucleases or endonucleases. As mentioned above, recent studies indicate that NMD might be initiated by endonucleases (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008; Eberle et al., 2009). This model does not imply that NMD is a passive mechanism that degrades mRNAs that are protected by ribosomes. Instead, it proposes an active role for UPF1: UPF1 triggers NMD by stimulating the

A

Early PTC

UPF1 (+) → eIF3

R3

P R1

60S

Start

STOP

40S

40S

40S

Cap 5'----------3'AAAAAAAAAAAAAAAAAA

Unstable translation circuit

NMD

B

Late PTC

UPF1 (+) → eIF3

R3

P R1

60S

Start

STOP

40S

40S

40S

Fast ribosome recycling

40S

40S

Cap 5'----------3'AAAAAAAAAAAAAAAAAA

Stable translation circuit

Normal mRNA stability
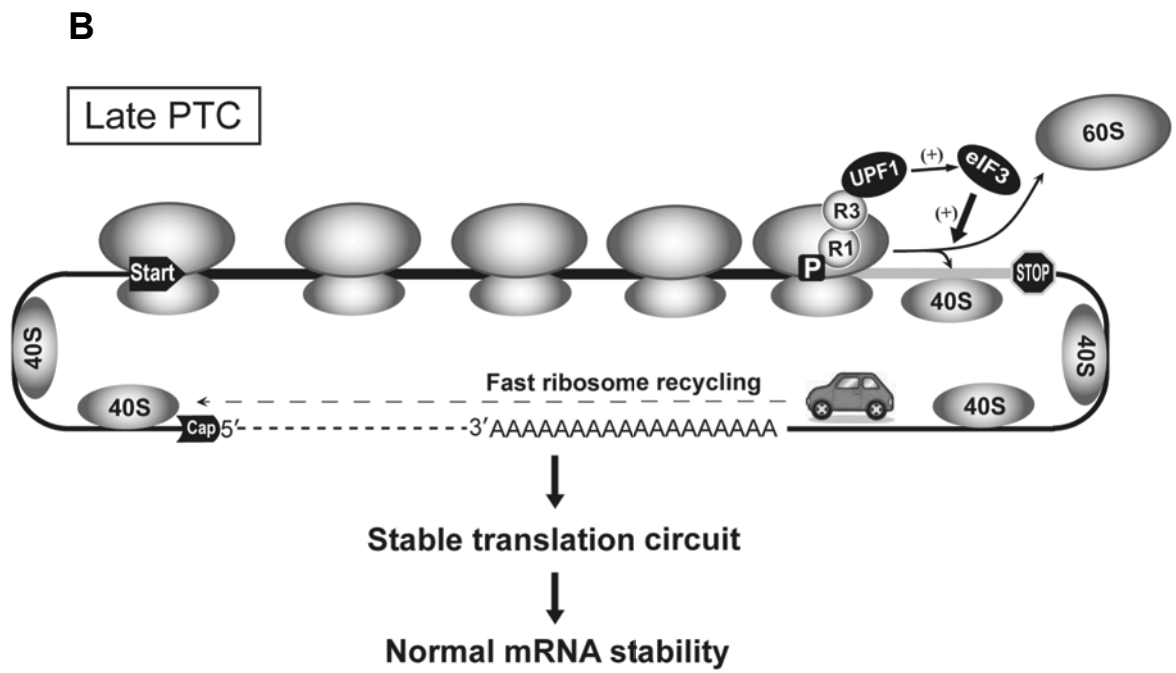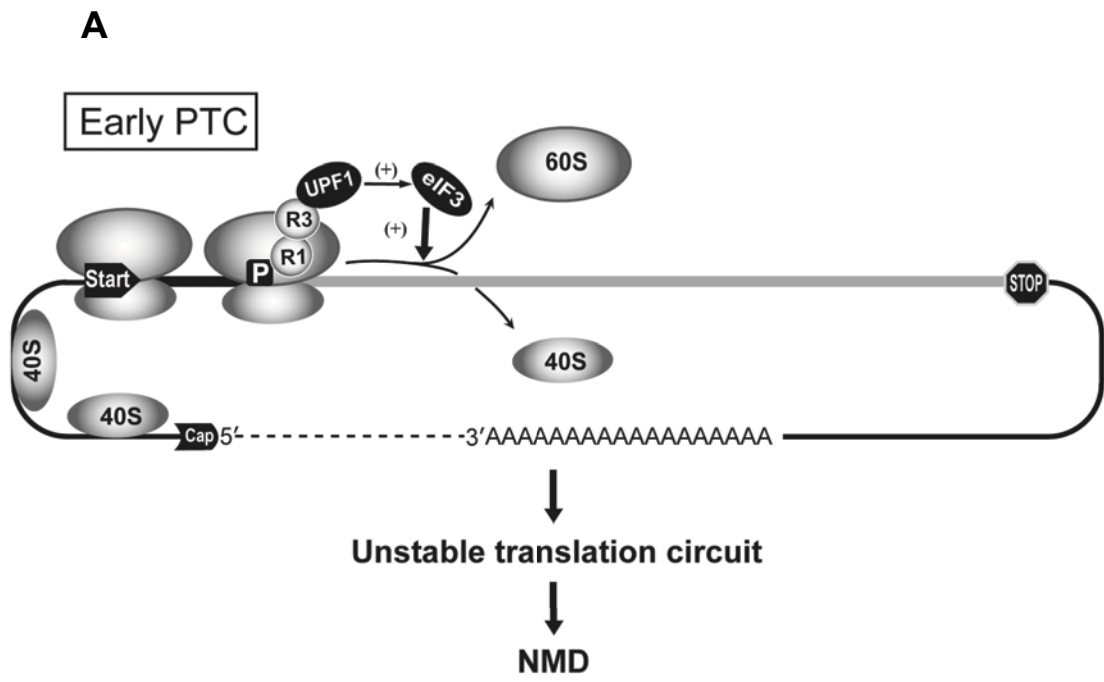
Figure 7.4 Ribosome release NMD model (Brogna and Wen, 2009).

(A) Following termination at early PTCs, the 80S splits into 60S and 40S subunits. UPF1 promotes the detachment of post-terminated ribosome from mRNA. Release of the ribosome subunits leaves the long region downstream of the PTC unprotected, destabilizing the mRNP and translation circuit, causing rapid mRNA decay.

(B) Termination at PTCs located near the normal 3′ UTR leaves only a short region that is not trafficked by translating ribosomes, which is not sufficient to destabilize the mRNP and translation circuit. Proximity to the normal 3′ UTR also promotes the recycling of 40S subunits, which remains associated with the mRNA and migrate to the 5′ end of the same mRNA. Both factors synergize to maintain normal mRNP stability.

release of ribosome subunits as the consequence of its interaction with eIF3. In UPF1 deleted strains, the sequence downstream of the PTC is still traversed by the ribosome subunits that fail to detach from the mRNAs. Ribosome release might be affected by depletion of the translation initiation factor eIF3. Recent studies proposed that eIF3 is required to release the ribosome subunits after translation termination (Pisarev et al., 2007). In agreement with our model, it has been reported that depletion of eIF3 affects NMD in both *S. cerevisiae* and human cells (Welch and Jacobson, 1999; Isken et al., 2008); recent studies also reported that UPF1 interacts with eIF3, and that phosphorylation of UPF1 suppresses translation initiation (Isken et al., 2008). Probably, the interaction between phosphorylated UPF1 and eIF3 prevents the joining of 40S and 60S into translation competent 80S.

7.6 Perspectives

This work clearly demonstrated that pre-mRNA splicing is linked to NMD in *S. pombe*. However, I don't yet know which are the *trans-acting* factors that mediate this linkage. The identification of these proteins would be extremely important to understand the mechanism that links splicing to NMD. It will also be important to further investigate the idea that NMD depend on *cis-acting* sequences like the DSE sequence proposed in the early studies in *S. cerevisiae*.

I believe that to further broaden our knowledge of NMD, it is essential to elucidate the molecular functions of UPF1. I found out that deletion of UPF1 impairs cell growth in *S. pombe* (4.5 hours/generation of *upf1Δ* strains compared to ~3 hours/generation of the wild-type strain in YES media, data not shown). Together with recent studies, this observation suggests that UPF1 might have an important function that is not related to NMD. In higher organism, UPF1 depletion delays the

cell cycle, (Dinman and Wickner, 1994; Azzalin and Lingner, 2006). The

identification of additional NMD factors and finding out the molecular functions of

UPF1 are clearly key priority for the future.

REFERENCES

Ahmed, F., Kumar, M., and Raghava, G. (2009). Prediction of polyadenylation signals in human DNA sequences using nucleotide frequencies. In Silico Biol *9*, 0012.

Alfa, C., Fantes, P., Hyams, J., and McLeod, M. (1993). Experiments with Fission Yeast: A Laboratory Course Manual (New York, Cold Spring Harbor Laboratory Press).

Altamura, N., Groudinsky, O., Dujardin, G., and Slonimski, P.P. (1992). NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-ligand motif and is involved in mitochondrial functions in Saccharomyces cerevisiae. J Mol Biol *224*, 575-87.

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*, 112-8.

Amrani, N., Ghosh, S., Mangus, D.A., and Jacobson, A. (2008). Translation factors promote the formation of two states of the closed-loop mRNP. Nature *453*, 1276-80.

Anders, K.R., Grimson, A., and Anderson, P. (2003). SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. Embo J *22*, 641-50.

Anderson, J.S., and Parker, R.P. (1998). The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. EMBO J *17*, 1497-506.

Aoufouchi, S., Yelamos, J., and Milstein, C. (1996). Nonsense mutations inhibit RNA splicing in a cell-free system: Recognition of mutant codon is independent of protein synthesis. Cell *85*, 415-422.

Aranda, A., and Proudfoot, N.J. (1999). Definition of transcriptional pause elements in fission yeast. Mol Cell Biol *19*, 1251-61.

Atkin, A.L., Altamura, N., Leeds, P., and Culbertson, M.R. (1995). The majority of yeast UPF1 co-localizes with polyribosomes in the cytoplasm. Mol Biol Cell *6*, 611-25.

Atkin, A.L., Schenkman, L.R., Eastham, M., Dahlseid, J.N., Lelivelt, M.J., and Culbertson, M.R. (1997). Relationship between yeast polyribosomes and Upf proteins required for nonsense mRNA decay. J Biol Chem *272*, 22163-72.

Ausubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and K., S. (1996). Current Protocols in Molecular Biology (New York, John Wiley & Sons).

Azzalin, C.M., and Lingner, J. (2006). The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. Curr Biol *16*, 433-9.

Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science *318*, 798-801.

Baek, D., and Green, P. (2005). Sequence conservation, relative isoform frequencies, and nonsense-mediated decay in evolutionarily conserved alternative splicing. Proc Natl Acad Sci U S A *102*, 12813-8.

Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., 3rd, 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*, 943-51.

Baker, K.E., and Parker, R. (2006). Conventional 3' end formation is not required for NMD substrate recognition in Saccharomyces cerevisiae. RNA *12*, 1441-5.

Baserga, S.J., and Benz, E.J., Jr. (1992). Beta-globin nonsense mutation: deficient accumulation of mRNA occurs despite normal cytoplasmic stability. Proc Natl Acad Sci U S A *89*, 2935-9.

Basi, G., Schmid, E., and Maundrell, K. (1993). TATA box mutations in the Schizosaccharomyces pombe nmt1 promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. Gene *123*, 131-6.

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

Belgrader, P., Cheng, J., Zhou, X., Stephenson, L.S., and Maquat, L.E. (1994). Mammalian nonsense codons can be cis effectors of nuclear mRNA half-life. Mol Cell Biol *14*, 8219-28.

Bhattacharya, A., Czaplinski, K., Trifillis, P., He, F., Jacobson, A., and Peltz, S.W. (2000). Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay. RNA *6*, 1226-35.

Bidou, L., Stahl, G., Hatin, I., Namy, O., Rousset, J.P., and Farabaugh, P.J. (2000). Nonsense-mediated decay mutants do not affect programmed -1 frameshifting. RNA *6*, 952-61.

Blencowe, B.J. (2006). Alternative splicing: new insights from global analyses. Cell *126*, 37-47.

Brogna, S. (1999). Nonsense mutations in the alcohol dehydrogenase gene of Drosophila melanogaster correlate with an abnormal 3 ' end processing of the corresponding pre-mRNA. RNA *5*, 562-573.

Brogna, S., Sato, T.A., and Rosbash, M. (2002). Ribosome components are associated with sites of transcription. Mol Cell *10*, 93-104.

Brogna, S., and Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. Nat Struct Mol Biol *16*, 107-13.

Brown, C.E., and Sachs, A.B. (1998). Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. Mol Cell Biol *18*, 6548-59.

Brumbaugh, K.M., Otterness, D.M., Geisen, C., Oliveira, V., Brognard, J., Li, X., Lejeune, F., Tibbetts, R.S., Maquat, L.E., and Abraham, R.T. (2004). The mRNA surveillance protein hSMG-1 functions in genotoxic stress response pathways in mammalian cells. Mol Cell *14*, 585-98.

Brune, C., Munchel, S.E., Fischer, N., Podtelejnikov, A.V., and Weis, K. (2005). Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. RNA *11*, 517-31.

Buhler, M., Paillusson, A., and Muhlemann, O. (2004). Efficient downregulation of immunoglobulin mu mRNA with premature translation-termination codons requires the 5'-half of the VDJ exon. Nucleic Acids Res *32*, 3304-15.

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

Buhler, M., Wilkinson, M.F., and Muhlemann, O. (2002). Intranuclear degradation of nonsense codon-containing mRNA. EMBO Rep *3*, 646-51.

Cali, B.M., and Anderson, P. (1998). mRNA surveillance mitigates genetic dominance in Caenorhabditis elegans. Mol Gen Genet *260*, 176-84.

Cali, B.M., Kuchma, S.L., Latham, J., and Anderson, P. (1999). smg-7 is required for mRNA surveillance in Caenorhabditis elegans. Genetics *151*, 605-16.

Caputi, M., Kendzior, R.J., Jr., and Beemon, K.L. (2002). A nonsense mutation in the fibrillin-1 gene of a Marfan syndrome patient induces NMD and disrupts an exonic splicing enhancer. Genes Dev *16*, 1754-9.

Carnahan, R.H., Feoktistova, A., Ren, L., Niessen, S., Yates, J.R., 3rd, and Gould, K.L. (2005). Dim1p is required for efficient splicing and export of mRNA encoding lid1p, a component of the fission yeast anaphase-promoting complex. Eukaryot Cell *4*, 577-87.

Carroll, S.M., Narayan, P., and Rottman, F.M. (1990). N6-methyladenosine residues in an intron-specific region of prolactin pre-mRNA. Mol Cell Biol *10*, 4456-65.

Cartegni, L., and Krainer, A.R. (2002). Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat Genet *30*, 377-84.

Carter, M.S., Doskow, J., Morris, P., Li, S., Nhim, R.P., Sandstedt, S., and Wilkinson, M.F. (1995). A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. J Biol Chem *270*, 28995-9003.

Carter, M.S., Li, S., and Wilkinson, M.F. (1996). A splicing-dependent regulatory mechanism that detects translation signals. Embo J *15*, 5965-75.

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*, 85-93.

Chan, C.C., Dostie, J., Diem, M.D., Feng, W., Mann, M., Rappsilber, J., and Dreyfuss, G. (2004). eIF4A3 is a novel component of the exon junction complex. RNA *10*, 200-9.

Chan, W.K., Huang, L., Gudikote, J.P., Chang, Y.F., Imam, J.S., MacLean, J.A., 2nd, and Wilkinson, M.F. (2007). An alternative branch of the nonsense-mediated decay pathway. Embo J *26*, 1820-30.

Chang, J.C., and Kan, Y.W. (1979). beta 0 thalassemia, a nonsense mutation in man. Proc Natl Acad Sci U S A *76*, 2886-9.

Chen, C.Y., and Shyu, A.B. (2003). Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Mol Cell Biol *23*, 4805-13.

Cheng, J., Belgrader, P., Zhou, X., and Maquat, L.E. (1994). Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance. Mol Cell Biol *14*, 6317-25.

Cheng, J., and Maquat, L.E. (1993). Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. Mol Cell Biol *13*, 1892-902.

Cheng, Z., Muhlrad, D., Lim, M.K., Parker, R., and Song, H. (2007). Structural and functional insights into the human Upf1 helicase core. Embo J *26*, 253-64.

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*, 77-87.

Cho, H., Kim, K.M., and Kim, Y.K. (2009). Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex. Mol Cell *33*, 75-86.

Christensen, A.K., and Bourne, C.M. (1999). Shape of large bound polysomes in cultured fibroblasts and thyroid epithelial cells. Anat Rec *255*, 116-29.

Christensen, A.K., Kahn, L.E., and Bourne, C.M. (1987). Circular polysomes predominate on the rough endoplasmic reticulum of somatotropes and mammotropes in the rat anterior pituitary. Am J Anat *178*, 1-10.

Clerici, M., Mourao, A., Gutsche, I., Gehring, N.H., Hentze, M.W., Kulozik, A., Kadlec, J., Sattler, M., and Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. EMBO J *28*, 2293-2306.

Clissold, P.M., and Ponting, C.P. (2000). PIN domains in nonsense-mediated mRNA decay and RNAi. Curr Biol *10*, R888-90.

Colot, H.V., Stutz, F., and Rosbash, M. (1996). The yeast splicing factor Mud13p is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. Genes Dev *10*, 1699-708.

Conti, E., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr Opin Cell Biol 17, 316-25.

Cui, Y., Dinman, J.D., Kinzy, T.G., and Peltz, S.W. (1998). The Mof2/Sui1 protein is a general monitor of translational accuracy. Mol Cell Biol *18*, 1506-16.

Cui, Y., Dinman, J.D., and Peltz, S.W. (1996). Mof4-1 is an allele of the UPF1/IFS2 gene which affects both mRNA turnover and -1 ribosomal frameshifting efficiency. Embo J *15*, 5726-36.

Cui, Y., Gonzalez, C.I., Kinzy, T.G., Dinman, J.D., and Peltz, S.W. (1999). Mutations in the MOF2/SUI1 gene affect both translation and nonsense-mediated mRNA decay. RNA *5*, 794-804.

Cui, Y., Hagan, K.W., Zhang, S., and Peltz, S.W. (1995). Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. Genes Dev *9*, 423-36.

Culbertson, M.R., and Neeno-Eckwall, E. (2005). Transcript selection and the recruitment of mRNA decay factors for NMD in Saccharomyces cerevisiae. RNA *11*, 1333-9.

Culbertson, M.R., Underbrink, K.M., and Fink, G.R. (1980). Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. Genetics *95*, 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*, 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*, 610-23.

Dahlseid, J.N., Lew-Smith, J., Lelivelt, M.J., Enomoto, S., Ford, A., Desruisseaux, M., McClellan, M., Lue, N., Culbertson, M.R., and Berman, J. (2003). mRNAs encoding telomerase components and regulators are controlled by UPF genes in Saccharomyces cerevisiae. Eukaryot Cell *2*, 134-42.

Danilevich, V.N., and Grishin, E.V. (2002). A new approach to the isolation of genomic DNA samples from yeast and fungi: preparation of DNA-containing cell envelopes and their use in PCR. Bioorg Khim *28*, 156-67.

Decker, C.J., and Parker, R. (2002). mRNA decay enzymes: decappers conserved between yeast and mammals. Proc Natl Acad Sci U S A *99*, 12512-4.

Decottignies, A., Sanchez-Perez, I., and Nurse, P. (2003). Schizosaccharomyces pombe essential genes: a pilot study. Genome Res *13*, 399-406.

Degot, S., Le Hir, H., Alpy, F., Kedinger, V., Stoll, I., Wendling, C., Seraphin, B., Rio, M.C., and Tomasetto, C. (2004). Association of the breast cancer protein MLN51 with the exon junction complex via its speckle localizer and RNA binding module. J Biol Chem *279*, 33702-15.

Dietz, H.C., and Kendzior, R.J., Jr. (1994). Maintenance of an open reading frame as an additional level of scrutiny during splice site selection. Nat Genet *8*, 183-8.

Dietz, H.C., McIntosh, I., Sakai, L.Y., Corson, G.M., Chalberg, S.C., Pyeritz, R.E., and Francomano, C.A. (1993). Four novel FBN1 mutations: significance for mutant transcript level and EGF-like domain calcium binding in the pathogenesis of Marfan syndrome. Genomics *17*, 468-75.

Dinman, J.D., and Wickner, R.B. (1994). Translational maintenance of frame: mutants of Saccharomyces cerevisiae with altered -1 ribosomal frameshifting efficiencies. Genetics *136*, 75-86.

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

Dower, K., and Rosbash, M. (2002). T7 RNA polymerase-directed transcripts are processed in yeast and link 3' end formation to mRNA nuclear export. RNA *8*, 686-97.

Eberle, A.B., Lykke-Andersen, S., Muhlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat Struct Mol Biol *16*, 49-55.

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

Enomoto, S., Glowczewski, L., Lew-Smith, J., and Berman, J.G. (2004). Telomere cap components influence the rate of senescence in telomerase-deficient yeast cells. Mol Cell Biol *24*, 837-45.

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Luhrmann, R. (2009). The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. Mol Cell 36, 593-608.

Falchuk, K.H., Mazus, B., Ulpino, L., and Vallee, B.L. (1976). Euglena gracilis DNA dependent RNA polymerase II: a zinc metalloenzyme. Biochemistry 15, 4468-75.

Ferraiuolo, M.A., Lee, C.S., Ler, L.W., Hsu, J.L., Costa-Mattioli, M., Luo, M.J., Reed, R., and Sonenberg, N. (2004). A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay. Proc Natl Acad Sci U S A *101*, 4118-23.

Ford, A.S., Guan, Q., Neeno-Eckwall, E., and Culbertson, M.R. (2006). Ebs1p, a negative regulator of gene expression controlled by the Upf proteins in the yeast Saccharomyces cerevisiae. Eukaryot Cell *5*, 301-12.

Forsburg, S.L., and Rhind, N. (2006). Basic methods for fission yeast. Yeast *23*, 173-83.

Fortes, P., Bilbao-Cortes, D., Fornerod, M., Rigaut, G., Raymond, W., Seraphin, B., and Mattaj, I.W. (1999). Luc7p, a novel yeast U1 snRNP protein with a role in 5' splice site recognition. Genes Dev *13*, 2425-38.

Fortes, P., Inada, T., Preiss, T., Hentze, M.W., Mattaj, I.W., and Sachs, A.B. (2000). The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. Mol Cell *6*, 191-6.

Frolova, L., Le Goff, X., Rasmussen, H.H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A.L., Celis, J.E., Philippe, M.*, et al.* (1994). A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. Nature *372*, 701-3.

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*, 334-41.

Fukuhara, N., Ebert, J., Unterholzner, L., Lindner, D., Izaurralde, E., and Conti, E. (2005). SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. Mol Cell *17*, 537-47.

Funakoshi, Y., Doi, Y., Hosoda, N., Uchida, N., Osawa, M., Shimada, I., Tsujimoto, M., Suzuki, T., Katada, T., and Hoshino, S. (2007). Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. Genes Dev *21*, 3135-48.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat Rev Mol Cell Biol *8*, 113-26.

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

Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P., and Izaurralde, E. (2003). Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. Embo J *22*, 3960-70.

Gehring, N.H., Kunz, J.B., Neu-Yilik, G., Breit, S., Viegas, M.H., Hentze, M.W., and Kulozik, A.E. (2005). Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Mol Cell *20*, 65-75.

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*, 939-49.

Gersappe, A., and Pintel, D.J. (1999). A premature termination codon interferes with the nuclear function of an exon splicing enhancer in an open reading frame-dependent manner. Mol Cell Biol *19*, 1640-1650.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-41.

Giorgi, C., Yeo, G.W., Stone, M.E., Katz, D.B., Burge, C., Turrigiano, G., and Moore, M.J. (2007). The EJC factor eIF4AIII modulates synaptic strength and neuronal protein expression. Cell *130*, 179-91.

Glavan, F., Behm-Ansmant, I., Izaurralde, E., and Conti, E. (2006). Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. Embo J *25*, 5117-25.

Gonzalez, C.I., Ruiz-Echevarria, 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*, 489-99.

Graber, J.H., McAllister, G.D., and Smith, T.F. (2002). Probabilistic prediction of Saccharomyces cerevisiae mRNA 3'-processing sites. Nucleic Acids Res *30*, 1851-8.

Green, R.E., Lewis, B.P., Hillman, R.T., Blanchette, M., Lareau, L.F., Garnett, A.T., Rio, D.C., and Brenner, S.E. (2003). Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. Bioinformatics *19 Suppl 1*, i118-21.

Grigull, J., Mnaimneh, S., Pootoolal, J., Robinson, M.D., and Hughes, T.R. (2004). Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. Mol Cell Biol *24*, 5534-47.

Grimson, A., O'Connor, S., Newman, C.L., and Anderson, P. (2004). SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans. Mol Cell Biol *24*, 7483-90.

Gross, S., and Moore, C.L. (2001). Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation. Mol Cell Biol *21*, 8045-55.

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*, e203.

Gudikote, J.P., Imam, J.S., Garcia, R.F., and Wilkinson, M.F. (2005). RNA splicing promotes translation and RNA surveillance. Nat Struct Mol Biol *12*, 801-9.

Hachet, O., and Ephrussi, A. (2004). Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. Nature *428*, 959-63.

Hampsey, M., Na, J.G., Pinto, I., Ware, D.E., and Berroteran, R.W. (1991). Extragenic suppressors of a translation initiation defect in the cyc1 gene of Saccharomyces cerevisiae. Biochimie *73*, 1445-55.

Harger, J.W., and Dinman, J.D. (2004). Evidence against a direct role for the Upf proteins in frameshifting or nonsense codon readthrough. RNA *10*, 1721-9.

He, F., Brown, A.H., and Jacobson, A. (1996). Interaction between Nmd2p and Upf1p is required for activity but not for dominant-negative inhibition of the nonsense-mediated mRNA decay pathway in yeast. RNA *2*, 153-70.

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*, 1580-94.

He, F., and Jacobson, A. (2001). Upf1p, Nmd2p, and Upf3p regulate the decapping and exonucleolytic degradation of both nonsense-containing mRNAs and wild-type mRNAs. Mol Cell Biol *21*, 1515-30.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast. Mol Cell *12*, 1439-52.

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., and Jacobson, A. (1993). Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1-mutant. Proc Natl Acad Sci U S A *90*, 7034-8.

Heitman, J., Movva, N.R., and Hall, M.N. (1991). Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. Science *253*, 905-9.

Helliwell, S.B., Wagner, P., Kunz, J., Deuter-Reinhard, M., Henriquez, R., and Hall, M.N. (1994). TOR1 and TOR2 are structurally and functionally similar but not identical phosphatidylinositol kinase homologues in yeast. Mol Biol Cell *5*, 105-18.

Hentze, M.W., and Kulozik, A.E. (1999). A perfect message: RNA surveillance and nonsense-mediated decay. Cell *96*, 307-10.

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*, 301-13.

Holbrook, J.A., Neu-Yilik, G., Hentze, M.W., and Kulozik, A.E. (2004). Nonsense-mediated decay approaches the clinic. Nat Genet *36*, 801-8.

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*, 16677-80.

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

Hosoda, N., Kobayashi, T., Uchida, N., Funakoshi, Y., Kikuchi, Y., Hoshino, S., and Katada, T. (2003). Translation termination factor eRF3 mediates mRNA decay through the regulation of deadenylation. J Biol Chem *278*, 38287-91.

Huntzinger, E., Kashima, I., Fauser, M., Sauliere, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA *14*, 2609-17.

Iborra, F.J., Jackson, D.A., and Cook, P.R. (2001). Coupled transcription and translation within nuclei of mammalian cells. Science *293*, 1139-42.

Ibrahim, H., Wilusz, J., and Wilusz, C.J. (2008). RNA recognition by 3'-to-5' exonucleases: the substrate perspective. Biochim Biophys Acta *1779*, 256-65.

Imataka, H., Gradi, A., and Sonenberg, N. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. EMBO J *17*, 7480-9.

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*, 218-23.

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

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*, 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.

Juneau, K., Miranda, M., Hillenmeyer, M.E., Nislow, C., and Davis, R.W. (2006). Introns regulate RNA and protein abundance in yeast. Genetics *174*, 511-8.

Kadlec, J., Guilligay, D., Ravelli, R.B., and Cusack, S. (2006). Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1. RNA *12*, 1817-24.

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*, 330-7.

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*, 355-67.

Kataoka, N., Diem, M.D., Kim, V.N., Yong, J., and Dreyfuss, G. (2001). Magoh, a human homolog of Drosophila mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. Embo J *20*, 6424-33.

Kataoka, N., Yong, J., Kim, V.N., Velazquez, F., Perkinson, R.A., Wang, F., and Dreyfuss, G. (2000). Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. Mol Cell *6*, 673-82.

Kaygun, H., and Marzluff, W.F. (2005). Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. Nat Struct Mol Biol *12*, 794-800.

Kebaara, B.W., and Atkin, A.L. (2009). Long 3'-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in Saccharomyces cerevisiae. Nucleic Acids Res.

Kertesz, S., Kerenyi, Z., Merai, Z., Bartos, I., Palfy, T., Barta, E., and Silhavy, D. (2006). Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucleic Acids Res *34*, 6147-57.

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*, 2545-56.

Kim, V.N., Kataoka, N., and Dreyfuss, G. (2001a). Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. Science *293*, 1832-6.

Kim, V.N., Yong, J., Kataoka, N., Abel, L., Diem, M.D., and Dreyfuss, G. (2001b). The Y14 protein communicates to the cytoplasm the position of exon-exon junctions. EMBO J *20*, 2062-8.

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*, 195-208.

Kobayashi, T., Funakoshi, Y., Hoshino, S., and Katada, T. (2004). The GTP-binding release factor eRF3 as a key mediator coupling translation termination to mRNA decay. J Biol Chem *279*, 45693-700.

Kozlov, G., Trempe, J.F., Khaleghpour, K., Kahvejian, A., Ekiel, I., and Gehring, K. (2001). Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. Proc Natl Acad Sci U S A *98*, 4409-13.

Kressler, D., de la Cruz, J., Rojo, M., and Linder, P. (1997). Fal1p is an essential DEAD-box protein involved in 40S-ribosomal-subunit biogenesis in Saccharomyces cerevisiae. Mol Cell Biol *17*, 7283-94.

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

Lackner, D.H., Beilharz, T.H., Marguerat, S., Mata, J., Watt, S., Schubert, F., Preiss, T., and Bahler, J. (2007). A network of multiple regulatory layers shapes gene expression in fission yeast. Mol Cell *26*, 145-55.

Lau, C.K., Diem, M.D., Dreyfuss, G., and Van Duyne, G.D. (2003). Structure of the Y14-Magoh core of the exon junction complex. Curr Biol *13*, 933-41.

Le, H., Tanguay, R.L., Balasta, M.L., Wei, C.C., Browning, K.S., Metz, A.M., Goss, D.J., and Gallie, D.R. (1997). Translation initiation factors eIF-iso4G and eIF-4B interact with the poly(A)-binding protein and increase its RNA binding activity. J Biol Chem *272*, 16247-55.

Le Hir, H., Gatfield, D., Braun, I.C., Forler, D., and Izaurralde, E. (2001a). The protein Mago provides a link between splicing and mRNA localization. EMBO Rep *2*, 1119-24.

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

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

Le Hir, H., Moore, M.J., and Maquat, L.E. (2000b). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. Genes Dev *14*, 1098-108.

LeBlanc, J.J., and Beemon, K.L. (2004). Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging. J Virol *78*, 5139-46.

Lee, S.I., Umen, J.G., and Varmus, H.E. (1995). A genetic screen identifies cellular factors involved in retroviral -1 frameshifting. Proc Natl Acad Sci U S A *92*, 6587-91.

Leeds, P., Peltz, S.W., Jacobson, A., and Culbertson, M.R. (1991). The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev *5*, 2303-14.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Mol Cell Biol *12*, 2165-77.

Legrain, P., and Rosbash, M. (1989). Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. Cell *57*, 573-83.

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

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol Cell *12*, 675-87.

Lemieux, C., and Bachand, F. (2009). Cotranscriptional recruitment of the nuclear poly(A)-binding protein Pab2 to nascent transcripts and association with translating mRNPs. Nucleic Acids Res.

Letunic, I., Goodstadt, L., Dickens, N.J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R.R., Ponting, C.P., and Bork, P. (2002). Recent improvements to the SMART domain-based sequence annotation resource. Nucleic Acids Res *30*, 242-4.

Lew, J.E., Enomoto, S., and Berman, J. (1998). Telomere length regulation and telomeric chromatin require the nonsense-mediated mRNA decay pathway. Mol Cell Biol *18*, 6121-30.

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

Li, B., Wachtel, C., Miriami, E., Yahalom, G., Friedlander, G., Sharon, G., Sperling, R., and Sperling, J. (2002). Stop codons affect 5' splice site selection by surveillance of splicing. Proc Natl Acad Sci U S A *99*, 5277-82.

Li, S., and Wilkinson, M.F. (1998). Nonsense surveillance in lymphocytes? Immunity *8*, 135-41.

Longman, D., Plasterk, R.H., Johnstone, I.L., and Caceres, J.F. (2007). Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes Dev *21*, 1075-85.

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

Lozano, F., Maertzdorf, B., Pannell, R., and Milstein, C. (1994). Low Cytoplasmic Messenger-RNA Levels of Immunoglobulin-Kappa Light-Chain Genes Containing Nonsense Codons Correlate with Inefficient Splicing. Embo J *13*, 4617-4622.

Luke, B., Azzalin, C.M., Hug, N., Deplazes, A., Peter, M., and Lingner, J. (2007). Saccharomyces cerevisiae Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. Nucleic Acids Res *35*, 7688-97.

Luo, M.L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., and Reed, R. (2001). Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. Nature *413*, 644-7.

Lykke-Andersen, J. (2002). Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. Mol Cell Biol *22*, 8114-21.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell *103*, 1121-31.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science *293*, 1836-9.

Ma, X.M., Yoon, S.O., Richardson, C.J., Julich, K., and Blenis, J. (2008). SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. Cell *133*, 303-13.

Maderazo, A.B., He, F., Mangus, D.A., and Jacobson, A. (2000). Upf1p control of nonsense mRNA translation is regulated by Nmd2p and Upf3p. Mol Cell Biol *20*, 4591-603.

Malapeira, J., Moldon, A., Hidalgo, E., Smith, G.R., Nurse, P., and Ayte, J. (2005). A meiosis-specific cyclin regulated by splicing is required for proper progression through meiosis. Mol Cell Biol *25*, 6330-7.

Man, O., and Pilpel, Y. (2007). Differential translation efficiency of orthologous genes is involved in phenotypic divergence of yeast species. Nat Genet *39*, 415-21.

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

Maquat, L.E. (1995). When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. RNA *1*, 453-65.

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

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*, 445-56.

Maquat, L.E., and Serin, G. (2001). Nonsense-mediated mRNA decay: insights into mechanism from the cellular abundance of human Upf1, Upf2, Upf3, and Upf3X proteins. Cold Spring Harb Symp Quant Biol *66*, 313-20.

Marcotrigiano, J., Lomakin, I.B., Sonenberg, N., Pestova, T.V., Hellen, C.U., and Burley, S.K. (2001). A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. Mol Cell *7*, 193-203.

Marintchev, A., and Wagner, G. (2005). eIF4G and CBP80 share a common origin and similar domain organization: implications for the structure and function of eIF4G. Biochemistry *44*, 12265-72.

Martineau, Y., Derry, M.C., Wang, X., Yanagiya, A., Berlanga, J.J., Shyu, A.B., Imataka, H., Gehring, K., and Sonenberg, N. (2008). Poly(A)-binding protein-interacting protein 1 binds to eukaryotic translation initiation factor 3 to stimulate translation. Mol Cell Biol *28*, 6658-67.

Matsuda, D., Hosoda, N., Kim, Y.K., and Maquat, L.E. (2007). Failsafe nonsense-mediated mRNA decay does not detectably target eIF4E-bound mRNA. Nat Struct Mol Biol *14*, 974-9.

Matsuo, Y., Asakawa, K., Toda, T., and Katayama, S. (2006). A rapid method for protein extraction from fission yeast. Biosci Biotechnol Biochem *70*, 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*, 1289-305.

Maundrell, K. (1993). Thiamine-repressible expression vectors pREP and pRIP for fission yeast. Gene *123*, 127-30.

Mazza, C., Ohno, M., Segref, A., Mattaj, I.W., and Cusack, S. (2001). Crystal structure of the human nuclear cap binding complex. Mol Cell *8*, 383-96.

McGlincy, N.J., and Smith, C.W. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem Sci *33*, 385-93.

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*, 134-40.

Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum Mol Genet *10*, 99-105.

Mendell, J.T., and Dietz, H.C. (2001). When the message goes awry: disease-producing mutations that influence mRNA content and performance. Cell *107*, 411-4.

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*, 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 mutes genomic noise. Nat Genet *36*, 1073-8.

Merz, C., Urlaub, H., Will, C.L., and Luhrmann, R. (2007). Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. RNA 13, 116-28.

Metzstein, M.M., and Krasnow, M.A. (2006). Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. PLoS Genet *2*, e180.

Michlewski, G., Sanford, J.R., and Caceres, J.F. (2008). The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. Mol Cell *30*, 179-89.

Mitrovich, Q.M., and Anderson, P. (2005). mRNA surveillance of expressed pseudogenes in C. elegans. Curr Biol *15*, 963-7.

Moldon, A., Malapeira, J., Gabrielli, N., Gogol, M., Gomez-Escoda, B., Ivanova, T., Seidel, C., and Ayte, J. (2008). Promoter-driven splicing regulation in fission yeast. Nature *455*, 997-1000.

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

Moriarty, P.M., Reddy, C.C., and Maquat, L.E. (1998). Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA. Mol Cell Biol *18*, 2932-9.

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

Morrison, M., Harris, K.S., and Roth, M.B. (1997). smg mutants affect the expression of alternatively spliced SR protein mRNAs in Caenorhabditis elegans. Proc Natl Acad Sci U S A *94*, 9782-5.

Morse, D.E., and Yanofsky, C. (1969). Polarity and the degradation of mRNA. Nature *224*, 329-31.

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

Muhlemann, O., Mock-Casagrande, C.S., Wang, J., Li, S., Custodio, N., Carmo-Fonseca, M., Wilkinson, M.F., and Moore, M.J. (2001). Precursor RNAs harboring nonsense codons accumulate near the site of transcription. Mol Cell *8*, 33-43.

Muhlrad, D., and Parker, R. (1999). Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. RNA *5*, 1299-307.

Nakamura, T.M., Morin, G.B., Chapman, K.B., Weinrich, S.L., Andrews, W.H., Lingner, J., Harley, C.B., and Cech, T.R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. Science *277*, 955-9.

Neu-Yilik, G., Gehring, N.H., Thermann, R., Frede, U., Hentze, M.W., and Kulozik, A.E. (2001). Splicing and 3' end formation in the definition of nonsense-mediated decay-competent human beta-globin mRNPs. Embo J *20*, 532-40.

Nott, A., Le Hir, H., and Moore, M.J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev *18*, 210-22.

Nott, A., Meislin, S.H., and Moore, M.J. (2003). A quantitative analysis of intron effects on mammalian gene expression. RNA *9*, 607-17.

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*, 1187-200.

Oliveira, C.C., and McCarthy, J.E. (1995). The relationship between eukaryotic translation and mRNA stability. A short upstream open reading frame strongly inhibits translational initiation and greatly accelerates mRNA degradation in the yeast Saccharomyces cerevisiae. J Biol Chem *270*, 8936-43.

Page, M.F., Carr, B., Anders, K.R., Grimson, A., and Anderson, P. (1999). SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast. Mol Cell Biol *19*, 5943-51.

Pal, M., Ishigaki, Y., Nagy, E., and Maquat, L.E. (2001). Evidence that phosphorylation of human Upf1 protein varies with intracellular location and is mediated by a wortmannin-sensitive and rapamycin-sensitive PI 3-kinase-related kinase signaling pathway. RNA *7*, 5-15.

Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature *427*, 753-7.

Pan, Q., Saltzman, A.L., Kim, Y.K., Misquitta, C., Shai, O., Maquat, L.E., Frey, B.J., and Blencowe, B.J. (2006). Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. Genes Dev *20*, 153-8.

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*, 1737-54.

Peltz, S.W., He, F., Welch, E., and Jacobson, A. (1994). Nonsense-mediated mRNA decay in yeast. Prog Nucleic Acid Res Mol Biol *47*, 271-98.

Perez-Canadillas, J.M. (2006). Grabbing the message: structural basis of mRNA 3'UTR recognition by Hrp1. Embo J *25*, 3167-78.

Pestova, T.V., Borukhov, S.I., and Hellen, C.U. (1998). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. Nature *394*, 854-9.

Pinto, I., Na, J.G., Sherman, F., and Hampsey, M. (1992). cis- and trans-acting suppressors of a translation initiation defect at the cyc1 locus of Saccharomyces cerevisiae. Genetics *132*, 97-112.

Pisarev, A.V., Hellen, C.U., and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. Cell *131*, 286-99.

Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501-12.

Pulak, R., and Anderson, P. (1993). mRNA surveillance by the Caenorhabditis elegans smg genes. Genes Dev *7*, 1885-97.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. RNA *11*, 1530-44.

Reichenbach, P., Hoss, M., Azzalin, C.M., Nabholz, M., Bucher, P., and Lingner, J. (2003). A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed. Curr Biol *13*, 568-74.

Rittinger, K., Budman, J., Xu, J., Volinia, S., Cantley, L.C., Smerdon, S.J., Gamblin, S.J., and Yaffe, M.B. (1999). Structural analysis of 14-3-3 phosphopeptide complexes identifies a dual role for the nuclear export signal of 14-3-3 in ligand binding. Mol Cell *4*, 153-66.

Rodriguez-Gabriel, M.A., Burns, G., McDonald, W.H., Martin, V., Yates, J.R., 3rd, Bahler, J., and Russell, P. (2003). RNA-binding protein Csx1 mediates global control of gene expression in response to oxidative stress. EMBO J *22*, 6256-66.

Rodriguez-Gabriel, M.A., Watt, S., Bahler, J., and Russell, P. (2006). Upf1, an RNA helicase required for nonsense-mediated mRNA decay, modulates the transcriptional response to oxidative stress in fission yeast. Mol Cell Biol *26*, 6347-56.

Romao, L., Inacio, A., Santos, S., Avila, M., Faustino, P., Pacheco, P., and Lavinha, J. (2000). Nonsense mutations in the human beta-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation. Blood *96*, 2895-901.

Ruiz-Echevarria, M.J., Gonzalez, C.I., and Peltz, S.W. (1998a). Identifying the right stop: determining how the surveillance complex recognizes and degrades an aberrant mRNA. Embo J *17*, 575-89.

Ruiz-Echevarria, M.J., and Peltz, S.W. (1996). Utilizing the GCN4 leader region to investigate the role of the sequence determinants in nonsense-mediated mRNA decay. Embo J *15*, 2810-9.

Ruiz-Echevarria, M.J., Yasenchak, J.M., Han, X., Dinman, J.D., and Peltz, S.W. (1998b). The upf3 protein is a component of the surveillance complex that monitors both translation and mRNA turnover and affects viral propagation. Proc Natl Acad Sci U S A *95*, 8721-6.

Sachs, A.B., and Varani, G. (2000). Eukaryotic translation initiation: there are (at least) two sides to every story. Nat Struct Biol *7*, 356-61.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Second Edition) (New York, Cold Spring Harbor Laboratory Press).

Sanford, J.R., Gray, N.K., Beckmann, K., and Caceres, J.F. (2004). A novel role for shuttling SR proteins in mRNA translation. Genes Dev *18*, 755-68.

Sayani, S., Janis, M., Lee, C.Y., Toesca, I., and Chanfreau, G.F. (2008). Widespread impact of nonsense-mediated mRNA decay on the yeast intronome. Mol Cell *31*, 360-70.

Schell, T., Kocher, T., Wilm, M., Seraphin, B., Kulozik, A.E., and Hentze, M.W. (2003). Complexes between the nonsense-mediated mRNA decay pathway factor human upf1 (up-frameshift protein 1) and essential nonsense-mediated mRNA decay factors in HeLa cells. Biochem J *373*, 775-83.

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

Schwartz, A.M., Komarova, T.V., Skulachev, M.V., Zvereva, A.S., Dorokhov, Y.L., and Atabekov, J.G. (2006). Stability of Plant mRNAs Depends on the Length of the 3
-Untranslated Region. Biochemistry (Mosc) *71*, 1377-84.

Serin, G., Gersappe, A., Black, J.D., Aronoff, R., and Maquat, L.E. (2001). Identification and characterization of human orthologues to Saccharomyces cerevisiae Upf2 protein and Upf3 protein (Caenorhabditis elegans SMG-4). Mol Cell Biol *21*, 209-23.

Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science *300*, 805-8.

Sheth, U., and Parker, R. (2006). Targeting of aberrant mRNAs to cytoplasmic processing bodies. Cell *125*, 1095-109.

Shi, H., and Xu, R.M. (2003). Crystal structure of the Drosophila Mago nashi-Y14 complex. Genes Dev *17*, 971-6.

Shibuya, T., Tange, T.O., Sonenberg, N., and Moore, M.J. (2004). eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. Nat Struct Mol Biol *11*, 346-51.

Simon, E., and Seraphin, B. (2007). A specific role for the C-terminal region of the Poly(A)-binding protein in mRNA decay. Nucleic Acids Res *35*, 6017-28.

Singh, G., Jakob, S., Kleedehn, M.G., and Lykke-Andersen, J. (2007). Communication with the exon-junction complex and activation of nonsense-mediated decay by human Upf proteins occur in the cytoplasm. Mol Cell *27*, 780-92.

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

Sun, X., and Maquat, L.E. (2002). Nonsense-mediated decay: assaying for effects on selenoprotein mRNAs. Methods Enzymol *347*, 49-57.

Szankasi, P., and Smith, G.R. (1996). Requirement of S. pombe exonuclease II, a homologue of S. cerevisiae Sep1, for normal mitotic growth and viability. Curr Genet *30*, 284-93.

Tarun, S.Z., Jr., and Sachs, A.B. (1996). Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. EMBO J *15*, 7168-77.

Tarun, S.Z., Jr., Wells, S.E., Deardorff, J.A., and Sachs, A.B. (1997). Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. Proc Natl Acad Sci U S A *94*, 9046-51.

Thakurta, A.G., Ho Yoon, J., and Dhar, R. (2002). Schizosaccharomyces pombe spPABP, a homologue of Saccharomyces cerevisiae Pab1p, is a non-essential, shuttling protein that facilitates mRNA export. Yeast *19*, 803-10.

Thakurta, A.G., Selvanathan, S.P., Patterson, A.D., Gopal, G., and Dhar, R. (2007). The nuclear export signal of splicing factor Uap56p interacts with nuclear pore-associated protein Rae1p for mRNA export in Schizosaccharomyces pombe. J Biol Chem *282*, 17507-16.

Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeier, C., Hentze, M.W., and Kulozik, A.E. (1998). Binary specification of nonsense codons by splicing and cytoplasmic translation. Embo J *17*, 3484-94.

Tuck, M.T., Wiehl, P.E., and Pan, T. (1999). Inhibition of 6-methyladenine formation decreases the translation efficiency of dihydrofolate reductase transcripts. Int J Biochem Cell Biol *31*, 837-51.

Urlaub, G., Mitchell, P.J., Ciudad, C.J., and Chasin, L.A. (1989). Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. Mol Cell Biol *9*, 2868-80.

Valentine, C.R. (1998). The association of nonsense codons with exon skipping. Mutat Res *411*, 87-117.

Wagner, E., and Lykke-Andersen, J. (2002). mRNA surveillance: the perfect persist. J Cell Sci *115*, 3033-8.

Wall, D., and Kaiser, D. (1999). Type IV pili and cell motility. Mol Microbiol *32*, 1-10.

Wang, J., Gudikote, J.P., Olivas, O.R., and Wilkinson, M.F. (2002a). Boundary-independent polar nonsense-mediated decay. EMBO Rep *3*, 274-9.

Wang, J., Hamilton, J.I., Carter, M.S., Li, S., and Wilkinson, M.F. (2002b). Alternatively spliced TCR mRNA induced by disruption of reading frame. Science *297*, 108-10.

Wang, W., Cajigas, I.J., Peltz, S.W., Wilkinson, M.F., and Gonzalez, C.I. (2006). Role for Upf2p phosphorylation in Saccharomyces cerevisiae nonsense-mediated mRNA decay. Mol Cell Biol *26*, 3390-400.

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*, 880-90.

Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Monch, K., Thoren, L.A., Nielsen, F.C., Jacobsen, S.E., Nerlov, C., and Porse, B.T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev *22*, 1381-96.

Welch, E.M., and Jacobson, A. (1999). An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA. Embo J *18*, 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*, 135-40.

Wen, J., and Brogna, S. (2008). Nonsense-mediated mRNA decay. Biochem Soc Trans *36*, 514-6.

Weng, Y., Czaplinski, K., and Peltz, S.W. (1996a). Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. Mol Cell Biol *16*, 5477-90.

Weng, Y., Czaplinski, K., and Peltz, S.W. (1996b). Identification and characterization of mutations in the UPF1 gene that affect nonsense suppression and the formation of the Upf protein complex but not mRNA turnover. Mol Cell Biol *16*, 5491-506.

Weng, Y., Czaplinski, K., and Peltz, S.W. (1998). ATP is a cofactor of the Upf1 protein that modulates its translation termination and RNA binding activities. RNA *4*, 205-14.

Wiegand, H.L., Lu, S., and Cullen, B.R. (2003). Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. Proc Natl Acad Sci U S A *100*, 11327-32.

Wittkopp, N., Huntzinger, E., Weiler, C., Sauliere, J., Schmidt, S., Sonawane, M., and Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol Cell Biol *29*, 3517-28.

Wittmann, J., Hol, E.M., and Jack, H.M. (2006). hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol Cell Biol *26*, 1272-87.

Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S.*, et al.* (2002). The genome sequence of Schizosaccharomyces pombe. Nature *415*, 871-80.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T.*, et al.* (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 Dev *23*, 1091-105.

Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., and Ohno, S. (2001). Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. Genes Dev *15*, 2215-28.

Yoine, M., Nishii, T., and Nakamura, K. (2006). Arabidopsis UPF1 RNA helicase for nonsense-mediated mRNA decay is involved in seed size control and is essential for growth. Plant Cell Physiol *47*, 572-80.

You, K.T., Li, L.S., Kim, N.G., Kang, H.J., Koh, K.H., Chwae, Y.J., Kim, K.M., Kim, Y.K., Park, S.M., Jang, S.K.*, et al.* (2007). Selective translational repression of truncated proteins from frameshift mutation-derived mRNAs in tumors. PLoS Biol *5*, e109.

Yun, D.F., and Sherman, F. (1995). Initiation of translation can occur only in a restricted region of the CYC1 mRNA of Saccharomyces cerevisiae. Mol Cell Biol *15*, 1021-33.

Yusupova, G.Z., Yusupov, M.M., Cate, J.H., and Noller, H.F. (2001). The path of messenger RNA through the ribosome. Cell *106*, 233-41.

Zhang, J., and Maquat, L.E. (1996). Evidence that the decay of nucleus-associated nonsense mRNA for human triosephosphate isomerase involves nonsense codon recognition after splicing. RNA *2*, 235-43.

Zhang, J., Sun, X., Qian, Y., LaDuca, J.P., and Maquat, L.E. (1998a). 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*, 5272-83.

Zhang, J., Sun, X., Qian, Y., and Maquat, L.E. (1998b). Intron function in the nonsense-mediated decay of beta-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. RNA *4*, 801-15.

Zhang, S., Ruiz-Echevarria, M.J., Quan, Y., and Peltz, S.W. (1995). Identification and characterization of a sequence motif involved in nonsense-mediated mRNA decay. Mol Cell Biol *15*, 2231-44.

Zhao, J., Hyman, L., and Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev *63*, 405-45.

Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature *407*, 401-5.

Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. (1995). Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. EMBO J *14*, 4065-72.

Zurlinden, A., and Schweingruber, M.E. (1997). Identification of a DNA element in the fission yeast Schizosaccharomyces pombe nmt1 (thi3) promoter involved in thiamine-regulated gene expression. J Bacteriol *179*, 5956-8.

APPENDIXES

Appendix I. The recipe for *E. coli* and *S. pombe* growth media

LB broth

Dissolve 10 g Bacto-tryptone, 5 g yeast extract, and 10 g NaCl in 800 mL dH$_2$O. Adjust pH to 7.5 with NaOH, bring the volume to 1L with dH$_2$O, transfer to a bottle and sterilize it by autoclaving at 121 °C for 15 minutes. Add antibiotics to cold media at the required concentration just before using it.

Agar-LB plates

Dissolve 10 g Bacto-tryptone, 5 g yeast extract, 10 g NaCl and 10 g agar are dissolved in 800 mL of dH$_2$O. Adjust pH to 7.5 with NaOH, bring the volume to 1 L with dH$_2$O, transfer to a bottle and sterilize it by autoclaving at 121 °C for 15 minutes. Cool the media to about 60 °C, add the required antibiotic (for example, 1 mL of 100 mg/mL ampincillin) and pour 25-30 ml/plate (9 cm Petri dishes). Plates can be stored at 4 °C.

SOC

Dissolve 20 g Bacto-tryptone, 5 g yeast extract, 0.6 g NaCl and 0.18 g KCl in 970 mL of dH$_2$O and sterilize by autoclaving at 121 °C for 15 minutes (before autoclaving, the media can be aliquoted into 20 mL glass bottles for convenience). Before using it, add the following (for 20 mL aliquot): 200 μL of 1 M MgCl$_2$, 200 μL of 1 M MgSO$_4$ and 240 μL of 30% glucose.

YES

Dissolve 5 g yeast extract, 225 mg each of the amino acid and nucleotides supplements (adenine, histidine, leucine, uracil and lysine hydrochloride), in 900 mL of dH$_2$O, transfer to a bottle and sterilize by autoclaving at 121 °C for 15 minutes. After autoclaving add 30 mL of sterile 30% glucose stock). The desired antibiotic is added to media just before use (for example the 100 μg/mL G418 for the screening of knockout strain by *KanMX6* cassette). Agar YES media was made by adding extra 2% agar, adding 3% glucose and desired antibiotics after autoclaving.

EMM without leucine

Dissolve 3 g potassium hydrogen phthalate, 2.2 g Na$_2$HPO$_4$, 5 g NH$_4$Cl and 225 mg of each of the amino acid and nucleotide supplements (adenine, histidine, uracil and lysine hydrochloride), in 875 mL of dH$_2$O and then sterilized by autoclaving at 121 °C for 15 minutes. After autoclaving, to cold media add 2% glucose, 1× salt stock, vitamin stock and mineral stock (see recipes below) (EMM agar media is made by adding 2% agar. After autoclaving, add 2% glucose.)

Stocks in media

1. 10× Glucose stock

300 g glucose dissolved in 1L dH$_2$O, autoclaved and stored at RT.

2. 50× Salt stock

52.5 g MgCl$_2$·6H$_2$O, 0.735 g CaCl$_2$·2H$_2$O, 50 g KCl and 2 g Na$_2$SO$_4$ in 1 L dH$_2$O. Autoclaved and stored at RT.

3. 1000× Vitamin stock

0.1 g pantothenic acid, 1 g nicotinic acid, 1 g inositol and 1 mg biotin in 100 mL dH$_2$O. Autoclaved and stored in dark environment.

4. 10000× Mineral stock

0.5 g boric acid, 0.4 g MnSO$_4$, 0.4 g ZnSO$_4$·7H$_2$O, 0.2 g FeCl$_2$·6H$_2$O, 40 mg molybdic acid, 0.1 g KI, 40 mg CuSO$_4$·5H$_2$O and 1 g citric acid in 100 mL dH2O. Autoclaved and stored in dark environment.

5. 1000× antibiotics stocks

100 mg/mL ampicillin and 50 mg/mL kanamycin for LB and 200 mg/L G418 and 100 mg/L hygromycin for YES. Filtration steriled and stored in -20 °C.

Appendix II. Primers, plasmids and yeast strains used in this study.

Table 1. Primers

Primer names	Primer sequences	Comments
gfp-ks	GGAGTCGAC**GGATCC**AATGGC TAGCAAAGGAGAAGAAC 3'	*BamH I*, cloning *GFP* ORF into pREP41
gfp-rev	AAGGAA**GGATCC**TTAGCAGCC AGATCCTTTGTATAG	*BamH I*, cloning *GFP* ORF into pREP41
SSP41	GGAATCCTGGCATATCATCA	Sequencing primer in *nmt41* promoter
RSP41	CGTAATATGCAGCTTGAATGG	Sequencing primer in *nmt1* terminator
N6-sens	GGAGTCGAC**GGATCC**AATGGC TAGCAAAGGA*TAA*GAAC	*BamH I*, PTC6 introduced
Sens-N27	GATGTTAATGGGCAC*TAA*TTTT CTGTCAGTG	PTC27 introduced
Rev-N27	CACTGACAGAAAA*TTA*GTGCC CATTAACATC	PTC27 introduced
Sens-N140	AACATTCTTGGGCAC*TAA*TTGG AATACAACT	PTC140 introduced
Rev-N140	AGTTGTATTCCAA*TTA*GTGCCC AAGAATGTT	PTC140 introduced
N6-ch-sens	CCAATGGCTAGCAAAGGAT	Confirmation of PTC6 mutagenesis
N27-ch-sens	TGGTGATGTTAATGGGCACT	Confirmation of PTC27 mutagenesis
N140-ch-sens	GGAAACATTCTTGGGCACT	Confirmation of PTC140 mutagenesis
Gavr-2-sens	TCGAC**GGATCC**AATG**CCTAGG** GCTAGCAAAGGAGAAGAAC	*Avr II* introduced at residue 2
PTC6avr-2-sens	TCGAC**GGATCC**AATG**CCTAGG** GCTAGCAAAGGA*TAA*GAAC	*Avr II* introduced at residue 2; PTC6 introduced
InHAeG-sens	CTTGAG**CCTAGG**TACCCATAC GATGTTCCTGA	*Avr II*, 147 bp, 291 bp and 402 bp DNA cloned
InHAeGL-rev	TTGGTT**CCTAGG**GGTCAGGGT GGTCACGAG	*Avr II*, 291 bp DNA cloned
InHAeGS-rev	TTTGTA**CCTAGG**CTCGACCAG GATGGGCAC	*Avr II*, 147 bp DNA cloned
HAeG402-rev	AACAAT**CCTAGG**GAAGAAGAT GGTGCGCTCC	*Avr II*, 402 bp DNA cloned
luc102 rev	ATATT**CCCGGG**CGCAACTGCA ACTCC	*XmaI*, 102 bp and 418 bp DNA insertion
luc102 sens	AACGA**CCCGGG**CTGAATACAA ATCACA	*XmaI*, 102bp DNA insertion

luc418 rev	ACTAT**CCCGGG**CCACACCCTT AGGTAACC	*XmaI*, 418 bp DNA insertion
Gavr2-243-rev	GGAA**GGATCC**TTA**CCTAGG**GC AGCCAGATCCTTTGTATAG	*Avr II* introduced immediately before stop codon
Gluc402-sens	ACTGC**CCTAGG**AATCACAGAA TCGTCGTATG	*Avr II,* 402 bp DNA introduced in *GFP*
Gluc402-rev	CCACT**CCTAGG**CACACCCTTA GGTAACCCA	*Avr II,* 402 bp DNA introduced in *GFP*
ade6-3Us	AAA**CCCGGG**CGACCATAGACA TAACTG	*Xma I*, replacement of the *nmt1* 3' UTR
ade6-3Ur	AAA**GAGCTC**GATGAACGCACT AATTTG	*Sac I*, replacement of the *nmt1* 3' UTR
Sgfp-sens	AGCGTTCAACTAGCAGACC	Sequencing primer in *GFP* coding region
Sgfp-rev	TGTAGTTCCCGTCATCTTTG	Sequencing primer in *GFP* coding region
GFP-Sh rev	TTCT**GGATCC**TTATGTCTTGTA GTTCCCGTCATC	Primer for short *GFP* with gfp ks
INubc4 sens	GTATGTAACTATTTAGTCTTGT GT	Intron introduced into *GFP* coding region
INubc4 rev	CTACGAATATGCTGTTAGTTAT C	Intron introduced into *GFP* coding region
5SP_SENS	GAACTACAAGACACATAT*CTA A*CTATTTAGTCTTGT	Mutagenesis of the 5' splice site of the intron
5SP_REV	AGACTAAATAG*TTAG*ATATGTG TCTTGTAGTTCCC	Mutagenesis of the 5' splice site of the intron
5sp-ch-sens	CGGGAACTACAAGACACATAT C	Confirmation of the mutagenesis of 5' splice site
IN3UTR-sens	AAAA**CCCGGG**GTATGTAACTA TTTAGTCTTGTGT	*Xma I,* Intron introducing into 3' UTR
IN3UTR-rev	AAAA**CCCGGG**CTACGAATATG CTGTTAGTTATC	*Xma I,* Intron introducing into 3' UTR
Gavr-120-sens	GTCAAGTTTGAAGGTGAT**CCT AGG**ACCCTTGTTAATAGA	*Avr II* introduced at residue 120
Gavr-120-rev	GATTCTATTAACAAGGGT**CCT AGG**ATCACCTTCAAACTT	*Avr II* introduced at residue 120
ypt3-sens	CCACC**GGATCC**ATGTGTCAAG AGGACGAATAC	*BamH I, ypt3* cloned into pREP41
ypt3N6-sens	CCACC**GGATCC**ATGTGTCAAG AGGAC*TAA*TACGATTATTTG	*BamH I,* PTC6 introduced
ypt3-rev	GTCCG**GGATCC**TAACAACATT GGGAAGAAGAC	*BamH I, ypt3* reverse transcribed and cloned into pREP41
ypt3N130-rev	AGATACAGCTCT*TTA*ATGTAGT AAATCCG	PTC130 introduced
ypt3N130-sens	GATTACTACAT*TAA*AGAGCTG TATCTACT	PTC130 introduced
KOS2s	GCCCAGAATACCCTCCTTGA	Confirmation of the *KanMX6*

		cassette in *S. pombe* genome
KOS2r	GCTTGATGGTCGGAAGAGGC	Confirmation of the *KanMX6* cassette in *S. pombe* genome
upf1ch-sens1	TAATCTATAGTACGACCAGCTGT	Confirmation of *upf1* deletion
upf1ch-rev1	GATGCATTTACTAGACAAAGAGG	Confirmation of *upf1* deletion
upf1ch-sens	CCGCTAAGCACCACATAA	Combination with kan inp1s to confirm the *upf1Δ*
upf2ch-sens	TCATACTGGAAAGGCTGCTA	Combination with kan inp1s to confirm the *upf2Δ*
kan-inp1s	CGTATGTGAATGCTGGTCG	Specific for *KanMX6* and *HphMX6* cassette
kan-inP1r	TATTCTGGGCCTCCATGTC	Specific for *KanMX6* and *HphMX6* cassette
rpl32-sens	GGCTGCTGTCAATATCAT	Amplifying *rpL32-2* gene
rpl32-rev	GTGACCTTTACACCGAGA	Amplifying *rpL32-2* gene
28rRNA-SENS	AGGCGTGGAGGTTTGTGAC	Amplifying the 28S rDNA
28rRNA-REV	TTAGATGACGAGGCATTTGG	Amplifying the 28S rDNA
exo2-sens	CCTTCCTTCGCCTCAATG	Confirmation of *XRN1* deletion
exo2-rev	ATGGTCCTGCTGCTCCTT	Confirmation of *XRN1* deletion
ura4ch-rev	CTGGGACAGCAATATCGTAC	Specific primer for *URA4* cassette
pabORF-sen	ACTCGCCGTTCTTTGGGTT	Confirmation of *pab1* deletion
pabORF-rev	GCAAGCATACGCTGGTTCA	Confirmation of *pab1* deletion
urach-rev1	ATGCTCCTACAACATTACCAC	Specific primer for *URA4* cassette
urach-sens1	TGGTAGGACAATACGGTAAGA	Specific primer for *URA4* cassette
RV_AD_dT	TAGAATTCAGCATTCGCTTCTTTTTTTTTTTTTTTTT**V**	Primer for reverse transcription; V, the mixture of C, G and A equally
RV_ADAPTOR	TAGAATTCAGCTTCGCTTC	Universal reverse primer to amplify the reverse transcription products
5SP_SENS	GAACTACAAGACACATAT**CTAA**CTATTTAGTCTTGT	Mutagenesis of the 5' splice site of the intron
5SP_REV	AGACTAAATAG**TTAG**ATATGTGTCTTGTAGTTCCC	Mutagenesis of the 5' splice site of the intron
5sp-ch-sens	CGGGAACTACAAGACACATATC	Confirmation of the mutagenesis of 5' splice site
ADHterm F	CTCTTATTGACCACACCTCTACC	Primer for screening the integration colonies

Leu1 Rev	GGTCATAAAGTTGAACGGATGTCG	Primer for screening the integration colonies
ADH_termF2	CGCTAGGGATAACAGGGTAA	Primer for screening the integration colonies
Leu1_Rev2	TGGAAGTTTCAGCTAACCAAG	Primer for screening the integration colonies
RNPS1-Sens	TCCTCTCAAAAGTTTAGTAATAAAATATAAACATTCTCAATTATGCTTTTTGTAATACTcggatccccgggttaattaa	Primer for RNPS1 deletion
RNPS1-Rev	TTCAAACTGGAAATTTGTTTTCACTTTTCTTTTTTATTCAGTAGGCCACAATTTGCGGGAgaattcgagctcgtttaaac	Primer for RNPS1 deletion
RNPS1ch-S	GTAATATGCGGCAAGGAGC	Primer for screening of RNPS1 deletion
RNPS1-ch-S2	CGGTATCCAGTTCGTCAAG	Primer for screening of RNPS1 deletion
eIF4A3-Sens	TTTCTCTCCTTCATTTTCCTCAAGTTGGGCTGACATAACATAACTACTACAACTTCCATTcggatccccgggttaattaa	Primer for eIF4AIII deletion
eIF4A3-Rev	TTTCCCCCAAATCTGTACAATAATCCAAAAAAAACCAGCTTTTTTTTTCCAATGGAAAATgaattcgagctcgtttaaac	Primer for eIF4AIII deletion
eIFch-sens	GTCTAATTCTAAGAAACCGCTAG	Primer for screening of eIF4AIII deletion
eIF4A-ch-S2	TCTATGCGTATGGCTATGAGAC	Primer for screening of eIF4AIII deletion
KD209	GT*GCCCATTAACAT*CACCATCTAATT*CAACAAGAAT*TGGGACAACT*CCAGT	
KD210	GTACAT*AACCTTCGGGCAT*GGCACTCTT*GAAAAGTCAT*GCCGTTTCAT*AT	
KD211	GATTCCAT*TCTTTTGTT*TGTCTGCCAT*GATGTATACAT*TGTGTGAGTT*ATA	FISH probes for GFP mRNA (Dower and Rosbash, 2002)
KD212	CCCAGCAGCT*GTTACAAACT*CAAGAAGGACCAT*GTGGTCT*CTCTTTTCGT*T	

In bold restriction enzyme recognition sites (introduced for cloning purpose). In bold italic are PTCs. In lower case are gene specific homologous sequences for gene knock-out. T* means amino-modified dT for Cy3 fluorochrome conjugation.

Table 2. Plasmids

Plasmids name	Resistance in *E.coli*	Marker in yeast	Comments
pFA6a-KanMX6	amp	KanMX6	KanMX6 cassette for knocking out a gene (Bahler et al., 1998)
<u>pREP41</u>	amp	*LEU2*	Episomal expression vector in *S. pombe* by *nmt41* promoter (Basi et al., 1993; Maundrell, 1993)
<u>pREP-*GFP*</u>	amp	*LEU2*	Derived from pREP41; GFP ORF cloned in *Bam HI*
pPTC6	amp	*LEU2*	Derived from pREP41; GFP ORF cloned in *Bam HI*, PTC6 introduced
pPTC27	amp	*LEU2*	Derived from pREP41; GFP ORF cloned in *Bam HI*, PTC27 introduced
pPTC140	amp	*LEU2*	Derived from pREP41; GFP ORF cloned in *Bam HI*, PTC140 introduced
<u>p*GFP*ivs</u>	amp	*LEU2*	Derived from pREP-*GFP,* intron DNA inserted in *Pml I* site
pPTC6ivs	amp	*LEU2*	Derived from p*GFP*-PTC6, intron DNA inserted in *Pml I* site
pPTC27ivs	amp	*LEU2*	Derived from p*GFP*-PTC27, intron DNA inserted in *Pml I* site
pPTC140ivs	amp	*LEU2*	Derived from p*GFP*-PTC140, intron DNA inserted in *Pml I* site
p*GFP*-ade6	amp	*LEU2*	Derived from pREP-*GFP, nmt1* 3' UTR replaced by *ade6* 3' UTR
pPTC6-ade6	amp	*LEU2*	Derived from p*GFP*-PTC6, *nmt1* 3' UTR replaced by *ade6* 3' UTR
pPTC27-ade6	amp	*LEU2*	Derived from p*GFP*-PTC27, *nmt1* 3' UTR replaced by *ade6* 3' UTR
pPTC140-ade6	amp	*LEU2*	Derived from p*GFP*-PTC140, *nmt1* 3' UTR replaced by *ade6* 3' UTR
pGFPivs-ade6	amp	*LEU2*	Derived from p*GFP*ivs, *nmt1* 3' UTR replaced by *ade6* 3' UTR
pPTC6ivs-ade6	amp	*LEU2*	Derived from p*GFP*ivs-PTC6, *nmt1* 3' UTR replaced by *ade6* 3' UTR
pPTC27ivs-ade6	amp	*LEU2*	Derived from p*GFP*ivs-PTC27, *nmt1* 3' UTR replaced by *ade6* 3' UTR

pPTC140ivs-ade6	amp	*LEU2*	Derived from p*GFP*ivs-PTC140, *nmt1* 3' UTR replaced by *ade6* 3' UTR
p*GFP* 3'ivs	amp	*LEU2*	Derived from pREP-*GFP,* intron DNA inserted in *Xma I* site
pPTC6 3'ivs	amp	*LEU2*	Derived from p*GFP*-PTC6, intron DNA inserted in *Xma I* site
pPTC27 3'ivs	amp	*LEU2*	Derived from p*GFP*-PTC27, intron DNA inserted in *Xma I* site
pPTC140 3'ivs	amp	*LEU2*	Derived from p*GFP*-PTC140, intron DNA inserted in *Xma I* site
p*GFP*-102	amp	*LEU2*	Derived from pREP-*GFP,* 102 bp DNA inserted in *Xma I* site
pPTC6-102	amp	*LEU2*	Derived from p*GFP*-PTC6, 102 bp DNA inserted in *Xma I* site
pPTC27-102	amp	*LEU2*	Derived from p*GFP*-PTC27, 102 bp DNA inserted in *Xma I* site
pPTC140-102	amp	*LEU2*	Derived from p*GFP*-PTC140, 102 bp DNA inserted in *Xma I* site
p*GFP*-418	amp	*LEU2*	Derived from pREP-*GFP,* 418 bp DNA inserted in *Xma I* site
pPTC6-418	amp	*LEU2*	Derived from p*GFP*-PTC6, 418 bp DNA inserted in *Xma I* site
pPTC27-418	amp	*LEU2*	Derived from p*GFP*-PTC27, 418 bp DNA inserted in *Xma I* site
pPTC140-418	amp	*LEU2*	Derived from p*GFP*-PTC140, 418 bp DNA inserted in *Xma I* site
p*GFP*-avrII3	amp	*LEU2*	Derived from pREP-*GFP, Avr II* site introduced at codon 3
pPTC6-avrII3	amp	*LEU2*	Derived from p*GFP*-PTC6, *Avr II* site introduced at codon 3
p*GFP*-S	amp	*LEU2*	Derived from p*GFP*-avrII3, 147 bp DNA inserted in *Avr II* site
pPTC6-S	amp	*LEU2*	Derived from pPTC6-avrII3, 147 bp DNA inserted in *Avr II* site
p*GFP*-L	amp	*LEU2*	Derived from p*GFP*-avrII3, 291 bp DNA inserted in *Avr II* site
pPTC6-L	amp	*LEU2*	Derived from pPTC6-avrII3, 291 bp DNA inserted in *Avr II* site
p*GFP*ivs-avrII120	amp	*LEU2*	Derived from p*GFP*ivs, *Avr II* site introduced at codon 120
pPTC140ivs-	amp	*LEU2*	Derived from pPTC140ivs, *Avr II* site

avrII120			introduced at codon 120
p*GFP*ivs-147	amp	*LEU2*	Derived from p*GFP*ivs-avrII120, 147 bp DNA inserted in *Avr II* site
pPTC140ivs-147	amp	*LEU2*	Derived from pPTC140ivs-avrII120, 147 bp DNA inserted in *Avr II* site
p*GFP*ivs-291	amp	*LEU2*	Derived from p*GFP*ivs-avrII120, 291 bp DNA inserted in *Avr II* site
pPTC140ivs-291	amp	*LEU2*	Derived from pPTC140ivs-avrII120, 291 bp DNA inserted in *Avr II* site
p*GFP*-avrII243	amp	*LEU2*	Derived from pREP-*GFP*, *Avr II* site introduced at codon 243
pPTC6-avrII243	amp	*LEU2*	Derived from p*GFP*-PTC6, *Avr II* site introduced at codon 243
pPTC27-avrII243	amp	*LEU2*	Derived from p*GFP*-PTC27, *Avr II* site introduced at codon 243
pPTC140-avrII243	amp	*LEU2*	Derived from p*GFP*-PTC140, *Avr II* site introduced at codon 243
p*GFP*-402	amp	*LEU2*	Derived from p*GFP*-avrII243, 402 bp DNA inserted in *Avr II* site
pPTC6-402	amp	*LEU2*	Derived from pPTC6-avrII243, 402 bp DNA inserted in *Avr II* site
pPTC27-402	amp	*LEU2*	Derived from pPTC27-avrII243, 402 bp DNA inserted in *Avr II* site
pPTC140-402	amp	*LEU2*	Derived from pPTC140-avrII243, 402 bp DNA inserted in *Avr II* site
pYPT3	amp	*LEU2*	Derived from pREP41; endogenous *YPT3* cloned in *Bam HI*
pYPT3N6	amp	*LEU2*	Derived from pREP41; endogenous *YPT3* cloned in *Bam HI*; Nonsense mutation introduced at codon 6
pYPT3N143	amp	*LEU2*	Derived from pREP41; endogenous *YPT3* cloned in *Bam HI*; Nonsense mutation introduced at codon 143
pcYPT3	amp	*LEU2*	Derived from pREP41; *YPT3* cDNA cloned in *Bam HI*
pcYPT3N6	amp	*LEU2*	Derived from pREP41; *YPT3* cDNA cloned in *Bam HI*; Nonsense mutation introduced at codon 6
pcYPT3N143	amp	*LEU2*	Derived from pREP41; *YPT3* cDNA cloned in *Bam HI*; Nonsense mutation introduced at codon 143
pDUAL-GFF41	amp	*LEU1,*	Integration plasmid, (Matsuyama et

		URA3	al., 2004)
pDUAL-nmt41	amp	*LEU1, URA3*	Derived from pDUAL-GFF41, *Bam HI* digestion to remove the DNA and self ligated
pDUAL-*GFP*	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP ORF cloned in *Bam HI*
pDUAL-PTC6	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP with nonsense mutation (codon 6) cloned in *Bam HI*
pDUAL-PTC27	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP with nonsense mutation (codon 27) cloned in *Bam HI*
pDUAL-N140	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP with nonsense mutation (codon 140) cloned in *Bam HI*
pDUAL-*GFP*ivs	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP ORF with an intron cloned in *Bam HI*
pDUAL-PTC6ivs	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP with a nonsense mutation (codon 6) and an intron cloned in *Bam HI*
pDUAL-PTC27ivs	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP with a nonsense mutation (codon 27) and an intron cloned in *Bam HI*
pDUAL-PTC140ivs	amp	*LEU1, URA3*	Derived from pDUAL-nmt41, GFP with a nonsense mutation (codon 140) and an intron cloned in *Bam HI*
pGFPivs-mut	amp	*LEU2*	Derived from p*GFP*ivs, 5' splice site mutated in GFP ORF
pPTC6ivs-mut	amp	*LEU2*	Derived from p*GFP*ivs-PTC6, 5' splice site mutated in GFP ORF
pPTC27ivs-mut	amp	*LEU2*	Derived from p*GFP*ivs-PTC27, 5' splice site mutated in GFP ORF
pPTC140ivs-mut	amp	*LEU2*	Derived from p*GFP*ivs-PTC140, 5' splice site mutated in GFP ORF
p*GFP*-sh	amp	*LEU2*	Derived from pREP41, short GFP ORF cloned in *Bam HI*
pPTC6-sh	amp	*LEU2*	Derived from pREP41, short GFP ORF with a nonsense mutation (codon 6) cloned in *Bam HI*

Underlined plasmids were shown their maps in Appendix III.

Table 3. *S. pombe* strains

Stock	Strain name	Genotype	Source
SPJK029	*upf2Δ*	*h⁻ leu1-32 ura4-D18 upf2::ura4+*	(Mendell et al., 2000)
SPJK033	MR3571	*h⁻ leu1-32 ura4-D18 ade6-M26 upf1::kanMX6*	(Rodriguez-Gabriel et al., 2006)
SPJK039	GP937	*h⁻ his3D leu1-23 ura4-D18 ade6-M216*	(Szankasi and Smith, 1996)
SPJK043	GP1255	*h⁺ leu1-32 ura4-D18 ade6-M216 exo2::ura4*	(Szankasi and Smith, 1996)
SPJK045	*sppab1Δ*	*h⁻ leu1-32 ura4-D18 ade6-M216 sppab1::ura4*	(Thakurta et al., 2002)
SPJK049	*rnps1Δ*	*h⁻ leu1-32 ura4-D18 ade6-M216 rnps1::kanMX6*	This study
SPJK051	*spbc3b9.08cΔ*	*h⁻ leu1-32 ura4-D18 ade6-M210 mago::kanMX6*	(Decottignies et al., 2003)
SPJK039-1	Int-GFP	*h⁻ leu1-32:: [GFP leu1⁺] ura4-D18 ade6-M216* (p*GFP*-wt integrated)	This study
SPJK039-2	Int-PTC6	*h⁻ leu1-32:: [GFP-PTC6 leu1⁺] ura4-D18 ade6-M216* (p*GFP*-PTC6 integrated)	This study
SPJK039-3	Int-PTC27	*h⁻ leu1-32:: [GFP-PTC27 leu1⁺] ura4-D18 ade6-M216* (p*GFP*-PTC27 integrated)	This study
SPJK039-4	Int-PTC140	*h⁻ leu1-32:: [GFP-PTC140 leu1⁺] ura4-D18 ade6-M216* (p*GFP*-PTC27 integrated)	This study
SPJK039-5	Int-GFP*ivs*	*h⁻ leu1-32:: [GFPivs leu1⁺] ura4-D18 ade6-M216* (p*GFPivs*-wt integrated)	This study
SPJK039-6	Int-PTC6*ivs*	*h⁻ leu1-32:: [GFPivs-PTC6 leu1⁺] ura4-D18 ade6-M216* (p*GFPivs*-PTC6 integrated)	This study
SPJK039-7	Int-PTC27*ivs*	*h⁻ leu1-32:: [GFPivs-PTC27 leu1⁺] ura4-D18 ade6-M216* (p*GFPivs*-PTC27 integrated)	This study
SPJK039-8	Int-PTC140*ivs*	*h⁻ leu1-32:: [GFPivs-PTC140 leu1⁺] ura4-D18 ade6-M216*	This study

|			
		(p*GFP_{ivs}*-PTC140 integrated)	
SPJK029-1	upf2-GFP	*h⁻ leu1-32::[GFP leu1⁺] ura4-D18 upf2::ura4+* (p*GFP*-wt integrated)	This study
SPJK029-2	upf2-PTC6	*h⁻ leu1-32::[GFP-PTC6 leu1⁺] ura4-D18 upf2::ura4+* (p*GFP*-PTC6 integrated)	This study
SPJK029-3	upf2-PTC27	*h⁻ leu1-32::[GFP-PTC27 leu1⁺] ura4-D18 upf2::ura4+* (p*GFP*-PTC27 integrated)	This study
SPJK029-4	upf2-GFP_{ivs}	*h⁻ leu1-32::[GFP_{ivs} leu1⁺] ura4-D18 upf2::ura4+* (p*GFP_{ivs}*-wt integrated)	This study
SPJK029-5	upf2-PTC6_{ivs}	*h⁻ leu1-32::[GFP_{ivs}-PTC6 leu1⁺] ura4-D18 upf2::ura4+* (p*GFP_{ivs}*-PTC6 integrated)	This study
SPJK029-6	upf2-PTC27_{ivs}	*h⁻ leu1-32::[GFP_{ivs}-PTC27 leu1⁺] ura4-D18 upf2::ura4+* (p*GFP_{ivs}*-PTC27 integrated)	This study

Appendix III. Plasmid maps.

Below are maps of 11 plasmids (Map 1 to Map 11). Map 1 and Map 2 show the diagrams of the two starting vectors used to generate all the NMD reporters, pREP41 and pDUAL-*nmt41* respectively. pREP-*GFP* (Map 3) represents the initial intronless reporters and GFP reporters with short PTC-3' end distance; p*GFP*-L (Map 4), GFP reporters with long AUG-PTC distance; p*GFP*-418 (Map 5), GFP reporters with long 3' UTR; p*GFP*-402 (Map 6), GFP reporters with long PTC-3' end distance; p*GFP*$_{ivs}$ (Map 7), intron-containing reporters and GFP$_{ivs}$ reporters with 5' splice-site mutation; p*GFP*$_{3'ivs}$ (Map 8), GFP reporters with 3' UTR intron; p*GFP*$_{ivs}$-291 (Map 9), GFP$_{ivs}$ reporters with long PTC-intron distance; pDUAL-*GFP* (Map 10), integrated GFP reporters (intronless and intron-containing); pcYPT3 (Map 11), the endogenous NMD reporters.
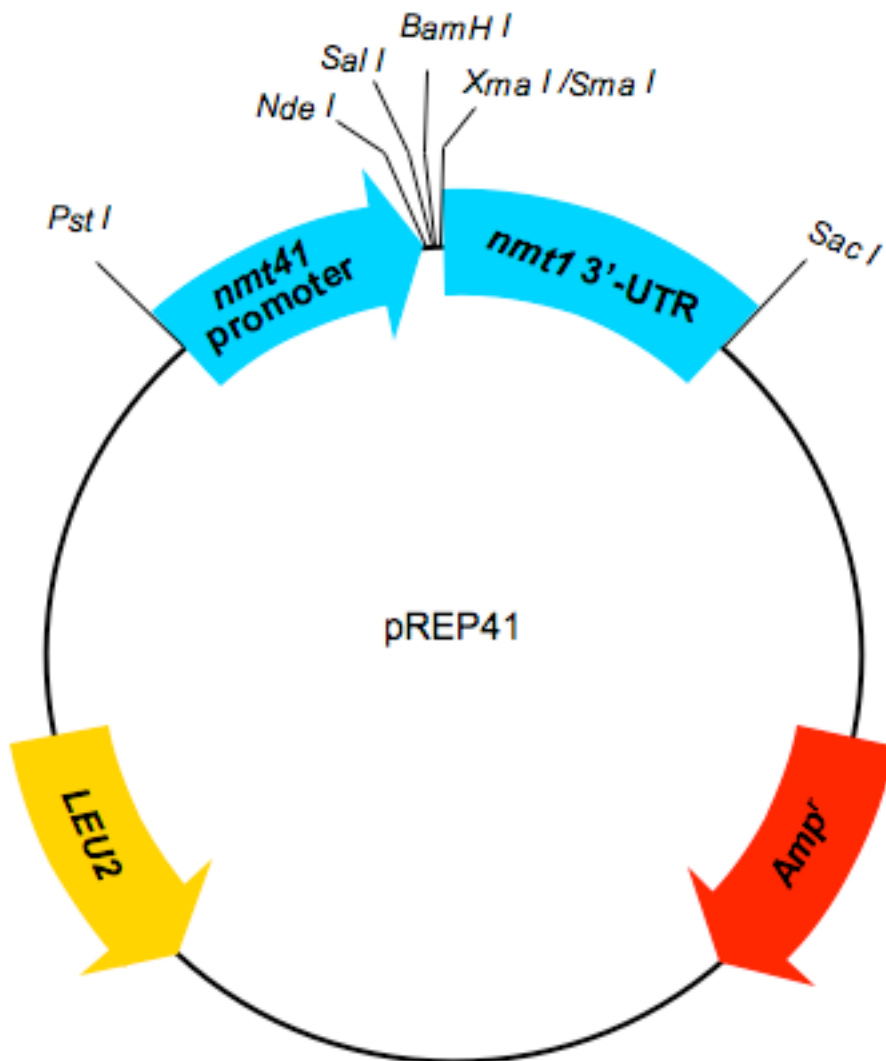
Map 1: pREP41.

The expression is driven by the *nmt41* promoter flanking the cloning sites. The *nmt41* artificial-promoter was attenuated from the endogenous *nmt1* promoter by mutagenesis in TATA region; nmt41 is middle-strength promoter suppressed by thiamine (Basi et al., 1993; Maundrell, 1993).
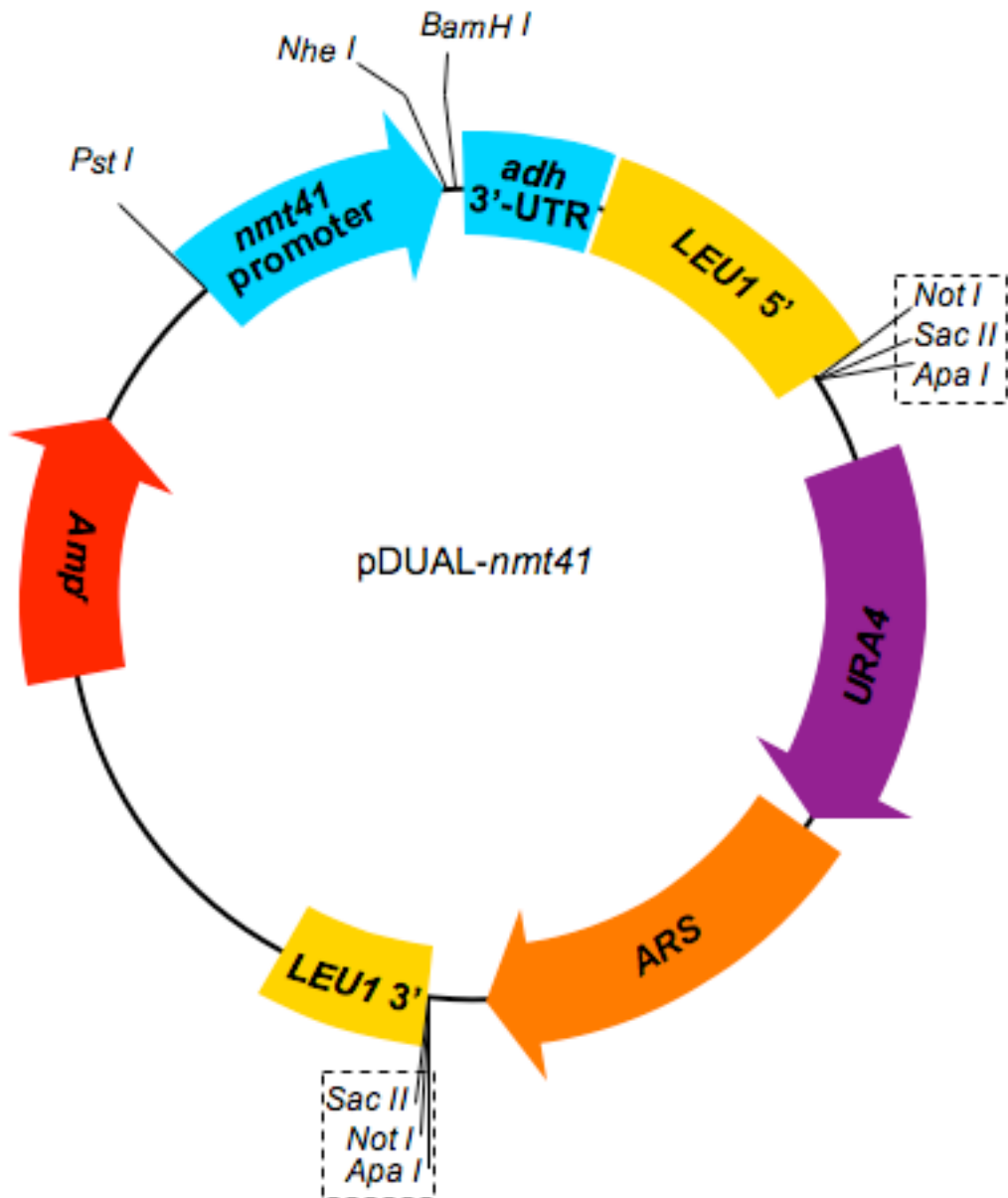
Map 2: pDUAL-*nmt41*

The expression is driven by *nmt41* promoter flanking the cloning site. *URA4* can be the complementary of uracil auxotroph in episomal expression. *LEU1* 5' half and 3' half were used to integrated DNA fragment in chromosome by homologues recombination. The enzyme sites in dashed rectangle are used to linearize the plasmid to facilitate integration (Matsuyama et al., 2004).
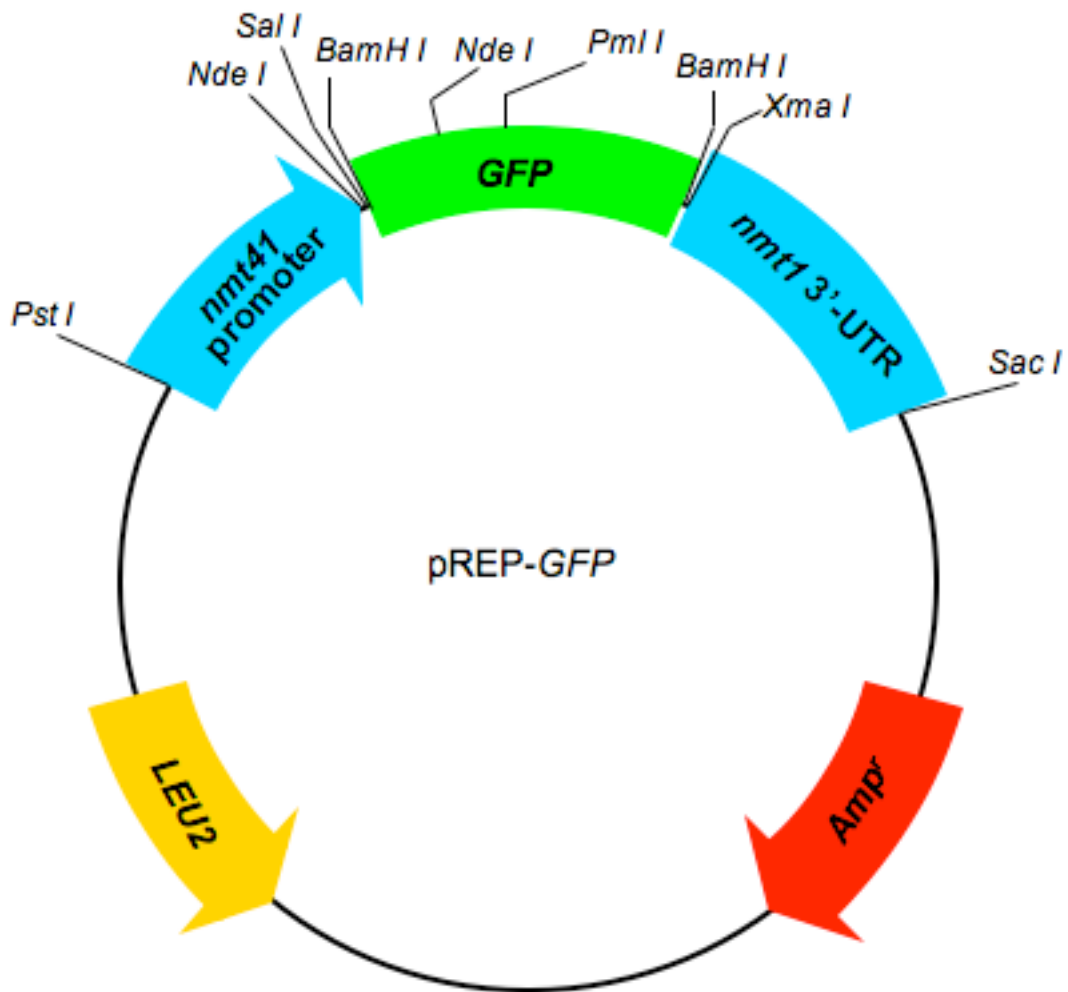
Map 3: pREP-*GFP* (parental plasmid for the intronless NMD reporters).

The GFP coding region was PCR amplified with primers gfp-ks and gfp-rev and cloned into the *Bam HI* site of pREP41. PTC6 was introduced directly by the primers sens-N6; PTC27 and PTC140 were introduced by overlap PCR with the primers sens-N27/rev-N27 and sens-N140/rev-N140, combined with gfp-ks and gfp-rev. The sequence of the primers is shown in Table 1, Appendix II.
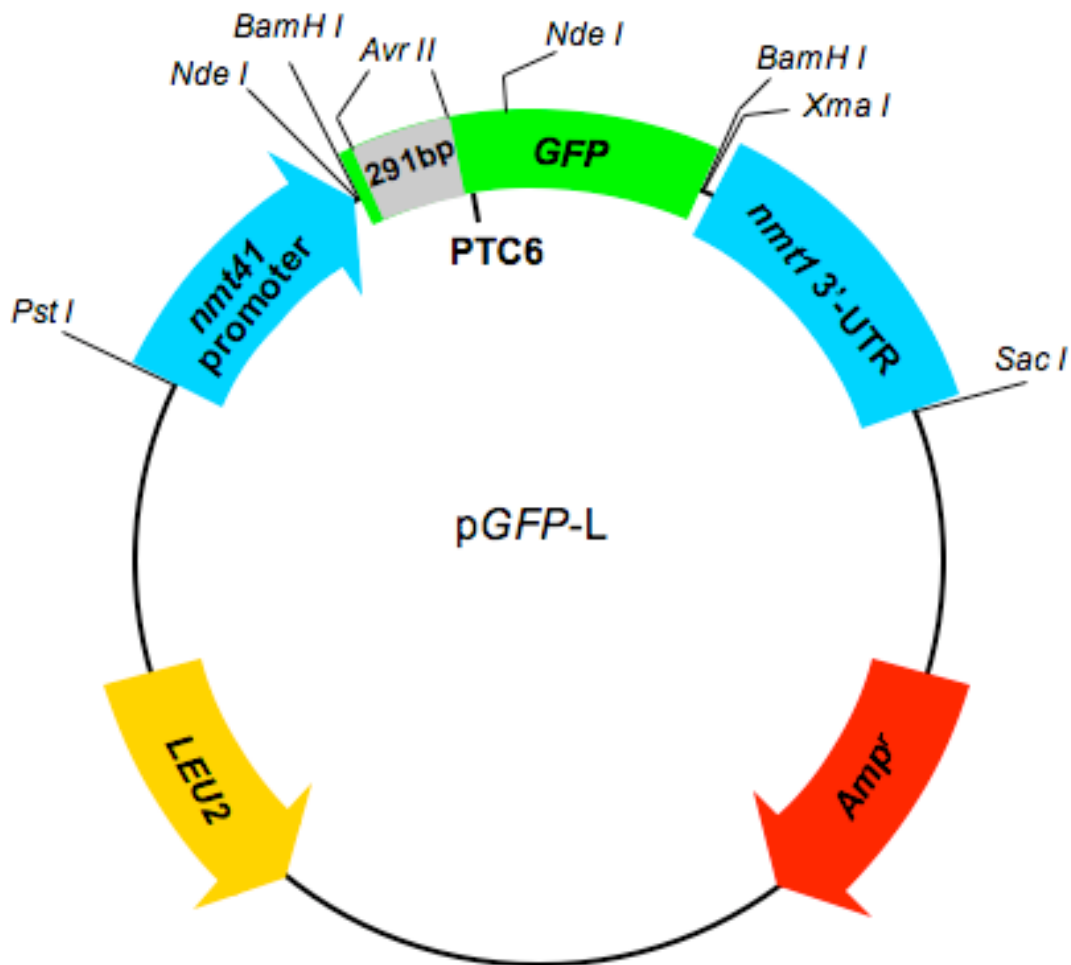
Map 4: p*GFP*-L (AUG-PTC lengthening constructs)

The 147 bp and 291 bp DNA fragments were PCR amplified from the EGFP coding region with primers InHAeG-sens, InHAeGS-rev and InHAeGL-rev, and cloned into a *Avr II* site. The *Avr II* site was introduced with primers Gavr-2-sens and PTC6avr-2-sens, and located immediately downstream of AUG. The primer sequences were listed in Table 1, Appendix II.
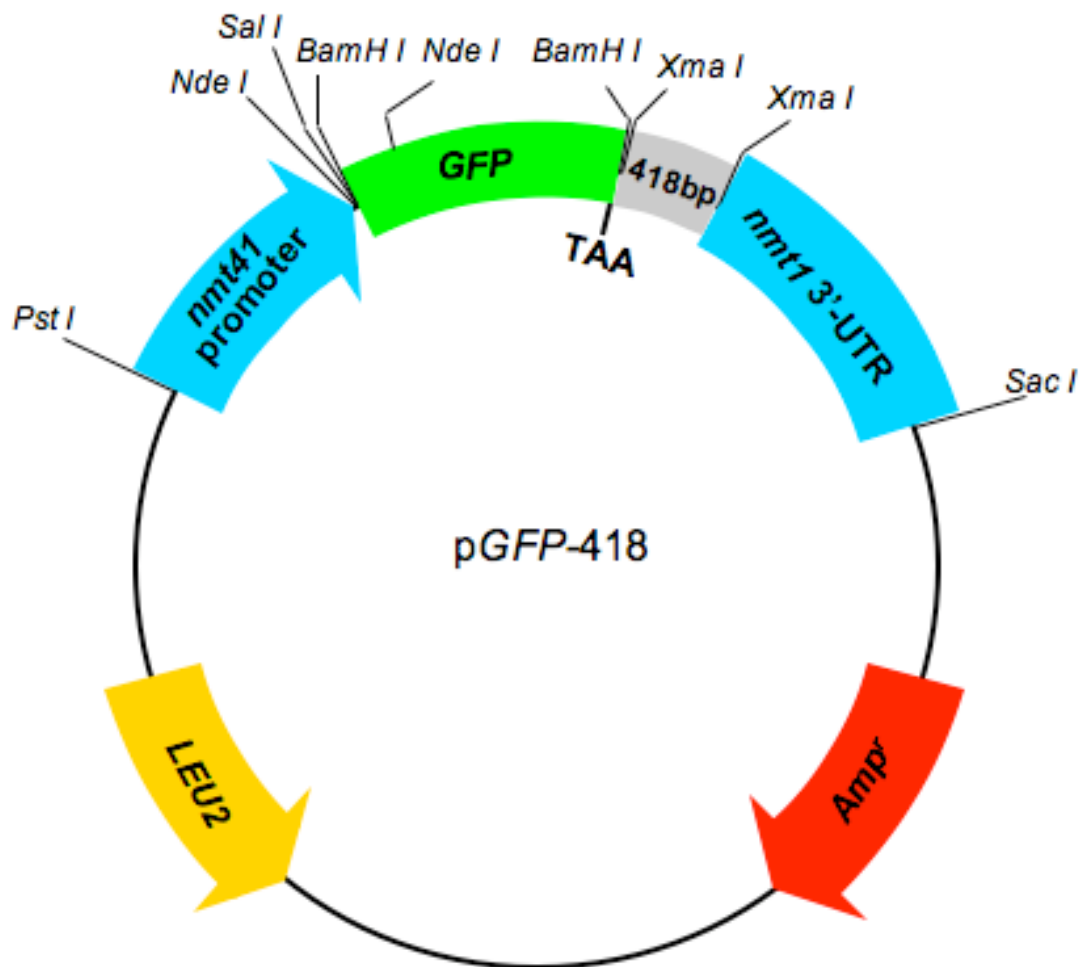
Map 5: p*GFP*-418 (3' UTR lengthening constructs)

A 418 bp of DNA fragments was PCR amplified from firefly luciferase coding region
with primers luc102-sens, luc102-rev and luc418-rev, and cloned into the *Xma I* site
of the intronless NMD reporters (pREP-GFP or derivatives, map 3). The *Xma l* is
located 9 nt downstream of the GFP normal stop codon. The primer sequences are
shown in Table 1, Appendix II.
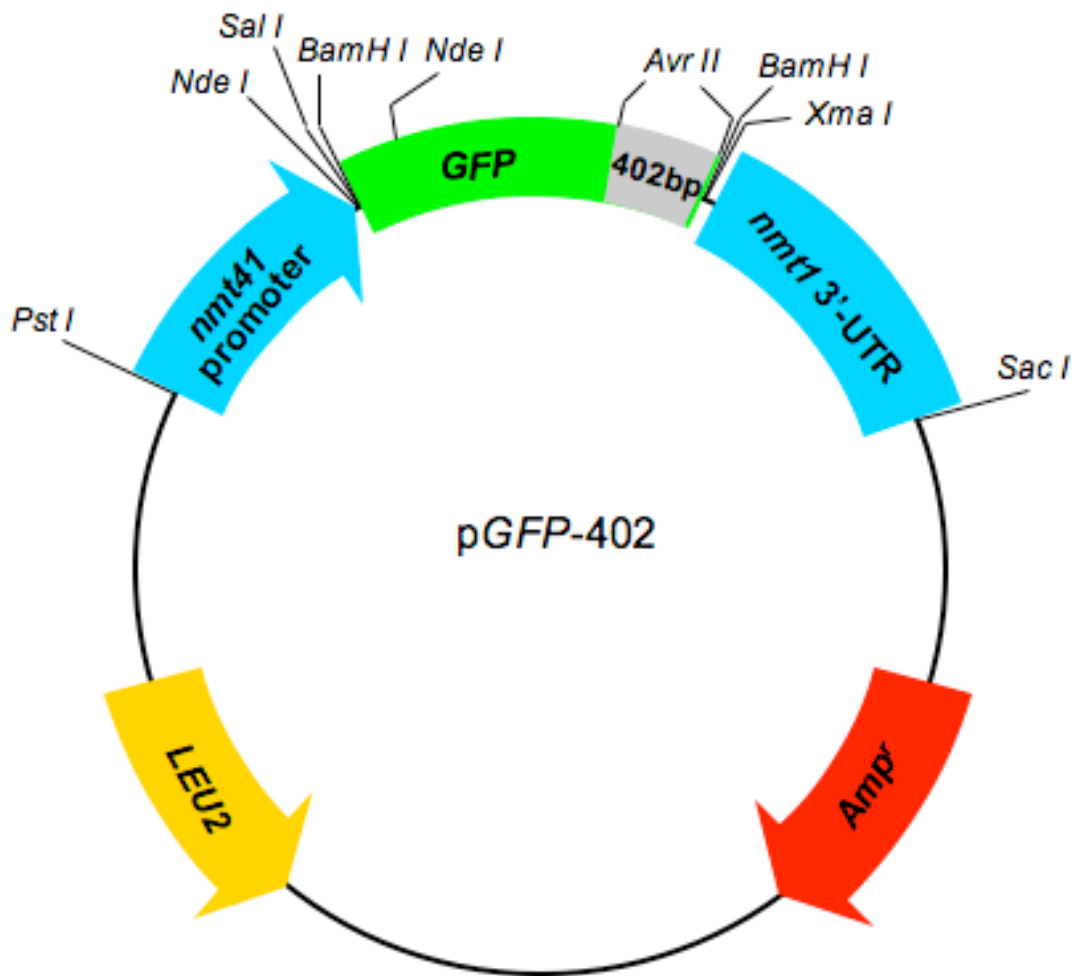
Map 6: p*GFP*-402 (PTC-3' end lengthening constructs)

A 402 bp DNA fragment was amplified from EGFP coding region with primers InHAeG-sens and HAeG402-rev and cloned into an *Avr II* site introduced in intronless constructs. The *Avr II* site was introduced by PCR the GFP coding region with primers GAvr243-rev and gfp-ks immediately upstream of the normal stop codon. The primer sequences are listed in Table 1, Appendix II.
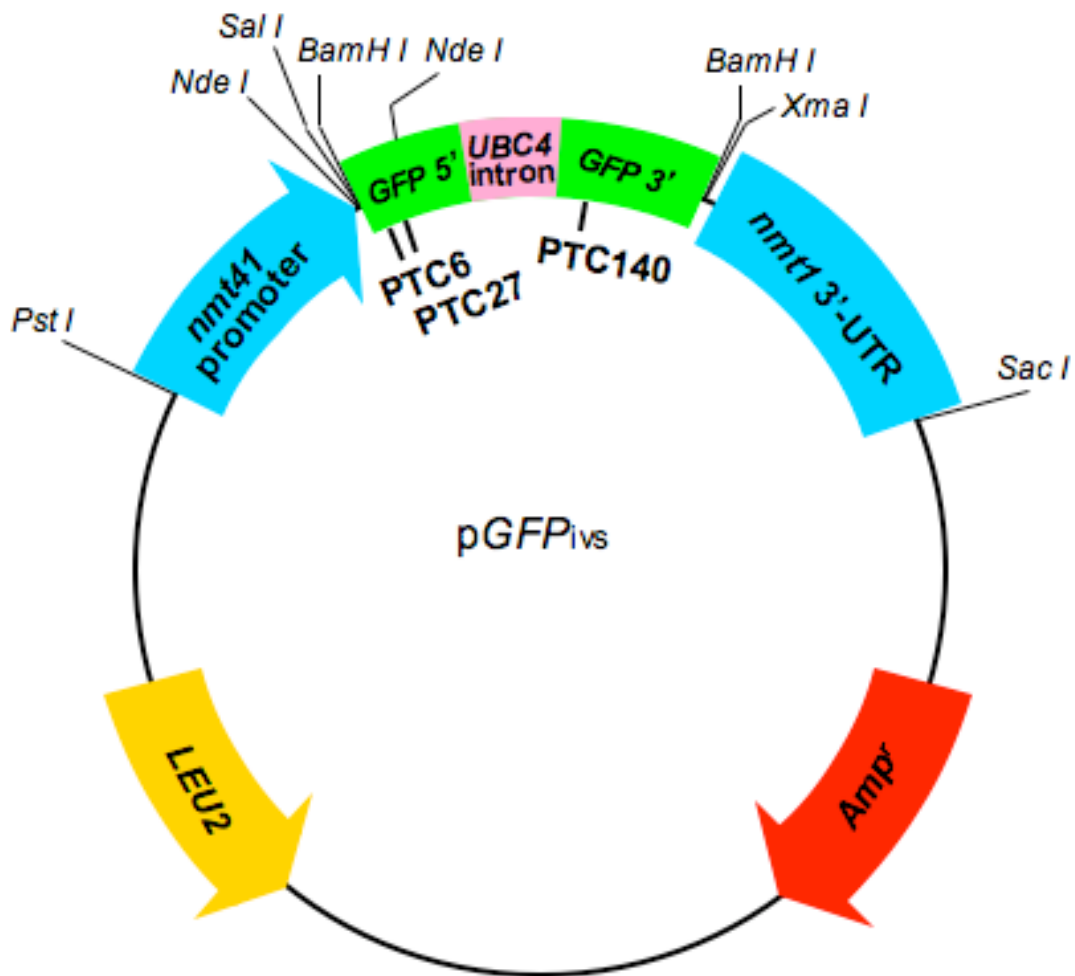
Map 7: p*GFP*ivs (intron-containing NMD reporters)

UBC4 intron was amplified with the primers INubc4 sens and INubc4 rev, and cloned into *Pml I* site then split GFP coding region at triplets position 110. GFP with intron were cloned into pREP41 and three nonsense mutations were introduced by overlap PCR as described in Materials and Methods, and Map 3. The primer sequences are listed in Table 1, Appendix II.
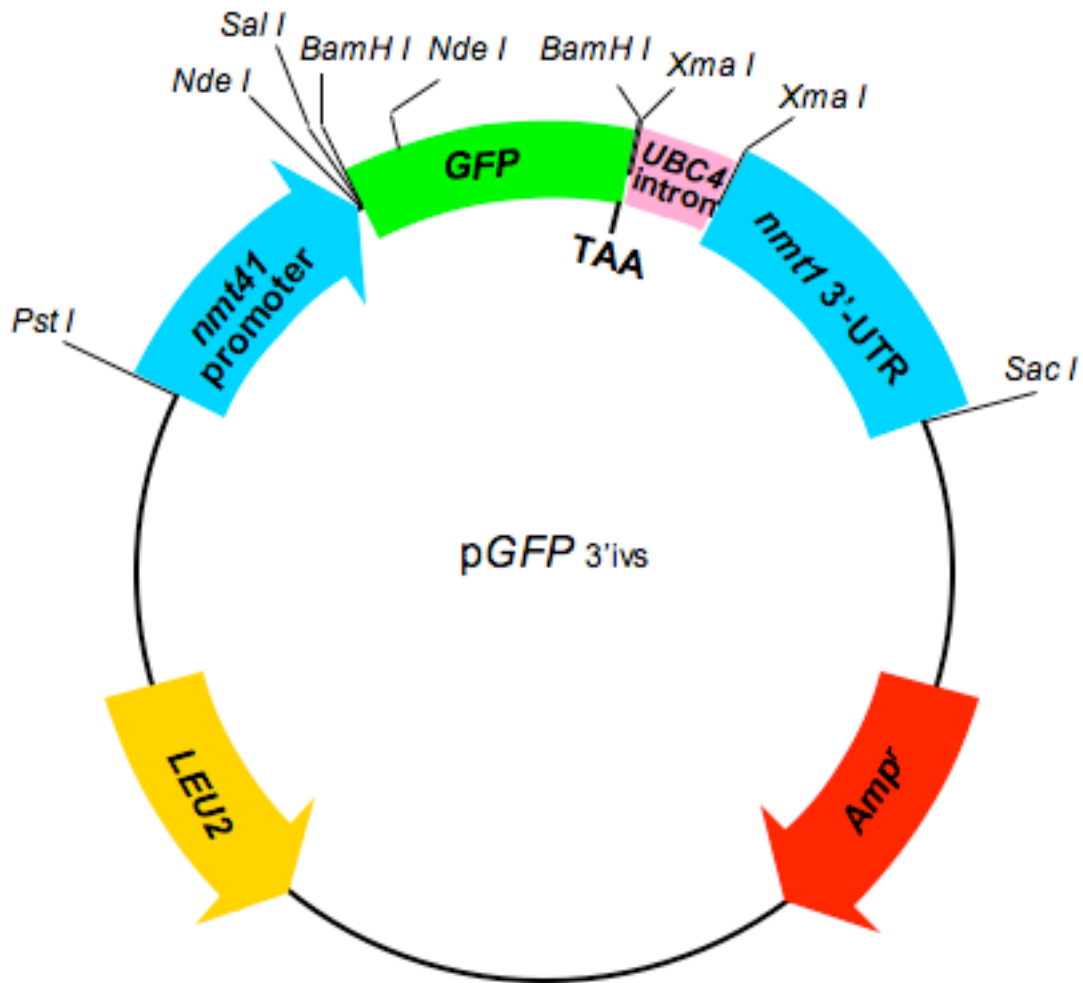
Map 8: p*GFP* 3'ivs (3' UTR intron reporters)

UBC4 intron was amplified with the primers IN3UTR-sens and IN3UTR-rev, and cloned into *Xma I* site of intronless NMD reporters. Then the intron is located 12 nt downstream of the GFP normal stop codon. The primer sequences are listed in Table 1, Appendix II.
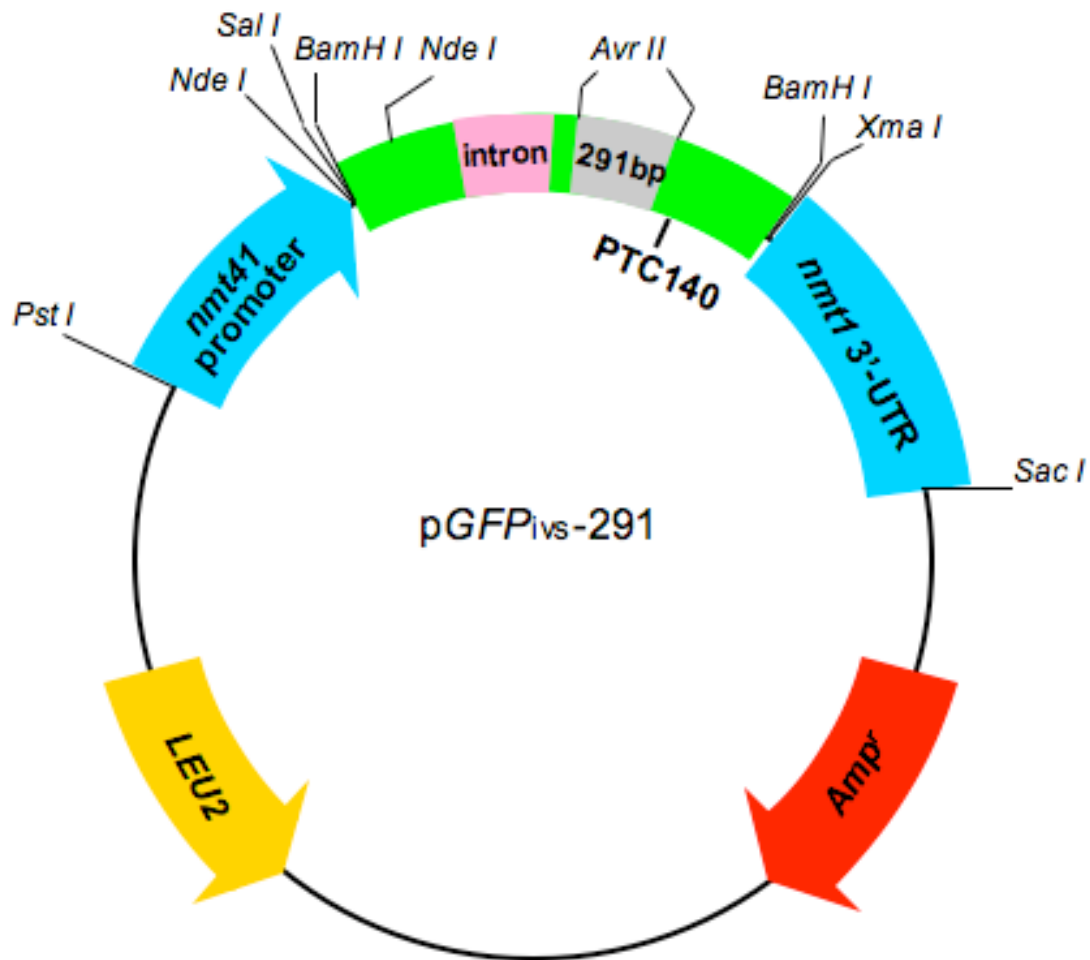
Map 9: p*GFP*ivs-291 (PTC-intron lengthening constructs)

The distance between the PTC and splicing was extended by inserting two different lengths of DNA fragments, 147 bp and 291 bp, which were amplified by PCR with the primers InHAeG-sens, InHAeGS-rev and InHAeGL-rev and cloned into *Avr II* site at triplets position 120 in the GFP coding region. The primer sequences are listed in Table 1, Appendix II.
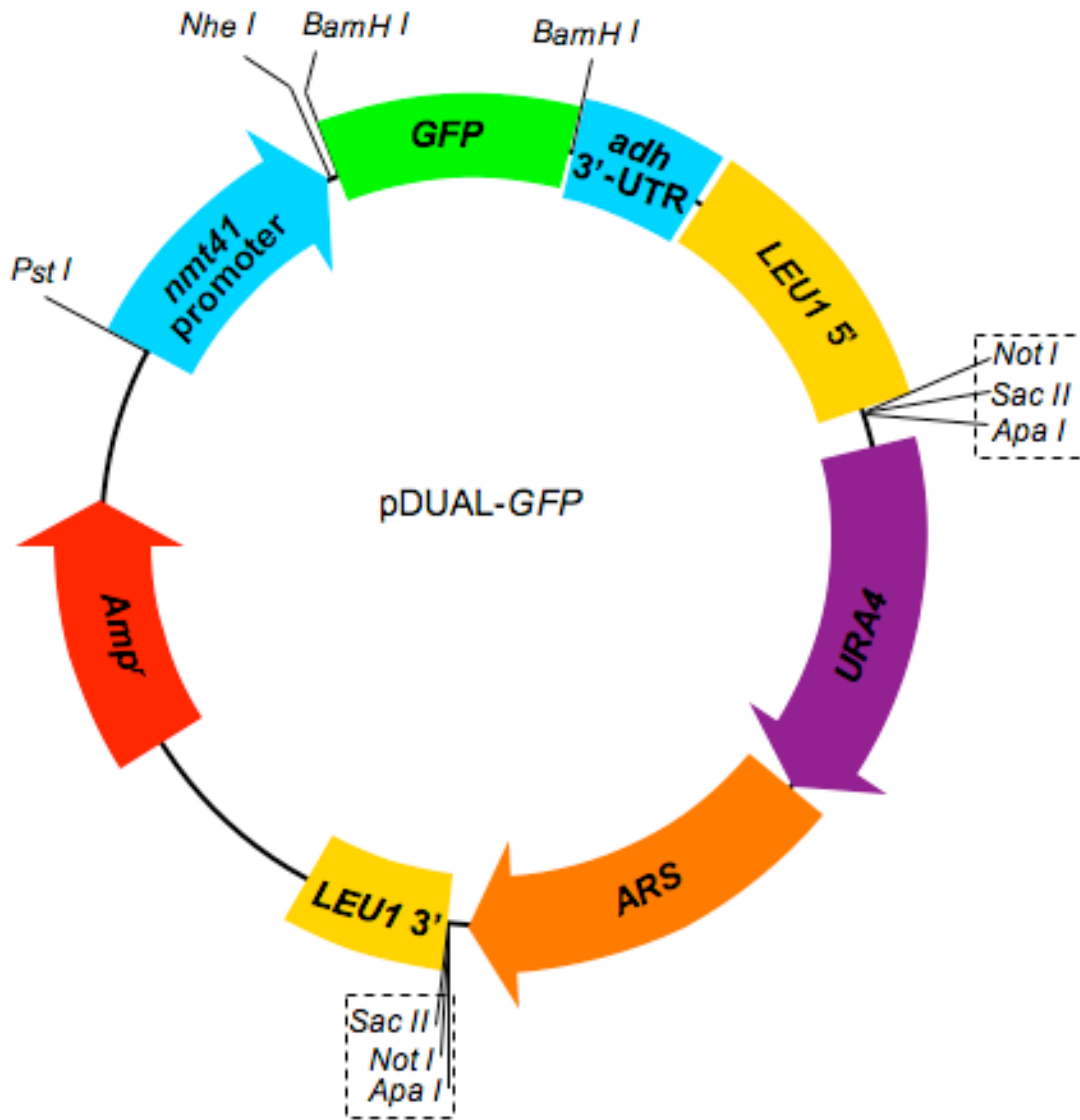
Map 10: pDUAL-*GFP* (integrated NMD reporters)

GFP reporters with/without intron were cloned to *BamH I* site of pDUAL-*nmt41* to make the integration of GFP reporters into *S. pombe* genome.
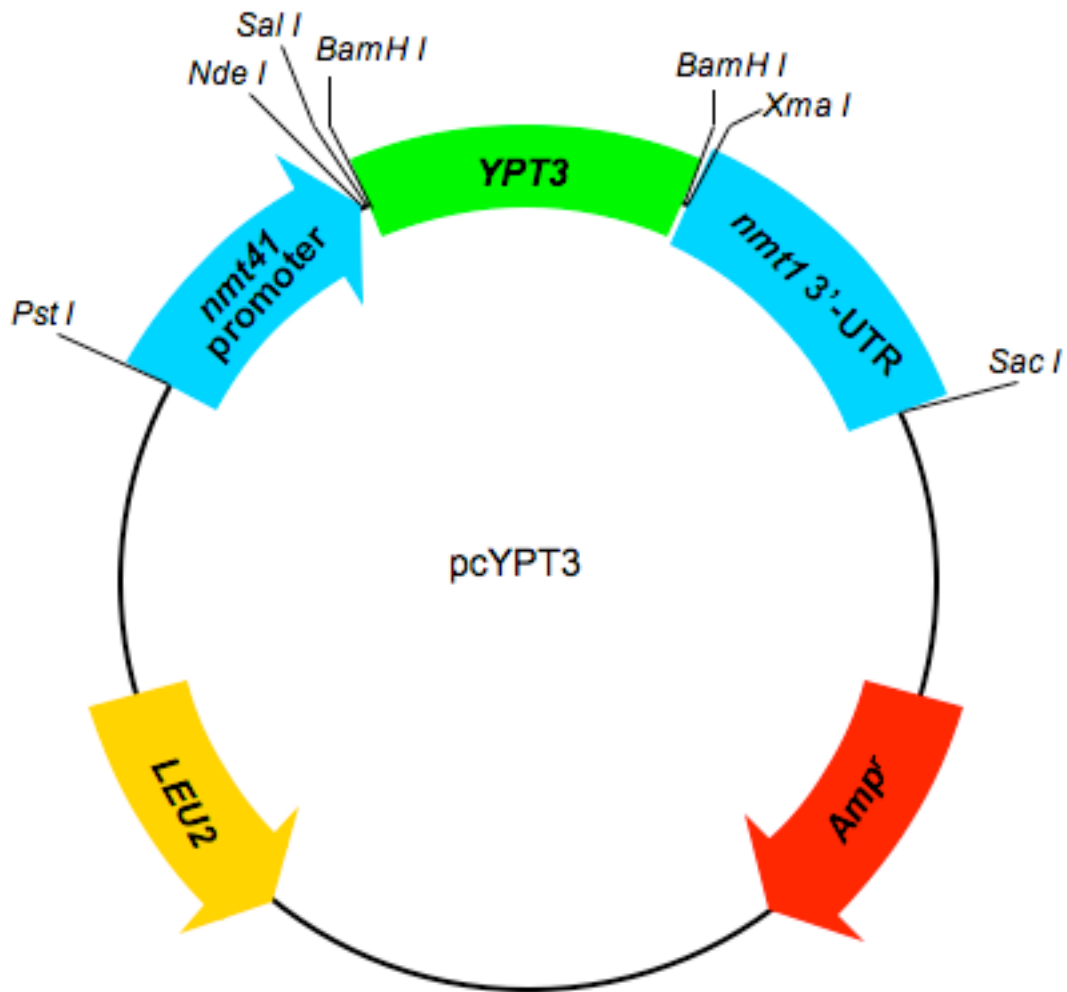
Map 11: pcYPT3 (endogenous NMD reporters)

The endogenous YPT3 and its cDNA were amplified with primers ypt3-sens and ypt3-rev and cloned into the *BamH I* site of pREP41. Then the nonsense mutations were introduced by PCR with the primers ypt3N6-sens (PTC6), ypt3N130-sens and ypt3N130-rev (PTC130). The primer sequences are listed in Table 1, Appendix II.

References

Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., 3rd, 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*, 943-51.

Basi, G., Schmid, E., and Maundrell, K. (1993). TATA box mutations in the Schizosaccharomyces pombe nmt1 promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. Gene *123*, 131-6.

Decottignies, A., Sanchez-Perez, I., and Nurse, P. (2003). Schizosaccharomyces pombe essential genes: a pilot study. Genome Res *13*, 399-406.

Dower, K., and Rosbash, M. (2002). T7 RNA polymerase-directed transcripts are processed in yeast and link 3' end formation to mRNA nuclear export. RNA *8*, 686-97.

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*, 1289-305.

Maundrell, K. (1993). Thiamine-repressible expression vectors pREP and pRIP for fission yeast. Gene *123*, 127-30.

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*, 8944-57.

Szankasi, P., and Smith, G.R. (1996). Requirement of S. pombe exonuclease II, a homologue of S. cerevisiae Sep1, for normal mitotic growth and viability. Curr Genet *30*, 284-93.

Thakurta, A.G., Ho Yoon, J., and Dhar, R. (2002). Schizosaccharomyces pombe spPABP, a homologue of Saccharomyces cerevisiae Pab1p, is a non-essential, shuttling protein that facilitates mRNA export. Yeast *19*, 803-10.